Aims and Scope

Neurospine provides spine clinicians and researchers with peer-reviewed articles on basic and clinical investigation of spine and spinal cord to enhance patient management, education, clinical or experimental research, and professionalism. The journal will consider submissions in areas on craniocervical to lumbosacral spine including the followings: neuroscience and pain research, bone and mineral research, disc and joint research, bio and industrial technology, pathophysiology, risk factors, symptomatology, imaging, treatment, rehabilitation of spine, spinal cord and peripheral nerve diseases. Specifically, basic and technology researches include the most influential research papers from all fields of science and technology, revolutionizing what physicians and researchers practicing the art of spinal neurosurgery worldwide know. Thus, we welcome valuable basic and translational technology research articles to introduce cutting-edge research of fundamental sciences and technology in clinical spinal neurosurgery. Clinical or basic research articles, review articles, case reports, technical notes, and letters to the editor written in English will be accepted.

About the Journal

Neurospine (ISO abbreviated journal name, Neurospine), the official journal of World federation of Neurosurgical Societies Spine Committee, is an international peer-reviewed open-access journal which published quarterly (last day of March, June, September, and December). It was first published in March 31, 2004 with Volume 1 and Number 1 with the name “Korean Journal of Spine,” and renamed as “Neurospine” in March issue, 2018.

Indexed In

Neurospine is indexed/tracked/covered by Science Citation Index Expanded (SCIE), PubMed, PubMed Central, Scopus, Directory of Open Access Journals (DOAJ), KoreaMed, CrossRef, KoMCI, and Google Scholar

Subscription Information

The Korean Spinal Neurosurgery Society will send Neurospine to certain relevant individuals and institutions free of charge. Full text is complimentary and available at http://www.e-neurospine.org. To order a subscription to Neurospine, please contact our editorial office.

Open Access

Neurospine is an open access peer-reviewed journal and launched in March 31, 2004. All articles published in Neurospine will be immediately and permanently free for everyone to read and download. All articles are distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Contact Neurospine

Editor-in-Chief Inbo Han

Publisher Korean Spinal Neurosurgery Society
Published by the Korean Spinal Neurosurgery Society
nd #407, Dong-A Villate 2 Town, 350, Seocho-daero, Seocho-gu, Seoul 06631, Korea
Tel: +82-2-585-5455 • Fax: +82-2-523-6812 • Email: ksns1987@gmail.com

Editorial Office
Department of Neurosurgery, CHA Bundang Medical Center, CHA University School of Medicine, 59 Yatap-ro, Bundang-gu, Seongnam 13496, Korea
Tel: +82-31-780-1924 • Fax: +82-31-780-5269 • Email: support@e-neurospine.org

Printed by Academya Publishing Co.
Room 2003, Daerung Techno Town 15-Cha, 401 Simin-daero, Dongan-gu, Anyang 14057, Korea
Tel: +82-31-389-8811 • Fax: +82-31-389-8817 • Email: journal@academya.co.kr • http://www.academya.co.kr

This work was supported by the Korean Federation of Science and Technology Societies (KOFST) grant funded by the Korean government.

© Copyright 2024 by the Korean Spinal Neurosurgery Society. All rights reserved.
Editorial Board

Seung Woo Suh  Korea University, Korea
Siravich Suvithayasiri  Chulabhorn Royal Academy, Thailand
Shih-Huang Tai  National Cheng Kung University Hospital, Taiwan
Keisuke Takai  Tokyo Metropolitan Neurological Hospital, Japan
Toshihiro Takami  Osaka City University Graduate School of Medicine, Japan
Lee Tan  University of California, USA
Fan Tao  Capital Medical University, China
Claudio E. Tatsui  The University of Texas MD Anderson Cancer Center, USA
Zafer Orkun Toktaş  Uskudar University, Türkiye
Tsung-Hsi Tu  Taipei Veterans General Hospital, Taiwan

Gianluca Vadalà  Campus Bio Medico University of Rome, Italy
Kota Watanabe  Keio University, Japan
Meng-Huang Wu  Taipei Medical University, Taiwan
Karim Wüertz-Kozak  Institute for Biomechanics, Switzerland
Satoshi Yamaguchi  University of low a Carver College of Medicine, Japan
Ei Yamamoto  Kindai University, Japan
Takao Yasuhara  Okayama University Graduate School of Medicine, Japan
Woon Tak Yuh  Hallym University, Korea
Takashi Yurube  Kobe University, Japan

Managing Editors

Seong Bae An  CHA University, Korea
David Anderson  The University of Sydney, Australia
Thomas J. Buell  Duke University, USA
Timo Michael Ecker  Lausanne University, Switzerland
Jung-Woo Hur  Catholic University, Korea
Junseok W. Hur  Korea University, Korea
Hyun Jun Jang  Yonsei University, Korea
Pius Kim  Chosun University, Korea
Sung Bum Kim  Kyung Hee University, Korea
Hemant Kumar  National Institute of Pharmaceutical Education and Research-Ahmedabad (NIPER-A), India
Chang Kyu Lee  Yonsei University, Korea

Assistant Editors

Seong Bae An  CHA University, Korea
David Anderson  The University of Sydney, Australia
Thomas J. Buell  Duke University, USA
Timo Michael Ecker  Lausanne University, Switzerland
Jung-Woo Hur  Catholic University, Korea
Junseok W. Hur  Korea University, Korea
Hyun Jun Jang  Yonsei University, Korea
Pius Kim  Chosun University, Korea
Sandro M. Krieg  Technical University of Munich, Germany
Jong Beom Lee  Chungbuk National University, Korea
Seung Jun Ryu  Eulji University, Korea
Michael Schwake  university Hospital Münster, Germany
Victor Egon Staartjes  University of Zurich, Switzerland
Martin N. Stienen  University Hospital Zurich, Switzerland
Corinna C. Zygourakis  Stanford University, USA

Statistical Editor

Eunyoung Lee  University of Texas, USA

Ethics Editors

Sun Huh  Hallym University, Korea
Young Yoo  Korea University, Korea

Social Media Editors

Peng-Yuan Chang  Taipei Veterans General Hospital, Taiwan
Suhun Lee  Pusan National University, Korea
Ikchan Jeon  Yeungnam University, Korea
John Shin  Harvard Medical School, Boston, USA

Manuscript Editor

Hwan Tae Ahn  JTS, Korea
Contents

Editorial

373  From the Editor-in-Chief: Featured Articles in the June 2024 Issue
    Inbo Han

Review Article


Commentary

    Karlo M. Pedro, Michael G. Fehlings

Original Article

404  Baseline Frailty Measured by the Risk Analysis Index and 30-Day Mortality After Surgery for Spinal Malignancy: Analysis of a Prospective Registry (2011–2020)
    Rachel Thommen, Christian A. Bowers, Aaron C. Segura, Joanna M. Roy, Meic H. Schmidt

Commentary

414  Commentary on “Baseline Frailty Measured by the Risk Analysis Index and 30-Day Mortality After Surgery for Spinal Malignancy: Analysis of a Prospective Registry (2011–2020)”
    Fon-Yih Tsuang

Original Article

416  Comparison of the Clinical Efficacy of Anabolic Agents and Bisphosphonates in the Patients With Osteoporotic Vertebral Fracture: Systematic Review and Meta-analysis of Randomized Controlled Trials
    Ikchan Jeon, Sung Bae Park, Bong Ju Moon, Miyoung Choi, Sung Uk Kuh, Jongtae Kim

Commentary

430  Commentary on “Comparison of the Clinical Efficacy of Anabolic Agents and Bisphosphonates in the Patients With Osteoporotic Vertebral Fracture: Systematic Review and Meta-analysis of Randomized Controlled Trials”
    Dae-Chul Cho
Original Article

432 The Utility and Feasibility of Smart Glasses in Spine Surgery: Minimizing Radiation Exposure During Percutaneous Pedicle Screw Insertion
Yoshiaki Hiranaka, Yoshiki Takeoka, Takashi Yurube, Takeru Tsujimoto, Yutaro Kanda, Kunihiko Miyazaki, Hiroki Ohnishi, Tomoya Matsuho, Masao Ryu, Naotsuki Kumagai, Koji Kuroshima, Ryosuke Kuroda, Kenichiro Kadakani

Commentary

440 Commentary on “The Utility and Feasibility of Smart Glasses in Spine Surgery: Minimizing Radiation Exposure During Percutaneous Pedicle Screw Insertion”
Wonghawat Liawrungrueang

Original Article

443 Radiological and Clinical Significance of Cervical Dynamic Magnetic Resonance Imaging for Cervical Spondylotic Myelopathy
Jun Jae Shin, Sun Joon Yoo, Tae Woo Kim, Jae-Young So, Won Joo Jeong, Mu Ha Lee, Joongkyun Shin, Yoon Ha

Commentary

455 Commentary on “Radiological and Clinical Significance of Cervical Dynamic Magnetic Resonance Imaging for Cervical Spondylotic Myelopathy”
Aydin Sinan Apaydin, Khoi Than

Review Articles

458 Assessing the Fractional Curve for Proper Management of Adult Degenerative Scoliosis

474 Artificial Intelligence in Spinal Imaging and Patient Care: A Review of Recent Advances
Sungwon Lee, Joon Yong Jung, Akaworn ahatthanatrakul, Jin-Sung Kim

487 Comparative Review of the Socioeconomic Burden of Lower Back Pain in the United States and Globally
Diana Chang, Austin Lui, Alisa Matsoyan, Michael M. Safaee, Henry Aryan, Christopher Ames

Original Articles

502 Radiographic and Clinical Outcomes of Transverse Process Hook Placement at the Proximal Thoracic Upper Instrumented Vertebra in Adult Spinal Deformity Surgery
Sang Hun Lee, Micheal Raad, Andrew H. Kim, David B. Cohen, Khaled M. Kebaish
510  Magnetic Resonance Imaging-Related Anatomic and Functional Parameters for the Diagnosis and Prognosis of Chiari Malformation Type I: A Systematic Review and Meta-analysis
Zairan Wang, Zhimin Li, Shiyuan Han, Xianghui Hu, Siyuan Pang, Yongning Li, Jun Gao

525  Are There Advantages in Cervical Intrafacetal Fusion With Minimal Posterolateral Fusion (PLF) Compared to Conventional PLF in Posterior Cervical Fusion?
Sun Woo Jang, Sang Hyub Lee, Jeong Kyun Joo, Hong Kyung Shin, Jin Hoon Park, Sung Woo Roh, Sang Ryong Jeon

536  Finite Element Analysis of Stress Distribution and Range of Motion in Discogenic Back Pain
Pyung Goo Cho, Seon Jin Yoon, Dong Ah Shin, Min Cheol Chang

544  Comparison of Transoral Anterior Jefferson-Fracture Reduction Plate and Posterior Screw-Rod Fixation in C1-Ring Osteosynthesis for Unstable Atlas Fractures
Mandi Cai, Yifeng Wu, Rencai Ma, Junlin Chen, Zexing Chen, Chenfu Deng, Xinzhao Huang, Xiangyang Ma, Xiaobao Zou

555  Knockdown of best1 Gene in Zebrafish Caused Abnormal Neuronal and Skeletal Development - A Subtype of Craniovertebral Junction Malformation?
Zhenlei Liu, Kang Li, Kai Wang, Lei Zhang, Shanhang Jia, He Wang, Fengzeng Jian, Hao Wu

565  Clinical and Radiological Outcomes in C2 Recapping Laminoplasty for the Pathologies in the Upper Cervical Spine
Dong Hun Kim, Jae Taek Hong, Jung Woo Hur, Il Sup Kim, Ho Jin Lee, Jong Beom Lee

575  A Comparative Study of 2 Techniques to Avoid Bone Cement Loosening and Displacement After Percutaneous Vertebroplasty Treating Unstable Kummell Disease
Jie Guo, Yesheng Bai, Liang Li, Jiangtao Wang, Yuhang Wang, Dingjun Hao, Biao Wang

588  Incidence and Survival of Patients With Malignant Primary Spinal Cord Tumors: A Population-Based Analysis
Huanbing Liu, Linnan Duan, Zhibin Li, Yuanhao Liu, Yubo Wang

596  Predictors of Persistent Postoperative Numbness Following Lumbar Fusion in Patients Older Than 75 Years: A Minimum 2-Year Follow-up
Takeru Tsujimoto, Masahiro Kanayama, Shotaro Fukada, Fumihiro Oha, Yukitoshi Shimamura, Yuichi Hasegawa, Tomoyuki Hashimoto, Kenichiro Kakutani, Takashi Yurube, Yoshiki Takeoka, Kunihiko Miyazaki, Norimasa Iwasaki

606  Biomechanical Analysis of Hybrid Artificial Discs or Zero-Profile Devices for Treating 1-Level Adjacent Segment Degeneration in ACDF Revision Surgery
Weishi Liang, Yihan Yang, Bo Han, Duan Sun, Peng Yin, Yong Hai

620  Using Machine Learning Models to Identify Factors Associated With 30-Day Readmissions After Posterior Cervical Fusions: A Longitudinal Cohort Study
Analyzing Large Language Models’ Responses to Common Lumbar Spine Fusion Surgery Questions: A Comparison Between ChatGPT and Bard
Siegmund Philipp Lang, Ezra Tilahun Joseph, Anayesia D. Gonzalez-Suarez, Robert Kim, Parastou Fatemi, Katherine Wagner, Nicolai Maldaner, Martin N. Stienen, Corinna Clio Zygourakis

Inhibition of Ferroptosis by Mesenchymal Stem Cell-Derived Exosomes in Acute Spinal Cord Injury: Role of Nrf2/GCH1/BH4 Axis
Yixin Chen, Bingfa Li, Jing Quan, Zhe Li, Yan Li, Yinbo Tang

Comparison of Single or Double Titanium Mesh Cage for Anterior Reconstruction After Total En Bloc Spondylectomy for Thoracic and Lumbar Spinal Tumors
Ao Leng, Qi Wang, Jiacong Li, Yu Long, Song Shi, Lingzhi Meng, Mingming Guo, Haolong Yu, Liangbi Xiang

The Quantitative Evaluation of Automatic Segmentation in Lumbar Magnetic Resonance Images
Yao-Wen Liang, Yu-Ting Fang, Ting-Chun Lin, Cheng-Ru Yang, Chih-Chang Chang, Hsuan-Kan Chang, Chin-Chi Ko, Toung-Hsi Tu, Li-Yu Fay, Jau-Ching Wu, Wen-Cheng Huang, Hsiang-Wei Hu, You-Yin Chen, Chao-Hung Kuo

A Nomogram for Predicting Overall Survival of Patients With Primary Spinal Cord Glioblastoma
Yao Wang, Qingchun Mu, Minfeng Sheng, Yanming Chen, Fengzeng Jian, Rujun Li

A Comparative Factor Analysis and New Magnetic Resonance Imaging Scoring System for Differentiating Pyogenic Versus Tuberculous Spondylodiscitis
Terdpong Tanaviriyachai, Patchara Pornsopunakorn, Kongtush Choowongkornol, Tada Virathepsuporn, Urawit Piyapromdee, Sarut Jongkittanakul, Weera Sudprasert, Siricha Wirawatpanakul

The Role of Spinal Cord Compression in Predicting Intraoperative Neuropsychological Monitoring Events in Patients With Kyphotic Deformity: A Magnetic Resonance Imaging-Based Study
Zhen Jin, Jie Li, Hui Xu, Zongshan Hu, Yanjie Xu, Ziyang Tang, Yong Qiu, Zhen Liu, Zexiang Zhu

Prediction of Screw Loosening After Dynamic Pedicle Screw Fixation With Lumbar Polyetheretherketone Rods Using Magnetic Resonance Imaging-Based Vertebral Bone Quality Score
Guozheng Jiang, Luchun Xu, Yukun Ma, Jianbin Guan, Yonglong Yang, Wenging Zhong, Wenhao Li, Shibao Zhou, JiaWei Song, Ningning Feng, Ziyu Qiu, Zeyu Li, YiShui Zhou, Letian Meng, Yi Qiu, Xing Yu

Comparison of Surgical Burden, Radiographic and Clinical Outcomes According to the Severity of Baseline Sagittal Imbalance in Adult Spinal Deformity Patients
Se-Jun Park, Jin-Sung Park, Dong-Ho Kang, Hyun-Jun Kim, Yun-Mi Lim, Chong-Suh Lee

Preliminary Clinical and Radiological Outcomes of the “No-Punch” Decompression Techniques for Unilateral Biportal Endoscopic Spine Surgery
Jwo-Luen Pao
Erratum

742  Erratum: Exploring IncRNA Expression Patterns in Patients With Hypertrophied Ligamentum Flavum
  Junling Chen, Guibin Zhong, Manle Qiu, Wei Ke, Jingsong Xue, Jianwei Chen
From the Editor-in-Chief: Featured Articles in the June 2024 Issue

Inbo Han
Department of Neurosurgery, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, Korea

In this editorial, we highlight the significant contributions of selected articles published in the June 2024 issue of *Neurospine*. These articles represent advancements in spine research, offering valuable insights into various aspects of spinal health, treatment modalities, and clinical outcomes.


Levett et al. [1] conducted a comprehensive review of the pharmacological treatment of degenerative cervical myelopathy (DCM). They synthesized clinical and experimental evidence on the efficacy of pharmacological agents in both humans and animal models. Despite some promising findings in animal studies, the translation to clinical practice remains challenging. The review underscores the potential for developing neuroprotective therapies to delay surgical intervention and optimize outcomes for patients with DCM.

**Article 2: “Baseline Frailty Measured by the Risk Analysis Index and 30-Day Mortality After Surgery for Spinal Malignancy: Analysis of a Prospective Registry (2011–2020)”**

Thommen et al. [2] analyzed data from a prospective surgical registry to evaluate the prognostic utility of baseline frailty, measured by the Risk Analysis Index, in predicting postoperative mortality among patients undergoing surgery for spinal malignancy. Their findings highlight the importance of preoperative frailty assessment in risk stratification and clinical decision-making, offering insights for improving patient outcomes in this challenging patient population.

**Article 3: “Comparison of the Clinical Efficacy of Anabolic Agents and Bisphosphonates in the Patients With Osteoporotic Vertebral Fracture: Systematic Review and Meta-analysis of Randomized Controlled Trials”**

Jeon et al. [3] conducted a systematic review and meta-analysis to compare the clinical efficacy of anabolic agents and bisphosphonates in patients with osteoporotic vertebral fractures (OVF). Their findings support the superiority of anabolic agents in preventing new OVF compared to bisphosphonates. However, they emphasize the need for additional large-scale studies to optimize treatment strategies for this patient population.
Article 4: “The Utility and Feasibility of Smart Glasses in Spine Surgery: Minimizing Radiation Exposure During Percutaneous Pedicle Screw Insertion”

Hiranaka et al. [4] investigated the utility and feasibility of smart glasses in spine surgery to minimize radiation exposure during percutaneous pedicle screw insertion. Their findings demonstrate the effectiveness of smart glasses in reducing radiation exposure without compromising procedural efficiency or accuracy, offering a promising approach to improving safety for spine surgeons.

Article 5: “Radiological and Clinical Significance of Cervical Dynamic Magnetic Resonance Imaging for Cervical Spondylotic Myelopathy”

Shin et al. [5] conducted a comprehensive study on the radiological and clinical significance of cervical dynamic magnetic resonance imaging (MRI) in patients with cervical spondylotic myelopathy. Their findings highlight the utility of dynamic MRI in evaluating cervical spinal motion and its correlation with clinical outcomes, providing valuable insights for optimizing treatment strategies and improving patient outcomes.

- **Conflict of Interest:** The author has nothing to disclose.

REFERENCES


Degenerative cervical myelopathy (DCM) is the leading cause of spinal cord dysfunction in adults, representing substantial morbidity and significant financial and resource burdens. Typically, patients with progressive DCM will eventually receive surgical treatment. Nonetheless, despite advancements in pharmacotherapeutics, evidence for pharmacological therapy remains limited. Health professionals from various fields would find interest in pharmacological agents that could benefit patients with mild DCM or enhance surgical outcomes. This review aims to consolidate all clinical and experimental evidence on the pharmacological treatment of DCM. We conducted a comprehensive narrative review that presents all pharmacological agents that have been investigated for DCM treatment in both humans and animal models. Riluzole exhibits effectiveness solely in rat models, but not in treating mild DCM in humans. Cerebrolysin emerges as a potential neuroprotective agent for myelopathy in animals but had contradictory results in clinical trials. Limaprost alfadex demonstrates motor function improvement in animal models and exhibits promising outcomes in a small clinical trial. Glucocorticoids not only fail to provide clinical benefits but may also lead to adverse events. Cilostazol, anti-Fas ligand antibody, and Jingshu Keli display promise in animal studies, while erythropoietin, granulocyte colony-stimulating factor and limaprost alfadex exhibit potential in both animal and human research. Existing evidence mainly rests on weak clinical data and animal experimentation. Current pharmacological efforts target ion channels, stem cell differentiation, inflammatory, vascular, and apoptotic pathways. The inherent nature and pathogenesis of DCM offer substantial prospects for developing neurodegenerative or neuroprotective therapies capable of altering disease progression, potentially delaying surgical intervention, and optimizing outcomes for those undergoing surgical decompression.

Keywords: Intervertebral disc degeneration, Cervical cord, Erythropoietin, limaprost-alfadex, Glucocorticoids, Riluzole
INTRODUCTION

Globally, degenerative cervical myelopathy (DCM) is the leading cause of spinal dysfunction in adults over the age of 55. Given an increasingly aging population, and a growing awareness about the condition, global health care systems should anticipate a rise in patients presenting with DCM. Therefore, DCM leads to various significant problems not only for the patients but also from the perspective of healthcare, society, and economy, affecting the quality of life of many patients, many of whom are still in the workforce.

Although the natural history of DCM in individual patients is variable and unpredictable, it is usually characterized by progression of the associated symptoms. However, the disease may demonstrate long quiescent periods, especially in patients with mild DCM. Thus, pharmacological management could be implemented in patients with mild and stable myelopathy or with paraclinical findings associated with the development or worsening of myelopathy to delay or halt the progression of the symptoms. Indeed, optimizing the pharmacologic management of DCM could also be attractive for elderly or high-risk surgical patients or patients who do wish to undergo surgery. As a principle, surgery should ideally become the last resort, whereas currently, it is often considered the main treatment.

Surgical decompression remains the gold standard in the management of moderate to severe cases of DCM that have not responded to conservative management, with the aim of relieving symptoms and preventing further injury. Nevertheless, despite appropriate surgical intervention, some patients may be left with permanent neurological disability due to preoperative irreversible injury to the spinal cord. Additionally, some patients may experience postoperative worsening due to neuroinflammation and reperfusion injury, leading to suboptimal recovery after decompression surgery. Therefore, the development of pharmacotherapy that could complement surgical decompression for DCM and optimize outcomes holds significant clinical importance.

However, despite such clinical potential, the evidence for nonsurgical management or adjuvant pharmacological therapy in DCM remains scarce. Nonetheless, this topic has piqued the interest of researchers from various fields, physicians from different specialties, and a large patient population. The objective of this study is to provide a detailed review of the current evidence regarding the pharmacological therapy of DCM. To our knowledge, this is the only review that comprehensively presents all pharmacological agents investigated for the treatment of DCM in both humans and animals.

METHODS

This narrative review was conducted by searching PubMed, Embase, Scopus, and Cochrane databases from inception to April 2022. The following keywords were used in various combinations: cervical spondylotic myelopathy, CSM, degenerative cervical myelopathy, DCM, spinal cord compression, neuroprotective, regenerative, neuroregenerative, reparative, neuroprotective, drugs, pharmaceutical, substance, medication, nonsteroidal, and steroids. English-language full-text studies were included if they involved a pharmacological component in the management of DCM in humans or animals. We chose the following exclusion criteria: publications that were not peer-reviewed, in vitro studies, articles that were not original studies (e.g., case reports, published congress abstracts, reviews), studies that did not report neurological/functional outcomes, and studies on medications or dietary supplements that were not clearly defined (e.g., active ingredients and their dose). Two authors (MG and MVA) screened all studies independently and in duplicate by title and abstract. Studies meeting all inclusion criteria were subsequently screened by their full-text in a similar fashion. The reference lists of the included articles were cross-referenced to identify additional articles. Data were extracted from the included studies in duplicate (JJL and MG) using a standardized charting template, which included the following information: author/year, study design, aim, population, drug/dosage/administration, outcomes, potential biases/limitations, and level of evidence (using Levels of Evidence for Therapeutic Studies from the Centre for Evidence-Based Medicine). For animal studies, the extracted data included animal model and pathology findings in addition to the aforementioned details. Studies were subsequently aggregated into their respective pharmacological agents, and were screened for dosages, adverse events, significant improvement in Japanese Orthopaedic Association (JOA) score, level of evidence (Table 1)/sample size, and strength of evidence.

RESULTS

1. Riluzole

Moon et al. demonstrated that riluzole administration (8 mg/kg intraperitoneal once a day [QD]) significantly improved gait performance and sensory symptoms in a rat model of DCM induced by a titanium screw-based chronic compression...
Table 1. Levels of evidence for therapeutic studies from the center for evidence-based medicine (http://www.cebm.net)

<table>
<thead>
<tr>
<th>Level</th>
<th>Type of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A</td>
<td>Systematic review (with homogeneity) of RCTs</td>
</tr>
<tr>
<td>1B</td>
<td>Individual RCT (with narrow confidence intervals)</td>
</tr>
<tr>
<td>1C</td>
<td>All or none study</td>
</tr>
<tr>
<td>2A</td>
<td>Systematic review (with homogeneity) of cohort studies</td>
</tr>
<tr>
<td>2B</td>
<td>Individual cohort study (including low quality RCT, e.g., &lt; 80% follow-up)</td>
</tr>
<tr>
<td>2C</td>
<td>“Outcomes” research; ecological studies</td>
</tr>
<tr>
<td>3A</td>
<td>Systematic review (with homogeneity) of case-control studies</td>
</tr>
<tr>
<td>3B</td>
<td>Individual case-control study</td>
</tr>
<tr>
<td>4</td>
<td>Case series and poor quality cohort and case-control study</td>
</tr>
<tr>
<td>5</td>
<td>Expert opinion without explicit critical appraisal or based on physiology bench research or “first principles”</td>
</tr>
</tbody>
</table>

RCT, randomized controlled trials.

sion device (p < 0.05). Karadimas et al.\(^\text{14}\) examined the effects of postoperative riluzole administration in a rodent model of DCM. Rats with DCM demonstrated a transient postoperative neurological deterioration similar to what is observed in some patients, suggesting that ischemia-reperfusion injury may occur after decompression surgery. Riluzole administration demonstrated a decrease in oxidative stress and postoperative decline of gait parameters after decompression surgery in the DCM rat model. Riluzole-treated rats displayed a significantly lower proportion of 8-oxoG DNA-positive cells (indicating oxidative damage) in vitro (p < 0.05).

Rajasekaran et al.\(^\text{17}\) conducted a clinical trial to evaluate the effectiveness of riluzole administration in patients with mild forms of DCM. Thirty patients with modified JOA (mJOA) scores ≥ 13 were recruited for this double-blinded, placebo-controlled randomized controlled trial. The study group was administered riluzole (50 mg orally twice a day [PO BID]) for a period of one month while the placebo group was administered vitamin B complex tablets. The patients were assessed with a new magnetic resonance imaging (MRI) and diffusion tensor imaging and clinical scores. There was no significant change in the clinical outcome scores and diffusion tensor indices of patients treated with riluzole as a standalone pharmacotherapeutic agent after one month. The strength of evidence\(^\text{14}\) of this study was considered moderate.

The most significant clinical trial investigating the efficacy of perioperative riluzole administration is the CSM-PROTECT trial.\(^\text{18}\) This international, double-blinded, randomized phase 3 clinical trial was conducted by 16 university-affiliated centers in Canada and the United States. Patients aged 18 to 80 years with moderate to severe DCM as characterized by mJOA scores of 8 to 14 were included in this trial. A total of 290 patients were randomized. Patients randomized to the study group were administered riluzole (50 mg PO BID) 14 days preoperatively and 28 days postoperatively. The primary endpoint was change in mJOA score from baseline to 6 months in the intention-to-treat population, defined as all patients who have undergone randomization and surgical decompression. There was no significant difference between the riluzole and placebo groups in mJOA scores at 6 months of follow-up (p = 0.14). Therefore, adjuvant treatment with riluzole in the perioperative setting did not improve functional recovery beyond decompression surgery alone for patients with moderate to severe DCM in this study.\(^\text{18}\) The strength of evidence of this study was considered moderate.

2. Cerebrolysin

Allam et al.\(^\text{19}\) conducted a prospective randomized control trial to evaluate the effect of cerebrolysin as a treatment modality for DCM. A group of 192 patients with moderate to severe DCM who refused surgery were subdivided blindly into 2 equal groups. The study group received 5 weekly parenteral injection of cerebrolysin (5 mL) for a total duration of 4 weeks while the control group received a placebo injection. Both groups also received a single daily dose of celecoxib during the treatment (200 mg PO daily). The JOA score was recorded at 1, 3, and 6 months. Over the 6-month study period, the mean JOA score of patients in the group that received the cerebrolysin injection improved from 11.5 ± 1.2 to 13.9 ± 1.3 while the mean JOA score of the placebo group increased from 11.3 ± 1.2 to 11.8 ± 1.23 with statistically significant differences when comparing the mean JOA recovery rate between the 2 groups at 1, 3, and 6 months (p < 0.0001). Also, no reported cases of neurologic deterioration over 6 months of follow-up were recorded in the cerebrolysin group. The strength of evidence of this study was considered high.

A prospective randomized control trial was conducted recently to determine the value of perioperative cerebrolysin administration.\(^\text{20}\) Sixty patients who underwent surgical decompression for DCM were divided into 2 groups. One group was administered a daily preparation of cerebrolysin for 21 days postoperatively while the second group was administered a placebo. The mJOA score, visual analogue scale (VAS) and an assessment of hand power and sensation were used to assess each group. The trial failed to identify any significant difference in
Pharmacological Treatment of DCM

Levett JJ, et al.

demonstrated that erythropoietin (EPO BID at a concentration of 60 μg/mL) resulted in a significant increase in forced locomotion capability, measured via forelimb stride length, in the treated rats when compared with the control group (p < 0.05).

A clinical trial was conducted by Sugawara et al. to investigate the potential benefits of limaprost alfadex in patients with DCM. A group of 21 patients with mild DCM managed nonoperatively were treated with an oral dose of 15 μg of limaprost alfadex daily for 3 months. The treatment resulted in an improvement in mJOA scores and grip and release count at 1 month that were maintained at 3 months (p = 0.017 and p = 0.001, respectively). The mean mJOA score improved by 1.30 points and the mean grip and release improved from 17.8 to 22.6 over the 3-month period. However, there was no control group in this trial. A phase 3 prospective randomized double-blinded clinical trial on the efficacy of oral limaprost administration following surgery for cervical myelopathy is currently being conducted by the Seoul National University Hospital (ClinicalTrials.gov Identifier: NCT02125981). The strength of evidence of this study was considered moderate.

3. Limaprost Alfadex

A study by Kurokawa et al. demonstrated that limaprost alfadex improved the motor function in a rodent model of cervical myelopathy. An expandable polymer was implanted under the C5–6 laminae of rats to develop compression-induced cervical myelopathy. Twice daily limaprost administration (300 μg/kg PO BID at a concentration of 60 μg/mL) resulted in a significant increase in forced locomotion capability, measured via forelimb stride length, in the treated rats when compared with the control group (p < 0.05).

A clinical trial was conducted by Sugawara et al. to investigate the potential benefits of limaprost alfadex in patients with DCM. A group of 21 patients with mild DCM managed nonoperatively were treated with an oral dose of 15 μg of limaprost alfadex daily for 3 months. The treatment resulted in an improvement in mJOA scores and grip and release count at 1 month that were maintained at 3 months (p = 0.017 and p = 0.001, respectively). The mean mJOA score improved by 1.30 points and the mean grip and release improved from 17.8 to 22.6 over the 3-month period. However, there was no control group in this trial. A phase 3 prospective randomized double-blinded clinical trial on the efficacy of oral limaprost administration following surgery for cervical myelopathy is currently being conducted by the Seoul National University Hospital (ClinicalTrials.gov Identifier: NCT02125981). The strength of evidence of this study was considered moderate.

4. Glucocorticoids

Vidal et al. conducted a study to assess the efficacy of perioperative methylprednisolone in enhancing neurological recovery and to evaluate its effect on the inflammatory response following decompression in an animal model of DCM. DCM was induced in a C57BL/6 mice model using an aromatic polyether material implanted underneath the C5–6 laminae to cause progressive compression of the spinal cord due to focal ossification. Decompressive surgery was conducted 12 weeks post-initial implantation. The mice in the trial group received one dose of methylprednisolone half an hour before surgical decompression and at 2 weeks after the decompression. This study demonstrated that methylprednisolone improved locomotor recovery without affecting the composition of circulating white blood cells (p < 0.05). Histological assessment of the spinal cords showed significant neuronal preservation (p < 0.05).

Human trials have examined the role of glucocorticoids as an adjunct to decompressive surgery. Blume et al. conducted a retrospective cohort study of patients undergoing posterior decompression and instrumentation of the cervical spine for DCM to investigate the effect of intraoperative dexamethasone. A 40-mg dose of intravenous (IV) dexamethasone was administered intraoperatively at the discretion of the senior surgeon. A total of 49 patients were recruited for the study and 25 patients received an intraoperative dose of dexamethasone. Patients were assessed pre- and postoperatively using the Neck Disability Index (NDI) and the mJOA score and there was no significant difference in the baseline scores between the 2 groups prior to surgery. No significant differences were observed between the 2 groups in terms of NDI and mJOA scores at follow-up. Furthermore, a significantly higher rate of wound infections was detected in the group that received intraoperative dexamethasone (p = 0.021). The strength of evidence of this study was considered moderate.

Lastly, a randomized controlled trial by Jeyamohan et al. aimed at comparing the effectiveness of intraoperative dexamethasone administration (0.2 mg/kg IV intraoperatively) on the incidence of postoperative swallowing and airway compromise also examined the effects on functional outcomes (including mJOA scores) and fusion rates. Patients who underwent multilevel anterior cervical discectomy and fusion (ACDF) were randomly assigned to receive intraoperative and postoperative doses of dexamethasone or normal saline and a placebo. The authors demonstrated that intraoperative administration of dexamethasone did not lead to a significant difference in mJOA scores. Moreover, dexamethasone administration significantly delayed fusion rates at 6 months (p = 0.048) without affecting the long-term fusion rates at 12 months (p = 0.57). The strength of evidence of this study was considered moderate.

5. Erythropoietin

A recent study by Tanaka et al. demonstrated that erythropoietin (EPO) improved the motor function in a rodent model of cervical myelopathy. An expandable polymer was implanted under the C5–6 laminae of rats to develop compression-induced cervical myelopathy. EPO administration started 8 weeks after the insertion of the polymer and motor function was assessed after surgery. Motor neurons and apoptotic cell death were eval-

https://doi.org/10.14245/ns.2448140.070
uated with immunohistochemistry. Results from this study demonstrated that rats treated with high-dose EPO maintained better motor function in strength (p < 0.0001). EPO also suppressed neuronal apoptotic cells and significantly prevented the loss of motor neurons (p < 0.0001).

A second study by Eryilmaz and Farooque investigated the therapeutic effects of the combination of EPO and methylprednisolone in the prevention of ischemia-reperfusion injury following decompression in patients with DCM. This randomized controlled study included 110 patients who underwent surgical decompression for DCM. The treatment group received 30 mg/kg of methylprednisolone and 3,000 U/kg of EPO intravenously 30 minutes prior to the start of their spinal decompression surgery while the control group only received 30 mg/kg of methylprednisolone without EPO. This study reported a statistically significant (p < 0.0001) increase in quality of life parameters, i.e., all dimensions of the World Health Organization Quality of Life assessment instrument (WHOQOL-100), in the group of patients who had also received EPO. Additionally, the JOA score improved significantly in the EPO group after surgery compared to the control group (p = 0.025). The EPO group had a preoperative JOA score of 10.25 ± 1.72 (control: 10.31 ± 2.05) and increased to 18.43 ± 2.81 (control: 15.06 ± 2.93) 3 months postoperatively. A significant decrease in the levels of interleukin (IL)-1β and IL-8 3 months after the treatment was also noted in the EPO group (p = 0.028 and p = 0.026, respectively). The strength of evidence of this study was considered high.

These recent findings suggest that EPO may prove beneficial in the management of DCM. Nonetheless, EPO’s therapeutic applications may be limited by its hematological adverse effects, such as red blood cell proliferation, high blood pressure, and prothrombotic properties. As a result, EPO derivatives that retain cytoprotective and neuroprotective effects with minimal erythropoietic activity have been developed.

6. Cilostazol

Yamamoto et al. investigated the neuroprotective effects of cilostazol on cervical myelopathy using a rat model of chronic cervical cord compression. Cord compression was induced using a chronic compression device of thin polyurethane sheets that gradually expanded over 48–72 hours by absorbing water after being surgically implanted under the C5–6 laminae. Cilostazol was orally administered (30 mg/kg PO) to the treatment group prior to the implantation of the device and continuing for the total trial period of 25 weeks, while the control group was administered vehicle solution under the same protocol. Results demonstrated that cilostazol preserved forepaw grip strength and forced running capability 25 weeks postimplantation of the device (p < 0.05). In addition, histopathological examination of the cervical spinal cords demonstrated that the drug helped preserve anterior horn motor neurons in the C5–6 spinal cord segments (p < 0.05).

7. Jingshu Keli

Using a rat model of cervical myelopathy, Yan et al. demonstrated that rats fed Jingshu Keli (JSDKL) 4.8 g/kg daily from day 7 to day 28 postoperatively recovered better gait performance than the control group (p < 0.001). Moreover, the active ingredients in JSDKL, ginsenoside Rb1 (GRb1), and notoginsenoside R1 (NGR1), decreased neuronal excitability through modulation of K+ (Kir) channels by reducing the frequency of action potentials and hyperpolarizing the resting membrane potential (p < 0.05). This led to decreased levels of mechanical and thermal pain at 21 days in the JSDKL group (p < 0.001 and p < 0.05, respectively). Although this study was conducted in a small group of rats (n = 40), further studies should elaborate on traditional Chinese herbs and other alternative medicines that can be beneficial in animal studies, and ultimately in humans.

8. Granulocyte Colony-Stimulating Factor

Yoshizumi et al. conducted a study on a rat spinal cord compression model. They administered granulocyte colony-stimulating factor (G-CSF) 15 mg/kg daily for 5 days subcutaneously to the treatment group and normal saline to the control group and found that the control group had significantly less motor neurons after treatment compared to the G-CSF group (p < 0.001). Moreover, measured using TUNEL (terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate-biotin nick end labeling) staining, G-CSF significantly decreased the number of apoptotic cells at 8 weeks (p < 0.05).

In a phase I and IIA prospective clinical trial, Sakuma et al. observed improved muscle function (p < 0.01), touch (p < 0.05), and pain sensation (p < 0.05) in a cohort of 15 patients with worsening symptoms of compressive cervical myelopathy. These patients were administered G-CSF 10 μg/kg once daily for 5 consecutive days. Mean JOA recovery rates at 1 and 6 months after administration were 49.9% ± 15.1% and 59.1% ± 16.3%, respectively, where recovery rates were defined as (postoperative score–preoperative score) × 100 (%). White blood cell count increased to more than 22,700 cells/mm³ after G-CSF therapy. No serious adverse events occurred during or after treatment. However, these results should be inter-

https://doi.org/10.14245/ns.2448140.070
preted with caution as this was an open-label trial without a control group. The strength of evidence of this study was considered very low.

The restorative function of G-CSF holds great promise as it promotes neural tissue repair and it reduces cell death, hence an ideal candidate for DCM patients. Use of G-CSF in the clinical setting is still in the early stages of research. Rigorous prospective trials are needed to assess the safety profile, efficacy, and optimal dosing regimens of G-CSF in humans.

9. Anti-Fas Ligand Antibody

Yu et al.\textsuperscript{12} performed a postmortem human tissue study to investigate the pathology and apoptotic mechanisms in human DCM and the therapeutic potential of anti-Fas ligand antibody in a mouse model. In the postmortem samples of patients with DCM, they showed increased Fas-mediated apoptosis of neurons and oligodendrocytes and an increase in inflammatory cells. To demonstrate the therapeutic potential of an anti-Fas ligand antibody, they administered anti-Fas ligand antibody (MFL3) at 50 mg intraperitoneally twice weekly for 4 weeks in Twy/Twy (tiptoe-walking-yoshimura) mice. Treatment decreased neural inflammation mediated by macrophages and activated microglia, glial scar formation and caspase-9 activation. Furthermore, the treatment promoted dramatic functional neurological recovery.

The antiapoptotic property of an anti-Fas ligand antibody represents a viable therapeutic direction in DCM patients. Although these results seem promising in attenuating proinflammatory pathways and neural degeneration in DCM, the recommended posology of anti-fas ligand antibody remains unknown as administration in DCM patients has yet to be conducted.

Characteristics of the included studies evaluating the pharmacological agents in humans and animals can be found in Tables 2 and 3, respectively.

DISCUSSION

Herein, we present the first comprehensive review of pharmacological agents for the management of mild DCM. Several findings were reported, including the following: (1) riluzole is not effective for treating mild DCM in humans, but only in rat models; (2) cerebrolysin shows potential as a neuroprotective agent for myelopathy, but its effectiveness in DCM patients should be further investigated; (3) limaprost alfadex has demonstrated motor function improvement in animal models and has promising results in a small clinical trial; (4) glucocorticoids not only failed to offer clinical benefits, but may have also led to adverse events; (5) cilostazol, anti-Fas ligand antibody, and Jingshu Keli have shown promise in animal studies, while EPO, G-CSF, and limaprost alfadex, in both animal and human studies. Overall, only weak clinical evidence and animal studies are available. Thus, these results have yet to be validated in large scale studies that provide a high level of evidence.

Research on the pharmacological management of patients with DCM might have multiple relevant clinical applications. Most patients present a progressive deterioration or a stepwise decline in their neurological function that may be characterized by significant periods of symptom stability.\textsuperscript{34} Also, previous reports have suggested similar results between patients with mild DCM treated conservatively and those with surgical decompression.\textsuperscript{35} Moreover, the AO Spine North America and Cervical Spine Research Society guidelines suggest offering either surgical intervention or a supervised trial of structured rehabilitation for patients with mild DCM, with surgery recommended in cases of neurological deterioration or lack of improvement.\textsuperscript{8} In addition, symptomatic radiculopathy, presence of MRI cervical cord hyperintensity, prolonged motor evoked potentials and somatosensory evoked potentials, and electromyographic findings of anterior horn cell lesions have been associated with development of myelopathy in asymptomatic patients.\textsuperscript{5} Furthermore, circumferential cord compression on axial MRI, a deformity of the spinal cord due to compression with an acute-angled lateral corner (one or both sides), abnormally increased range of preoperative neck and head motion, lower segmental lordotic angle, segmental instability, and reduced diameter of the cerebrospinal fluid column have been correlated with further worsening of myelopathy.\textsuperscript{5} Thus, neuroprotective substances that prevent the progression of the disease or promote neuro-regeneration might be a reasonable alternative in the future for the aforementioned categories of patients (i.e., patients with findings associated with development or worsening of myelopathy or patients with mild and stable myelopathy). However, it is important to stress that pharmacological treatment is far yet from being a main treatment for DCM, especially for patients with moderate or severe DCM, where surgical intervention is recommended based on moderate quality evidence and a strong recommendation.\textsuperscript{8}

Furthermore, surgical management remains indicated in patients with moderate to severe forms of DCM.\textsuperscript{8} However, some patients may present neurological deterioration after surgical decompression. Up to 11.6% of patients who undergo decompression for DCM may experience deterioration of their neuro-
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Potential biases/limitations</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rajasekaran et al., 2016</td>
<td>Double-blinded, placebo-controlled randomized controlled trial</td>
<td>To evaluate the effectiveness of Riluzole as a pharmacotherapeutic treatment option for early cervical myelopathy</td>
<td>Intervention: 15 patients received Riluzole Control: 15 patients received vitamin B</td>
<td>Riluzole 50 mg PO BID for 1 month</td>
<td>Clinical scores: modified JOA, Nurick-grading, SF-12, Neck Disability Index; diffusion tensor imaging datametrics: apparent-diffusion coefficient, fractional anisotrophy, volume ratio, anisotrophy, Eigen vectors</td>
<td>No significant change in modified JOA scores or other clinical scores between groups, diffusion tensor imaging datametrics did not show statistically significant changes</td>
<td>1B</td>
</tr>
<tr>
<td>Fehlings et al., 2021</td>
<td>Multicentre, double-blind, placebo-controlled, randomized, phase 3 trial</td>
<td>To investigate whether riluzole enhances outcomes in patients undergoing decompression surgery for DCM</td>
<td>Intervention: 141 patients undergoing decompression for DCM received riluzole Control: 149 patients undergoing decompression surgery for DCM received a placebo</td>
<td>Riluzole 50 mg PO BID for 14 days preoperatively and then for 28 days postoperatively</td>
<td>No significant change in modified JOA score at 6 months (p = 0.14). Increased serious adverse events in the riluzole group compared to the control group (n = 43 vs. n = 34, respectively). Most common adverse events were neck or arm or shoulder pain, arm paraesthesia, dysphagia, and worsening of myelopathy</td>
<td>Heterogeneity of spinal cord injury etiology in DCM; insensitivity and interpretation of statistical significance in the outcome instruments (i.e., modified JOA scale and Nurick grade); poor generalizability of findings to other populations; &lt; 80% 1-year follow-up</td>
<td>2B</td>
</tr>
<tr>
<td>Allam et al., 2018</td>
<td>Prospective randomized study</td>
<td>To evaluate the effect of cerebrolysin as a conservative modality on DCM patients</td>
<td>Intervention: 96 patients received cerebrolysin Control: 96 patients received placebo IM injection Both groups received celecoxib 200 mg as a single, after-meal daily dose for 4 weeks</td>
<td>Cerebrolysin 5 mL IM QD for 5 days/week for 4 weeks</td>
<td>At 1 month, myelopathy improved in 92% of patients in group I and 52% in group II. At 6 months, the improvement was seen in 87% of patients in group I and 33% in group II. The mean JOA recovery rate was significantly higher in group I compared to group II at all time points (1, 3, and 6 months) (p &lt; 0.0001)</td>
<td>Treatment was not uniquely cerebrolysin, rather a combination of cerebrolysin and celecoxib; lack of blinding of the drug provider and the main investigator</td>
<td>1B</td>
</tr>
<tr>
<td>Sharma et al., 2022</td>
<td>Prospective randomized controlled trial</td>
<td>To analyze the role of cerebrolysin in patients with DCM managed with surgery</td>
<td>Intervention: 30 patients received cerebrolysin Control: 30 patients received a placebo</td>
<td>Cerebrolysin 5 mL IV QD diluted in 100 mL 0.9% NaCl over 30 minutes for 21 days postoperatively</td>
<td>Both groups showed significant improvement in mJOA and VAS scores at 3 weeks, 3 months, 6 months, and 1 year postoperatively (p &lt; 0.01), but no significant difference between the groups. The cerebrolysin group showed significant improvement in hand function at 1 year compared to the placebo (p = 0.03)</td>
<td>Small sample size, single-center study, long duration of IV therapy, social determinants influencing the sex ratio</td>
<td>2B</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Potential biases/limitations</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugawara et al., 2009</td>
<td>Prospective study</td>
<td>To examine the effect of oral administration of limaprost alfadex on myelopathy symptoms in patients with mild cervical spinal canal stenosis</td>
<td>21 Patients with mild spondylotic cervical spinal canal stenosis without any improvements after oral administration of NSAID drugs, muscle relaxant, and/or vitamin B12 for at least 2 months before referral to oral administration of limaprost alfadex 15 μg PO QD</td>
<td>JOA score significantly improved from 14.0 to 15.0 at 1 month (p = 0.022) and 15.2 at 3 months (p = 0.009)</td>
<td>Grip and release count significantly improved from 17.8 to 21.4 at 1 month (p = 0.017) and 22.6 at 3 months (p = 0.001)</td>
<td>Observational study – no study group, short follow-up</td>
<td>2B</td>
</tr>
<tr>
<td>Blume et al., 2018</td>
<td>Retrospective study</td>
<td>To investigate the effect of intraoperative dexamethasone on wound healing, complications, and clinical outcome in patients with posterior surgery for DCM</td>
<td>Intervention: 25 patients underwent posterior instrumentation – decompression and received dexamethasone Control: 24 patients also underwent posterior instrumentation – decompression, but did not receive dexamethasone Dexamethasone 40 mg IV intraoperatively</td>
<td>No significant differences between groups in pre- and postoperative findings, complications, neurologic symptoms, and follow-up (NDI and modified JOA score). There was a higher rate of wound healing complications in the dexamethasone group (p = 0.021)</td>
<td>Retrospective design, small sample size, different lengths of follow-up, exclusion of ventral decompressive surgical cases, selection bias in dexamethasone administration</td>
<td>2B</td>
<td></td>
</tr>
<tr>
<td>Jeyamohan et al., 2015</td>
<td>Prospective, randomized, double-blinded, controlled trial</td>
<td>To determine if perioperative dexamethasone use improves perioperative dysphagia and airway edema</td>
<td>Intervention: 56 patients underwent multilevel anterior cervical reconstruction and received dexamethasone Control group: 56 patients also underwent multilevel anterior cervical reconstruction, but received saline Dexamethasone 0.2 mg/kg IV intraoperatively Postoperative doses: Dexamethasone 0.06 mg/kg IV Q6H for the first 24 hours</td>
<td>No significant differences in the myelopathy scores, axial pain scores, extremity pain scores, ODI, or SF-12 scores (either mental or physical summary component) 6, 12, and 24 months. Severity of dysphagia in the postoperative period up to 1 month was significantly lower in the steroid group (p = 0.027). Airway difficulty and need for intubation trended toward significance in the placebo group (p = 0.057). Fusion rates at 6 months were significantly lower in the steroid group, but lost significance at 12 months (p = 0.048 and p = 0.57, respectively)</td>
<td>Nonstandardized dexamethasone-dosing schedule; subjectivity of functional outcome swallowing scale score; steroid treatment deviations; inherent bias towards steroid treatment; short follow-up; &lt; 80% 1-year follow-up</td>
<td>2B</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### Table 2. Summary of studies investigating the pharmacological treatment of DCM in humans (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Potential biases/limitations</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eryilmaz et al., 2021</td>
<td>Prospective randomized controlled trial</td>
<td>To investigate the therapeutic effects of combined erythropoietin and methylprednisolone therapy on ischemia-reperfusion injury to the spinal cord and its effects on interleukin-1 beta (IL-1β), IL-1 receptor antagonist (IL-1RA), and IL-8 (IL-8) levels</td>
<td>Intervention: 55 patients received both methylprednisolone and erythropoietin and 55 patients received only methylprednisolone</td>
<td>Methylprednisolone: 30 mg/kg IV 30 minutes preoperatively spinal cord decompression Erythropoietin: 3,000 U/kg IV 30 minutes preoperatively spinal cord decompression</td>
<td>Three months after treatment, the observation group showed a significant improvement in JOA scores (p = 0.025) and 40-point rating method scores (p = 0.019) compared to the control group. Three months after treatment, the observation group had higher S-100β levels (p = 0.041), lower neuron-specific enolase levels (p = 0.032), lower IL-1β levels (p = 0.026), higher IL-1RA levels (p = 0.021), and lower IL-8 levels (p = 0.028) compared to the control group. The observation group had higher scores in all dimensions of the World Health Organization Quality of Life assessment instrument (WHOQOL-100), indicating better quality of life compared to the control group (p &lt; 0.001 for all dimensions)</td>
<td>No comparator group that did not receive methylprednisolone; no justification of sample size or selection of dosages; short follow-up</td>
<td>1B</td>
</tr>
<tr>
<td>Sakuma et al., 2011</td>
<td>Prospective clinical trial (Phase I and IIa)</td>
<td>To evaluate the safety and efficacy of neuroprotective therapy using G-CSF for patients with worsening symptoms of compression myelopathy</td>
<td>15 Patients with worsening symptoms of compression myelopathy Intervention: 5 patients received the 5 μg dose (followed by decompression surgery) and 10 patients received the 10 μg dose (followed by decompression surgery in 9 patients) Control group: NA (open-label study design)</td>
<td>G-CSF 5 μg/kg IV QD or 10 μg/kg IV QD for 5 consecutive days</td>
<td>G-CSF administration suppressed the progression of myelopathy in all 15 patients. Muscle (p &lt; 0.01), touch (p &lt; 0.05), and pain (p &lt; 0.05) improvement were observed in all patients receiving G-CSF 10 μg/kg QD at 6 months. Mean JOA recovery rates at 1 and 6 months after administration in the 10 μg group were 49.9% ± 15.1% and 59.1% ± 16.3%, respectively. White blood cell count increased to more than 22,700 cells/mm³ after G-CSF therapy. No serious adverse events occurred during or after treatment</td>
<td>Absence of a control group; open-label design</td>
<td>4</td>
</tr>
</tbody>
</table>

DCM, degenerative cervical myelopathy; PO, orally; BID, twice a day; JOA, Japanese Orthopaedic Association; IM, intramuscular; QD, once a day; IV, intravenous; NSAID, nonsteroidal anti-inflammatory drug; NS, not specified; mJOA, modified JOA; VAS, visual analogue scale; Q6H, every 6 hours; ODI, Oswestry Disability Index; SF-12, 12-item Short Form health survey; G-CSF, granulocyte colony-stimulating factor.
Table 3. Summary of studies investigating the pharmacological treatment of DCM in animals

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Animal model</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Pathology</th>
<th>Potential biases/limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moon et al.,15 2014</td>
<td>Experimental animal research</td>
<td>To examine the effects of riluzole on neurobehavioral outcomes, neuropathic pain, and tissue preservation in a rat model of chronic cervical spinal cord compression (DCM)</td>
<td>Sham group (n = 6), control group (n = 18), riluzole group (n = 17)</td>
<td>Female Sprague-Dawley rats that underwent C6–7 laminectomy and implantation of the rod of a chronic compression device into the C2 and the T2 spinous processes</td>
<td>Riluzole 8 mg/kg intraperitoneal QD or control solution (30% 2-hydroxypropyl-β-cyclodextrin) was initiated 1 week postoperatively and continued for 7 weeks</td>
<td>Control group showed increased sensitivity of mechanical allodynia (via von Frey filament testing) compared to the sham group at weeks 2–8 (p = 0.003 for week 2, p &lt; 0.001 for weeks 3–8) Riluzole group had decreased sensitivity compared to the control group at weeks 2, 6, 7, and 8 (p = 0.012, p = 0.025, p = 0.039, p &lt; 0.001, respectively) Control group demonstrated decreased latency of thermal hyperalgesia (via tail flick test) compared to the sham group (p = 0.007) Riluzole group showed increased latency compared to the control group (p = 0.006) Gait assessed using CatWalk system: Riluzole group had increased swing phase duration in forelimbs and hindlimbs compared to the control group (p &lt; 0.05). Riluzole group showed increased paw intensity in forelimbs and hindlimbs compared to the control group (p &lt; 0.05). Riluzole group exhibited increased hindlimb swing speed compared to the control group (p &lt; 0.05)</td>
<td>Immunohistochemical analysis revealed decreased phosphorylated NR1 and NR2B positive cells in the dorsal horns of the riluzole group compared to the control group (p &lt; 0.001) Riluzole administration reduced microglia activation in the dorsal horns compared to the control group (p = 0.001). Riluzole treatment led to decreased scar tissue area and increased preserved gray matter area compared to the control group (p &lt; 0.05)</td>
<td>The experimental design did not allow for investigation of the potential effects on the induction of thermal hyperalgesia after the gradual spinal cord compression The quantification of the immunofluorescence involves certain limitations The results only allowed to make an associative inference regarding the relationship between pNR1 expression and pNR2B expression changes in the microglia phenotype</td>
</tr>
</tbody>
</table>

(Continued)
### Table 3. Summary of studies investigating the pharmacological treatment of DCM in animals (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Animal model</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Pathology</th>
<th>Potential biases/limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karadi- mas et al., 2015</td>
<td>Experimental animal research &amp; retrospective review of prospective clinical trial (AOSpine North America CSM study)</td>
<td>Investigate the role of ischemia-reperfusion injury and the use of riluzole in improving DCM outcomes</td>
<td>Animal model: Five groups of rats (sham, control, riluzole alone, decompression alone, riluzole+ decompression) Clinical trial: 278 DCM patients</td>
<td>Female Sprague-Dawley rats in which a piece of aromatic polyether was inserted underneath the C6 lamina This aromatic polyether serves as a scaffold for the precipitation of inorganic salts, leading to controlled, progressively increased pressure on the cervical spinal cord</td>
<td>Riluzole 8 mg/kg intraperitoneal QD started at 4 weeks after material implantation. Surgical decompression took place 6 weeks after material implantation, and riluzole administration continued until 2 weeks after decompression</td>
<td>Riluzole treated rats did not exhibit significant declines in gait parameters in the first week after decompression surgery. Riluzole treatment significantly improved forelimb stride length, forepaw initial contact, and regularity index parameters compared to the decompression-only group (p &lt; 0.05). Riluzole administration reduced the proportion of preserved neurons expressing oxidative DNA damage in the rat spinal cord (p &lt; 0.05). Riluzole-treated rats displayed a significantly lower proportion of 8-oxoG DNA-positive cells (indicating oxidative damage) in vitro (p &lt; 0.05). Riluzole reduced depolarization of the mitochondrial membrane potential in vitro (p &lt; 0.05). Rats receiving combined decompression surgery and riluzole treatment showed significantly improved forelimb stride length compared to decompression surgery alone (p &lt; 0.05). Combined treatment significantly improved coordination between forelimbs and hindlimbs compared to the control, riluzole, and decompression groups (p &lt; 0.05). Rats treated with the combination approach had superior motor neuron preservation in the cervical spinal cord compared to decompression alone (p &lt; 0.05). Riluzole or decompression alone significantly decreased below-level neuropathic pain in DCM rats (p &lt; 0.05). Combination of decompression surgery and riluzole administration led to significantly lower microglial activation in the lumbar dorsal horns compared to either treatment alone (p &lt; 0.05).</td>
<td>Immunohistochemistry for choline acetyltransferase, neuronal nuclei, protein kinase C–γ, glial fibrillary acidic protein, Iba-1; Double immunofluorescence for 8-oxoG DNA and neuronal nuclei; JC-1 staining for mitochondrial membrane potential</td>
<td>Small sample size of each group One of the authors was the Chairman of AOSpine North America, a not-for-profit foundation, which served as the sponsor of the study</td>
</tr>
</tbody>
</table>
Table 3. Summary of studies investigating the pharmacological treatment of DCM in animals (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Animal model</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Pathology</th>
<th>Potential biases/limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kurokawa et al., 2011</td>
<td>Experimental animal research</td>
<td>To explore the possibility of limaprost alfadex for DCM</td>
<td>Group A: Sham operation without permanent cord compression (receiving distilled water 5 mL/kg BID (n = 6))</td>
<td>Male Wistar rats, in which a sheet of expandable urethane-compound polymer was inserted in the C5 and C6 sublaminal space</td>
<td>Limaprost alfadex 300 μg/kg PO BID at a concentration of 60 μg/mL</td>
<td>Rats with chronic spinal cord compression demonstrated latent and progressive deterioration in forced locomotion capability 6 to 11 weeks after the induction of compression. Specifically, forelimb stride length decreased significantly (p &lt; 0.05) in the decompression group compared to baseline values. Rats with the compression treated with limaprost alfadex retained the ability to perform the forced exercise. Notably, the limaprost alfadex treatment group showed no significant decrease in forelimb stride length, indicating preserved forced locomotion capability compared to the decompression-only group (p &lt; 0.05)</td>
<td>Spinal cord harvested for motor neuron counts</td>
<td>Small sample size of each group. Limaprost alfadex provided by Ono Pharmaceutical Co., Osaka, Japan.</td>
</tr>
<tr>
<td>Vidal et al., 2018</td>
<td>Experimental animal research</td>
<td>To assess the efficacy of perioperative methylprednisolone in enhancing neurological recovery and evaluate its effect on the inflammatory response following decompression for DCM</td>
<td>Intervention: 18 mice received decompression with methylprednisolone treatment (30 mg/kg) Control: 18 mice received compression with saline treatment</td>
<td>C57BL/6 mice with induced DCM (polyether material implanted underneath C5–6 laminae)</td>
<td>Methylprednisolone 30 mg/kg IV. One dose given 30 minutes before decompression and another dose at 2 weeks postdecompression</td>
<td>Improved locomotor recovery and reduced motor complications following methylprednisolone treatment (p &lt; 0.05) (measured using CatWalk system). Preservation of neurons in the spinal cord with methylprednisolone treatment compared to the control group (p &lt; 0.05). Modest reduction in parenchymal inflammation with methylprednisolone treatment</td>
<td>Cervical spinal cord homogenates analyzed for cytokines (interleukin [IL]-1β, IL-1β, TNF-α, IL-4, IL-10, IL-6) using Luminex assay. Immunohistochemistry used for glial (Iba1, glial fibrillary acidic protein) and neuronal (neuronal nuclei, oligodendrocyte transcription factor 2) cell markers</td>
<td>Composition of peripheral white blood cells between humans and mice is different. Thus, the inflammatory response after surgical decompression may differ between mice and humans. DCM patients may have other comorbidities, including cardiovascular disease and diabetes, which are not present in the animal model. Therefore, the potential clinical translation of this work to DCM patients will need to control for the relevant side-effects of steroids.</td>
</tr>
</tbody>
</table>
### Table 3. Summary of studies investigating the pharmacological treatment of DCM in animals (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (# in groups)</th>
<th>Animal model</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Pathology</th>
<th>Potential biases/ limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanaka et al., 2019</td>
<td>Experimental animal research</td>
<td>To evaluate the effect of human recombinant EPO on a rat model of spinal cord compression-induced cervical myelopathy</td>
<td>Sham group (n = 12); Vehicle group (compression+ normal saline; n = 12); low-dose EPO group (compression+ EPO low dose; n = 12); and high-dose EPO group (compression+ EPO high dose; n = 12)</td>
<td>Male Wistar rat model of spinal cord compression-induced cervical myelopathy (expandable polymer implanted underneath C5–6 laminae)</td>
<td>Human recombinant EPO administered subcutaneously: low-dose EPO group received 500 IU/kg QD; high-dose EPO group received 5,000 IU/kg QD of human recombinant EPO. Administration started at 8 weeks postoperatively and continued until 16 weeks</td>
<td>High-dose EPO significantly maintained motor function in the compression groups. Strength improved in the high-dose EPO group throughout the period of EPO administration from 9 to 16 weeks postoperatively (p &lt; 0.001). EPO significantly prevented the loss of motor neurons and decreased neuronal apoptotic cells (p &lt; 0.001). The number of synaptophysin-positive axons was significantly higher in the high-dose group compared to the vehicle group at 10 and 16 weeks postoperatively (p &lt; 0.005). High-dose EPO significantly lowered EPO-receptor-positive anterior horn cells (p &lt; 0.005). EPO significantly decreased TUNEL-positive cells and Caspase-3-positive cells compared to the vehicle group (p &lt; 0.001 and p &lt; 0.001, respectively). High-dose EPO significantly reduced APP-positive cells in the white matter (p &lt; 0.05)</td>
<td>H&amp;E, neuronal nuclei, choline acetyltransferase, glial fibrillary acidic protein, allopahycocyanin, EPO receptor, 5-HT, growth associated protein 43, synaptophysin, and amyloid precursor protein staining were performed to assess motor neurons, glial cells, and axonal markers. TUNEL and Caspase-3 staining were used to investigate apoptotic cell death</td>
<td>Clinically there are several known side-effects EPO that make the continuous administration of EPO over a long period seem unrealistic</td>
</tr>
</tbody>
</table>
| Yama-moto et al., 2014 | Experimental animal research | To investigate the neuroprotective effect of cilostazol on DCM | 40 male Wistar rats (group A: sham operation+ vehicle, n = 7; group B: sham operation+ cilostazol, n = 7; group C: compression+ vehicle, n = 13; group D: compression+ cilostazol, n = 13) | Male Wistar rat model of chronic spinal cord compression-induced cervical myelopathy (expandable polymer implanted underneath C5–6 laminae) | Cilostazol 30 mg/kg PO QD for 25 weeks | Preservation of forepaw grip strength: group D had significantly higher grip strength compared to group C at 7 weeks and thereafter (p < 0.005); Preservation of forced running capability: group D showed no decrease in locomotion, while group C exhibited progressive deterioration starting at 18 weeks (p < 0.05); Preservation of anterior horn motor neurons: group D had a significantly lower loss of motor neurons (7.1%) compared to group C (34.4%) (p < 0.05); Decreased number of TUNEL-positive apoptotic cells: group D (compression+ Cilostazol) showed significantly lower numbers of apoptotic cells in both gray and white matter compared to group C (compression+ vehicle) | Histopathological examination of cervical spinal cords | The results suggested that the acute effects of surgery were more significant than the inflation of the polymer sheet, which was to occur only in the compression groups after surgery. This study was supported by a grant-in-aid and the provision of cilostazol from Otsuka Pharmaceutical Co. The expandable polymer was provided courtesy of Sanyo Chemical Industries, Ltd. | **(Continued)**
**Table 3. Summary of studies investigating the pharmacological treatment of DCM in animals (Continued)**

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population (n in groups)</th>
<th>Animal model</th>
<th>Drug, dosages, and administration</th>
<th>Outcome(s)</th>
<th>Pathology</th>
<th>Potential biases/limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yan et al., 2019</td>
<td>Experimental animal research</td>
<td>To investigate the effects of Jingshu Keli on DCM</td>
<td>Behavioral tests: 40 rats for gait analysis, divided into 5 groups (control, model (underwent DCM modeling surgery), and 3 Jingshu Keli treatment groups (1.2 g/kg, 2.4 g/kg, and 4.8 g/kg))</td>
<td>Sprague-Dawley rats with DCM induced by compression. A plastic monofilament fishing line was passed cranially from the C6–7 to the C4–5 interlaminar space and was secured on the dorsal aspect of the laminae at C5 and C6</td>
<td>Jingshu Keli 1.2 g/kg or 2.4 g/kg or 4.8 g/kg PO QD starting from day 7 postoperatively and continued until day 28</td>
<td>Gait Performance (28th day): Jingshu Keli 4.8 g/kg group showed a significant improvement in gait compared to the model group (p &lt; 0.001). Mechanical Pain (14th and 21st day): Jingshu Keli 2.4 g/kg and 4.8 g/kg groups showed a significant increase in paw withdrawal threshold compared to the model group (p &lt; 0.01 on day 14, p &lt; 0.001 on day 21). Thermal Pain (14th and 21st day): Jingshu Keli 4.8 g/kg group showed a significant increase in paw withdrawal threshold compared to the model group (p &lt; 0.05 on day 14, p &lt; 0.01 on day 21). Neuronal Excitability: Jingshu Keli, GRb1, and NGR1 significantly reduced the frequency of action potentials by 38.5%, 27.2%, and 25.9%, respectively, and hyperpolarized the resting membrane potential by 15.0%, 13.8%, and 12.1%, respectively. The effects were mediated through modulation of Kir channels (p &lt; 0.05).</td>
<td>In cell culture, brainstem neurons were prepared from newborn Sprague-Dawley rats. Whole cell patch clamp recordings were used to study the action potentials and K+ currents (KV and Kir) in the cultured cells. Fluorescence immunostaining was conducted to study the Kir3.1 protein in the cells.</td>
<td>Small sample size of each group. Whether the phosphorylation of Kir3.1 is the reason or result for JSKL promotion on Kir currents remains debatable, uncommon animal DCM model.</td>
</tr>
<tr>
<td>Yoshimi et al., 2016</td>
<td>Experimental animal research</td>
<td>To explore the potential of G-CSF as a pharmacologic treatment for DCM</td>
<td>36 Rats were divided into 3 groups: group A (sham operation+ normal saline), group B (cord compression+normal saline), and group C (cord compression+G-CSF)</td>
<td>Wistar rats in which a sheet of expandable urethane-compound polymer was implanted between the C5–6 laminae. The volume of the sheet expands by absorbing tissue water, reaches 230% of the original volume, and remains constant</td>
<td>G-CSF 15 mg/kg QD or normal saline administered subcutaneously 5 days a week</td>
<td>In the prevention experiment, G-CSF preserved motor functions throughout 26 weeks, significantly decreased the number of apoptotic cells at 8 weeks (p &lt; 0.05). In the treatment experiment, G-CSF administration from 8 weeks after surgery temporarily restored motor function to a level equal to the sham group. Motor neuron count: group B (compression+normal saline) had significantly fewer motor neurons compared to groups A (sham+normal saline) and C (compression+G-CSF) (p &lt; 0.001).</td>
<td>Cervical spinal cords were examined histopathologically using H&amp;E staining. TUNEL staining was performed to assess apoptotic cell death at 8 weeks after surgery.</td>
<td>Small sample size of each group. Recombinant human G-CSF was provided by Chugai Pharmaceutical Co., Ltd., Tokyo, Japan. Clinically there are several known side-effects which make the continued administration of G-CSF for CSM over long periods seem unrealistic.</td>
</tr>
<tr>
<td>Study</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population (# in groups)</td>
<td>Animal model</td>
<td>Drug, dosages, and administration</td>
<td>Outcome(s)</td>
<td>Pathology</td>
<td>Potential biases/limitations</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>-----</td>
<td>-------------------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Yu et al., 2011</td>
<td>Experimental animal research &amp; postmortem tissue study</td>
<td>To investigate the pathology and apoptotic mechanisms in human DCM and the therapeutic potential of anti-Fas ligand antibody in a mouse model</td>
<td>Human DCM group: 6 patients (61–89 years old) with DCM, 6 patients with motor weakness, sensory disturbances, and spinal cord compression</td>
<td>Twy/Twy mice (NPPS gene mutation) which develop extradural calcified deposits at C2–3, spinal cord compression and progressive spinal cord dysfunction</td>
<td>Anti-Fas ligand antibody (MFL3) 50 mg intraperitoneally twice weekly for 4 weeks in Twy/Twy mice</td>
<td>Human DCM: Severe anterior horn atrophy, neuronal loss, axonal loss, myelin pallor, and gliosis observed in compressed epicentre. Decreased motor neurons (MAP2-positive) and axonal density (NF200-positive) in compressed region. Increased Fas-positive neurons, Fas ligand-positive neurons, and apoptotic cells (TUNEL-positive) in compressed region compared to controls (p = 0.002, p = 0.001, and p &lt; 0.05, respectively)</td>
<td>Autopsies performed within 1–30 hours postmortem. Morphological assessment with H&amp;E, Luxol fast blue, and immunohistochemistry (Fas, Fas ligand, MAP2, CD68, Iba1, TUNEL) in human DCM. Luxol fast blue and H&amp;E staining used to identify atrophy, neuronal loss, axonal loss, and myelin pallor</td>
<td>Small sample size of each group, lack of randomized control group, potential biases from postmortem interval and tissue storage, limited generalizability to other populations, lack of long-term follow-up</td>
</tr>
</tbody>
</table>

DCM, degenerative cervical myelopathy; QD, once a day; TNF, tumor necrosis factor; PO, orally; BID, twice a day; EPO, erythropoietin; TUNEL, terminal deoxynucleotidyl transferase-mediated deoxyuridine triphosphate-biotin nick end labeling; H&E, haematoxylin and eosin; GRb1, ginsenoside Rb1; NGRI, notoginsenoside R1; G-CSF, granulocyte colony-stimulating factor; Twy/Twy, tiptoe-walking-yoshimura.
logical function in the immediate postoperative period. In addition, data from the recent AO Spine North America CSM study demonstrated that 9.3% of patients exhibited postoperative functional decline and that 44% of operated patients were left with significant neurological deficits 6 months after surgery. Therefore, adjunct medical therapy that could improve or optimize the outcomes of patients undergoing surgical decompression would be valuable.

The pathophysiology of DCM is complex and diverse. Multiple mechanisms are thought to be responsible for the neuronal loss, axonal degeneration and myelin impairment seen in DCM. Mechanical compression of the spinal cord can lead to inflammation, ischemia and apoptosis resulting in spinal cord dysfunction. In addition, histological and immunohistochemical studies in rodents have demonstrated that immediate neurological decline after decompression may be initiated by an ischemia-reperfusion injury and immune reaction in the spinal cord. Rapidly released cytokines perpetuate the inflammatory response after decompression. Furthermore, rat spinal cords demonstrate an increase in oxidative stress after decompression. Fig. 1 summarizes the common mechanisms of action of the abovementioned medications.

Riluzole (2-amino-6(trifluoromethoxy)benzothiazole; Rilutek, Sanofi-Aventis Inc., Paris, France) is a sodium channel/glutamate blocker from the benzothiazide group that is the first and only drug approved for the management of amyotrophic lateral sclerosis (ALS) in the United States. Studies have suggested a key role of sodium and glutamate mediated cellular injury in models of spinal cord compression. Therefore, it was hypothesized that riluzole could slow the neurodegeneration process of motor neurons via sodium channels and by decreasing glutamate mediated excitotoxicity. Cerebrolysin is a mixture of peptides derived from enzymatic lysis of porcine brains products. Its pharmacodynamics are similar to endogenous neurotrophic factors, and it can readily cross the blood-brain barrier. Cerebrolysin has demonstrated its neuroprotective potential through its action on cellular structural integrity and neurogenesis and has been studied in patients with neurological disorders such as dementia, stroke, and traumatic brain injury. Limaprost alfadex is an oral prostaglandin E1 analog that

---

**Fig. 1.** Target mechanisms of current pharmacological agents used for the treatment of degenerative cervical myelopathy. EPO, erythropoietin; G-CSF, granulocyte colony-stimulating factor; JSKL, Jingshu Keli; NGR1, notoginsenoside R1; GrB1, ginsenoside Rb1. 

Pharmacological Treatment of DCM

Levett JJ, et al.

has vasodilatory and antithrombotic properties. Its use was approved for the treatment of lower extremity ischemic symptoms and neurogenic claudication secondary to lumbar spinal stenosis. Glucocorticoids are a potent class of anti-inflammatory and immunosuppressive drugs. Their action on the glucocorticoid receptor mediates multiple pathways that suppress the inflammatory response. Steroids are commonly used in the management of metastatic spinal cord compression. Therefore, it was hypothesized that steroids could help mitigate the potential postsurgical inflammatory response and spinal cord reperfusion injury following decompressive surgery in patients with DCM.

EPO is a glycoprotein cytokine secreted mainly by the kidney in response to hypoxia that stimulates erythropoiesis. In addition, EPO protects tissue from ischemia and reperfusion injury, has antiapoptotic effects and improves regeneration after injury. In the past 2 decades, several studies have demonstrated its neuroprotective benefits in cerebral infarction, brain trauma, and acute spinal cord injuries. In hypoxic conditions, endogenous EPO is secreted by astrocytes in response to low oxygen partial pressure and may act on neurons as an important paracrine neuroprotective mediator of ischemic preconditioning. Locally produced EPO may also protect the neural tissue by promoting angiogenesis.

Cilostazol (6- [4-(1-cyclohexyl-1H-tetrazol-5-yl) butoxy]-3,4-dihydro-2-(1H)-quinolinone) is a potent phosphodiesterase inhibitor that suppresses platelet aggregation and acts as a direct arterial vasodilator. Its use has been approved in the U.S. Food and Drug Administration and in many Asian countries for the management of intermittent claudication. Moreover, nerve conduction velocity and blood flow were significantly higher with the administration of cilostazol to a canine model with cauda equina compression. In addition, cilostazol has demonstrated its neuroprotective effect by reducing the size of ischemic brain infarction in a rat model of cerebral ischemia through inhibition of apoptotic and oxidative cell death. JSKL is a traditional Chinese herbal formula. Previous research demonstrated that some active ingredients of the formula could alleviate pain through the modulation of ion channels. Its active components, NGR1, and GRβ1, can suppress voltage-gated K+ channels and activate chloride channels, respectively. G-CSF is a cytokine that promotes differentiation of cells in the neutrophil lineage. It mobilizes bone marrow cells to the peripheral circulation. It has shown to restore damaged spinal cord tissue and it has recovered neural function in rats and mice. Fas, also known as Fas antigen or CD95, is a cell surface protein involved in cell death. When it binds to its ligand, FasL, caspases are activated to promote apoptosis. Elevated expression of Fas was associated with neural apoptosis and release of proinflammatory cytokines in a rat model.

Some studies on the pharmacotherapy of DCM did not evaluate any neurologic/functional outcome, but some findings which might be worthy of further investigation. To characterize the effect of glucocorticoids on the postsurgical systemic inflammatory response, Demura et al. investigated the relationship between perioperative steroid administration and IL-6 serum levels in patients with cervical myelopathy treated by laminoplasty. The study concluded that preoperative administration of dexamethasone attenuates the systemic inflammatory reaction to surgery, as demonstrated by decreased postoperative IL-6 levels. Neurotrophic factors play an important role in cell survival and have an antiapoptotic activity on neurons. Uchida et al. have conducted studies on the applications of neurotrophins in the treatment of DCM. Most of their research was completed using a Twy mouse; a naturally existing mutant rodent that develops spontaneous calcification at the C1–2 vertebral level, which mimics significant compression of the spinal cord between C2 and C3 segments with aging. The above research team successfully achieved adenovirus vector (Adv) mediated transfer of the brain-derived neurotrophic factor (BDNF) to the spinal accessory motor neurons between C1 and C3. They also demonstrated that the number of anterior horn neurons was significantly higher in the Adv-BDNF-transfected mice compared to the control. Two years later, the same team repeated their adenovirus-mediated retrograde transfer with the neurotrophin-3 gene. Once again, mice transfected with the Adv-NT-3 gene showed enhanced survival of anterior horn neurons in the Twy mice chronic cord compression model. Further immunohistochemical studies in 2012 revealed that there was a significant decrease in apoptosis and an increased presence of neurons and oligodendrocytes in the spinal cords of the Adv-BDNF transfected mice.

Several other substances have been investigated in relevant conditions and it may be reasonable to investigate their use regarding DCM. Through a rodent model of compressive thoracic myelopathy, Holly et al. demonstrated that rats fed a diet rich in docosahexaenoic acid and curcumin (DHA-Cur) maintained significantly higher tissue concentrations of spinal cord BDNF both at the level of compression and in the region of lumbar enlargement than those that did not receive dietary treatment. An animal model of DCM underwent either a diet rich in DHA and curcumin or a standard Western diet. The omega-3 fatty acid DHA has demonstrated therapeutic poten-

https://doi.org/10.14245/ns.2448140.070
tial secondary to its effects in decreasing inflammation and providing building material to plasma membranes and to its effects on overall neuronal function.\textsuperscript{69} Curcumin is a naturally occurring chemical compound found in turmeric, which has anti-inflammatory and antioxidant properties.\textsuperscript{70,71} Together, they might have beneficial effects on neuronal function. Gait analysis revealed significantly improved function in the DHA-Cur group. Moreover, levels of syntaxin-3 were elevated, and levels of lipid peroxidation (4-HNE) were decreased in the DHA-Cur group, suggesting the neural repair potential of this dietary regimen in DCM.

It has been suggested that autophagy promotes neuronal survival under hypoxic conditions.\textsuperscript{72} The expression levels of p62, ubiquitinated proteins, and LC3 in mice models of spinal cord compression were examined by Tanabe et al.\textsuperscript{72} and p62 was expressed in neurons, axons, astrocytes, and oligodendrocytes under hypoxic stress. These results suggested according to the authors that autophagy induction can be another potential therapeutic target for patients with spinal cord compression or DCM that has not been further investigated so far.

The most common mechanisms implicated in the pathobiology of DCM encompass apoptosis, inflammation, and vascular changes.\textsuperscript{70} Considering the vascular nature of DCM, one study\textsuperscript{73} examined if renin-angiotensin system inhibitors or other anti-hypertensors were associated with preoperative functional status and imaging markers of spinal cord compression. In their retrospective study of 266 patients, 37 patients were taking angiotensin-II receptor blockers, 44 were taking angiotensin-converting enzyme inhibitors, and 61 were taking other medications. Patients with hypertension presented with poor preoperative functional status measured using mJOA and Nurick scores (p < 0.01). Moreover, patients with hypertension who were treated with renin-angiotensin system inhibitors (specifically, angiotensin-II receptor blockers) had decreased T2-weighted signal intensity change compared to the untreated patients without hypertension (p = 0.04).

Ossification of the posterior longitudinal ligament (OPLL) is a disease that can lead to DCM. Recent studies have demonstrated that the use of H2-receptor antagonists, that are primarily used in the treatment of gastroesophageal reflux disease, may also be effective in treating heterotopic ossification.\textsuperscript{74,75} The therapeutic potential of famotidine was evaluated with a mice model for OPLL.\textsuperscript{76} The results of this study showed that administration of famotidine suppressed the progression of the ossification and reduced mortality when administered early in the development of the ossification. Furthermore, Liu et al.\textsuperscript{77} examined the suppression by famotidine of osteogenic differentiation in mesenchymal stem cells in patients with OPLL. Patients with DCM and OPLL were treated with famotidine and their ligamentum flavum was later collected during cervical spine surgery. Famotidine at a dose of 100 nM for 3 weeks strongly suppressed mRNA production of multiple osteogenic markers by the mesenchymal stem cells.

While the discussed treatments have potential benefit in the treatment of DCM, they are associated with side-effects and adverse events that should be thoroughly considered. For instance, Riluzole may cause dizziness, gastrointestinal disturbances, hepatotoxicity, and asthenia in ALS.\textsuperscript{78} Moreover, its administration requires careful monitoring since it can induce hypersensitivity reactions.\textsuperscript{79} Limaprost, an oral prostaglandin E1 analog, has been associated with gastrointestinal disturbances, flushing, and vertigo.\textsuperscript{80} Its vasodilatory effects require careful monitoring in patients with underlying cardiovascular conditions. The use of glucocorticoids, notably dexamethasone in the context of DCM, is accompanied by immunosuppression, delayed wound healing, blood glucose elevations, and a decrease in bone mineral density.\textsuperscript{81,82} Its administration significantly delayed fusion rates at 6 months in a randomized controlled trial of patients undergoing ACDF.\textsuperscript{24} While EPO exhibits antiapoptotic effects, its administration increases the risk of hypertension, thrombosis, hyperviscosity, and anaphylactic reactions.\textsuperscript{83} Although Cilostazol has demonstrated neuroprotective effects in animal models, its systemic vasodilatory mechanisms can lead to serious adverse events, including headaches, palpitations, and diarrhea.\textsuperscript{84} Its antithrombotic properties increase bleeding risks, which can be a concern in the perioperative setting of a patient with DCM. The adverse events reported in the included human studies are outlined in Table 4. Overall, the risk-benefit ratio of the pharmacological interventions is a complex interplay of their therapeutic benefits against their side-effects profiles. While animal models provide valuable insight into the neuroprotective properties of these drugs, the translation of these findings to human subjects requires rigorous clinical trials to ascertain their efficacy and safety profiles. Moreover, the majority of findings in human subjects stem from clinical trials with short follow-up or small sample size. To properly inform clinical decisions for the pharmacological treatment of DCM, larger clinical trials with diverse populations are warranted.

The balance of evidence surrounding riluzole’s efficacy remains uncertain. While preclinical studies, such as those by Moon et al.\textsuperscript{15} and Karadimas et al.\textsuperscript{16} have shown neurological
<table>
<thead>
<tr>
<th>Study</th>
<th>Dose</th>
<th>Adverse events</th>
<th>Significant difference improvement in JOA score</th>
<th>Level of evidence†/ sample size</th>
<th>Strength of evidence‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPO (Eryilmaz et al.⁹)</td>
<td>EPO: 3,000 U/kg IV 30 minutes preoperatively spinal cord decompression Methylprednisolone: 30 mg/kg IV 30 minutes preoperatively spinal cord decompression</td>
<td>NS</td>
<td>Yes</td>
<td>1B/n = 110</td>
<td>High</td>
</tr>
<tr>
<td>Limaprost alfadex (Sugawara et al.²³)</td>
<td>15 μg PO QD for 3 months</td>
<td>Vertigo, disequilibrium, lightheadedness</td>
<td>Yes</td>
<td>2B/n = 21</td>
<td>Moderate</td>
</tr>
<tr>
<td>Riluzole (Rajasekaran et al.¹⁷/ Fehlings et al.¹⁸ [CSM-Protect])</td>
<td>50 mg PO BID for 1 month</td>
<td>NS</td>
<td>No*</td>
<td>1B/n = 30</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>50 mg PO BID for 14 days preoperatively and then for 28 days postoperatively</td>
<td>Neck or arm or shoulder pain, arm paraesthesia, dysphagia, and worsening of myelopathy</td>
<td>No*</td>
<td>2B/n = 290</td>
<td>Moderate</td>
</tr>
<tr>
<td>Cerebrolysin (Allam et al.¹⁹/ Sharma et al.²⁰)</td>
<td>5 mL IM QD for 5 days/week for 4 weeks</td>
<td>Headache, dizziness, rash</td>
<td>Yes</td>
<td>1B/n = 192</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>5 mL IV QD diluted in 100 mL 0.9% NaCl over 30 minutes for 21 days postoperatively</td>
<td>Headache &amp; dizziness</td>
<td>No</td>
<td>2B/n = 60</td>
<td>Moderate</td>
</tr>
<tr>
<td>G-CSF (Sakuma et al.²³)</td>
<td>5 μg/kg IV QD or 10 μg/kg IV QD for 5 consecutive days</td>
<td>Surgical site infection</td>
<td>Yes**</td>
<td>4/n = 15</td>
<td>Very low</td>
</tr>
<tr>
<td>Glucocorticoids (Blume et al.²²/ Jeyamohan et al.²⁴)</td>
<td>40 mg IV intraoperatively</td>
<td>Wound healing complications</td>
<td>No*</td>
<td>2B/n = 49</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>0.2 mg/kg IV intraoperatively Postoperative doses: dexamethasone 0.06 mg/kg IV Q6H for the first 24 hours</td>
<td>Delayed fusion rates at 6 months</td>
<td>No*</td>
<td>2B/n = 112</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

JOA, Japanese Orthopaedic Association; EPO, erythropoietin; NS, not specified; PO, orally; QD, once a day; BID, twice a day; IM, intramuscular; IV, intravenous; Q6H, every 6 hours 6. *Modified JOA score used instead of JOA score. **No p-value reported. †No mention of loss to follow-up. ‡Pilot study. §From the Centre for Evidence-Based Medicine, http://www.cebm.net. (see Table 1). ¶Definition of levels of evidence (LoE) and overall strength of evidence (SoE). Global Spine J 2015;5:262.
improvements in rodent models, these findings have not been consistently replicated in human trials. Most notably, the CSM-PROTECT trial and the double-blinded, placebo-controlled randomized controlled trial by Rajasekaran et al. did not demonstrate clinical neurological improvement following administration of riluzole. Thus, its clinical applicability remains uncertain, requiring further investigation.

The optimal clinical use of cerebrolysin in the context of DCM remains controversial. Allam et al. demonstrated significant improvement in JOA scores over a 6-month period for the nonoperative management of DCM. However, Sharma et al. did not find significant differences in perioperative JOA and VAS scores. There was improvement in hand function at one year. Considering the conflicting results in these randomized controlled trials, careful consideration must be taken whether cerebrolysin is administered in a perioperative or nonoperative context. Future studies should elaborate on the optimal conditions of its administration in DCM patients.

Consistent findings across human studies indicate either improvement or worsening of functional outcomes with each medication. However, studies using glucocorticoids for the management of neurological symptoms presented conflicting results. Two of these studies were used in the perioperative setting and observed no significant difference in myelopathy neurological symptoms scores between groups administered with the glucocorticoid (dexamethasone). However, the third study used a glucocorticoid (methylprednisolone) in combination with EPO preoperatively and found a significant improvement in JOA score and 40-point rating scale compared to the control group. Thus, the effects of the glucocorticoid could be masked by the EPO in this study. Glucocorticoids suggest improvement in locomotor recovery postdecompression surgery in animal models. However, such neurological benefits lack applicability in the clinical setting, as demonstrated by a retrospective cohort study and a randomized controlled trial. Glucocorticoids delayed fusion rates at 6 months and increased surgical wound infection rates. Considering the systemic side-effects associated with glucocorticoids and its inability to provide neurological improvement to DCM patients, caution is advised, especially in patients managed operatively.

Limaprost alfadex improved motor function in rodent models and in a phase 3 randomized double-blinded clinical trial. The results from the latter study should be interpreted with caution considering the small sample size and the absence of control group. The current evidence supports continued investigation (a phase 3 double-blind randomized control trial is currently underway [ClinicalTrials.gov Identifier: NCT02125981]).

EPO has demonstrated neuroprotective effects in both animal and human studies. Clinical trials showed improved quality of life outcomes and JOA scores when combined with methylprednisolone following decompression surgery in DCM patients. The combination with methylprednisolone is a confounding factor, however, regarding the efficiency of EPO per se.

Cilostazol has shown significant preservation of motor function and neural tissue integrity in preclinical studies. These findings suggest that this pharmacological agent holds promise for DCM in animal models. To our knowledge, there are no clinical trials in human subjects that evaluate its efficacy, thus lacking in clinical applicability. Similarly, JSKL exhibits neuroprotective and analgesic effects in animal models, but lacks evidence in humans. Anti-Fas ligand antibody also lacks trials in humans. Only a postmortem and a preclinical mouse study of anti-Fas ligand antibody demonstrated a reduction in inflammation and an improvement in neurological recovery.

G-CSF restored damaged spinal cord tissue and recovered neural function in rats and mice. It was demonstrated in a phase I and IIa clinical trial that G-CSF improved motor and sensory outcomes in patients with worsening symptoms of compressive cervical myelopathy. However, this was an open-label trial without a control group and a small sample size (cohort of 15 patients).

Of the treatment regimens that demonstrated significant improvement in mJOA score, administration of EPO with methylprednisolone was supported by level 1B (individual randomized controlled trial with narrow confidence interval) and high strength of evidence, limaprost alfadex by level 2B (cohort study with small sample size) and moderate strength of evidence. G-CSF was supported by level 4 (case series) and very low strength of evidence. Regarding cerebrolysin, one study demonstrated significant improvement in mJOA score, administration of EPO with methylprednisolone was supported by level 1B (individual randomized controlled trial with narrow confidence interval) and high strength of evidence, limaprost alfadex by level 2B (cohort study with small sample size) and moderate strength of evidence. The pharmacological agents that did not demonstrate improvement in mJOA score were supported by either level 1B (riluzole) or 2B (riluzole, cerebrolysin, glucocorticoids) studies with moderate strength of evidence. The pharmacological agents that did not demonstrate improvement in mJOA score were supported by either level 1B (riluzole) or 2B (riluzole, cerebrolysin, glucocorticoids) studies with moderate strength of evidence. The pharmacological agents that did not demonstrate improvement in mJOA score were supported by either level 1B (riluzole) or 2B (riluzole, cerebrolysin, glucocorticoids) studies with moderate strength of evidence. The pharmacological agents that did not demonstrate improvement in mJOA score were supported by either level 1B (riluzole) or 2B (riluzole, cerebrolysin, glucocorticoids) studies with moderate strength of evidence. The pharmacological agents that did not demonstrate improvement in mJOA score were supported by either level 1B (riluzole) or 2B (riluzole, cerebrolysin, glucocorticoids) studies with moderate strength of evidence.
The level of evidence of the included studies was classified according to the Levels of Evidence for Therapeutic Studies from the Centre for Evidence-Based Medicine (https://sebm.net). The appraisal is outlined in Table 4 with the criteria for levels of evidence in Table 1. We summarized the level and strength of evidence and the functional outcome improvement of the discussed pharmacological agents in Fig. 2.

LIMITATIONS AND FUTURE DIRECTION

Despite the promise of the pharmacological treatment options for DCM, several limitations hinder the widespread adoption of these treatments. First, half of the studies investigating therapeutic drugs for the treatment of DCM are in animal models. Although this is a pivotal step before pursuing clinical trials in humans, the current studies in humans only include small sample sizes and/or short follow-up, which compromised the reliability of the results. While animal models provide a setting for preliminary experimentation of new pharmacological agents, the results from these studies present several clinical drawbacks. Pharmacokinetics differ between humans and animals, making it more challenging to evaluate potential side-effects in animal studies. Moreover, the pathophysiology of DCM differs between humans and animals, including variations in the composition of peripheral leukocytes. Additionally, the presence of comorbidities can confound results in human trials, considerations for which are generally less extensive in animal studies. The generalizability of lower quality evidence presents an obstacle in translating scientific findings into clinical practice. That said, these results should be interpreted with caution due to the potential presence of significant intrinsic bias within each study. Also, the variation in study settings is significant; some studies present results in the perioperative setting, while others do not take surgery into consideration. By aggregating these studies, we aimed to provide a comprehensive overview of the available evidence for clinicians to appraise and tailor to their practice. In addition, these studies in humans often lack control groups or standardized comparators. Second, the studies investigating the pharmacological treatment of DCM have different endpoints since the medication is initiated at different periods (e.g., pre-operative, postoperative, etc.) and different kinds of evaluation. Third, the safety profiles of the available drugs and their interactions are not fully elucidated. As demonstrated in preclinical studies, intrinsic patient characteristics can alter the therapeutic mechanism of a given medication. Lastly, the lack of standardized treatment protocols and guidelines for the medical management of DCM impedes evidence-based practices. Addressing these limitations will be critical to translating experimental findings to clinical practice for patients with DCM.

Our review highlights the heterogeneity in clinical outcomes observed in DCM studies. Future trials should aim to standardize clinical endpoints and incorporate patient-reported outcomes to facilitate a comprehensive evaluation of these interventions. Considering the extensive array of potential drugs available for the management of DCM, maintaining consistent outcomes will enable more robust comparisons among the drugs. Also, considering the potential variations in how these drugs interact with DCM patients, there is a need for rigorous pharmacoki-
netic and pharmacodynamic studies to understand the inter-
play of genetic predispositions and drug response. Only after
the efficacy and safety profiles of these drugs have been estab-
lished in large trials should standardized treatment protocols
and guidelines be developed through consensus among experts
and evidence-based practice.

Current guidelines\(^8\) recommend the surgical management of
patients with moderate (mJOA 12–14) to severe (mJOA ≤ 11)
DCM, while there is controversy surrounding the management
of patients with mild DCM (mJOA 15–17). Considering that
the pathophysiology of DCM is so diverse, several treatment
modalities have been conceived. That said, other than standard
pain management and physical rehabilitation, no standardized
guidelines on the nonoperative management of DCM exist. The
discussed pharmacological agents were evaluated either in the
perioperative or the nonoperative setting. In the nonoperative
setting, limaprost alfadex (15 μg PO QD for 3 months; level 2B
and moderate strength of evidence) and cerebrolysin (5 mL in-
tramuscularly QD for 5 days/wk for 4 weeks; level 1B and high
strength of evidence) could be recommended. EPO combined
with methylprednisolone (3,000 U/kg IV EPO and 30 mg/kg
IV of methylprednisolone 30 minutes preoperatively decom-
pression; level 1B and high strength of evidence) could be rec-
commended in the perioperative setting, while its effects in the
nonoperative setting have not been studied.

For patients with mild DCM, it would be of interest to know
whether pharmacologic treatment delays or prevents the need
for surgery. Furthermore, it is important to elaborate on the
postoperative effects of these drugs on patients. Exploring wheth-
er the pharmacologic treatment under investigation delays the
clinical manifestations of DCM in patients with concordant ra-
diologic or neurophysiologic findings is of significant interest.

CONCLUSION

This article presents a comprehensive review of the current
status of evidence regarding the pharmacological management
of DCM. Efforts aim at addressing inflammatory, vascular, and
apoptotic pathways. EPO (combined with methylprednisolone),
limaprost alfadex and G-CSF report neurological improvements
in patients with DCM. Studies administering riluzole or cere-
bolysin present conflicting results. Glucocorticoids should be
avoided as they increase infection rates and delay fusion. Anti-
Fas ligand antibody, cilostazol, and JSKL have demonstrated
promising results only in animal models. Further translational
research ought to be conducted under multidisciplinary collab-
orations utilizing molecules that have demonstrated therapeutic
potential in animal models. Additionally, robustly designed
clinical studies are imperative to thoroughly explore the clinical
outcomes associated with the aforementioned medications.

NOTES

Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This study received no specific grant from
any funding agency in the public, commercial, or not-for-profit
sectors.

Author Contribution: Conceptualization: MG; Data cura-
tion: J JL, MG, SM, MVA; Formal analysis: J JL, MG; Methodol-
ogy: J JL, MG, CS, MHW, JDG, PJ, JAO, G K, AKD; Project ad-
ministration: MG, CS, MHW, JDG, PJ, JAO, G K, AKD; Visual-
ization: J JL, SM, NAS, MVA; Writing – original draft: J JL, MG,
SM, WAM, NAS, MVA; Writing – review & editing: CS, MHW,
JDG, PJ, JAO, G K, AKD.

ORCID
Jordan J Levett: 0000-0002-1999-2017
Miltiadis Georgiopoulos: 0000-0001-8656-1412
Andreas K. Demetriades: 0000-0002-2004-9448

Acknowledgements
“With profound grief and sorrow, we regret to inform you
that our co-author and friend Dr. Georgios Klironomos passed
away unexpectedly on August 24, 2023, at the young age of
48 years. Before I present a short professional biography of Dr.
Klironomos, I want to highlight that Dr. Klironomos was not
just an outstanding neurosurgeon, but a real gem as a person
with a kind personality and a smile that always touched our
hearts. He had been a dear friend and a mentor, eager to help
and listen, and a role model to many of us as a father, a husband,
and a great doctor-neurosurgeon. He was born in a small vil-
vage in a small country, i.e., Talanta in Greece, but made it to some
of the best institutions and hospitals worldwide. Despite his ex-
ceptional career, he would always tell me that nothing else can
be important without the happiness and love of our families.
It is not an exaggeration to claim that everyone loved and ad-
mired George. He was very charismatic but also a very simple
person at the same time. Surely his lovely daughter and wife are
deeply sad, but I am sure that they are also greatly proud of
George. Personally, I can not yet accept the reality when it comes
to George’s passing. I wish I could hear his voice again greeting
me in a typical Greek manner befitting close friends.”
Dr. Klironomos was an Assistant Professor of neurosurgery at the Donald and Barbara Zucker School of Medicine at Hofstra/Northwell in Hempstead, NY, and an attending Neurosurgeon at Northwell Health, NY, mainly at the South Shore University Hospital, with a subspecialization in neuro-oncology, skull base, and cerebrovascular neurosurgery. He was awarded as the top neurological surgeon in Long Island, New York for 2023. Before that, he was a Clinical Assistant Professor in Neurosurgery at the University of Pittsburgh. Moreover, he received extensive fellowship training at a high level in cerebrovascular and skull base neurosurgery at North Shore/Lenox Hill University Hospital in New York and in neurooncology and skull base (incl. endoscopic) neurosurgery at Toronto Western Hospital of the University of Toronto in Canada. Additionally, he also had other positions throughout his career, including research roles, and participated in many publications in major journals. He had a Ph.D. and an M.Sc. degree from the University of Patras, Greece, where he also completed his undergraduate studies in medicine with the 2nd highest grade among his class.

The whole team would like to honor George for his contribution to neurosurgery and medicine and offer our sincere condolences to his family.

REFERENCES

14. Definition of levels of evidence (LoE) and overall strength of evidence (SoE). Global Spine J 2015;5:539.
18. Fehlings MG, Badhiwala JH, Ahn H, et al. Safety and effica-
28. Doggrell SA. A neuroprotective derivative of erythropoietin
26. Eryilmaz F, Farooque U. The efficacy of combined medica
19. Allam AFA, Abotakia TAA, Koptan W. Role of Cerebrolysin
15. Costantini S, Malerba L, Rovini P, et al. Riluzole treatment
10. Yamamoto S, Kurokawa R, Kim P. Cilostazol, a selective
colony-stimulating factor improves motor function in rats
colony-stimulating factor improves motor function in rats
cients notoginsenoside R1 and ginsenoside Rb1 alleviate the
symptoms of cervical myelopathy through Kir3.1 mediated mechanisms. CNS Neurol Disord Drug Targets 2019;18:
631-42.
colony-stimulating factor improves motor function in rats
therapy using granulocyte colony-stimulating factor for pa
tients with worsening symptoms of compression myelopa
thy, part I: a phase I and IIa clinical trial. Eur Spine J 2012;
lotic myelopathy: conservative versus surgical treatment af
0. Halvorsen CM, Lied B, Harr ME, et al. Surgical mortality
and complications leading to reoperation in 318 consecutive
posterior decompressions for cervical spondylotic myelopa
ative cervical myelopathy and the physiology of recovery
myelopathy: insights into its pathobiology and molecular
riluzole inhibition of glutamate release from rat cerebral
cortex nerve terminals (synaptosomes). Neuroscience 2004;
125:191-201.
effects of sodium channel blockers after spinal cord injury:
improved behavioral and neuroanatomical recovery with
5. Masliah E, Diez-Tejedor E. The pharmacology of neurotroph
ic treatment with Cerebrolysin: brain protection and repair
to counteract pathologies of acute and chronic neurological


Karlo M. Pedro\textsuperscript{1,2,3}, Michael G. Fehlings\textsuperscript{1,2,3}

\textsuperscript{1}Division of Neurosurgery, Krembil Neuroscience Centre, University Health Network, Toronto, ON, Canada
\textsuperscript{2}Institute of Medical Science, University of Toronto, Toronto, ON, Canada
\textsuperscript{3}Division of Neurosurgery and Spine Program, Department of Surgery, University of Toronto, Toronto, ON, Canada

Degenerative cervical myelopathy (DCM) is the commonest form of spinal cord impairment in adults worldwide.\textsuperscript{1} While surgery is largely effective in improving spinal cord function and reducing pain, many patients do harbour residual neurological impairment and/or have residual neuropathic pain.\textsuperscript{2} The RECODE-DCM project, supported by AO Spine, identified that the development of neuroprotective and neuroregenerative treatments for DCM was an important priority.\textsuperscript{3} Given this, the recent article by Levett et al.,\textsuperscript{4} which reviewed emerging pharmacological approaches for DCM, is timely. The authors reviewed 18 studies based on a literature review up to April 2022, focusing on pharmacological interventions for DCM. They reported that erythropoietin (EPO) (with methylprednisolone) showed promising results based on a randomized controlled study utilizing a three-month quality of life (QoL) outcome. The sodium-glutamate antagonist riluzole, based on a Phase 3 multicenter randomized controlled trial (RCT), failed to achieve significance in enhancing the modified Japanese Orthopaedic Association (mJOA) score, largely because of ceiling effects in the outcome measure, but did show signal in a number of secondary outcomes including neck and arm pain. Moreover, as summarized below, a new reanalysis of the cervical spondylotic myelopathy (CSM)-PROTECT RCT using a global statistical approach—an article which came out after the review by Leveltt et al., has shown that riluzole confers a significant benefit to patients with DCM undergoing decompressive surgery.\textsuperscript{5}

Narrative reviews serve as important repositories of distilled knowledge, but synthesizing information and making recommendations based on this type of literature should be approached with caution. The strength of evidence is more appropriately synthesized through meta-analysis or systematic reviews rather than narrative expositions. In this article, several critical factors require closer examination. For instance, the limaprost study, a prospective,
nonrandomized controlled trial, and the cerebrolysin study, which presents conflicting results from Allam et al.\textsuperscript{4} and Sharma et al.,\textsuperscript{7} should receive a lower strength of evidence than the riluzole study. The EPO+methylprednisolone study was graded the highest in terms of evidence; however, this rating should be downgraded due to the inadequacy of the control group (i.e., methylprednisolone only), as it lacked a placebo group (i.e., no intervention) necessary to discern the true treatment effect. Similar concerns apply to the cerebrolysin study by Allam et al.,\textsuperscript{4} where the treatment combined cerebrolysin and celecoxib, making it impossible to assess the pure effect of the intervention with only a single control group. Additionally, except for the cerebrolysin, glucocorticoid, and riluzole studies, all other clinical trials did not include a formal sample size computation. This raises uncertainties about whether the study populations were sufficient to detect true effect sizes or were merely chosen for convenience. Moreover, the representation of evidence in an inverted pyramid in Fig. 2 distorts the overall picture by combining clinical and pre-clinical studies, the latter of which are not covered by the GRADE approach.

Adverse effect assessments are crucial when evaluating medical interventions for DCM. While extensively reported in the CSM-PROTECT trial, it was noticeably absent in the studies on limparost and EPO+methylprednisolone. The lack of comprehensive adverse effect data in these studies renders a thorough evaluation of both intervention efficacy and safety impossible. Short snippets of these limitations were mentioned in the narrative review but did not adequately reflect the caution necessary when interpreting these results.

The application of the GRADE approach entails evaluating the most critical outcome, which determines the overall quality of evidence. For instance, if QoL is prioritized, the study on EPO provides the best evidence, whereas the cerebrolysin study offers the highest evidence of treatment effect when the mJOA score is considered crucial. This issue presents a fundamental challenge in spinal cord injury trials, as there is no consensus on the most appropriate and comprehensive outcome scale to represent recovery after surgery. Although the mJOA is commonly used for assessing outcomes in DCM, its limitations are well documented. Notably, its tendency to reach a ceiling effect in mild DCM and its inability to represent the multidimensional aspects of recovery make it inadequate as a primary trial endpoint. For example, neck pain, now recognized as a top recovery priority in DCM patients, is not adequately captured by the solitary use of the mJOA.\textsuperscript{8} Similarly, QoL and functional recovery remain unrepresented within the criteria of the mJOA.

This challenging issue must be addressed using innovative solutions and alternative analytical approaches. The global statistical test has shown promise in neurology trials by effectively representing multiple outcomes scales in a cohesive analysis.\textsuperscript{9} This nonparametric test consolidates various outcome measures, providing a comprehensive evaluation of the treatment effect between drug and placebo groups and computes a global summary measure. Its recent application in the reanalysis of the CSM-PROTECT trial demonstrates its feasibility and increased power to detect treatment effects, especially when the treatment positively impacts multiple outcome scales.\textsuperscript{10} Alternatively, longitudinal clustering techniques using unsupervised machine learning algorithms (e.g., k-means or hierarchical clustering), growth modelling, or latent class analysis may reveal symptom clusters that better capture the heterogeneity of the DCM population. Implementing these models using big data in DCM shows promise in identifying patient phenotypes with more granular features than those described using the mJOA scale.\textsuperscript{11}

In our opinion, this narrative review effectively summarizes the current state of knowledge on potential therapeutic options for DCM. However, the heterogeneous outcomes, varying follow-up periods, and the limitation imposed by current outcome scales in DCM necessitate caution when synthesizing data from diverse clinical trials. Emerging therapies, including neuroregenerative strategies (e.g., cell, growth, or small molecule-based) and neuromodulatory techniques (e.g., spinal cord, transcranial or neuromuscular stimulation), also warrant evaluation alongside pharmacologic interventions.\textsuperscript{3} Potential genetic modifiers of treatment response, such as the presence of the ApoE4 allele, will enable stratified analysis of interventions in the future.\textsuperscript{12} As emphasized in this article, two emerging concepts are anticipated to have a major impact on spine surgery clinical trials: (1) the adaptation of a multivariable approach in assessing outcomes in spine trials and (2) secondary analysis using a global statistical approach to reveal latent beneficial treatment effects. The latter is exemplified by a secondary reanalysis of the CSM-PROTECT trial, in which riluzole has shown a significant global benefit in patients with DCM undergoing surgical treatment.\textsuperscript{5} Based on these new data, clinicians may wish to consider the use of riluzole in the management of patients with DCM undergoing surgery to enhance neurological outcomes and reduce pain.

- **Conflict of Interest:** MGF is supported by the Robert Campeau Family Foundation/Dr. C.H. Tator Chair in Brain and Spinal Cord Research. Except for that, the authors have no conflicts to declare.
REFERENCES


Baseline Frailty Measured by the Risk Analysis Index and 30-Day Mortality After Surgery for Spinal Malignancy: Analysis of a Prospective Registry (2011–2020)

Rachel Thommen1,2, Christian A. Bowers1, Aaron C. Segura2, Joanna M. Roy3, Meic H. Schmidt2

1School of Medicine, New York Medical College, Valhalla, NY, USA
2Bowers Neurosurgical Frailty and Outcomes Data Science Lab, Sandy, UT, USA
3Topiwala National Medical College, Mumbai, India

Objective: To evaluate the prognostic utility of baseline frailty, measured by the Risk Analysis Index (RAI), for prediction of postoperative mortality among patients with spinal malignancy (SM) undergoing resection.

Methods: SM surgery cases were queried from the American College of Surgeons – National Surgical Quality Improvement Program database (2011–2020). The relationship between preoperative RAI frailty score and increasing rate of primary endpoint (mortality or discharge to hospice within 30 days, “mortality/hospice”) were assessed. Discriminatory accuracy was assessed by computation of C-statistics (with 95% confidence interval [CI]) in receiver operating characteristic (ROC) curve analysis.

Results: A total of 2,235 cases were stratified by RAI score: 0–20, 22.7%; 21–30, 11.9%; 31–40, 54.7%; and ≥ 41, 10.7%. The rate of mortality/hospice was 6.5%, which increased linearly with increasing RAI score (p < 0.001). RAI was also associated with increasing rates of major complication, extended length of stay, and nonhome discharge (all p < 0.05). The RAI demonstrated acceptable discriminatory accuracy for prediction of primary endpoint (C-statistic, 0.717; 95% CI, 0.697–0.735). In pairwise ROC comparison, RAI demonstrated superiority versus modified frailty index-5 and chronological age (p < 0.001).

Conclusion: Preoperative frailty, as measured by RAI, is a robust predictor of mortality/hospice after SM surgery. The frailty score may be applied in clinical settings using a user-friendly calculator, deployed here: https://nsgyfrailtyoutcomeslab.shinyapps.io/spinalMalignancyRAI/.

Keywords: Frailty, Risk Analysis Index, Spinal tumor, Metastatic, Spinal oncology, National Surgical Quality Improvement Program

INTRODUCTION

Spinal malignancies (SMs) are often associated with significant morbidity and reduced quality of life from neurological deficits and/or fracture-associated pain. These patients commonly present with pain and/or focal neurologic deficits. Spine metastases were found to be the most common osseous destination for secondary malignancy (70%).1,2 Recent literature reported 15.67% of patients diagnosed with a solid tumor have spinal metastasis; 9.6% of patients with SM will develop epidural spinal cord compression and 12.6% will experience pathologic compression fractures.3 Although primary central nervous
system malignancies are less common, 2.9% occur in the spinal cord or cauda equina regions.\(^4\) Despite a low incidence of primary spinal neoplasms (0.77 per 100,000), their incidence peaks at the 75–84 year age group (1.80 per 100,000 in individuals), which is relevant to research studying frailty’s impact on SM surgery.\(^5\)

There is an increased risk of developing primary malignancy with increasing age, and people are living longer thanks to advancements in medical science and cancer management. Due to improved rates of survival of cancer patients, there is a higher incidence of older patients presenting with a greater cancer disease burden with spinal metastasis, in addition to the peak incidence of primary SMs that accompanies increasing patient age until the 75–84 year group.\(^6\) Therefore, it is becoming even more important to accurately measure preoperative risk and quantify frailty as a measure of physiologic reserve, as surgical outcomes prediction cannot be based on age alone for risk stratification.\(^7-10\)

The majority of previous frailty studies in the neurosurgical literature have used the 11- or 5-modified frailty index (mFI), the mFI-11 or mFI-5, respectively. The Risk Analysis Index (RAI) is a robust frailty index originally developed and validated in surgical populations.\(^11-13\) The RAI is uniquely versatile for both clinical prospective application with a patient-centered questionnaire and for large retrospective database analysis. However, the generalizability of RAI to neurosurgical spinal oncology is currently unknown.

The objective of the present study is to evaluate the predictive ability of the RAI for 30-day mortality and other adverse postoperative outcomes in patients undergoing surgical intervention for a SM using data from a large multicenter, clinical surgical registry representing over 700 hospitals across 49 U.S. states and 11 countries.

**MATERIALS AND METHODS**

1. **Study Design**

   The present study, designed as a secondary analysis of a prospective surgical registry, was conducted in accordance with STROBE (Strengthening the Reporting of Observational studies in Epidemiology) guidelines.

2. **Data Source**

   The patient population was derived from the American College of Surgeons-National Surgical Quality Improvement Program (ACS-NSQIP) database. The ACS-NSQIP is a prospective, peer-controlled, validated database for quantifying 30-day surgical outcomes. NSQIP data are prospectively entered at each participating hospital by ACS-trained surgical clinical reviewers to improve consistency and reliability. The present study was performed under the data user agreement of the ACS with our institution and was approved by Institutional Review Board of University of New Mexico (Study ID 21-315). Given the de-identified nature of the information in NSQIP database, patient consent was neither sought nor required.

3. **Participants**

   The malignant spinal tumor patient population was queried from the NSQIP (2011–2020) using a combination of Current Procedural Terminology (CPT) codes and International Classification for Disease (ICD), 9th and 10th Revision, Clinical Modification codes (Table 1). Initial query using exclusively primary CPT codes yielded 7,851 records. The postoperative diagnoses fields in NSQIP (“PODIAG”) further narrowed the search to 2,235 records using prespecified ICD codes (Table 1).

4. **Predictor Variables**

   The primary predictor of interest was preoperative frailty, measured by the RAI score. The RAI was originally developed and validated for compatibility with the ACS-NSQIP database.\(^12\) RAI has been previously utilized to predict outcomes in patients undergoing spine surgery.\(^14,15\) The revised RAI score, as described by Arya et al.,\(^12\) was computed using standard NSQIP variables: age, sex, disseminated cancer (“DISCANC”), weight loss (“WTLOSS”), renal failure (“RENAFAIL; DIALYSIS”), congestive heart failure (“HXCHF”), shortness of breath (“DYSPEANEA”), functional status (“FNSTATUS2”), and living status (“TRANST” = nursing home or chronic care facility). The mFI-5, a readily utilized risk index in neurooncological literature, was computed and analyzed for comparative purposes alongside RAI.\(^7,16-18\) mFI-5 has a 5-point maximum score, with 1 point given to each positive variable. mFI-5 score was calculated from NSQIP using the cumulative score of nonindependent functional status (“FNSTATUS2”), diabetes mellitus (“DIABETES”), chronic obstructive pulmonary disease (“HXCOPD”), and congestive heart failure (“HXCHF”).

5. **Primary and Secondary Outcomes**

   The primary outcome for this study was mortality within 30 days of surgery and/or discharge to hospice, which was calculated using the NSQIP variables “DOPERTOD” (days from operation to death for mortality) and “DISCHDEST” (hospital discharge destination). The secondary outcomes were non-
home discharge disposition (NHD), extended length of hospital stay (ELOS), major complication occurrence, and Clavien-Dindo grade IV complication occurrence (CD IV). CD IV complications include life-threatening complications that require intensive care unit management. Major complication occurrences were defined by one or more instances of prolonged intubation (defined as greater than 48 hours), unplanned reintubation, sepsis, septic shock, pneumonia, deep venous thrombosis/thrombophlebitis, pulmonary embolism, cerebrovascular accident/stroke with neurological deficit, acute renal failure, myocardial infarction, cardiac arrest requiring cardiopulmonary resuscitation, superficial surgical site infection (SSI), deep incisional SSI, organ space SSI, or wound dehiscence.

6. Statistical Methods

Statistical analyses were performed with a combination of R Project for Statistical Computing ver. 4.2.1 (The R Foundation, Vienna, Austria) (https://www.R-project.org/), IBM SPSS Statistics ver. 28.0 (IBM Co., Armonk, NY, USA), GraphPad Prism ver. 9.0 (GraphPad Software Inc., La Jolla, CA, USA), and MedCalc for Windows ver. 19.4 (MedCalc Software, Ostend, Belgium). Statistical significance was set a priori to alpha of 0.05. The R gtsummary package was utilized to generate descriptive statistics for the spine tumor surgery cohort. RAI was analyzed as a continuous variable and categorized into a base “nonfrail” category of 0–20 and subsequent ascending 5-point bins (e.g., 21–25, 26–30). The binning scheme was defined a priori and arbitrary in nature to avoid scoring biases associated with tiers. The Cochran-Armitage trend test was utilized to determine significance of proportional trends (frailty score bin vs. binary outcomes). Discriminatory accuracy of frailty score for primary outcome was quantified using C-statistics (with 95% confidence interval [CI]) and interpreted using standard criteria per Hosmer-Lemeshow. A C-statistic > 0.7 was indicative of adequate discriminatory accuracy. The relative discriminatory superiority of RAI (vs. increasing chronological age and mFI-5) was assessed using pairwise receiver operating characteristic (ROC) comparison (DeLong) tests. The R, packages rms and shiny

Table 1. List of CPT, ICD-9-CM, and ICD-10-CM codes used to identify SM surgery cases from the NSQIP database 2011–2020

<table>
<thead>
<tr>
<th>Coding system</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>63275-63280</td>
<td>Laminectomy for biopsy/excision of extradural spinal neoplasm</td>
</tr>
<tr>
<td></td>
<td>63280-63283</td>
<td>Laminectomy for biopsy/excision of intradural extramedullary spinal neoplasm</td>
</tr>
<tr>
<td></td>
<td>63285-63287</td>
<td>Laminectomy for biopsy/excision of intradural intramedullary spinal neoplasm</td>
</tr>
<tr>
<td></td>
<td>63290</td>
<td>Laminectomy for biopsy/excision of combined intradural/extradural spinal neoplasm</td>
</tr>
<tr>
<td></td>
<td>63300-63303</td>
<td>Vertebral corpectomy for excision of extradural spinal neoplasm</td>
</tr>
<tr>
<td></td>
<td>63304-63307</td>
<td>Vertebral corpectomy for excision of intradural spinal neoplasm</td>
</tr>
<tr>
<td>ICD-9-CM</td>
<td>170.2</td>
<td>Malignant neoplasm of vertebral column excluding sacrum and coccyx</td>
</tr>
<tr>
<td></td>
<td>170.6</td>
<td>Malignant neoplasm of pelvic bones, sacrum, and coccyx</td>
</tr>
<tr>
<td></td>
<td>192.2</td>
<td>Malignant neoplasm of spinal cord</td>
</tr>
<tr>
<td></td>
<td>192.3</td>
<td>Malignant neoplasm of spinal meninges</td>
</tr>
<tr>
<td></td>
<td>198.3-198.5</td>
<td>Secondary malignant neoplasm of brain and spinal cord</td>
</tr>
<tr>
<td></td>
<td>733.13</td>
<td>Pathological fracture of vertebrae</td>
</tr>
<tr>
<td></td>
<td>239.7</td>
<td>Neoplasm of uncertain behavior other parts of central nervous system</td>
</tr>
<tr>
<td>ICD-10-CM</td>
<td>C41.2</td>
<td>Malignant neoplasm of vertebral column</td>
</tr>
<tr>
<td></td>
<td>C72</td>
<td>Malignant neoplasm of spinal cord</td>
</tr>
<tr>
<td></td>
<td>C79.49</td>
<td>Malignant neoplasm of other parts of central nervous system</td>
</tr>
<tr>
<td></td>
<td>C79.51</td>
<td>Secondary malignant neoplasm of bone</td>
</tr>
<tr>
<td></td>
<td>D43.4</td>
<td>Neoplasm of uncertain behavior of spinal cord</td>
</tr>
<tr>
<td></td>
<td>M48.50X</td>
<td>Pathological fracture in neoplastic disease, unspecified site</td>
</tr>
<tr>
<td></td>
<td>M84.68</td>
<td>Pathological fracture in other disease, other site</td>
</tr>
</tbody>
</table>

were used to generate an interactive Risk Analysis Index and Malignant Spinal Tumor Surgery Outcomes calculator found at the following website: https://nsgyfrailtyoutcomeslab.shinyapps.io/spinalMalignancyRAI/.

Table 2. Demographics, clinical characteristics, and baseline frailty in patients undergoing surgery for SM, subgroup by primary endpoint (mortality/hospice), ACS-NSQIP 2011–2020 (N = 2,235)

<table>
<thead>
<tr>
<th>Variable</th>
<th>All cohort</th>
<th>Primary endpoint (mortality/hospice)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients</td>
<td>2,235</td>
<td>2,099</td>
<td>136</td>
</tr>
<tr>
<td>Age (yr), median (IQR)</td>
<td>61 (52–70)</td>
<td>61 (52–70)</td>
<td>66 (58–73)</td>
</tr>
<tr>
<td>Female sex (biological)</td>
<td>865 (39)</td>
<td>822 (39)</td>
<td>43 (32)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>1,476 (66)</td>
<td>1,381 (66)</td>
<td>95 (70)</td>
</tr>
<tr>
<td>Black</td>
<td>218 (9.8)</td>
<td>204 (9.7)</td>
<td>14 (10)</td>
</tr>
<tr>
<td>Asian</td>
<td>98 (4.4)</td>
<td>94 (4.5)</td>
<td>4 (2.9)</td>
</tr>
<tr>
<td>Other</td>
<td>443 (20)</td>
<td>420 (20)</td>
<td>23 (17)</td>
</tr>
<tr>
<td>Hispanic ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>162 (7.2)</td>
<td>156 (7.4)</td>
<td>6 (4.4)</td>
</tr>
<tr>
<td>Body mass index (kg/m²), median (IQR)</td>
<td>27.0 (23.5–30.5)</td>
<td>27.0 (23.5–30.5)</td>
<td>26.0 (22.0–30.0)</td>
</tr>
<tr>
<td>Acute care transfer</td>
<td>423 (19)</td>
<td>385 (18)</td>
<td>38 (28)</td>
</tr>
<tr>
<td>RAI, median (IQR)</td>
<td>34 (26–42)</td>
<td>34 (26–42)</td>
<td>37 (34–41)</td>
</tr>
<tr>
<td>RAI categorical bins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–20</td>
<td>507 (23)</td>
<td>498 (24)</td>
<td>9 (6.6)</td>
</tr>
<tr>
<td>21–25</td>
<td>173 (7.7)</td>
<td>169 (8.1)</td>
<td>4 (2.9)</td>
</tr>
<tr>
<td>26–30</td>
<td>93 (4.2)</td>
<td>90 (4.3)</td>
<td>3 (2.2)</td>
</tr>
<tr>
<td>31–35</td>
<td>635 (28)</td>
<td>605 (29)</td>
<td>30 (22)</td>
</tr>
<tr>
<td>36–40</td>
<td>587 (26)</td>
<td>540 (26)</td>
<td>47 (35)</td>
</tr>
<tr>
<td>41–45</td>
<td>172 (7.7)</td>
<td>145 (6.9)</td>
<td>27 (20)</td>
</tr>
<tr>
<td>46–50</td>
<td>51 (2.3)</td>
<td>41 (2.0)</td>
<td>10 (7.4)</td>
</tr>
<tr>
<td>≥51</td>
<td>17 (0.8)</td>
<td>11 (0.5)</td>
<td>6 (4.4)</td>
</tr>
<tr>
<td>mFI-5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1,079 (48)</td>
<td>1,031 (49)</td>
<td>48 (35)</td>
</tr>
<tr>
<td>1</td>
<td>807 (36)</td>
<td>759 (36)</td>
<td>48 (35)</td>
</tr>
<tr>
<td>2</td>
<td>301 (13)</td>
<td>272 (13)</td>
<td>29 (21)</td>
</tr>
<tr>
<td>3</td>
<td>44 (2.0)</td>
<td>37 (1.8)</td>
<td>7 (5.1)</td>
</tr>
<tr>
<td>4</td>
<td>4 (0.2)</td>
<td>0 (0)</td>
<td>4 (2.9)</td>
</tr>
</tbody>
</table>

Values are presented as number of patients (%) unless otherwise indicated. SM, spinal malignancy; ACS-NSQIP, American College of Surgeons – National Surgical Quality Improvement Program; IQR, interquartile range; RAI, Risk Analysis Index; mFI-5, modified frailty index-5.

2. RAI and Mortality Within 30 Days of Operation
The rate of primary outcome (30-day mortality/hospice) was 6.1% (N = 136). This was significantly associated with increasing frailty by RAI (Fig. 1). RAI demonstrated adequate discriminatory accuracy for prediction of the primary endpoint in ROC

RESULTS

1. Participants and Descriptive Statistics
There were 2,235 patients that underwent SM surgery and 507 were RAI 0–20, 266 were RAI 21–30, 1,222 were RAI 31–40, and 240 were RAI ≥41. Baseline demographics, clinical characteristics, and baseline frailty status are summarized in Table 2. The median (interquartile range) age was 61 (52–70) and 39% were female.

https://doi.org/10.14245/ns.2347120.560
Analysis (C-statistic, 0.717; 95% CI, 0.697–0.735) (Fig. 2). On pairwise ROC comparison, RAI score had superior discriminatory accuracy compared to mFI-5 and increasing chronological age (both p < 0.001, DeLong test, Fig. 2).

3. Complications, ELOS, and NHD

There were 181 major complication occurrences (8.1%), 126 CD IV occurrences (5.6%), 937 NHDs (42%), 571 ELOS (26%), within 30 days. Increasing RAI score (in stepwise 5-point bins) was statistically significantly associated with increasing occurrence rate of major complication, CD IV, ELOS, and NHD (Table 3, Fig. 3).

**DISCUSSION**

In this analysis of 2,235 SM patients undergoing surgical treatment in a prospective surgical registry, increasing frailty, as measured by the RAI, provided discriminatory accuracy for prediction of postoperative mortality/hospice discharge. Increasing RAI scores were also associated with the other secondary outcomes consisting of major complications, CD IV, NHD, and ELOS. Comparative ROC curve analysis demonstrated the superior discrimination of the RAI compared to the mFI-5 and

**Table 3.** Postoperative complications, discharge outcomes, and 30-day outcomes after surgery for SM stratified by RAI frailty score, ACS-NSQIP 2011–2020 (N = 2,235)

<table>
<thead>
<tr>
<th>Variable</th>
<th>All cohort</th>
<th>RAI category</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patient</td>
<td>2,235</td>
<td>507</td>
<td>266</td>
</tr>
<tr>
<td>Postoperative major complication</td>
<td>181 (8.1)</td>
<td>29 (5.7)</td>
<td>19 (7.1)</td>
</tr>
<tr>
<td>Clavien-Dindo IV complication</td>
<td>126 (5.6)</td>
<td>17 (3.4)</td>
<td>15 (5.6)</td>
</tr>
<tr>
<td>Extended length of stay</td>
<td>571 (25.5)</td>
<td>89 (17.6)</td>
<td>63 (23.7)</td>
</tr>
<tr>
<td>Nonhome discharge disposition</td>
<td>937 (41.9)</td>
<td>146 (28.8)</td>
<td>119 (44.7)</td>
</tr>
<tr>
<td>Mortality within 30 days</td>
<td>136 (6.1)</td>
<td>9 (1.8)</td>
<td>7 (2.6)</td>
</tr>
<tr>
<td>Mortality or hospice</td>
<td>146 (6.5)</td>
<td>9 (1.8)</td>
<td>8 (3.0)</td>
</tr>
</tbody>
</table>

Values are presented as number of patients (%).
SM, spinal malignancy; RAI, Risk Analysis Index; ACS-NSQIP, American College of Surgeons – National Surgical Quality Improvement Program.
chronological age in prediction of the primary endpoint. These results may improve informed consent and surgical decision-making by enabling more accurate preoperative surgical risk prediction.

Few studies have similarly examined the use of frailty tools for outcome prediction in SM patients. Many studies have described the importance of capturing frailty for preoperative risk assessment, moreover accurate frailty measurement and prediction is exceedingly crucial for SM patients who are often among the most frail due to their metastatic disease or SMs, advanced chronologic age, prior radiotherapy, and the necessity of surgical intervention despite these factors. Hersh et al. established that the mFI-5 and Metastatic Spinal Tumor Frailty Index (MSTFI) were superior to the widely used Charlson Comorbidity Index, supporting the well-established tenet that increasing frailty, and not just comorbidity accounting, robustly predict outcomes. A recent meta-analysis determined mFI-5 correlates better with postoperative complication prediction, while MSTFI predicts mortality better. However, other scholars found the MSTFI had poor discrimination in predicting mortality and postoperative complications for SM patients in a single center machine learning study. Another retrospective cohort study from a single institution similarly concluded that the mFI-5 did not accurately predict outcomes in surgically treated SM patients. In a series of surgical SM patients treated at a single comprehensive cancer center, Ehresman et al. analyzed specific preoperative comorbidities to create a web-based calculator predicting non-home discharge disposition. Preoperative albumin values, emergency status, and increasing complexity of the surgical procedure are established variables known to affect postoperative outcomes. In addition to functional dependence, knowledge of the primary site of cancer, number, and site of metastases play a role in risk stratification in patients with metastatic spinal tumors. However, these variables do not quantify the baseline physiological reserve of a patient (i.e., frailty) even in conjunction with the mFI-5. Hall and colleagues adapted the Minimum Data Set Mortality Risk Index-Revised as a tool for the prospective or retrospective calculation of frailty. The 14-variables captured by RAI to create a frailty score have been recalibrated and validated across multiple surgical specialties.

Prior studies have robustly demonstrated the predictive value of mFI-5 for poor postoperative outcomes in various surgical populations, including SM. However, the mFI-5 is limited as a unidimensional scale, merely consisting of 4 comorbidities plus a binary functional assessment of independent versus dependent status (fully dependent or partially dependent). The RAI incorporates multiple domains of frailty by assessing comorbidities, functional status, nutritional baseline, mental status, and transfer status. The RAI’s more comprehensive assessment of physiologic reserve in patients with SM is reflected by its superior discrimination in mortality/hospice prediction (compared to mFI-5) in patients with SMs. This study also adds support to the previous literature that also refutes the notion that frailty simply increases in proportion with chronological age. Rather, frailty and chronological age need to be independently assessed preoperatively to accurately predict risk associated with any potential surgical intervention. As such, the predictive ability of RAI plays a vital role in this cohort where the prevalence of primary SMs increases with age and improved rates of cancer survival predispose to the risk of future metastases. This study complements the growing body of literature validating the use of RAI in predicting adverse postoperative outcomes in neurosurgical spine patient populations.

Frailty assessment at clinic visits could guide preoperative counseling and prehabilitation efforts to prevent adverse outcomes in elective spine surgery procedures. The ease of calculating frailty scores could play a role in the extent of their incorporation into routine clinical practice. For instance, scores like the Hospital Frailty Risk Score (HFRS) are calculated from over 100 ICD-10 codes. Although it is unclear as to how well HFRS is currently being utilized in clinical practice, one might imagine that calculating frailty scores based on HFRS might...
prove challenging.\textsuperscript{50} RAI, on the other hand, can be calculated using a 14-item questionnaire by the caregiver. Prior literature has described its ease of calculation in prospective cohorts of patients undergoing spine surgery.\textsuperscript{50,56}

While not the primary endpoint, the RAI was also associated with ELOS, NHD, major complications, and CD IV complications. These adverse outcomes impact the patient and their family, but also increase the cost and burden on the healthcare system.\textsuperscript{57} The elevated risk of poor outcomes with higher RAI scores highlights the importance of preoperative counseling with patients and families for more precise risk-benefit estimation.\textsuperscript{58,59} As our analysis demonstrates the superiority of RAI over mFI-5 in terms of discriminatory accuracy for predicting mortality and since increasing RAI is associated with all the other adverse outcomes, clinicians may consider implementing RAI scoring during the preoperative surgical evaluation of SM patients.

RAI score was modified according to criteria described in prior studies to compute the score without the preoperative cognitive decline variable since this variable is no longer available in the NSQIP database.\textsuperscript{20} The NSQIP only supplies a single ICD diagnosis field called "PODIAG" (postoperative diagnosis) and thus ICD codes cannot be utilized to define diagnoses other than the primary surgical diagnosis. There are also limitations related to specific outcomes of interest for the surgical treatment of SM. The NSQIP does not record postoperative functional outcomes such as urinary/bowel incontinence or other neurological deficits. Spinal cord compression is typically an indication for emergent/urgent surgical decompression and restoration. Information on quality of life or symptom improvement after surgery would add value to future research as it relates to RAI frailty indexing preoperatively. Additionally, variables critical in neurooncology research are not recorded in the NSQIP such as tumor size, surgical approach, and extent of resection. Furthermore, ICD coding is inherently difficult in distinguishing primary from metastatic neoplasms (Table 1). Thus, the authors opted to include all "malignant" spinal tumors. Further prospective research using qualitative data may provide better insight into discrete outcomes between these 2 groups that was unable to be performed in our retrospective analysis. Similarly, the MSTFI is not available using NSQIP, which we could not compare in this study to retrospective RAI. However, further prospective RAI studies can compare MSTFI in this patient population. Despite these limitations, this study is robust, with a large sample size of 2,235 patients, providing the necessary statistical power to analyze the prognostic utility of the RAI score.

CONCLUSION

Preoperative frailty, measured by the RAI score, was a predictor of postoperative mortality and/or hospice discharge in patients undergoing resection of SMs. Increased RAI scores are not only highly predictive of 30-day mortality or discharge to hospice, but are also associated with major complications, CD IV complications, ELOS, and NHD. Of note, the RAI demonstrated superior discrimination as compared to mFI-5 and age for outcome prediction. The frailty analysis may be translated to the bedside with a user-friendly application. The present work provides a foundation for frailty research in patients with SMs, including ongoing prospective clinical studies at the authors' institution.

NOTES

\textbf{Conflict of Interest:} In order to comply with the Hospital Participation Agreement (HPA) that is agreed to between the ACS and participating sites, facility identifiers as well as geographic information regarding the case have been removed. The HPA stipulates that the ACS does not identify participating sites. Site identification could be possible even with blinded identifiers through advanced statistics. A stipulation of access to the PUF is completion of the Data Use Agreement that strictly prohibits attempts to identify hospitals, health care providers, or patients. The authors have nothing to disclose.

\textbf{Funding/Support:} This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

\textbf{Author Contribution:} Conceptualization: RT, CAB, ACS, MHS; Data curation: RT, ACS; Formal analysis: RT, ACS; Methodology: CAB, ACS, JMR; Visualization: CAB, ACS; Writing – original draft: RT, CAB; Writing – review & editing: RT, ACS, JMR, MHS.

\textbf{ORCID}

Rachel Thommen: 0000-0002-3286-4466
Christian A. Bowers: 0000-0001-6299-503X
Aaron C. Segura: 0000-0003-0735-6977
Joanna M. Roy: 0000-0002-4829-1869
Meic H. Schmidt: 0000-0003-2259-9459

\textbf{REFERENCES}

1. Wewel JT, O'Toole JE. Epidemiology of spinal cord and col-
2. Sutcliffe P, Connock M, Shyangdan D, et al. A systematic re-
view of evidence on malignant spinal metastases: natural 
history and technologies for identifying patients at high risk 
of vertebral fracture and spinal cord compression. Health 
miology of spinal metastases, metastatic epidural spinal cord 
compression and pathologic vertebral compression fractures 
in patients with solid tumors: a systematic review. J Bone 
Oncol 2022;35:100446.
report: primary brain and other central nervous system tu-
ors diagnosed in the United States in 2012–2016. Neuro 
Oncol 2019;21(Suppl 5):v1-100.
5. Schellinger KA, Propp JM, Villano JL, et al. Descriptive epi-
demiology of primary spinal cord tumors. J Neurooncol 
tics 2018: GLOBOCAN estimates of incidence and mortal-
ity worldwide for 36 cancers in 185 countries. CA Cancer J 
7. Kazim SF, Dicpinigaitis A, Bowers C, et al. Frailty status is a 
more robust predictor than age of spinal tumor surgery out-
comes: a NSQIP analysis of 4,662 patients. Neurospine 2022; 
on spine surgery: systematic review on 10 years clinical stud-
9. Hersh AM, Pennington Z, Hung B, et al. Comparison of 
frailty metrics and the Charlson Comorbidity Index for pre-
dicting adverse outcomes in patients undergoing surgery for 
80 years and older: risk stratification using the modified 
validation of the Risk Analysis Index for measuring frailty 
in surgical populations. JAMA Surg 2017;152:175-82.
validation of the Risk Analysis Index: a surgical frailty as-
ient frailty and postoperative mortality across multiple non-
nonendoscopic approaches to single-level lumbar spine de-
compression: propensity score-matched comparative analy-
sis and frailty-driven predictive model. Neurospine 2023;20: 
119-28.
by risk analysis index and adverse discharge outcomes after 
adult spine deformity surgery: analysis of 3104 patients from 
a prospective surgical registry (2011-2020). Spine J 2023;23: 
739-45.
ified frailty index: an effective predictor of mortality in brain 
of baseline frailty status and age with postoperative morbidi-
ty and mortality following intracranial meningioma resec-
18. Dicpinigaitis AJ, Hanft S, Cooper JB, et al. Comparative as-
sociations of baseline frailty status and age with postopera-
tive mortality and duration of hospital stay following meta-
static brain tumor resection. Clin Exp Metastasis 2022;39: 
303-10.
measured by risk analysis index predicts complications and 
poor discharge outcomes after Brain Tumor Resection in a 
22. Swets JA. Measuring the accuracy of diagnostic systems. 
23. The R Project for Statistical Computing [Internet]. The R 
www.r-project.org/.
24. Harrell FE Jr. rms: regression modeling strategies [Internet]. 
2024 [cited 2022 Aug 22]. Available from: https://CRAN.R-
project.org/package=rms.
framework for R [Internet]. 2022 [cited 2022 Aug 21]. Avail-
bable from: https://CRAN.R-project.org/package=shiny.
of frailty in predictive modeling of short-term outcomes in 
the surgical management of metastatic tumors to the spine. 
27. Clegg A, Young J, Ciliffe S, et al. Frailty in elderly people. Lan-
cet 2013;381:752-62.


Commentary on “Baseline Frailty Measured by the Risk Analysis Index and 30-Day Mortality After Surgery for Spinal Malignancy: Analysis of a Prospective Registry (2011–2020)”

Fon-Yih Tsuang1,2

1Division of Neurosurgery, Department of Surgery, National Taiwan University Hospital, Taipei City, Taiwan
2Spine Tumor Center, National Taiwan University Hospital, Taipei City, Taiwan

In the past decades, thanks to medical advances, the survival of cancer patients has improved, and the incidence of the spine metastasis has also increased. Numerous studies have focused on predicting the survival of spine metastasis patients, leading to the development of several prediction systems such as Tomita score, revised Tokuhashi score, modified Bauer score, New England Spine Metastasis Scores, SORG nomogram… etc.1 While surgery for spine metastasis primarily aims for the mechanical stabilization and neurological restoration, the patient’s postoperative survival is a crucial factor that should be considered preoperatively.

In addition to survival, frailty has garnered increasing interests.2 Frailty is a recognized geriatric syndrome that predicts poor outcomes and is characterized by and age-related and precipitous decline in function. It also indicates a reduced ability to recover from physiological stresses. For instance, in a surgical cohort of spine metastasis from non-small cell lung cancer, aggressive surgery showed longer overall survival and progression-free survival to ambulatory patients compared to palliation surgery, but it also demonstrated higher mortality rates at the 30 and 90 days.3 Assessing frailty is therefore critical in the decision-making process for surgery.4 The Modified Frailty Index (mFI), a cumulative and weighted deficit model developed by Rockwood et al.,5 is considered an appropriate risk stratification tool in spine surgery.4 However, in the context of the spine metastasis surgery, the mFI does not account for the impact of systemic treatment—chemotherapy, disease burden, and performance status—ECOG, which are important factors in determined how well a patient can tolerate the surgery.2

Nutritional status is also related to survival and complications following spine metastasis surgery.6 Besides survival prediction and frailty evaluation, nutrition plays an important role in the postoperative course and should be considered in treatment plan, though it is not a part of 5-factor mFI (mFI-5) or 11-factor mFI evaluations. Consequently, new tools are being developed. The Metastatic Spinal Tumor Frailty Index, which includes nine independent variables such as malnutrition and surgical invasiveness,7 has been frequently studied but lack predictive validity. The Hospital Frailty Risk Score has shown positive predic-
tive validity, but it is not clinical feasible as it is measured based on administrative data. The H2-FAILS score, which includes serum albumin level for nutrition evaluation and performance status, demonstrated better prediction of 30-day mortality than the mFI-5. The Risk Analysis Index (RAI) for spinal malignancy surgical outcomes has shown superiority in prediction of mortality/hospice rates versus mFI-5 and chronological age.

Compared to other frailty tools, RAI involves more performance and nutritional evaluations, such as exertional tachypnea, dependency of daily living, weight loss, and appetite.

There is no perfect prediction system for survival evaluation of spine metastasis patients, nor is there a perfect tool for frailty. Most tools are not proven accurate enough for clinical utility, and no frailty tools have been investigated for reliability. Frailty is only a part of preoperative evaluation and should not be the sole factor in deciding whether to proceed with surgery. The extent of surgical intervention and invasiveness should be balanced with the frailty status. An international retrospective cohort study showed that the spine metastasis patients have equal satisfaction with surgery whether they survive less and more than three months.

Survival prediction is challenging and continually evolving. Frailty represents the physiologic resilience of a patient to tolerate surgery. In conclusion, for spine metastasis patients, frailty encompasses comorbidities, performance status, nutrition, disease burden, and the impact of systemic treatment. There is no perfect frailty tool yet and frailty alone should not be used to refuse surgery for spine metastasis patients. Patient preference and satisfaction also matter. Further studies are needed to clarify how frailty should be evaluated, to validate these evaluations, and to determine whether nutrition supplementation and prehabilitation are effective interventions to reverse frailty.

*Conflict of Interest:* The author has nothing to disclose.

**REFERENCES**

Comparison of the Clinical Efficacy of Anabolic Agents and Bisphosphonates in the Patients With Osteoporotic Vertebral Fracture: Systematic Review and Meta-analysis of Randomized Controlled Trials

Ikchan Jeon1*, Sung Bae Park2*, Bong Ju Moon3, Miyoung Choi4, Sung Uk Kuh5*, Jongtae Kim6

1Department of Neurosurgery, Yeungnam University Hospital, Yeungnam University College of Medicine, Daegu, Korea
2Department of Neurosurgery, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul National University College of Medicine, Seoul, Korea
3Department of Neurosurgery, Spine and Spinal Cord Institute, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Korea
4Clinical Evidence Research, Division of Health Technology Assessment Research, National Evidence-Based Healthcare Collaborating Agency, Seoul, Korea
5Department of Neurosurgery, Spine and Spinal Cord Institute, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea
6Department of Neurosurgery, Incheon St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Incheon, Korea

Objective: We investigated the clinical efficacy of anabolic agents compared with bisphosphonates (BPs) for the incidence of new osteoporotic vertebral fracture (OVF) and fracture healing of OVF in the patients with OVF via meta-analyses of randomized controlled trials (RCTs).

Methods: Electronic databases, including PubMed, Embase, and Cochrane Library were searched for published RCTs till December 2022. The RCTs that recruited participants with osteoporosis at high-/very high-risk of fracture (a history of osteoporotic vertebral or hip fracture) or fresh OVF were included in this study. We assessed the risk of bias on every included RCTs, estimated relative risk (RR) for the incidence of new OVF and fracture healing of OVF, and overall certainty of evidence. Meta-analyses were performed by Cochrane review manager (RevMan) ver. 5.3. Cochrane risk of bias 2.0 and GRADEpro/GDT were applied for evaluating methodological quality and overall certainty of evidence, respectively.

Results: Five hundred eighteen studies were screened, and finally 6 eligible RCTs were included in the analysis. In the patients with prevalent OVF, anabolic agents significantly reduced the incidence of new OVF (teriparatide and romosozumab vs. alendronate and risdronate [RR, 0.57; 95% confidence interval, 0.45–0.71; p < 0.00001; high-certainty of evidence]; teriparatide vs. risdronate [RR, 0.50; 95% confidence interval, 0.37–0.68; p < 0.0001; high-certainty of evidence]). However, there was no evidence of teriparatide compared to alendronate in fracture healing of OVF (RR, 1.23; 95% confidence interval, 0.95–1.60; p = 0.12; low-certainty of evidence).

Conclusion: In the patients with prevalent OVF, anabolic agents showed a significant superiority for preventing new OVF than BPs, with no significant evidence for promoting fracture healing of OVF. However, considering small number of RCTs in this study, additional studies with large-scale data are required to obtain more robust evidences.

Keywords: Teriparatide, Romosozumab, Bisphosphonate, Osteoporosis, Vertebral fracture, Meta-analysis
INTRODUCTION

Osteoporotic vertebral fracture (OVF) is among the most common fragile fractures, affecting 30% to 50% of individuals over the age of 50. The presence of at least one OVF substantially increases the risk of future OVFs, more than quadrupling it within a two-year period. Repeated vertebral fractures and severe vertebral collapse associated with osteoporosis can lead to spinal deformity. In addition to causing chronic pain, severe spinal deformity impairs gastrointestinal and respiratory functions, resulting in reduced daily activities and a lower quality of life. Furthermore, when bone union at the fracture site is delayed, pseudarthrosis can occur, accompanied by persistent pain. Neurological issues such as delayed myelopathy, if they result from delayed union, may necessitate surgery in some cases. Therefore, it is crucial to treat patients with OVF early to minimize vertebral collapse, facilitate early bone union, and prevent pseudarthrosis, as well as to reduce the incidence of new OVFs.

For patients at a higher risk of subsequent fractures with continuous significant bone loss, current evidence and recent guidelines increasingly support the use of anabolic agents promoting bone formation as the first-line treatment. Anabolic agents rapidly reduce the risk of fractures, especially in the first year following a fracture, and significantly enhance clinical outcomes. The primary anabolic agents widely used in recent times are teriparatide and romosozumab. Teriparatide is effective in decreasing the incidence of new OVFs in postmenopausal women with severe osteoporosis, which can be achieved by preferentially promoting the differentiation of preosteoblasts into osteoblasts, stimulating existing osteoblasts to form new bone, and decreasing osteoblast apoptosis. In previous clinical trials, romosozumab exhibited a rapid and significant decrease in the incidence of new OVFs, along with an increase in bone mineral density (BMD) compared with the control group. Romosozumab, as an anticalcston monoclonal antibody, has a dual effect of enhancing bone formation and inhibiting its resorption by blocking the sclerostin pathways. Sclerostin, a molecule derived from osteocytes and encoded by the SOST gene, has been discovered to regulate bone turnover by inhibiting osteoblastogenesis and bone formation. It does so by blocking the Wnt signaling pathways, which play a crucial role in bone formation and morphogenesis.

Bisphosphonates (BPs), as antiresorptive agents that work by inhibiting osteoclast-mediated bone resorption, continue to be widely considered and used as one of the treatment options for patients with osteoporosis. Nevertheless, recent head-to-head trials comparing BPs with anabolic agents have shown the superiority of anabolic agents in reducing the risk of fractures. New guidelines now recommend initial treatment with anabolic agents for patients at imminent or very high-risk of fractures. Some relevant meta-analyses have demonstrated the effectiveness of anabolic agents in reducing new OVFs compared to BPs in postmenopausal osteoporosis patients. However, unfortunately, there is significant heterogeneity in the participants and outcomes among the included studies in meta-analyses, and other randomized controlled trials (RCTs) on this subject have been published without providing conclusive results. There is still a lack of relevant and comprehensive meta-analyses with RCTs comparing the clinical efficacy between anabolic agents and BPs for reducing the incidence of new OVFs in the patients with OVF, commonly referred to as subsequent OVFs. Furthermore, the previous comparative studies of clinical efficacy between the 2 drugs for the fracture healing of fresh OVF shows heterogeneous and unclear conclusions. We think that the reliable recommendations in clinical fields related with the use of anabolic agents and BPs in the patients with OVF are required via scientific verification process.

In this study, we conducted a systematic review and meta-analysis of RCTs to determine whether anabolic agents are superior to BPs for preventing of new OVFs and promoting fracture healing of OVF in the patients with OVF.

MATERIALS AND METHODS

1. Protocol

This meta-analysis was conducted in accordance with the recommendations of the Cochrane Handbook for Systematic Reviews of Interventions and was reported following the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

2. Search Strategy

Three researchers conducted a systematic search of major electronic databases (PubMed, Embase, and the Cochrane Library) using a carefully designed search strategy for relevant studies published in English up to December 2022. We aimed to gather RCTs that compared the effects of anabolic agents and BPs on the incidence of new OVFs or fracture healing of OVF in patients with osteoporosis at high-risk or very high-risk of fracture. We established the search terms included keywords found in the titles, abstracts, or MeSH (medical subject headings) terms in each database’s search engine.
Anabolic Agents Versus Bisphosphonate on OVF

Jeon I, et al.

It takes into account various...value significantly lesser than 50%.

In cases of high heterogeneity that...value greater than 50% indicated...statistics, where an I^2...fracture, secondary osteoporosis, or did not report results...being a very high fracture probability based on country-specific thresholds. Additionally, individuals who have recently experienced a fracture (e.g., within the past 12 months), sustained a fracture while on approved osteoporosis therapy, incurred fractures due to medications causing skeletal harm, encountered multiple fractures, have a very low BMD T-score (e.g., ≤-3.0), are at high risk for falls or have a history of injurious falls, or exhibit a very high fracture probability per FRAX (e.g., >30% major osteoporotic fracture, >4.5% hip) are categorized as osteoporosis at very high-risk fracture.

However, studies that recruited patients with traumatic vertebral fracture, secondary osteoporosis, or did not report results in dichotomous data (i.e., patient-years, etc.), were excluded. Post hoc analyzed RCTs were also included, with taking care of duplicated data input. Disagreements between reviewers were resolved by discussion or, if unresolved, by consultation with librarians and a statistic expert.

4. Data Extraction

The basic characteristics of each study were independently extracted by 3 authors using a structured table that included information on study design, the number of participants, interventions, comparisons, and outcomes. The primary outcomes assessed in this study were the development of new OVFs and fracture healing of OVF. The validity of extracted data was reviewed by the other authors.

5. Risk of Bias Assessment of Studies

In this study, 3 authors independently evaluated the methodological quality of the RCTs using the Cochrane risk of bias 2 tool. The tool assessed the risk of bias as high, low, or unclear across several criteria; including random sequence generation, allocation concealment, blinding of participants and personnel to the study protocol, blinding of outcome assessment, incomplete outcome data, selective reporting, and other potential sources of bias. Disagreements were addressed by consensus with the involvement of a third review author. RCTs that were found to have a high risk of bias in more than one key domain were categorized as high risk, while RCTs with a low risk of bias in all key domains were considered low risk. RCTs that did not fit either of these categories were categorized as having an unclear risk of bias. Assessment of publication bias was attempted through the use of funnel plots.

6. Data Synthesis

The relative risk (RR) and its corresponding 95% confidence intervals (CIs) were employed to assess the impact of interventions for RCTs, with p-values less than 0.05 considered statistically significant. Data analysis was conducted using RevMan 5.3 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014) to estimate a pooled effect through a fixed-effect model. Heterogeneity among the studies was assessed using the I^2 statistics, where an I^2 value greater than 50% indicated significant heterogeneity. In cases of high heterogeneity that remained unexplained, sensitivity analysis was planned, taking into account factors such as subjects, interventions, or outcomes. However, sensitivity analysis was skipped in low heterogeneity with I^2 value significantly lesser than 50%.

7. Assessment of Certainty of Evidence

In assessing the overall certainty of evidence, we utilized the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) framework, and we employed the GRADEpro/GDT (Guideline Development Tool) software, which is accessible at https://gradepro.org/. GRADE represents a widely accepted and transparent framework for summarizing evidence and offers a systematic approach for formulating clinical practice recommendations. It takes into account various
factors such as study design, risk of bias, inconsistency, indirectness, imprecision, and other relevant considerations. The final determination of the overall certainty of evidence according to GRADE falls into one of 4 categories: very low, where the true effect is likely markedly different from the estimated effect; low, where the true effect may be markedly different from the estimated effect; moderate, where the true effect is likely close to the estimated effect; and high, where the true effect closely aligns with the estimated effect. To address imprecision, we calculated the Optimal Information Size (OIS), the minimum sample size required in a single and adequately powered study to evaluate the effects of intervention in the general population, using a Sample Size Calculator, available at https://www.stat.ubc.ca/~rollin/stats/ssize/n2.html.

RESULTS

1. Search Selection and Characteristics

The PRISMA flow diagram (Fig. 1) illustrates the selection and exclusion process of the studies. Initially, a total of 518 studies were screened (204 from PubMed, 224 from Embase, and 90 from Cochrane Library). Subsequently, 259 of these studies underwent a full-text assessment. Following a thorough examination of the full texts, 6 RCTs involving 3,642 and 3,655 patients with osteoporosis at high-risk fracture treated with anabolic agents and BPs, respectively, were ultimately included in this meta-analysis. Among the included RCTs, 4 of them compared the effects of anabolic agents (3 of teriparatide and 1 of romosozumab) with BPs (2 of alendronate and 2 of risedronate) in terms of the incidence of new OVFs. The other 2 RCTs compared the effects on fracture healing in OVF between teriparatide and alendronate. The detailed characteristics of these 6 RCTs are summarized in Table 1.

2. Risk of Bias Assessment

Fig. 2 provides a summary of the details regarding the risk of bias. In total, 3 RCTs were categorized as having a low risk of bias, while 2 RCTs were deemed to have a high risk of bias. All RCTs exhibited adequate random sequence generation, managed incomplete outcome data, and avoided selective reporting. However, there was uncertainty in 3 RCTs regarding appropriate allocation concealment, one RCT in blinding of outcome assessment, and one RCT in other bias, respectively. Blinding of participants and personnel assessments showed high-risk in 2 RCTs and unclear in one RCT. We assessed publication bias using funnel plots in each analysis (Figs. 3–5). Although sufficient analyses could not be performed due to small number of RCTs included in this study, there was no suspicious evident for pub-
Table 1. General characteristics of the included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Journal</th>
<th>Participants</th>
<th>Sample size and medication</th>
<th>Medication period</th>
<th>Age (yr)</th>
<th>Follow-up</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendler et al. [19]</td>
<td><em>Lancet</em></td>
<td>Postmenopausal women &gt; 45 years with BMD T-score ≤ 1.5 and either ≥ 2 moderate or a severe vertebral fracture</td>
<td>N = 680, teriparatide (20 μg, daily SC) N = 680, risedronate (35 mg, weekly PO)</td>
<td>24 Months</td>
<td>72.6 ± 8.8</td>
<td>71.6 ± 8.6</td>
<td>24 Months New OVF (morphometric fracture-radiograph)</td>
</tr>
<tr>
<td>Hadji et al. [20]</td>
<td><em>Osteoporosis</em></td>
<td>Postmenopausal women ≥ 45 years with BMD T-score ≤ -2.0 and ≥ 1 moderate vertebral fracture</td>
<td>N = 360, teriparatide (20 μg, daily SC) N = 350, risedronate (35 mg, weekly PO)</td>
<td>18 Months</td>
<td>70.5 ± 8.8</td>
<td>71.6 ± 8.1</td>
<td>18 Months New OVF (morphometric fracture-radiograph)</td>
</tr>
<tr>
<td>Geusens et al. [21]</td>
<td><em>Bone</em></td>
<td>Postmenopausal women with BMD T-score ≤ -2.5 and either ≥ 1 moderate or severe vertebral fracture, or BMD T-score of ≤ -2.0 and either ≥ 2 moderate or a severe vertebral fracture</td>
<td>N = 2,046, romosozumab (210 mg, monthly SC) N = 2,047, alendronate (70 mg, weekly PO)</td>
<td>12 Months</td>
<td>74.4 ± 7.5</td>
<td>74.2 ± 7.5</td>
<td>12 Months New OVF (morphometric fracture-radiograph)</td>
</tr>
<tr>
<td>Hagino et al. [22]</td>
<td><em>Osteoporosis</em></td>
<td>Postmenopausal women ≥ 75 years with osteoporosis at high fracture risk (BMD T-score &lt; -3.3 and either ≥ 2 vertebral fractures, a grade 3 vertebral fracture, or hip fracture)</td>
<td>N = 489, teriparatide (56.6 μg, weekly SC) N = 496, alendronate (5 mg daily PO, 35 mg weekly PO, or 900 μg every 4 weeks IV)</td>
<td>72 Weeks</td>
<td>81.4 ± 4.5</td>
<td>81.5 ± 4.7</td>
<td>72 Weeks New OVF (morphometric fracture-radiograph)</td>
</tr>
<tr>
<td>Shigenobu et al. [23]</td>
<td><em>Bone Reports</em></td>
<td>Patients (43; 5 men and 38 women) with fresh osteoporotic vertebral compression fracture</td>
<td>N = 19, teriparatide (56.6 μg, weekly SC) N = 24, alendronate (35 mg weekly PO, risedronate 17.5 mg weekly PO, or 75 mg monthly PO)</td>
<td>12 Weeks</td>
<td>80.2</td>
<td>75.6</td>
<td>12 Weeks Fracture healing of OVF (CT scan)</td>
</tr>
<tr>
<td>Ikeda et al. [24]</td>
<td><em>Journal of Bone and Mineral Metabolism</em></td>
<td>Women with fresh osteoporotic vertebral fractures</td>
<td>N = 48, teriparatide (56.6 μg, weekly SC) N = 48, alendronate (35 mg, weekly PO)</td>
<td>12 Weeks</td>
<td>78.9</td>
<td>80.3</td>
<td>12 Weeks Fracture healing of OVF (radiograph)</td>
</tr>
</tbody>
</table>

BP, bisphosphonate; BMD, bone mineral density; SC, subcutaneous; PO, per os (oral); OVF, osteoporotic vertebral fracture; IV, intravenous; CT, computed tomography.
Fig. 2. Risk of bias summary (A) and graph (B).

Fig. 3. Forest and funnel plots for comparing anabolic agents versus bisphosphonates of the incidence of new osteoporotic vertebral fracture in the patients with osteoporotic vertebral and hip fracture. (A) Anabolic agents versus bisphosphonates. (B) Teriparatide versus bisphosphonates. (C) Romosozumab versus alendronate. M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

https://doi.org/10.14245/ns.2347256.628  www.e-neurospin.org  421
Fig. 4. Forest and funnel plots for comparing anabolic agents versus bisphosphonates of the incidence of new osteoporotic vertebral fracture in the patients with osteoporotic vertebral fracture. (A) Anabolic agents versus bisphosphonates. (B) Teriparatide versus bisphosphonates. (C) Romosozumab versus alendronate. M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

Fig. 5. Forest and funnel plots for comparing anabolic agents versus bisphosphonates of the fracture healing of osteoporotic vertebral fracture. M-H, Mantel-Haenszel; CI, confidence interval; df, degrees of freedom.

Application bias considering narrowly located on the top of the plot with symmetric distribution on both side of graph.

3. Efficacy Outcomes

1) Incidence of new OVF

The incidence of new OVF data in the patients with prevalent osteoporotic vertebral or hip fracture was taken from 4 RCTs with including 3,575 and 3,583 patients in anabolic agents (3 teriparatide and 1 romosozumab) and BPs (2 alendronate and 2 risedronate), respectively. There was a significant decrease in the incidence of new OVF (RR, 0.59; 95% CI, 0.49–0.72; p < 0.00001; I² = 0%) (Fig. 3A). The analysis with only teriparatide among the anabolic agents was performed with 3 RCTs including 1,529 of teriparatide and 1,526 patients in teriparatide and BPs (1 alendronate and 2 risedronate), respectively. Teriparatide also showed significant decrease in the incidence of new OVF compared to BPs (RR, 0.56; 95% CI, 0.44–0.72; p < 0.00001; I² = 3%) (Fig. 3B). In the analysis with only romosozumab by one RCT including 2,046 of romosozumab and 2,047 of alendronate, there was a significant decrease in the incidence of new
OVF compared to alendronate (RR, 0.65; 95% CI, 0.46–0.90; p = 0.01) (Fig. 3C).

The incidence of new OVF data in the patients with only prevalent OVF was taken from 3 RCTs with including 3,009 and 3,004 patients in anabolic agents (2 teriparatide and 1 romosozumab) and BPs (1 alendronate and 2 risedronate), respectively. One RCT that could not measure the incidence of new OVF according to prevalent OVF was excluded, and the selected data with prevalent OVF from the remaining RCTs were used to evaluate the incidence of new OVF according to the presence of prevalent OVF. There was a significant decrease in the incidence of new OVF (RR, 0.57; 95% CI, 0.45–0.71; p < 0.00001; I² = 0%) (Fig. 4A). The analysis with only teriparatide among the anabolic agents was performed with 2 RCTs including 1,040 of teriparatide and 1,030 patients in teriparatide and BPs (2 risedronate), respectively. Teriparatide also showed significant decrease in the incidence of new OVF compared to BPs (RR, 0.50; 95% CI, 0.37–0.68; p < 0.0001; I² = 0%) (Fig. 4B). In the analysis with only romosozumab by 1 RCT including 1,969 of romosozumab and 1,964 of alendronate, there was a significant decrease in the incidence of new OVF compared to alendronate (RR, 0.65; 95% CI, 0.46–0.90; p = 0.01) (Fig. 4C).

The detailed data of the incidence of new OVF are summarized in Table 2.

Table 2. Summary of findings in new osteoporotic vertebral fracture and fracture healing

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Comparison (composition of RCTs)</th>
<th>Incidence of events/sample size</th>
<th>OR (95% CI)/heterogeneity</th>
<th>Certainty of evidence</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence of new OVF</td>
<td>Osteoporosis with prevalent osteoporotic vertebral or hip fracture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Anabolic agents vs. bisphosphonates (3 teriparatide+1 romosozumab vs. 2 alendronate+2 risedronate)</td>
<td>151/3,575 255/3,583 RR = 0.59 (95% CI, 0.49–0.72), p &lt; 0.00001/I² = 0%</td>
<td>High</td>
<td>Important</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teriparatide vs. bisphosphonates (3 teriparatide vs. 1 alendronate+2 risedronate)</td>
<td>96/1,529 170/1,526 RR = 0.56 (95% CI, 0.44–0.72), p &lt; 0.00001/I² = 3%</td>
<td>High</td>
<td>Important</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romosozumab vs. bisphosphonate (1 romosozumab vs. 1 alendronate)</td>
<td>55/2,046 85/2,047 RR = 0.65 (95% CI, 0.46–0.90), p = 0.01/NA</td>
<td>NA</td>
<td>Important</td>
<td></td>
</tr>
<tr>
<td>Osteoporosis with prevalent osteoporotic vertebral fracture</td>
<td>Anabolic agents vs. bisphosphonates (2 teriparatide+1 romosozumab vs. 1 alendronate+2 risedronate)</td>
<td>110/3,009 193/3,004 RR = 0.57 (95% CI, 0.45–0.71), p &lt; 0.00001/I² = 0%</td>
<td>High</td>
<td>Important</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Teriparatide vs. bisphosphonates (2 teriparatide vs. 2 risedronate)</td>
<td>55/1,040 108/1,030 RR = 0.50 (95% CI, 0.37–0.68), p &lt; 0.0001/I² = 0%</td>
<td>High</td>
<td>Important</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romosozumab vs. bisphosphonate (1 romosozumab vs. 1 alendronate)</td>
<td>55/1,969 85/1,964 RR = 0.65 (95% CI, 0.46–0.90), p = 0.01/NA</td>
<td>NA</td>
<td>Important</td>
<td></td>
</tr>
<tr>
<td>Fracture healing</td>
<td>Teriparatide vs. bisphosphonates (2 teriparatide vs. 2 alendronates)</td>
<td>46/67 40/72 RR = 1.23 (95% CI, 0.95–1.60), p = 0.12/I² = 35%</td>
<td>Low</td>
<td>Important</td>
<td></td>
</tr>
</tbody>
</table>

RCT, randomized controlled trial; OR, odds ratio; CI, confidence interval; OVF, osteoporotic vertebral fracture; RR, relative risk.

2) Fracture healing of OVF

The fracture healing of OVF data was taken from 2 RCTs with including 67 and 72 patients in teriparatide and alendronate of groups with osteoporosis at very high fracture risk, respectively. There was no statistically significant difference in the fracture healing of OVF between teriparatide and alendronate (RR, 1.23; 95% CI, 0.95–1.60; p = 0.12; I² = 35%) (Fig. 5). The detailed data of fracture healing of OVF are summarized in Table 2.

4. Assessment of Certainty of Evidence - GRADE

The GRADEpro/GDT analyses regarding the overall quality of evidence were conducted in the incidence of new OVF and fracture healing of OVF. In the incidence of new OVF, the GRADEpro/GDT analyses were applied depending on the type of prevalent osteoporotic fracture and use of medication, respectively. There were high-quality of evidences between anabolic agents (teriparatide and romosozumab) to BPs regardless of the type of prevalent osteoporotic fracture. The OISs in each analysis are approximately 1,004 and 978, respectively. Similarly, the quality of evidences between teriparatide and BPs were also showed high-quality of evidences regardless of the type of prevalent osteoporotic fracture, with an OISs of about 525 and 421 in each analysis. However, when evaluating the quality of evi-
dence for the fracture healing of OVF between teriparatide and alendronate, the assessment resulted in a low-quality of evidence. This was primarily due to the limited number of participants (214 of OIS) and the heterogeneity of the CIs between the included RCTs. However, the assessment of certainty of evidence using GRADE was not conducted for subgroup analyses that included only one RCT. The detailed data are summarized in Table 2.

DISCUSSION

BPs, which were first introduced in the 1990s, have long been the cornerstone of osteoporosis treatment. In the United States, there are 4 available BPs (alendronate, ibandronate, risedronate, and zoledronate), and among these, 3 (alendronate, risedronate, and zoledronate) are supported by robust evidence for their broad-spectrum effectiveness in preventing fractures and are considered as first-line treatment options.31-33 They are effective in reducing the risk of vertebral, nonvertebral, and hip fractures. However, ibandronate has shown effectiveness primarily in preventing vertebral fractures.31 A recent network meta-analysis revealed that BPs can reduce the risk of vertebral, nonvertebral, and hip fractures by percentages ranging from 33% to 62%, 16% to 21%, and 27% to 40% (excluding ibandronate in nonvertebral and hip fractures), respectively.31,32 Consequently, the AACE guideline still recommends the use of alendronate and ibandronate for individuals with osteoporosis at high-risk fracture, while zoledronate is recommended specifically for those with osteoporosis at very high-risk fracture.33 Nevertheless, data from the Danish national health registry indicate that 14.6% of patients continue to be identified as remaining at high-risk fracture despite being compliant with BP treatment.33 Although BPs are still considered effective for reducing the risk of osteoporotic fractures in osteoporosis at high-risk fracture based on the available evidence, it may be worth considering the potential benefits of anabolic agents over BPs, especially for patients who have already experienced prevalent osteoporotic fractures. This consideration can be supported by the superior efficacy of anabolic agents compared with BPs (especially teriparatide in the patients with prevalent OVF) for reducing the incidence of new OVF in this meta-analysis of RCTs.

Several previous meta-analyses have directly compared the effects of teriparatide and BPs in reducing the incidence of subsequent vertebral fracture. In a meta-analysis comparing teriparatide to alendronate, there was no significant difference in the occurrence of subsequent vertebral fractures.30 While other meta-analyses have reported the superiority of teriparatide over BPs in reducing the incidence of subsequent vertebral fractures,31-33 however, these findings come with certain limitations including relatively higher heterogeneity and participant duplication across the included studies. In this study, we reviewed the effect of anabolic agents including teriparatide and romosozumab as well as teriparatide over BPs depending on the kinds of prevalent osteoporotic fractures in the participants. We tried to identify the real effect of the medications for the reducing the incidence of new OVF in the patients with prevalent OVF (subsequent OVF). Notably, the included vast majority of postmenopausal women with osteoporosis at high-risk or very high-risk fracture had prevalent vertebral fractures, and there was only a small portion of participants in a single RCT who have prevalent hip fracture with no vertebral fracture.34 The results showed statistically significant superiorities of anabolic agents and teriparatide over BPs with high-quality evidence regardless of the prevalent osteoporotic fractures. However, among them, teriparatide showed most powerful efficacy over risedronate for reducing subsequent OVF with 0.5 of RR and high-quality evidence in the patients with prevalent OVF. Consequently, among anabolic agents, teriparatide showed the superiority over BPs for reducing new OVF after osteoporotic fractures, and it has been found to be more effective for preventing subsequent OVF in patients with prevalent OVF. Unfortunately, it is important to note that assessing the effect of romosozumab with meta-analysis was impossible due to the lack of head-to-head RCTs. In the literature, there is still a limited number of meta-analyses comparing the efficacy of romosozumab, and it is difficult to estimate the exact effect of romosozumab on BPs due to the combination of BP and placebo as the control group.35,36 Nevertheless, in this study, we guess a similar effect in reducing the incidence of new OVF between teriparatide and romosozumab based on the given low heterogeneity with I² = 0% observed between the included RCTs.

We were very surprised that there were very limited number of RCTs comparing anabolic agents and BPs in the prevention of new OVF and fracture healing of OVF in this study. Therefore, we should be more careful to interpret results and obtain the scientific evidence in the meta-analysis using a small number of RCTs. The certainty of evidence was evaluated using the GRADE framework, and meaningful results were obtained for each element, such as study design, risk of bias, inconsistency, indirectness, and imprecision. All RCTs included in this study involved patients with osteoporosis at high-risk or very-high risk fracture. Higher baseline fracture risk is indeed a significant...
contributor to overall fracture risk. However, all RCTs have consistently indicated that the reduction in the incidence of new OVF with anabolic agents is significantly independent. This is supported by the presence of statistically significant RRs with low heterogeneity ($I^2 = 0\%-3\%$) across the included RCTs in all analyses related with new OVF. Heterogeneity is a crucial consideration, especially in small sized meta-analyses. When the results of existing studies of a treatment are homogeneous or nearly homogeneous, there is a reasonable expectation that the treatment will have a similar effect when applied to new subjects. Conversely, when the results are highly heterogeneous, predicting the effect of the treatment on new subjects becomes challenging unless the reasons for the heterogeneity are well-understood. Typically, in meta-analyses comparing studies, definitive conclusions about heterogeneity are often difficult to reach. However, the strengths of our study lie in the very low heterogeneity observed and the sufficient sample sizes included in each analysis for the incidence of new OVFs with head-to-head RCTs. This reduced the potential for type 1 error and addressed the limitation of a small number of RCTs.

The methodology to define vertebral fractures varies between and within meta-analyses. Vertebral fractures are one of the most common skeletal fractures, and two-thirds to three-quarters of vertebral fractures are not recognized at the time of their clinical occurrence and require spinal imaging to be detected. Epidemiologic studies of vertebral fractures have focused primarily on radiographic vertebral fractures, and delineating the prevalence and incidence between clinical and radiological vertebral fractures is complicated (in this meta-analysis, they are named clinical and morphometric fractures, respectively). Additionally, it may be further worse as the lack of consensus as to exactly what changes within a vertebra on spine imaging warrant a diagnosis of vertebral fracture, such that some aspects of the epidemiology of vertebral fractures may depend somewhat on the chosen definition of vertebral fracture. Clinical fractures are confirmed on the imaging studies with the occurrence of related symptoms such as back pain. Usually, the incidence of clinical fractures is lower compared to radiological fractures due to the variety of subjective symptoms. Among the included RCTs in this study, 1 RCT presented the overall annual incidences of clinical and radiological fractures with 17.34% and 2.17%, respectively. There was no statistically significant difference in the incidence of new OVF defined as clinical fracture between teriparatide and alendronate due to the extremely low incidence of clinical fracture. The methodological variety is the important point should be carefully considered to conduct meta-analysis and interpret the results of it. In this study, we attempted to objectively measure the incidence of new OVF based on morphometric fracture through serial radiographs with excluding the data of clinical fracture. Nevertheless, the definition of new OVF according to the degree and shape of the fracture is unclear between the RCTs and its effect on the analyses may not be completely excluded.

Currently, there are no approved drug treatments specifically designed to promote fracture healing, despite the availability of several drugs for preventing osteoporotic fractures. Teriparatide or parathyroid hormone (PTH) analogue increases bone mass and reduces bone loss, thereby promoting bone formation. While animal studies have provided support for this hypothesis, the evidence regarding its application in humans is less conclusive. Some studies suggest that the administration of PTH analogues has a beneficial impact on fracture healing, whereas others report no discernible effect on fracture healing rates. In the prior meta-analyses, the evidence supporting the idea that teriparatide enhances fracture healing was not significant, primarily due to RCTs characterized by high heterogeneity, low-quality evidence, and the inclusion of various types of fractures. Notably, subgroup analysis comparing teriparatide to BPs in 2 RCTs showed no significant superiority. Additionally, there is no standardized method for confirming fracture healing, and approaches vary between individual studies. In 2 RCTs included this study, they applied different methods, computed tomography and radiograph, to determine fracture healing. Furthermore, researches focusing on vertebral fracture healing through RCTs has been extremely limited, and, to date, no meaningful meta-analysis on this subject exists. The 2 RCTs included in this meta-analysis aimed to assess the efficacy of fracture healing at 12 weeks after newly developed OVF between teriparatide and alendronate. However, considering the variation in standards used by physicians to make decisions about fracture healing, limited number of participants, and the overall low-quality of evidence, our finding are inconclusive and suggest the need for further researches.

Numerous prospective observational studies have consistently shown that prevalent OVFs are associated with subsequent OVFs, prevalent radiographic vertebral fracture is associated with a 4-fold increase in subsequent radiographic vertebral fractures. In a view point of clinical significance for reducing subsequent OVFs, it is important to evaluate the clinical efficacy of anabolic agents compared to BP at high-risk fracture with prevalent OVFs. This meta-analysis has strengths, as it features RCTs characterized by a preplanned parallel comparison between anabolic agents.
agents and BPs for the incidence of new OVF and the fracture healing of OVF as primary outcomes in the homogeneous group of patients with osteoporosis at high-risk or very high-risk fracture. Additionally, we conducted subgroup analyses based on the presence of prevalent OVFs and specific types of anabolic agents and BPs, with long-term follow-ups lasting at least 12 months and ensuring the avoidance of duplication by excluding studies that involved similar participants from the same research subject. However, it is important to acknowledge several limitations in this study. First, despite the subgroup analyses depending on the type of anabolic agents and prevalent osteoporotic fractures, some subgroup analyses were based on the inclusion of only 1 or 2 RCTs, which has limited ability to ascertain their real significance and reliability. Second, variations in the duration of follow-up and different administration routes/dosage of drugs among the included RCTs may have influenced the final outcomes. Third, the cost-effectiveness of anabolic agents compared with BPs should be confirmed to suggest the use of anabolic agents in the patients with osteoporosis at high-risk fracture. Fourth, as previously mentioned, the lack of well-defined criteria and method to confirm the efficacy of anabolic agents in the analysis of new OVF and fracture healing. Considering these limitations, our results may require cautious and conservative interpretation in real clinical field. Additional studies with large-scale data are required to obtain more robust evidences.

**CONCLUSION**

Anabolic agents demonstrated a significant advantage in preventing new OVF compared to BPs with high-quality evidence in patients with osteoporosis at high-/very high-risk of fractures. Particularly, there was notable significant efficacy of anabolic agents compared to BPs for the prevention of subsequent OVFs. However, considering small number of RCTs in this study, our results may require cautious and conservative interpretation and additional studies with large-scale data are mandatory to obtain more robust evidences.

**NOTES**

**Supplementary Material:** Supplementary material can be found via https://doi.org/10.14245/ns.2347256.628.

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This work was supported by a grant from the 2022 Korean Spinal Neurosurgery Society research fund.

**REFERENCES**


Anabolic Agents Versus Bisphosphonate on OVF

Jeon I, et al.

Commentary on “Comparison of the Clinical Efficacy of Anabolic Agents and Bisphosphonates in the Patients With Osteoporotic Vertebral Fracture: Systematic Review and Meta-analysis of Randomized Controlled Trials”

Dae-Chul Cho
Department of Neurosurgery, Kyungpook National University Hospital, School of Medicine, Kyungpook National University, Daegu, Korea

Osteoporotic vertebral fractures (OVFs) are a clinically important problem and are becoming more frequent as the aging population continues to increase. Current evidence and new treatment guidelines now recommend initial treatment with anabolic agents for patients at imminent or very high risk of fractures.

This meta-analysis conducted to determine whether anabolic agents, including teriparatide and romosozumab, are superior to bisphosphonates (BPs) in preventing new OVFs and promoting fracture healing in patients with OVFs. Six randomized controlled trials (RCTs) involving 3,642 and 3,655 patients with osteoporosis at high risk of fracture, treated with anabolic agents and BPs, respectively, were included in this meta-analysis.

This study reveals that anabolic agents, specifically teriparatide and romosozumab, significantly reduce the incidence of new OVFs compared to BPs. This is evidenced by high-certainty evidence and low heterogeneity across the included RCTs. The relative risk (RR) of 0.57 for teriparatide and romosozumab versus alendronate and risedronate underscores the superior efficacy of anabolic agents in preventing subsequent fractures. On the contrary, there was no statistically significant difference in the fracture healing of OVF between teriparatide and alendronate (RR, 1.23; 95% confidence interval, 0.95–1.60; p = 0.12; I² = 35%). These results align with current guidelines advocating for the use of anabolic agents in patients at high or very high risk of fractures, highlighting their rapid and significant impact on bone formation and fracture prevention.

Notably, the authors tried to overcome the limitations of meta-analysis including relatively higher heterogeneity and participant duplication across the included studies. Given the increasing prevalence of osteoporosis and its associated complications, this study addresses a critical need for evidence-based treatment strategies, taking into account various factors such as study design, risk of bias, inconsistency, indirectness, imprecision, and other relevant considerations.

The authors emphasize that we should be more careful when interpreting results and ob-
taining the scientific evidence in meta-analysis using a small number of RCTs. Although this study is inconclusive due to the variations in standards used by physicians to make decisions about fracture healing, the limited number of participants, and the overall low quality of evidence, this article is expected to provide valuable insights into the relative effectiveness of anabolic agents versus BPs in managing OVF.

The significant reduction in fracture incidence achieved with anabolic agents, particularly teriparatide, supports their use as a first-line treatment in this patient population. While additional research is needed to address the study's limitations and explore further clinical applications, the current findings represent a crucial step forward in optimizing osteoporosis management and improving patient outcomes. This study is a valuable addition to the literature and provides a strong impetus for the continued advancement of anabolic therapies in osteoporosis treatment.

- **Conflict of Interest:** The author has nothing to disclose.

**REFERENCES**

The Utility and Feasibility of Smart Glasses in Spine Surgery: Minimizing Radiation Exposure During Percutaneous Pedicle Screw Insertion

Yoshiaki Hiranaka, Yoshiki Takeoka, Takashi Yurube, Takeru Tsujimoto, Yutaro Kanda, Kunihiko Miyazaki, Hiroki Ohnishi, Tomoya Matsuo, Masao Ryu, Naotoshi Kumagai, Kohei Kuroshima, Ryosuke Kuroda, Kenichiro Kakutani

Department of Orthopaedic Surgery, Kobe University Graduate School of Medicine, Kobe, Japan

Objective: Spine surgeons are often at risk of radiation exposure due to intraoperative fluoroscopy, leading to health concerns such as carcinogenesis. This is due to the increasing use of percutaneous pedicle screw (PPS) in spinal surgeries, resulting from the widespread adoption of minimally invasive spine stabilization. This study aimed to elucidate the effectiveness of smart glasses (SG) in PPS insertion under fluoroscopy.

Methods: SG were used as an alternative screen for fluoroscopic images. Operators A (2-year experience in spine surgery) and B (9-year experience) inserted the PPS into the bilateral L1–5 pedicles of the lumbar model bone under fluoroscopic guidance, repeating this procedure twice with and without SG (groups SG and N-SG, respectively). Each vertebral body's insertion time, radiation dose, and radiation exposure time were measured, and the deviation in screw trajectories was evaluated.

Results: The groups SG and N-SG showed no significant difference in insertion time for the overall procedure and each operator. However, group SG had a significantly shorter radiation exposure time than group N-SG for the overall procedure (109.1 ± 43.5 seconds vs. 150.9 ± 38.7 seconds; p = 0.003) and operator A (100.0 ± 29.0 seconds vs. 157.9 ± 42.8 seconds; p = 0.003). The radiation dose was also significantly lower in group SG than in group N-SG for the overall procedure (1.3 ± 0.6 mGy vs. 1.7 ± 0.5 mGy; p = 0.023) and operator A (1.2 ± 0.4 mGy vs. 1.8 ± 0.5 mGy; p = 0.013). The 2 groups showed no significant difference in screw deviation.

Conclusion: The application of SG in fluoroscopic imaging for PPS insertion holds potential as a useful method for reducing radiation exposure.

Keywords: Smart glasses, Pedicle screw, Fluoroscopy, Augmented reality, Radiation exposure

INTRODUCTION

Spine surgeons are at risk of radiation exposure owing to the use of fluoroscopy during surgery, fracture or dislocation repositioning, and administering block injections. Exposure to ionizing radiation is associated with various health risks, including an increased risk of cancer and skin disorders of the fingers.1,4 There is an increasing concern about radiation exposure, especially in the field of spine surgery, due to the widespread use of minimally invasive spine stabilization,5 which frequently requires percutaneous pedicle screw (PPS) insertion under fluoroscopic guidance.6,7 Protective measures, including lead aprons, thyroid collars, and gloves, can reduce radiation exposure but cannot eliminate its risk. In addition, prolonged use of these protective gear can lead to physical discomfort and fatigue, directly affecting the surgeon’s performance. Therefore,
Usefulness of Smart Glasses in Spine Surgery

Hiranaka Y, et al.

developing new methods to reduce radiation exposure is crucial in spinal surgery. Attention has shifted to wearable technology as a relatively easily implementable solution to reduce radiation exposure in spinal surgeries.

Wearable technology, represented by smart glasses (SG), has entered the medical field and is expected to significantly impact surgery in various specialties. Since the launch of Google Glass (Google Inc., Mountain View, CA, USA) in 2013, various types of SG have been released and have become commonly used for medical purposes, including education, surgery navigation, and monitoring vital signs. MOVERIO (EPSON Co., Ltd., Tokyo, Japan) (Fig. 1)—another pair of SG—displays images from the monitoring screen to a wearable display. One valuable application of SG is in surgical procedures performed under fluoroscopic guidance. The frequent diversion of a surgeon’s attention from the operative field to the fluoroscopic monitor can decrease procedural accuracy. Therefore, SG can display fluoroscopic images on wearable displays, and surgeons can perform procedures while keeping their eyes on the operative field. This technology has been reported to reduce radiation exposure and improve screw insertion accuracy under fluoroscopic guidance in trauma surgery. However, no reports have addressed the impact of SG on reducing radiation exposure and improving screwing accuracy in the spinal region.

We hypothesized that PPS insertion with SG would reduce surgeons’ radiation exposure and improve screw insertion accuracy. Therefore, this pilot study aimed to elucidate the efficacy and feasibility of SG for PPS insertion under fluoroscopic guidance.

MATERIALS AND METHODS

1. Devices

This study does not include information on human tissues, materials, or patients. Therefore, ethical approval was not required. This study used the MOVERIO model BT-30E (Fig. 1) as the SG. This model has a wearable binocular HD 1,280 × 720 pixel display. The COREVISION 3D (FUJIFILM Co., Ltd., Tokyo, Japan) fluoroscopy system was used in this study. The SG and fluoroscopic monitor were connected using a cable with a high-definition multimedia interface port to project the monitor’s screen onto the wearable display in real time without any noticeable time lag. The Erisma-LP MIS (Clariance Inc., Beauvais, France) PPS systems were used in this study.

2. Evaluation of PPS Insertion

First, operator A, who had 2 years of experience in spinal surgery, conducted the experiment. In total, 10 PPSs were inserted into 5 bilateral L1–5 vertebrae in one lumbar model bone (Sawbone, Sawbones Inc., Malmo, Sweden) under fluoroscopic guidance with SG (group SG). The optimal diameters and lengths of the inserted screws were selected based on the model bone’s profile. The lumbar model bone was covered with a soft cloth to simulate the surgical situation more accurately, ensuring that the model bone remained hidden from view.

Similarly, 10 PPSs were inserted into another model bone using the conventional technique while watching the fluoroscopic images.
monitor without using SG (group N-SG). The operator repeated this procedure twice, and 10 vertebrae with 20 PPSs were evaluated with and without SG. Operator B, with 9-year experience in spinal surgery, also performed the same series of procedures as operator A. The time required to insert 2 PPSs into each vertebra (insertion time), radiation exposure time, and radiation dose were measured. Computed tomography (CT) was performed after PPS insertion and axial or sagittal slices of the multiplanar reconstruction CT images were used to assess the presence or absence of screw deviation. The direction and amount of screw deviation were evaluated using the grading system (type A–I) by Abul-Kasim et al.14 (Table 1). Those meeting the criteria of type A and G (acceptably placed screw in either axial or sagittal images) were defined as having no deviation, while all others were defined as having a deviation (Fig. 3A and B).

### Table 1. The grading system used for the assessment of screw placement

<table>
<thead>
<tr>
<th>Different types of misplacement</th>
<th>Axial images (A–F)</th>
<th>Sagittal images (G–I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Acceptably placed screw</td>
<td>G: Acceptably placed screw</td>
<td></td>
</tr>
<tr>
<td>B: MCP (medial cortical perforation) grade 1</td>
<td>H: FP (foraminal perforation) grade 1</td>
<td></td>
</tr>
<tr>
<td>C: MCP grade 2</td>
<td>I: EPP (endplate perforation) grade 1</td>
<td></td>
</tr>
<tr>
<td>D: LCP (lateral cortical perforation) grade 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E: LCP grade 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F: ACP (anterior cortical perforation of the vertebral body) grade 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Detailed definitions**

- **MCP**
  - Grade 0: Acceptable placement; screw within the pedicle medullary canal or minimal breach of medial cortex
  - Grade 1: Partially medialized screw
  - Grade 2: Totally medialized screw

- **LCP**
  - Grade 0: Acceptable placement; screw within the pedicle medullary canal or minimal breach of lateral cortex
  - Grade 1: Partially lateralized screw
  - Grade 2: Totally lateralized screw

- **ACP**
  - Grade 0: Acceptable placement; screw tip within the vertebral body

---

### 3. Illustration of PPS Insertion Using SG

The operator wore a radiation protection apron, neck cover, and gloves during the procedures. When wearing the SG, the operators could perform the procedures without diverting their gaze from the surgical field (Fig. 4A). Fig. 4B illustrates the operator’s actual view of PPS insertion when using the SG. Lowering the gaze allows the operator to view the surgical field, and maintaining the gaze level lets the operator view the fluoroscopic image projected on the wearable displays. This enables the operators to confirm the fluoroscopic image and the surgical field without moving their heads. However, in the conventional method of PPS insertion without SG, the operator has to shift their gaze away from the surgical field to check the fluoroscopic monitor, necessitating head movement during the procedure (Fig. 4C).

### 4. Statistical Analysis

The collected data were statistically analyzed using IBM SPSS.
Fig. 4. Illustration of percutaneous pedicle screw insertion. (A) When wearing the smart glasses (SG) and operating, the gaze is always directed towards the surgical field, minimizing the need to move the head up and down. (B) Field-of-view while wearing the smart glasses. Surgeons can observe the fluoroscopic image or surgical field with minimal eye movements. The operator can instantly refer to the fluoroscopic image intraoperatively without moving the head. (C) Without wearing the SG, it is necessary to divert attention from the surgical field when looking at the fluoroscopic monitor.

Statistics ver. 20.0 (IBM Co., Armonk, NY, USA). The groups SG and N-SG overall and by operator comparisons of the insertion time, radiation exposure time, and radiation dose were performed using an unpaired t-test. Similarly, the chi-square test was used to compare the 2 groups for the PPS’s deviation from the pedicle. All statistical tests were 2-sided, and statistical significance was set at p < 0.05.

RESULTS

Overall, the 2 groups showed no significant difference in insertion time (SG: 503.5 ± 144.7 seconds, N-SG: 549.4 ± 119.8 seconds; p = 0.28); however, radiation exposure time was significantly shorter, and radiation dose was significantly lower in group SG than group N-SG (SG: 109.1 ± 43.5 seconds, N-SG: 150.9 ± 38.7 seconds; p = 0.003, SG: 1.3 ± 0.6 mGy, N-SG: 1.7 ± 0.5 mGy; p = 0.023, respectively) (Fig. 5A). For operator A, there was no significant difference in insertion time between the 2 groups (SG: 485.2 ± 116.6 seconds, N-SG: 516.4 ± 124.6 seconds; p = 0.57), whereas radiation exposure time was significantly shorter, and radiation dose was also significantly lower in group SG than group N-SG (SG: 100.0 ± 29.0 seconds, N-SG: 157.9 ± 42.8 seconds; p = 0.003, SG: 1.2 ± 0.4 mGy, N-SG: 1.8 ± 0.5 mGy; p = 0.013, respectively). For operator B, the 2 groups showed no significant differences in insertion time (SG: 521.8 ± 172.9 seconds, N-SG: 582.4 ± 111.1 seconds; p = 0.37), radiation exposure time (SG: 118.1 ± 54.5 seconds, N-SG: 143.9 ± 34.9 seconds; p = 0.23), and radiation dose (SG: 1.4 ± 0.7 mGy, N-SG: 1.6 ± 0.5 mGy; p = 0.45) (Fig. 5B).

There was no significant difference in insertion accuracy between the 2 groups’ overall procedure or by operator comparisons; however, there was a tendency for less deviation from the pedicle in group SG for operator B (p = 0.24) (Table 2). Regarding the details of the deviation types (Table 1), in the case of operator A, group SG had 3 instances of type B, 1 instance each of type C, D, and E, while group N-SG had 4 instances of type B and 2 instances of type C. In the case of operator B, group SG had 2 instances of type B, and group N-SG had 3 instances of type B and 3 instances of type I.

DISCUSSION

This study highlights that using SG during PPS insertion can significantly reduce radiation exposure compared with the conventional method. Furthermore, in the case of operator B, who had less experience with spinal surgeries, a greater reduction in radiation exposure was demonstrated with the use of SG.

Wearing SG minimizes the need to divert attention from the surgical field, whereas conventionally, surgeons have to check the fluoroscopic monitor during the procedure. With SG, surgeons can observe the fluoroscopic image or surgical field with minimal eye movement. This enabled stable screw insertion procedures, leading to a reduction in the fluoroscopy time of each scan, and consequently resulting in a decrease in both the total fluoroscopy time and radiation dose. Additionally, there are several reasons why SG contributed to the reduction in time under fluoroscopic operation. Firstly, the use of the wearable displays equipped in SG has been reported to reduce posture

https://doi.org/10.14245/ns.2448090.045
discomfort. This could potentially decrease fatigue during continuous screw insertion, as in our study. Furthermore, the images on the wearable displays are closer than those on a fluoroscopic monitor, allowing for a clearer view. This could have been particularly beneficial in scenarios requiring detailed image examination, such as verifying the pedicle.

SG application has helped improve the accuracy of guidewire insertion in femoral fracture surgeries and reduce the radiation.
The utility of SG has been reported in the trauma field; however, there are limited studies on the spinal region and no reports addressing radiation exposure reduction. Navigation systems can also reduce radiation exposure in spine surgeries; however, they are not commonly used because of their high cost and the need for additional skin incisions to place the navigation reference. Furthermore, using navigation systems may protect the surgeon and paramedical staff from radiation exposure but not to the patient. Mendelsohn reported that the radiation exposure of patients undergoing spine surgery under intraoperative CT-based navigation was approximately 2.7 times higher than that under fluoroscopic guidance. On the other hand, SG have the advantage of not requiring registration like navigation systems, allowing for immediate use without radiation exposure during nonsurgical procedures. In conclusion, SG constitute a low-cost and easy-to-implement option for reducing radiation exposure for medical staff and patients.

The 3 principles of radiation safety are time (minimizing the time spent near the radiation source), distance (maximizing the distance from the radiation source), and shielding (using appropriate shielding devices). Therefore, performing procedures as far as possible from the irradiation field while not hindering the operator’s skills is necessary in addition to using shielding devices such as lead glasses, thyroid protectors, aprons, and radiation-reducing gloves. Lead goggles exhibited the highest shielding effect when the line-of-sight was directed toward the main scattering source. This indicates that the shielding effect was maximized when the gaze was focused on the surgical field while confirming the anteroposterior image. Consequently, when diverting the gaze from the surgical field to view the fluoroscopy monitor, there is an increased radiation exposure to the lens. Therefore, using SG to keep the operator’s gaze on the surgical field, especially during surgeries that confirm the anteroposterior image, can efficiently enhance the lens’s radiation protection. In actual clinical practice, continuous fluoroscopy is primarily utilized in the anteroposterior view during PPS insertions. This context could further highlight the importance of SG, especially if they were equipped with x-ray shielding capabilities. However, “time” is the most effective factor among the 3 principles, and unnecessary irradiation should be avoided as much as possible. As most of the radiation exposure to the operator comes from scattered radiation from the patient, measures to reduce patient exposure, such as minimizing fluoroscopy time, are often effective in reducing operator exposure. This adheres to the ALARA (As Low As Reasonably Achievable) principle, which aims to minimize a patient’s radiation exposure as much as possible. “Time” is a parameter that can be shortened depending on the ingenuity employed during the procedure. This study demonstrated that SG application effectively reduces radiation exposure time, providing a significant benefit in minimizing the health damage caused by radiation exposure. Reducing surgical time through skill improvement is also essential in minimizing radiation exposure time. This study revealed that the effect of SG usage on reducing radiation exposure was greater for operator A, who was less experienced in spinal surgery. Inexperienced surgeons, who had to repeatedly alternate their focus between the surgical field and the fluoroscopic monitor, could achieve stable surgeries using SG without moving their heads. Therefore, using SG could bridge the gap in surgeon skill levels, which should be particularly beneficial for less-experienced surgeons.

This study has some limitations. First, the model bones could not perfectly replicate human bone structures and we were unable to use materials that closely simulate the tension and realism of living tissues to cover the model bones. This might have affected the experimental results. However, we regard this study’s results as a milestone for future clinical trials. Second, the number of experiments was limited. Therefore, the challenges encountered during a single screw insertion significantly influenced the results. In the case of operator B, there was considerable variability in insertion time, radiation dose, and radiation exposure.

### Table 2. The comparison of screw accuracy between groups SG and N-SG overall and by operator

<table>
<thead>
<tr>
<th>Group</th>
<th>Deviation (+)</th>
<th>Deviation (−)</th>
<th>Accuracy (%)</th>
<th>Total</th>
<th>Deviation (+)</th>
<th>Deviation (−)</th>
<th>Accuracy (%)</th>
<th>Total</th>
<th>Deviation (+)</th>
<th>Deviation (−)</th>
<th>Accuracy (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>8</td>
<td>32</td>
<td>80</td>
<td>40</td>
<td>6</td>
<td>14</td>
<td>70</td>
<td>20</td>
<td>2</td>
<td>18</td>
<td>90</td>
<td>20</td>
</tr>
<tr>
<td>N-SG</td>
<td>12</td>
<td>28</td>
<td>70</td>
<td>40</td>
<td>6</td>
<td>14</td>
<td>70</td>
<td>20</td>
<td>6</td>
<td>14</td>
<td>70</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>60</td>
<td>75</td>
<td>p = 0.44</td>
<td>12</td>
<td>28</td>
<td>70</td>
<td>p = 1.00</td>
<td>8</td>
<td>32</td>
<td>80</td>
<td>p = 0.24</td>
</tr>
</tbody>
</table>

SG, smart glasses; N-SG, non-SG.

---

exposure time, especially in group SG. This variability might have masked the true significance of SG. therefore, further trials may be required to obtain sophisticated data on accuracy, reproducibility, and learning curve. Finally, the SG used in this study (MOVERIO) could not be used with radiation protection goggles, reducing the lens’ protection against exposure. Low doses of radiation exposure can cause late-onset radiation cancer; therefore, using lead-lined goggles to protect the lens has been well-established for performing procedures under fluoroscopy. However, there is a report on SG that can be attached to goggles, and this is a useful approach that can be implemented quickly. In addition, Dorey et al. have reported on the usefulness of SG with lead-shielded lenses, and further research is necessary on the development and clinical application of SG equipped with shielding functions in the lenses.

CONCLUSION

Using SG during PPS insertion can significantly reduce radiation exposure compared with the conventional method. SG application helps minimize potential harm to healthcare professionals by reducing the time spent near the radiation source. SG application is a low-cost, easy-to-implement option for reducing radiation exposure during spinal surgery.

NOTES

Conflict of Interest: Yoshiki Takeoka and Kohei Kuroshima endowed course (Surgical Spine, Inc., Tokyo, Japan and SMI, Inc., Nagoya, Japan). Except for that, the other authors have nothing to disclose.

Funding/Support: This study was supported by Johnson & Johnson, Inc., Tokyo, Japan.

Author Contribution: Conceptualization: YH, YT, RK, KKakutani; Data curation: YH, YT, HO, TM, MR, NK, KK, RK, KKakutani; Formal analysis: YH, YT, TT, YK, KM, HO, TM, MR, NK, KKuroshima, RK, KKakutani; Funding acquisition: TY; Methodology: YH, YT, RK, KKakutani; Project administration: YH, YT, TY, TT, YK, KM, HO, TM, MR, NK, KKuroshima, RK, KKakutani; Visualization: YH; Writing – original draft: YH, YT; Writing – review & editing: YT, TT, YK, KM, HO, TM, MR, NK, KKuroshima, RK, KKakutani.

ORCID

Yoshiaki Hiranaka: 0000-0002-7479-7345
Yoshiki Takeoka: 0000-0001-5360-5618

Takashi Yurube: 0000-0002-3007-361X
Takeru Tsujimoto: 0000-0002-1516-3107
Yutaro Kanda: 0000-0001-5538-5516
Kunihiro Miyazaki: 0000-0001-7047-5595
Hiroki Ohnishi: 0009-0002-9305-6291
Tomoya Matsuou: 0009-0004-4824-5768
Masao Ryu: 0009-0008-4133-2489
Naotoshi Kumagai: 0000-0002-0961-2631
Kohei Kuroshima: 0000-0003-3327-0712
Kenichiro Kakutani: 0000-0002-4006-037X

REFERENCES

Commentary on “The Utility and Feasibility of Smart Glasses in Spine Surgery: Minimizing Radiation Exposure During Percutaneous Pedicle Screw Insertion”

Wongthawat Liawrungrueang

Department of Orthopaedics, School of Medicine, University of Phayao, Phayao, Thailand

Scientific knowledge used to medicine to aid in diagnosis, prevention, treatment, and innovation is referred to as medical technology. It does this by creating tools, machines, and pharmaceuticals using engineering and biotechnology methods. The manufacturing of equipment and techniques utilized in the medical field such as augmented reality (AR)-assisted real-time visualization of spine surgery, neuromonitoring systems, robotics-assisted surgery, robotic-assisted pedicle screw placement, and intraoperative navigation systems is specifically referred to when discussing spinal medical technology.

Augmented and mixed-reality technologies are included in smart glasses (SG) for spine surgery, giving surgeons access to real-time imaging, guidance, and patient information. With the use of these glasses, the surgeon may plan and navigate surgery more efficiently by the image on the wearable displays are closer than those on a fluoroscopic monitor, allowing for a clearer view and reducing radiation exposure during percutaneous pedicle screw (PPS) insertion. This study examines the potential and usefulness of SG in spine surgery. Adoption of SG offers a possible way to reduce related health concerns, since radiation exposure to spine surgeons during fluoroscopy-guided treatments increases.

The MOVERIO SG manufactured by Epson Co., Ltd. (Tokyo, Japan) are a series of wearable AR devices designed for various applications. The latest MOVERIO smart glass delivers an engaging AR experience through quality QHD (quad high definition) or 3-dimensional (3D) images. Its binocular and lightweight see-through display also keeps you aware of your surroundings while you are viewing your content. The objective of the research, which employed operators with varying degrees of experience, was to assess how much SG reduced radiation exposure and increased procedural accuracy. Operators alternated between SG and traditional approaches to direct the insertion of PPS into lumbar model bones under fluoroscopic supervision, using the MOVERIO SG model BT-30E and the COREVISION 3D fluoroscopy system. The SG and non-SG groups’ insertion times did not differ significantly, according to the data. However, especially for less experienced operators, the use of SG considerably decreased the duration and amount of radiation exposure. Additionally, deviation studies showed that SG did not impair the precision of screw insertion.

The introduction of SG addressed critical concerns regarding radiation exposure in spine surgery.
surgery, in line with the principles of minimizing time near radiation sources, maximizing distance, and using shielding devices. By projecting real-time fluoroscopic images into wearable displays, SG enabled surgeons to maintain focus on the operative field, minimizing the need for head movement and reducing fatigue compared to traditional methods. Furthermore, SG facilitated clearer image visualization, potentially enhancing procedural accuracy. While acknowledging limitations such as the use of model bones and a small sample size, the study underscores the potential of SG integration in spinal surgeries to enhance safety and optimize outcomes. Future research could explore SG integration with radiation protection goggles and validate findings in larger clinical settings.

In future direction, SG is poised to transform spine surgery with their potential applications. The application of AR guidance augments the surgeon’s field of view with digital information, hence improving surgical precision. Experts can offer real-time instruction throughout difficult operations through re-

Table 1. Summary of the advancements in current surgical smart glass in the field of spine surgery shows great potential and is continuously developing

<table>
<thead>
<tr>
<th>Product</th>
<th>Company</th>
<th>Description of surgical glass</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>HoloLens&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Microsoft</td>
<td>A headset offering mixed-reality experiences and precise guidance for surgical procedures operated hands-free</td>
<td>Enables a nuanced division between the real and digital realms through an immersive mixed-reality encounter. Utilizing holographic overlays enhances accuracy and reduces errors in surgery. Enables users to engage with holograms and complete tasks using hands-free technology, eliminating the need for physical input.</td>
</tr>
<tr>
<td>M-Series Smart Glasses&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Vuzix</td>
<td>Lightweight and ergonomic smart glasses packed with various features for augmented reality experiences</td>
<td>These devices are characterized by a comfortable and lightweight design, perfect for long-term use. High-quality display and optics, they offer clear visuals and an enhanced user interface.</td>
</tr>
<tr>
<td>MOVERIO Smart Glasses&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Epson</td>
<td>Smart glasses equipped with augmented reality capabilities, featuring a binocular display for immersive experiences</td>
<td>Enhances realism and offers depth awareness through its binocular display, creating an immersive experience. Provides flexible applications with interactive 3-dimensional models that are beneficial for training, teaching, and amusement.</td>
</tr>
<tr>
<td>Atheer AiR Glasses&lt;sup&gt;13&lt;/sup&gt;</td>
<td>At hear</td>
<td>Gesture-controlled smart glasses with augmented reality interface, offering real-time guidance for surgical procedures</td>
<td>The user experience and efficiency are improved with intuitive gesture controls that make interaction smooth and easy to use. With the help of real-time augmented reality guidance, surgeons may perform surgeries more accurately and with better patient outcomes.</td>
</tr>
<tr>
<td>X2 Smart Glasses&lt;sup&gt;14&lt;/sup&gt;</td>
<td>ThirdEye Gen</td>
<td>Lightweight smart glasses with a wide field of view, providing an extensive augmented reality experience</td>
<td>A large field of view improves immersion and makes it easier for people to interact with digital information. Well-known for their comfortable and lightweight designs, they can be used for extended periods of time in a variety of settings.</td>
</tr>
<tr>
<td>HMT Smart Glasses&lt;sup&gt;15&lt;/sup&gt;</td>
<td>RealWear</td>
<td>Ruggedized smart glasses designed for hands-free operation in industrial environments, featuring a voice-operated interface</td>
<td>Industrial-grade durability guarantees reliability in tough work environments such as manufacturing, construction, and field services. Hands-free usage is made possible by voice-activated interfaces, enhancing both efficiency and user experience.</td>
</tr>
<tr>
<td>Xvision&lt;sup&gt;16&lt;/sup&gt;</td>
<td>Augmedics</td>
<td>Surgical navigation system utilizing augmented reality for precise real-time guidance during procedures</td>
<td>By offering accurate real-time guidance, it improves surgical accuracy and lowers the possibility of complications. Additionally, it increases safety by reducing errors and enhancing results by giving surgeons vital information and visualizations during surgeries.</td>
</tr>
<tr>
<td>Caduceus S&lt;sup&gt;17&lt;/sup&gt;</td>
<td>SURGLASSES Inc.</td>
<td>Augmented reality display shows real-time info to surgeons, hands-free for focus and enables remote collaboration. Customizable interface for personal support, sterile, ergonomic design for surgical settings. Training modules enhance skills through immersive simulations</td>
<td>Enhanced visuals and hands-free operation promote accurate decision-making and focus. Real-time collaboration offers instant support, while personalized interfaces and clean design meet hygiene standards. Immersive training aids in skill development.</td>
</tr>
</tbody>
</table>

https://doi.org/10.14245/ns.2448568.284
mote help. They also facilitate training and education through immersive experiences and live streaming. Integration with surgical navigation systems ensures accurate feedback on instrument positioning. Patient-specific planning optimizes surgical approaches, while enhanced communication features streamline teamwork in the operating room. As technology advances, SG promise to revolutionize spine surgery, improving outcomes and patient care. My summary of the advancements in current surgical smart glass in the field of spine surgery shows great potential and is continuously developing (Table 1).

This commentary on this study highlights the promising impact of SG in reducing radiation exposure and enhancing procedural efficiency in spine surgeries. It presents a viable and cost-effective solution to bolster surgical safety, thereby reducing health risks for patients and healthcare professionals alike. However, it's essential for spine surgeons to possess the necessary knowledge and skills to conduct surgeries autonomously, without solely depending on robotics or computer assistance. Emergency situations demand swift and informed decision-making, calling for the expertise and proficiency of physicians. It's crucial to acknowledge that complications may arise unexpectedly.

- **Conflict of Interest:** The author has nothing to disclose.

**REFERENCES**

Radiological and Clinical Significance of Cervical Dynamic Magnetic Resonance Imaging for Cervical Spondylotic Myelopathy

Jun Jae Shin¹, Sun Joon Yoo¹, Tae Woo Kim², Jae-Young So³, Won Joo Jeong⁴, Mu Ha Lee⁵, Joongkyum Shin⁶,⁷, Yoon Ha⁶,⁸

¹Department of Neurosurgery, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Korea
²Department of Neurosurgery, Inje University Sanggye Paik Hospital, Inje University College of Medicine, Seoul, Korea
³Department of Neurosurgery, National Health Insurance Service Ilsan Hospital, Goyang, Korea
⁴Department of Hospital Medicine, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Korea
⁵Department of Neurosurgery, Gangnam Severance Hospital, Yonsei University School of Medicine, Seoul, Korea
⁶Department of Neurosurgery, Yonsei University College of Medicine, Seoul, Korea
⁷University of Pacific, Stockton, CA, USA
⁸POSTECH Biotech Center, Pohang University of Science and Technology, Pohang, Korea

Objective: The study compared the morphometric changes of the cervical spinal cord using dynamic magnetic resonance imaging (MRI) in patients with cervical spondylotic myelopathy (CSM) and assessed the correlation with kinematic changes, cord cross-sectional area (CSA), and high signal intensity (SI) on T2-weighted imaging (T2WI).

Methods: Patients with CSM were evaluated through dynamic MRI for sagittal and axial CSA changes of the cervical cord, cerebrospinal fluid (CSF) reserve ratio, degree of cord impingement, cord compression rate, range of motion (ROM), and severity of SI on T2WI. The degree of cord impingement was evaluated using the Muhle grading system. Clinical outcomes were assessed using Japanese Orthopaedic Association scoring and Nurick grade.

Results: The study included 191 patients (113 males) with a mean age of 55.34 ± 12.09 years. The lowest sagittal CSF reserve ratio and cord occupation rate were observed during extension. Cord impingement and SI change were more prevalent in extension-positioned MRI. Preoperative cervical ROM was greater in patients with intensely high SI change. There was no difference between ROM on dynamic radiographs and dynamic MRI. Preoperative cervical ROM was greater in patients with intensely high SI change.

Conclusion: Dynamic MRI is useful for evaluating neck movement. Patients with high SI had greater ROM before surgery but worse outcomes after. Neck extension exacerbated cervical stenosis and cord compression compared to flexion, and cervical spinal motion contributed to the severity of CSM. Cervical spinal motion should be carefully evaluated, particularly in hyperextension, to prevent worsening of CSM.

Keywords: Cervical myelopathy, Dynamic MRI, Cervical motion, Signal intensity, Anterior decompression, Posterior decompression

INTRODUCTION

Magnetic resonance imaging (MRI) is essential to diagnose cervical spondylotic myelopathy (CSM), as it can accurately determine the relationship between the disc and the spinal cord (SC) and the presence of intramedullary signal change. High signal intensity (SI) on T2-weighted imaging (T2WI) has been regarded as an important prognostic factor and was found to
correlate with the severity of CSM. Studies have demonstrated a high signal change in T2WI, cord compression, and neurological outcomes with only static MRI. Additionally, some pathologies can be found to reflect dynamic strain of the cervical spine when invisible cases using static imaging.

Dynamic mechanical factors are reported as the cause of the occurrence and exacerbation of cervical myelopathy. Dynamic MRI is beginning to be clinically applied for the diagnosis and treatment of cervical spine diseases, especially CSM, which is reported to cause and worsen symptoms through cervical motion. However, there is still no consensus on the significance of changes in SI severity and dynamic MRI findings. Furthermore, the impact of dynamic factors on the clinical outcomes of cervical myelopathy needs to be better understood.

In this study, we investigated prospectively dynamic MRI in preoperative neck flexion-extension positions in cervical myelopathic patients. We compared the morphometric changes of cord compression on dynamic MRI, changes in the SI on T2WI, and parameters of the cervical spine in patients with CSM. The objective of the study was to analyze the relationship between preoperative intramedullary signal change on MRI, dynamic factors, and surgical outcomes.

**MATERIALS AND METHODS**

1. Patient Populations

This prospective cohort study included 233 patients with CSM caused by cervical disc protrusion, bony spur, or ossification of the posterior longitudinal ligament (OPLL) who underwent surgery at Yongin Severance Hospital and Inje University Sanggye Paik Hospital between June 2018 and December 2022. Among these patients, 32 patients were excluded due to congenital anomaly, trauma, previous history of cervical fusion, or ankylosing spondylitis (Fig. 1). This study was approved by the Institutional Review Board of Yonsei College of Medicine, Yongin Severance (IRB No. 9-2023-0222). A diagnosis of CSM was assigned with radiological confirmation by MRI, and the diagnosis was determined when one or more upper motor neuron domains were involved (e.g., spasticity, hyperreflexia, positive Babinski sign), based on neurological examination.

2. Radiologic Assessment

1) Plain radiographs

Preoperative lateral standing plain radiographs were obtained in neutral, flexion, and extension positions. The following parameters were evaluated: C2–7 angle (the angle created by a line parallel to the inferior end plate of the C2 body and a line parallel to that of the C7 body was measured on neutral, flexion, and extension position). The C2–7 angle was measured in the neutral position, flexion position, and extension position. Cervical range of motion (ROM) was measured as the difference of C2–7 angle values between the extension and flexion positions. The dynamic parameter of C2–7 ROM was calculated using the following formula: C2–7 ROM (°) = (extension C2–7 Cobb lordotic angle) – (flexion C2–7 Cobb lordotic angle).

2) Patient positioning and MRI protocol

All patients underwent high-resolution MRI using the 3.0T Signa MRI unit (GE HealthCare, Chicago, IL, USA) or 3.0T Skyra unit (Siemens, Munich, Germany). For cervical dynamic MRI, the patient’s neck was positioned in extension and flexion using a custom-made cushion placed under the head and shoulders (Fig. 2). A routine neutral examination was first performed in the supine position. The patients were tolerable to the neck extension and flexion position under the observation of a physician. There was no standard predetermined flexion-extension position to avoid neurologic problems. Neutral MRI obtained T1- and T2-weighted sequences in sagittal and axial views; only T2-weighted images were obtained by dynamic MRI.

3) Signal intensity grading

Increased SI referred to a high-intensity area compared with the adjacent isointensity portion of the SC in both sagittal and axial planes. We defined the increased SI at the narrowest level of the SC as “grade 0 (G0)” if no intramedullary high SI appeared on the T2WI, as “grade 1 (G1)” if there was a predominantly
faint and indistinct border, and as “grade 2 (G2)” if there was a predominantly intense and well-defined border.2,3

4) Cervical stenosis grading

Cervical stenosis was evaluated using the classification of Muhle et al.,8 which has 4 stages: stage 0, normal width of the spinal canal, and no signs of anterior and posterior subarachnoid space narrowing; stage 1, partial obliteration of the anterior or posterior subarachnoid space or of both; stage 2, complete obliteration of the anterior or posterior subarachnoid space or of both; and stage 3, anterior or posterior cord impingement or both. Muhle grade was determined by evaluating multi-positional MRI at the most compressed lesion. The degree of cervical stenosis was evaluated using Muhle classification on T2WI in 3 kinetic positions: neutral, flexion, and extension (axial and sagittal).

5) Cross-sectional area

The cross-sectional area (CSA) of the SC was measured at the greatest compressed levels. The CSA was obtained at the affected level during flexion, neutral position, and extension (Fig. 3). The CSA of the SC and canal on the midsagittal plane was measured at the area between a line crossing the SC at the lower endplate of C2 and the lower endplate of C7. A comparison of

Fig. 2. Positions of the cervical spine for the dynamic magnetic resonance imaging (MRI). During cervical dynamic MRI, the patient’s neck was alternately positioned in extension and flexion using a custom-made cushion under the head and shoulders. Flexion-positioned MRI (A), neutral-positioned MRI (B), and extension-positioned MRI (C).

Fig. 3. Measurement technique of morphometric parameters on T2-weighted magnetic resonance imaging (MRI) during flexion, neutral, and extension. Flexion MRI in sagittal section (A), neutral MRI in sagittal section (B), extension MRI in sagittal section (C), flexion MRI in axial section at C4–5 (D), neutral MRI in axial section at C4–5 (E), and extension MRI in axial section at C4–5 (F). Black dashed line indicates the SC area, black solid line indicates the SC plus CSF area of compression level, and white dashed line indicates the C2–7 angle. CSF, cerebrospinal fluid; SC, spinal cord.
morphometric parameters at compression levels in all 3 positions on axial T2WI was performed, including the SC area, cerebrospinal fluid (CSF) area, and CSF reserve ratio (CSF/CSF+SC). The degree of SC compression changes in the cervical flexion-extension state was compared and analyzed to measure the CSA using software (ZeTTA PACS, TaeYoung Soft Co., Ltd., Gwacheon, Korea). CSA measurements were conducted 2 times, and the mean was used to minimize intraobserver differences.

6) Diameter
On axial and sagittal T2WI MRI, the SC occupation rate in the dural sac was calculated using the following formula: (diameter of the SC)/(diameter of the dural sac) × 100. The measurements were performed at the level of maximum SC compression. The level with the smallest SC area was selected for patients with multisegmental involvement.

3. Assessment of Clinical Outcomes
Clinical outcomes were assessed using the Japanese Orthopaedic Association (JOA) scoring system and Nurick grade. Data were obtained for all patients preoperatively and for a minimum of 24 months postoperatively. The recovery ratio was calculated using the following formula:

JOA recovery ratio (%) = (postoperative JOA score–preoperative JOA score)/(17 (full score)–preoperative JOA score)

4. Treatments
The surgical treatments were determined based on various factors such as the number of affected levels, patient comorbidities, cervical alignment, instability, occupying ratio, and surgeon preference. Various surgical approaches such as anterior, posterior, or combined operations were chosen to decompress affected levels. Follow-up examinations were conducted at 3-month, 6-month, 12-month, or 24-month intervals after the surgery to investigate the clinical correlation between postoperative neurological outcomes and radiographic imaging findings.

5. Statistical Analysis
All data are expressed as mean ± standard deviation or percentage. For the demographic data, the means of continuous and categorical data were compared by Student t-test and chi-square test, respectively. Repeated measures analysis of variance was performed. Post hoc Scheffé test was used to compare the mean results on neutral, flexion, and extension MRI at the compression levels. Radiological parameters measured by 2 observers were analyzed using the Cohen κ coefficient for categorical variables and intraclass correlations (ICCs; 2-way mixed model with consistency agreement; 95% confidence interval) for continuous variables. Cohen κ coefficient and ICC values were categorized as poor agreement (0.00–0.20), fair agreement (0.21–0.40), moderate agreement (0.41–0.60), good agreement (0.61–0.80), and very good agreement (0.81–1.0). Patients were classified based on the severity of SI on T2WI in a neutral position, such as none, faint, or intense SI. The stage of stenosis based on the classification of Muhle et al. was considered a quantitative variable. All statistical analyses were performed using MedCalc ver. 22.014 (MedCalc, Mariakerke, Belgium), and p-values < 0.05 were considered to indicate statistical significance.

RESULTS
Clinical and radiographic data were recorded from 191 patients

Table 1. Patient demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>55.34 (29–79)</td>
</tr>
<tr>
<td>Sex, male:female</td>
<td>113:78</td>
</tr>
<tr>
<td>Preoperative clinical assessment</td>
<td></td>
</tr>
<tr>
<td>JOA</td>
<td>11.07 ± 2.41 (5–15)</td>
</tr>
<tr>
<td>Nurick</td>
<td>2.56 ± 0.98 (1–5)</td>
</tr>
<tr>
<td>2-Year postoperative clinical assessment</td>
<td></td>
</tr>
<tr>
<td>JOA</td>
<td>14.89 ± 1.75 (9–17)</td>
</tr>
<tr>
<td>Recovery rate (%)</td>
<td>85.49 ± 13.21 (40–100)</td>
</tr>
<tr>
<td>Nurick</td>
<td>1.63 ± 0.87 (0–4)</td>
</tr>
<tr>
<td>Most compressed levels</td>
<td></td>
</tr>
<tr>
<td>C3–4</td>
<td>36 (18.8)</td>
</tr>
<tr>
<td>C4–5</td>
<td>50 (26.2)</td>
</tr>
<tr>
<td>C5–6</td>
<td>77 (40.3)</td>
</tr>
<tr>
<td>C6–7</td>
<td>28 (14.7)</td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
</tr>
<tr>
<td>Anterior</td>
<td>113 (59.16)</td>
</tr>
<tr>
<td>Posterior</td>
<td>69 (36.13)</td>
</tr>
<tr>
<td>Combined anterior-posterior</td>
<td>9 (4.71)</td>
</tr>
<tr>
<td>Pathology</td>
<td></td>
</tr>
<tr>
<td>Disc</td>
<td>141 (73.8)</td>
</tr>
<tr>
<td>Osteophyte</td>
<td>12 (6.3)</td>
</tr>
<tr>
<td>OLF</td>
<td>9 (4.7)</td>
</tr>
<tr>
<td>OPLL</td>
<td>29 (15.2)</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation (range) or number (%). JOA, Japanese Orthopaedic Association; OLF, ossification of ligament flavum; OPLL, ossification of the posterior longitudinal ligament.
(113 males and 78 females) who underwent anterior decompression (n = 113), posterior cervical decompression (n = 69), including cervical laminoplasty (n = 43) or laminectomy with fusion (n = 26), and combined anterior-posterior surgery (n = 9) for CSM and completed at least 24 months of follow-up. The average follow-up period was 49.8 ± 13.4 months (25–64 months). The mean age at the time of surgery was 55.34 ± 12.09 years (range, 29–79 years). Baseline patient characteristics, demographics, and surgical information are listed in Table 1.

1. Radiological Findings According to Neck Motion

The sagittal CSA of the cervical cord was smaller when the neck was extended (516.95 ± 106.20 mm²) compared with when it was in a neutral position (534.04 ± 86.96 mm²) or flexed (560.70 ± 90.53 mm²). Similarly, the sagittal CSA of CSF was smaller in the extension position (294.53 ± 101.02 mm²) compared with the neutral (353.50 ± 106.61 mm²) and flexion positions (349.43 ± 123.80 mm²). The sagittal CSF reserve ratio also showed significant differences between flexion-extension neck motion and the neutral posture (p = 0.004). However, there were no significant changes in the axial CSA of the SC, CSF area, or CSF reserve ratio during neck movement (Table 2).

In addition, the diameter of the sagittal SC and cord occupation rate at the compression level were significantly different depending on neck movement, with the lowest diameter observed during the extended posture. The sagittal cord occupation rate was lowest during extension (69.07% ± 20.62%) posture compared with neutral (74.36% ± 18.07%) and flexion (78.02% ± 17.31%) posture (p < 0.001). The smallest diameter of the axial SC was observed in the extension posture compared with flexion and neutral postures (p = 0.015). During neck motion, there was no difference in the diameter of the dural sac, and the axial cord occupancy rate was also similar (Table 2).

2. Signal Intensity, Compression Level, and Cervical Stenosis With Neck Motion

The severity of intramedullary high SI on T2WI changed with neck movement. Out of 191 patients, 70 (36.65%) showed no intramedullary SI (G0) on sagittal T2WI in a neutral position, 47 patients had faint intramedullary SI (G1), and 74 had intense SI change (G2). However, in the extension posture, 3 of 70 patients who had no SI in the neutral position exhibited faint SI (G1). The percentage of high SI (G1 and G2) increased in the extension posture (124 of 191, 64.92%) compared with the neutral posture (121 of 191, 63.35%), while that of the high SI decreased in the flexion posture (118 of 191, 61.78%) (Fig. 4). Our results showed that 4.3% of patients with no SI in a neutral position displayed SI changes in an extension MRI, and 6.4% of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Flexion</th>
<th>Neutral</th>
<th>Extension</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagittal CSA at compression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC (mm²)</td>
<td>560.70 ± 90.53</td>
<td>534.04 ± 86.96</td>
<td>516.95 ± 106.20</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>CSF area (mm²)</td>
<td>349.43 ± 123.80</td>
<td>353.50 ± 106.61</td>
<td>294.53 ± 101.02</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>CSF reserve ratio (%)</td>
<td>37.63 ± 9.63</td>
<td>39.34 ± 8.71</td>
<td>35.89 ± 9.06</td>
<td>0.004*</td>
</tr>
<tr>
<td>Axial CSA at compression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC (mm²)</td>
<td>68.95 ± 15.18</td>
<td>67.60 ± 15.44</td>
<td>68.03 ± 15.92</td>
<td>0.637</td>
</tr>
<tr>
<td>CSF area (mm²)</td>
<td>79.77 ± 33.89</td>
<td>77.33 ± 30.78</td>
<td>76.51 ± 35.26</td>
<td>0.594</td>
</tr>
<tr>
<td>CSF reserve ratio (%)</td>
<td>51.55 ± 13.08</td>
<td>51.60 ± 13.37</td>
<td>50.50 ± 14.31</td>
<td>0.759</td>
</tr>
<tr>
<td>Sagittal diameter at compression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC (mm)</td>
<td>5.04 ± 1.29</td>
<td>4.70 ± 1.26</td>
<td>4.59 ± 1.55</td>
<td>0.006*</td>
</tr>
<tr>
<td>Dural sac (mm)</td>
<td>6.45 ± 0.72</td>
<td>6.33 ± 0.69</td>
<td>6.61 ± 0.82</td>
<td>0.002*</td>
</tr>
<tr>
<td>Cord occupation rate (%)</td>
<td>78.02 ± 17.31</td>
<td>74.36 ± 18.07</td>
<td>69.07 ± 20.62</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Axial diameter at compression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC (mm)</td>
<td>5.97 ± 1.02</td>
<td>5.67 ± 1.05</td>
<td>5.48 ± 1.49</td>
<td>0.015*</td>
</tr>
<tr>
<td>Dural sac (mm)</td>
<td>10.31 ± 1.96</td>
<td>10.19 ± 1.97</td>
<td>10.04 ± 2.25</td>
<td>0.472</td>
</tr>
<tr>
<td>Cord occupation rate (%)</td>
<td>59.49 ± 12.43</td>
<td>55.60 ± 9.61</td>
<td>56.64 ± 16.39</td>
<td>0.096</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. CSA, cross-sectional area; SC, spinal cord; CSF, cerebrospinal fluid.

*p < 0.05, statistically significant differences.
patients with SI in a neutral position displayed no SI in a flexion MRI. Compared with the neutral position, the extension-positioned MRI increased compression levels in 122 of 191 (63.87%) patients. On the other hand, flexion-positioned MRI decreased compression levels in 16 of 191 patients (8.38%). Moreover, upon analysis of Muhle classification, the grade of cervical stenosis increased in the extension posture compared with the neutral and flexion postures (Table 3).

### 3. Radiological and Clinical Findings According to the Severity of Signal Intensity

Patients who had intensely high SI on T2WI before surgery had significantly lower preoperative JOA scores and recovery ratios (10.09% ± 2.33% and 79.82% ± 15.26%, respectively) compared with those who did not have SI changes (12.09% ± 2.13% and 92.93% ± 6.35%, respectively). The mean preoperative cervical ROM was more prominent in patients with intensely high SI change compared with those with no SI, with values of 45.04° ± 18.90° in SI G2, 35.05° ± 15.63° in SI G1, and 30.34° ± 12.31° in SI G0 (p < 0.001) (Table 4). Moreover, there was no difference between ROM on dynamic radiographs and dynamic MRI (p = 0.3296).

### 4. Interobserver Reliability

For the CSA assessment, interobserver reliability was classi-
fied as "very good agreement" (ICC between 0.85 and 0.95) at the compressed level and in the neutral, flexion, and extension positions (Table 5). The interobserver reliability for cord and dural sac diameters was classified as "very good agreement" (ICC between 0.81 and 0.86) at the compressed level and in the neutral, flexion, and extension positions. Details are presented in Table 5. The interobserver reliability was found to be very good, with a Cohen $\kappa$ coefficient of 0.99 for the intramedullary SI grade on sagittal T2WI and 0.99 for the detection of the Muhle grade of cervical stenosis.

**DISCUSSION**

Several previous researchers used quantitative MRI to measure the SC's transverse area or the diameters of compressed lesions. They reported that the spinal canal becomes narrower when the neck is extended. However, no previous prospective research analyzed changes in the CSAs of the SC and subarachnoid space or changes in intramedullary SI in the sagittal plane using dynamic MRI. The present study compared the morphometric measurements of the SC in kinematic motion among patients with CSM and investigated the relationship between cervical motion, myelopathic symptoms, and the severity of SI. The results revealed that the morphometric SC was more compressed during neck extension than during neck flexion or neutral positions. The sagittal plane of T2WI on magnetic resonance was more effective than axial T2WI in detecting changes in the CSA of the cord and CSF reserve ratio. The diameter of the cord became smaller during neck extension on axial and sagittal MRI. The severity of SI on T2WI was more prominent in extension-positioned MRI than in flexion-positioned MRI. Additionally, patients with a higher grade of SI had a more comprehensive range of neck motion before the surgery and experienced an unfavorable recovery after surgery.

Hyperextension of the neck causes canal narrowing by inducing buckling ligamentum flavum and the laminae, and the cord impinges anteriorly against a disc or bony spur. They will induce a higher intrinsic pressure, increasing axial tension and potential ischemic injury. This "pincer effect" repeatedly aggravates SC compression, leading from mild myelopathic symptoms to
severe myelopathy. A previous report showed the narrowing of the spinal canal by 17%–29% in extension MRI, but canal diameter measurements were performed in only 2 cases. We measured the CSA on T2WI, and our results showed 8.57% narrowing of the CSA of the spinal canal on the sagittal plane on T2WI MR by neck extension but no change of the spinal canal on the axial plane during extension position. During flexion, the cervical SC is stretched and more anterior in the canal. Uchida et al. reported increased SC compression during neck flexion in patients with cervical myelopathy. In contrast to previous reports, flexion-positioned MRI may reflect a state without SC compression, as the spinal canal is enlarged by the neck flexion position. Watanabe et al. measured the SC pressure at the C5–6 level in 20 patients with CSM. The authors found that compressive forces to the dura at the stenotic level were low in neutral and flexion positions but increased in neck extension. The pressure increased to 23.6 ± 7.5 mmHg with neck extension and decreased to 5.3 ± 2.7 mmHg with flexion. In this study, we found that the spinal canal sagittal CSA increased by 2.54% in a flexion position compared with a neutral position as analyzed on MRI. In cadaveric studies, it was determined that flexion stretched the cervical SC while extension loosened it. Thus, the subarachnoid space should be considered a buffer for cord compression rather than the spinal canal. Muhle et al. reported shortening of the subarachnoid space during neck extension. In this study, we measured the CSF reserve as a subarachnoid space, including values of both SC and CSF during neck motion. We observed a significant difference in the CSF reserve ratio between flexion and extension, with a smaller ratio in extension than in flexion. The sagittal plane of T2WI was more effective than axial T2WI in detecting changes in the CSA of the cord and CSF reserve ratio.

Signal change on T2WI at the level of cord compression is an important prognostic factor that also correlates with the severity of CSM. Some pathologies may not be visible on static MRI. Zeitoun et al. stated that extension-positioned MRI did not help to identify intramedullary SI change due to significant cervical canal stenosis and cord compression. Flexion-positioned MRI permits better intramedullary high SI visualization on a T2-weighted sequence. Our present findings are not in agreement with these previous observations. There is a correlation between the severity of cord compression and the SI change, as reported in the literature. Patients with more advanced spondylosis had significantly more stenosis at dynamic positions compared with those with less advanced disease. We showed a significant increase in cord impingement in extension (87 of 191 patients, 45.5%) versus flexion (36 of 191 patients, 18.8%) in CSM, which is consistent with the findings of other studies. We observed an increase in the prevalence of intramedullary SI change in extension (64.92%) versus flexion (61.78%). We found that 4.3% of patients with no SI in a neutral position changed into SI change in an extension MRI and 6.4% of patients with SI in a neutral position changed into no SI in a flexion MRI. Based on our results, extension-positioned MRI provides a more reliable evaluation of high intramedullary SI than the neutral and flexion positions.

The increase in SI in the neck extension position has several possible explanations. The intramedullary signal changes on T2WI were presumed to indicate myelomalacia or cord gliosis secondary to long-standing compression of the SC. Previous experimental studies support the notion that chronic compression of the SC leads to diminished blood flow, and that ischemia to the cord is the pathophysiological mechanism of cervical compressive myelopathy. The presence of a high SI lesion on T2WI MRI reflects cord edema. In the intermediate stage, a signal change reflects cystic necrosis of the central gray matter after prolonged cord edema. Ramanaukas et al. reported that, in the early and intermediate stages, the SC exhibited high SI on T2WI, whereas at a later stage, the SC manifested low SI on T1-weighted imaging (T1WI) and high SI on T2WI. In our study, the extension-positioned MRI presented a narrower CSA of the compressed cord and a lower CSF reserve ratio than the flexion- or neutral-positioned MRI, and compression of the SC aggravates ischemia of the cord. In reversible intramedullary signal changes, SC compression is aggravated during neck extension, causing the cord SI change to intensify; conversely, during neck flexion, SC compression improves slightly, and the resulting intramedullary SI change could mask the spinal SI change. An intramedullary SI change might become irreversible if this phenomenon occurs repeatedly over time, such as a low SI on a T1WI MRI. Patients with preoperatively low SI on T1WI MRI, that is, a snake eye appearance, had poor neurological outcomes after decompression.

Dynamic MRI can reveal detailed compression levels, but static MRI might not show cord compression. We found that compression levels increased by 63.87% when patients underwent an extension-positioned MRI. SC compression might be observed in asymptomatic patients on neutral MRI, and not every compression level is clinically significant.
in compression levels found in extension MRI should be considered when determining whether surgical treatment and surgical level expansion are necessary. As it is difficult to statistically show the influence of dynamic MRI on surgical decision-making in this study, we have included example cases in the supplementary figures (Supplementary Figs. 1–3). In the future, we plan to investigate how changes in dynamic MRI caused by neck movement affect changes in surgical treatment strategies. A multicenter expert opinion study is needed to investigate the clinical relevance and effectiveness of dynamic magnetic resonance images.

Previous studies showed that the anteroposterior diameter of the dural sac and SC is shorter during extension than flexion. Machino et al. confirmed that the anteroposterior diameter of the dural sac and SC in patients with CSM is significantly shorter than that in asymptomatic subjects. This study has shown that the sagittal spinal canal diameter and cord occupancy rate differ significantly during neck motion. The compressed cord diameter decreases in the extension position rather than in flexion. Accurately assessing the difference in diameters on the axial view can be challenging in cases of severe cord compression.

Stretch and shear forces are a leading cause of myelopathy, supported by evidence from various experimental models, including neural injury, tethered cord syndrome, and diffuse axonal injury. Studies have shown that segmental instability and mobility of the cervical spine play a significant role in the onset and prognosis of CSM. However, there is limited data on the relationship between cervical motion, high SI, and the severity of myelopathy symptoms. There are several grading systems for SI changes in the SC, but most focus on 2 intensity types: faint/fuzzy or intense/well-defined. In the present study, there was a correlation between greater neck ROM and more severe SI changes. With respect to neurological outcomes based on the severity of SI, the patients with no signal changes on T2WI had greater improvement in the JOA recovery ratio than patients with faint or intense SI changes on T2WI.

Dynamic MRI has benefits in identifying missed pathologies not visible on static MRI, such as changes in compressed levels and SI grade. In the following cases, we recommend routine or additional dynamic MRI for more benefits. Dynamic MRI has the potential to enable an early diagnosis when patients exhibit signs of myelopathy without severe cord compression or SI changes in static MRI by allowing the clinicians to assess changes in SC compression according to neck movement. Dynamic MRI can be helpful in planning the proper surgical position, and surgical decompression can be performed carefully when a severely compressed lesion is fully understood. Surgeons need to prescribe dynamic MRI to evaluate compression lesions and determine the appropriate surgical approach and levels. Dynamic MRI can also be helpful in assessing severe spondylosis in elderly patients with cervical myelopathy to determine which of the multilevel compression lesions caused by a bony spur or disc protrusion is the most compressed. On the other hand, dynamic MRI has the potential risk of symptom exacerbation, such as developing weakness in the upper or lower extremities, spasticity, or gait disturbance. As the potential benefits of dynamic MRI must be weighed against the risks of symptom worsening, dynamic MRI should be performed with careful consideration in patients with severe compressive myelopathy who have disc protrusions, osteophyte formation, hypertrophied ligamentum flavum, cervical canal stenosis, or segmental instability. In addition, it is crucial to consider the cost-effectiveness and time requirements of dynamic MRI. In Korea, a dynamic MRI can cost an additional USD 300 and take an extra 20 minutes compared with a routine cervical MRI.

In clinical practice, it is essential to avoid neck hyperextension exercises, which can exacerbate cervical conduction abnormalities and lead to severe cervical myelopathy. Before conducting a dynamic MRI, it is recommended that the patient’s neck flexion and extension motion be practiced. If patients develop paralysis of their upper or lower extremities, worsening of neurological deficits, or intolerable numbness in both hands during the neck motion trial, a dynamic MRI would not be indicated because it dangerously exacerbates the symptoms. Also, patients with intense SI on static MRI must carefully perform flexion-and-extension-positioned MRIs. Dynamic MRI should only be conducted during the daytime, and the patient should be closely monitored by the doctor throughout the test. We received permission from the patients before we conducted dynamic MRIs, and they made it clear that they would call for a halt to the procedure immediately if any symptoms, such as numbness or weakness, worsened. During our study period, 2 patients had to discontinue the dynamic MRI due to claustrophobia or panic symptoms, even after being given sedatives. However, none of the patients experienced neurological deficits that worsened during the dynamic MRI.

Dynamic MRI is a diagnostic tool that can reveal pathological phenomena that might not be visible in static MRI. However, it can worsen neurological deficits due to severe SC compression in the extended posture. Therefore, it should never be used in cases of traumatic instability, and it is contraindicated in cases of traumatic spinal hematoma or spontaneous spinal epidural...
hematoma.

This study had several limitations. The study included a small number of patients and did not consider other cervical alignment parameters, such as T1 slope, TIA, neck tilt, C2–7 sagittal vertical axis, or rotational motion. We included surgical cases and did not evaluate nonoperative cases. We focused on the degree of morphological cord compression visible on MRI using quantitative measurements in cervical myelopathic patients who needed surgery. Moreover, we only used T2-weighted images for data measurements as T1WI was not performed on dynamic motion due to high costs and the national insurance policy. The degrees of neck motion were not standardized for dynamic MRI due to pain or neurological deficits limiting neck motion. Additionally, the study was performed using dynamic supine MRI, which does not resemble physiological status in the upright position. However, our results showed no significant difference in ROM between cervical dynamic x-ray and dynamic MRI. We plan to perform further research to investigate the numerical value of signal change on T1WI and T2WI in a large number of patients with cervical myelopathy. Nonetheless, our current findings may be helpful in considering the decision of surgical levels and approaches for patients with CSM.

**CONCLUSION**

Dynamic MRI is a valuable tool for evaluating the physiological state during neck movement. We found that dynamic MRI identified the severity of cord compression and the prevalence of SI change and detected missed lesions more effectively than static MRI. The severity of intramedullary high SI is more prominent in an extended position. Patients with a higher grade of SI had a wider range of neck motion before surgery and experienced a less favorable recovery after surgery. The degree of cervical stenosis and cord compression on neck extension increased, and cervical spinal motion contributed to the severity of CSM, particularly in a hyperextension posture, which may help avoid aggravating myelopathic symptoms.

**NOTES**

**Supplementary Materials:** Supplementary Figs. 1-3 can be found via https://doi.org/10.14245/ns.2448166.083.

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This study was supported by a faculty research grant of Yonsei University College of Medicine (6-2021-0136).

**ORCID**

Jun Jae Shin: 0000-0002-1503-6343
Sun Joon Yoo: 0000-0002-6161-1847
Tae Woo Kim: 0000-0002-7457-8804
Jae-Young So: 0000-0001-8100-0763
Won Joo Jeong: 0009-0003-6029-1786
Mu Ha Lee: 0000-0003-2344-8118
Joongkyum Shin: 0000-0002-1907-3766
Yoon Ha: 0000-0002-3775-2324

**REFERENCES**


https://doi.org/10.14245/ns.2448166.083


Supplementary Fig. 1. Multilevel cord compression during extension. A 73-year-old male was diagnosed with cervical spondy-otic myelopathy due to a herniated cervical disc and bony spur at C5–6–7. (A) Flexion-positioned magnetic resonance imaging (MRI) showed cord compression decrease at C5–6–7 and no effacement of the subarachnoid space. (B) Neutral-positioned MRI showed partial obliteration of the anterior subarachnoid space at C5–6–7. (C) Extension-positioned MRI showed severe cord impingement at C3–4–5–6–7 and complete cord compression at C3–4–5. (D) We performed cervical selective laminoplasty and fusion C3–4–5–6–7, and his upper and lower weakness improved after the surgery.
Supplementary Fig. 2. Mild cord compression in neutral-positioned magnetic resonance imaging (MRI). A 55-year-old female was diagnosed with cervical spondylotic myelopathy due to a herniated cervical disc and bony spur at C5–6. (A) Flexion-positioned MRI showed mild cord compression decrease and no effacement of the subarachnoid space at C5–6. (B) Neutral-positioned MRI showed a partial anterior subarachnoid space at C5–6. (C) Extension-positioned MRI showed severe cord impingement at C5–6 and intense intramedullary SI (G2). (D) We performed anterior cervical discectomy and fusion C5–6, and she experienced an improvement in her arm and leg weakness and numbness.
Supplementary Fig. 3. A pathological lesion with no visible compression lesion in neutral-positioned magnetic resonance imaging (MRI). A 57-year-old male had undergone anterior cervical discectomy and fusion (ACDF) with a standalone cage on C3–4 10 years ago and had upper extremities weakness and ataxic gait aggravation one year previously. He was diagnosed with cervical spondylotic myelopathy, but no cord compression was found. (A) Flexion-positioned MRI showed SI at C3–4 and C5–6 but no cord compression. (B) Neutral-positioned MRI did not show any cord compression. (C) Extension-positioned MRI showed severe cord impingement at C4–5 and an increase in the intramedullary SI (G2) at C5–6. (D) We performed ACDF C4–5–6, and he showed improved gait disturbance and weakness in the upper extremities after surgery.
Commentary on “Radiological and Clinical Significance of Cervical Dynamic Magnetic Resonance Imaging for Cervical Spondylotic Myelopathy”

Aydin Sinan Apaydin¹, Khoi Than²

¹Department of Neurosurgery, Karabuk University School of Medicine, Karabuk, Turkey
²Department of Neurosurgery, Duke University School of Medicine, Durham, NC, USA

Cervical spondylotic myelopathy (CSM) is a progressive, degenerative disease of the cervical spinal column that causes a great reduction in quality of life and increased morbidity in patients. A multitude of degenerative changes occur in CSM that result in spinal cord compression and neurologic deficits, and accurate diagnosis and prompt surgical treatment are essential for favorable patient outcomes. All neurosurgeons and spine surgeons need to have up-to-date information on this complex pathology and current diagnostic and treatment strategies.

We have read with great interest the prospective work of Shin et al.¹ on the clinical and radiological significance of cervical dynamic magnetic resonance imaging (MRI) for CSM. MRI is the gold standard imaging modality to diagnose CSM; in particular, T2-weighted MRI very clearly delineates the relationship between the spinal cord and surrounding structures. This prospective study provides readers with invaluable data regarding the efficacy of dynamic cervical MRI in showing the signal intensity changes at different neck positions. The authors found that neck extension significantly reduces the spinal cord diameter in the sagittal plane at compressed levels. Also, the smallest diameter of the spinal cord on an axial plane is seen in an extension posture. Furthermore, the authors demonstrated that the number of compressed levels increased in a majority of patients (63.87%) in extension posture, and Muhle classification grade cervical stenosis also increased in extension posture compared with neutral and flexion postures.

In addition, the authors associated the signal intensity on preoperative imaging with lower recovery ratios. In their study, a higher range of motion was found in patients with higher signal intensity change in the extended posture, which had unfavorable outcomes postoperatively; in other words, cervical spine stenosis worsens in an extended posture. The authors have postulated that segmental instability of the cervical spine can increase the shear and tear forces on the spinal cord and can contribute to myelopathy.¹

Plain radiographs and computed tomography (CT) scans still have important roles in the radiological work-up of patients with CSM. The anterior-posterior and lateral radiographs can provide fast and reliable results on patients’ spinal alignment, but they cannot create a
3-dimensional (3D) visualization and assessment of the paraspinous muscles, ligaments, and nerve roots of the cervical spine. The CT scan can provide a 3D visualization of the cervical spinal column and is the gold standard for diagnosing ossification of the posterior longitudinal ligament, which can cause similar presenting symptoms as CSM. CT can also provide the preoperative assessment of the transverse foramen and intervertebral foramen for surgical planning. Also, CT myelography can be used to evaluate the spinal cord when an MRI scan is not applicable.

We agree that dynamic MRI can be a valuable tool in a surgeon’s arsenal for evaluating CSM, and it can provide valuable insight into patients’ conditions where static MRI cannot. Tykocki et al. demonstrated the dynamic stretch-associated injury through dynamic MRI and found narrowing in the spinal canal on extension with reduced cerebrospinal fluid ratio compared to flexion. They also found that the spinal cord area was smaller in the flexion position due to compression by anterior osteophytes and discs. This is also backed by several studies showing the efficacy of dynamic MRI. Makhchoune et al. also concluded that the spinal canal narrowed in the extension posture more than in the flexion posture. They also compared CT and MRI scans which showed statistical significance in extension and flexion postures signifying the diagnostic value of dynamic MRI. Kolcun et al. concluded that dynamic MRI showed cord compression by soft tissues, bony spurs of spondylisis, and mild segmental listhesis that may be overlooked by dynamic x-ray films.

Recently, it has been hypothesized that static MRI scans are not the most diagnostically sensitive MRI modality available. Shin et al., in their present study, demonstrated that dynamic MRI was able to demonstrate the degree of stenosis quantitatively where static MRI scans fall short, prompting its usefulness in CSM for diagnosis and presurgical planning. Although promising, there is still the need for multicenter randomized controlled trials directly comparing static and dynamic MRI and other imaging modalities. One way to design this study would be to include patients with clinical and plain radiographic findings suggestive of CSM, including patients aged between 18–80, and excluding those with previous cervical spinal surgeries. All patients would undergo plain radiographs on antero-posterior and lateral views and static MRIs would also be taken. Subsequently, patients will be randomly assigned into 2 groups: one group will only have static MRI, while a randomly selected group will additionally undergo dynamic MRI preoperatively. Dynamic MRI will be conducted in extension, neutral, and flexion postures to measure spinal parameters such as spinal canal diameter and T2 signal intensity. Preoperative scores, including the Neck Disability Index and Japanese Orthopaedic Association scores, will be recorded. These data, along with postoperative scores and follow-up MRI findings at 3, 6, and 12 months, will be compared between the 2 groups to evaluate the diagnostic sensitivity and clinical utility of static versus dynamic MRI in diagnosing CSM, aiding in surgical planning, and patient outcomes. The data acquisition and blinding of the clinicians and radiologists would be a crucial step to reduce biases. Also one of the limitations of this study will be the financial burden on the hospital and the length of the statistical MRI procedure. Therefore seeking funding will be necessary. Questionnaires can be distributed among physicians working on the project to measure the feasibility of dynamic MRI. Also, the training of neurosurgeons and radiologists training are important factor in while examining imaging findings.

Xu et al. described a protocol for a randomized prospective trial with patients divided into 3 groups based on their baseline clinical scores and static MRI images. But their protocol differed as they planned dynamic MRIs for all patients and surgeons planned 2 different surgical plans depending either on static or dynamic MRIs. They will choose the surgical plan randomly based on a number system.

In conclusion, CSM is an important cause of neurologic deficits and morbidity, with radiographic imaging as one of the pillars of patient success in treatment. Dynamic cervical MRI gave promising results in multiple studies, waiting to be proven once again in randomized controlled trials.

- **Conflicts of Interest:** The authors have nothing to disclose.

**REFERENCES**

and dynamic magnetic resonance imaging. World Neurosurg 2018;114:e317-22.
Assessing the Fractional Curve for Proper Management of Adult Degenerative Scoliosis

Seth C. Ransom1,*, Zach Pennington1,*, Nolan J. Brown2, Shane Shahrestani3, Jessica Ryvlin4, Ali Shoustari1, John C. Hagen5, Anthony L. Mikula1, Nikita Lakomkin1, Luis D. Diaz-Aguilar2, Benjamin D. Elder1, Joseph A. Osorio2, Martin H. Pham2

1Department of Neurologic Surgery, Mayo Clinic, Rochester, MN, USA
2Department of Neurological Surgery, University of California, San Diego, CA, USA
3Department of Neurosurgery, Cedars-Sinai Medical Center, Los Angeles, CA, USA
4Department of Neurological Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA, USA
5Department of Radiology, Baylor College of Medicine, Houston, TX, USA

Adult degenerative scoliosis (ADS) is a coronal plane deformity often accompanied by sagittal plane malalignment. Surgical correction may involve the major and/or distally-located fractional curves (FCs). Correction of the FC has been increasingly recognized as key to ameliorating radicular pain localized to the FC levels. The present study aims to summarize the literature on the rationale for FC correction in ADS. Three databases were systematically reviewed to identify all primary studies reporting the rationale for correcting the FC in ADS. Articles were included if they were English full-text studies with primary data from ADS (≥ 18 years old) patients. Seventy-four articles were identified, of which 12 were included after full-text review. Findings suggest FC correction with long-segment fusion terminating at L5 increases the risk of distal junctional degeneration as compared to constructs instrumenting the sacrum. Additionally, circumferential fusion offers greater FC correction, lower reoperation risk, and shorter construct length. Minimally invasive surgery (MIS) techniques may offer effective radiographic correction and improve leg pain associated with foraminal stenosis on the FC concavity, though experiences are limited. Open surgery may be necessary to achieve adequate correction of severe, highly rigid deformities. Current data support major curve correction in ASD where the FC concavity and truncal shift are concordant, suggesting that the FC contributes to the patient’s overall deformity. Circumferential fusion and the use of kickstand rods can improve correction and enhance the stability and durability of long constructs. Last, MIS techniques show promise for milder deformities but require further investigation.

Keywords: Adult spinal deformity, Fractional curve, Radiography, Adult degenerative scoliosis, Spine surgery, Neurosurgery

INTRODUCTION

Adult spinal deformity (ASD) comprises both degenerative spinal deformity (de novo deformity) and residual deformity from adolescent idiopathic scoliosis. One type of de novo ASD – adult degenerative scoliosis (ADS) – affects over 30% of elderly Americans.1,3 Patients with ADS most commonly present with lumbosacral radicular pain caused by stenosis of the foramina on the concavity of the lumbosacral curve (present in up to 97% of patients).1,2 For these patients with isolated radicular pain, limited neural element decompression via a laminoforaminotomy may represent a reasonable and effective treatment approach. However, for those patients whose symptoms can be attributed to their global spinal malalignment (e.g., decreased physical ac-
tivity tolerance and chronic fatigability of the paraspinal and proximal leg musculature), definitive therapy involves surgical correction of the sagittal and coronal alignments via long-segment instrumented fusion. 17

When correcting the coronal deformity in these patients, restoration of overall coronal alignment may require treatment of both the major and fractional curves (FCs). 6 The FC is defined as the compensatory curve located caudal to the major curve, most commonly at the lumbosacral junction (L4–S1). 6 The L4, L5, and S1 nerve roots exiting on the FC concavity are the most common precipitants of radicular pain in patients with ADS; therefore, many experts recommend correction of any FCs exceeding 15°. 7,10 In fact, prior work suggests that failure to correct a coexisting FC in ADS patients reduces the likelihood that global coronal balance will be successfully restored and maintained in the postoperative period. 8,11–13

Related to the concept of FC correction is whether the fusion construct should extend to the sacrum. For patients with adolescent idiopathic scoliosis (which becomes adult idiopathic scoliosis later in life), termination at L3 or L4 is common as an attempt to preserve motion segments.14 However, in de novo (degenerative) ADS, the lumbar FC generally contributes to radicular pain generation and may contribute to the overall deformity. In such cases, the construct must include the FC, necessitating instrumentation to the sacrum. However, there remains debate regarding the degree to which the FC drives the overall deformity, and consequently whether it needs to be included in the fusion construct. The present systematic review aims to address these questions, specifically focusing on: (1) characterizing the importance of preoperative FC assessment as it relates to surgical planning, (2) comparing surgical outcomes for ADS patients undergoing deformity correction surgery according to whether or not the FC was corrected, and (3) to discuss future directions relevant to ongoing investigations regarding the importance of FC correction in patients with ADS.

MATERIALS AND METHODS

1. Search Strategy

In accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, the PubMed, Ovid and Cochrane databases were queried for studies published between January 2010 and June 2023 focusing on FC management in patients with ADS.25 All 3 databases were searched using the following keywords: (“spine deformity” OR “complex spinal deformity” OR “scoliosis” OR “degenerative scoliosis”) AND (“fractional curve” OR “fractional lumbosacral curve” OR “lumbosacral hemi curve” OR “compensatory lumbosacral curve” OR “secondary lumbosacral curve” OR “minor lumbosacral curve”) AND (“measurement” OR “analysis” OR “radiography” OR “imaging” OR “assessment”).

2. Study Selection Process

Two investigators (SCR, NJB) independently screened each article according to title and abstract. When points of disagreement regarding study inclusion arose, they were resolved by a third reviewer (MHP). Full-text English translations of each article identified on title and abstract screen were obtained and screened for inclusion in the final analysis. Studies were included if they: (1) presented primary data, (2) described a coherent methodology for assessing the FC or FC and deformity-related symptomatology, (3) included only adults (age ≥ 18 years) with degenerative scoliosis, and (4) reported relevant data enabling characterization of the role that FC management plays in corrective surgery for patients with double curves (comprised of the major and FCs) in the setting of ADS.

3. Data Extraction and Analysis

The following information was extracted from included studies (when available): year of publication and surname of first author, study sample size (number of patients), method of FC quantification, surgical details (when surgery was performed), relevant patient inclusion and exclusion criteria implemented by each study, findings relevant to FC correction (and level of statistical significance when indicated), and any conclusions made regarding the importance of correcting the FC in patients with double curves secondary to ADS. Although descriptive statistics were reported, data were too heterogeneous to allow for a quantitative meta-analysis.

4. Quality and Risk of Bias Assessment

Given the retrospective, nonrandomized nature of the included studies, assessment of potential sources of bias was conducted in order to provide a fair assessment of the strength of the evidence informing the conclusions made in the present study. Accordingly, we assessed the quality and Risk of Bias (ROB) of each study using the Cochrane Risk Of Bias In Non-randomized Studies-of Interventions (ROBINS-I) tool. Overall, studies were mostly in the moderate-to-high ROB category, with 5 studies comprising high ROB and 7 studies qualifying as moderate ROB. Finally, one study was assigned low ROB.
3. FC as a Radiographic Predictor of Outcomes

Seven of the studies presented data correlating FC curve size with the severity of the major curve, extent of operative correction required for a successful outcome, global coronal imbalance, and the severity of lumbosacral radiculopathy. Pugely et al. examined 48 patients evaluated for ADS at a single center. The cohort was divided into those presenting with isolated back pain, those presenting with combined low back and anterior thigh/knee (femoral) pain, and those presenting with combined low back pain and sciatica. Those presenting with sciatica more commonly had symptomatic foraminal stenosis within the concavity of the FC, unless the structural curve lay in the lower lumbar spine (apex below L3), in which case symptoms derived from foraminal stenosis in the structural (main) curve concavity. In both patients with sciatica and femoral pain, foraminal size was significantly lower than in patients with lumbago alone. It thus follows that radicular pain down the back of the leg could suggest foraminal stenosis within the FC concavity as the driver of clinical presentation. The authors concluded that precise localization of the source of leg pain in ADS patients is a critical component of operative planning as it affects decisions related to determining which specific levels should be targeted by operative intervention.

4. FC Analysis and Surgical Planning

Eight articles focused on the impact of including the major/minor curve in devising the construct and the need for pelvic fixation. Amara et al. described their single-institutional experience in which they treated 99 patients with ADS. All included patients demonstrated lumbar radiculopathy ipsilateral to any FCs > 10° at the L4–S1 levels, specifically. Patients were divided into 3 groups—those undergoing correction of the FC only (N = 27), those undergoing T10-pelvis fusion with correction of the both the fractional and major coronal curves (n = 46), and those undergoing fusion to the upper thoracic spine with correction of the major coronal curve only (n = 26). Those undergoing fusion to the upper thoracic spine were noted to have significantly larger major curve angles than patients who were not fused to the upper thoracic levels. Furthermore, the authors found that patients in the group undergoing FC correction alone had significantly lower surgical morbidity, reduced complications, and the lowest likelihood to require inpatient rehabilitation facility placement relative to the other groups. Across all 3 groups, there was no significant difference in coronal balance postoperatively. This suggests that FC correction may be indicated for select patients, namely those presenting with a
<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Method of FC measurement (Cobb angle)</th>
<th>Surgery type</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
<th>Study sample size</th>
</tr>
</thead>
</table>
| Zhang et al.20   | Single-institution retrospective | Angle between superior endplate of L4 and the line formed by the pedicles of S1 | PSIF of > 5 segments ending at L5–S1 with facetectomy and osteotomy (decompression and TLIF were performed if anterior support was needed or to relieve spinal stenosis) | 1. Primary spinal deformity correction  
2. Instrumented fusion via posterior-only approach | 1. Fusion levels < 5;  
2. history of hip or knee arthroplasty;  
3. absolute discrepancy of leg length > 20 mm | N = 101 |
| Amara et al.23   | Single-institution retrospective | The curve below the major curve of thoracolumbar or lumbar scoliosis. Inclusion criterion: Cobb angle between L3–S1 > 10° | PSIF with 1–3 interbody fusions (ALIF/LLIF/TLIF) at the FC | 1. FC > 10°  
2. Low back or extremity pain ipsilateral to FC concavity  
3. Treatment of FC with interbody fusion  
4. Preop and postop long-standing radiographs  
5. > 1-year follow-up | NR | N = 78 (1 level = 19;  
2 levels = 36;  
3 levels = 23) |
| Amara et al.20   | Single-institution retrospective | The curve below the major curve of a lumbar or thoracolumbar scoliosis measured via Cobb angle; only Cobb angle > 10° considered FC | PSIF of L4–S1 (FC) versus T10-pelvis (LT) versus T2–4 to pelvis (UT) | 1. FC from L4–S1 > 10°  
2. Radiculopathy ipsilateral to the concavity of FC  
3. Pre and postop radiography studies  
4. > 1-year follow-up | 1. Previous lumbar fusion surgery | N = 99 (FC = 27;  
LT = 46, UT = 26) |
| Chou et al.22    | Multicenter retrospective study | Coronal Cobb angle of fractional curve | PSIF vs. cMIS | 1. > 18 years of age  
2. Minimum of 3 levels fused  
3. Minimum 2-year follow-up  
4. FC > 10°  
5. At least one of the following: SVA ≥ 5 cm, PT ≥ 20°, lumbar Cobb angle ≥ 20°, or a PI-LL ≥ 10° | 1. Hybrid open posterior surgery with interbody fusion | N = 118 (open = 79;  
cMIS = 39) |
| Brown et al.17   | Single-institution retrospective | Angle between the line connecting the superior iliac alae and the line formed by the pedicles of H | PSIF to L5 | 1. Fusion extending above T12  | 1. Need for decompression at L5–S1  
2. Pre-existing L5–S1 deformity (not including isolated degeneration at L5–S1) | N = 16 |
| Yagi et al.21    | Single-institution retrospective | Coronal Cobb method | Combined single-rod anterior fusion and short PSIF to sacrum (hybrid) versus long PSIF with anterior release (control) | 1. Thoracic and thoracolumbar/lumbar curves (> 80°)  
2. Nonprogressive thoracic deformity (> 30° flexibility)  
3. Fractional curve (with segmental instability, stenosis or facet arthrosis) or degenerative disc disease | 1. Osteoporosis  
2. Revision surgery | N = 66 (33 per group) |

(Continued)
### Table 1. Design and sample description of 12 studies assessing preoperative and/or postoperative fractional curve of adult degenerative scoliosis patients (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type</th>
<th>Method of FC measurement (Cobb angle)</th>
<th>Surgery type</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
<th>Study sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manwaring et al.24 2014</td>
<td>Single-institution retrospective</td>
<td>NR</td>
<td>Staged cMIS with versus without L5–S1 TLIF</td>
<td>1. Treatment of ADS with at least 2-level MIS LLIF procedure</td>
<td>1. Hybrid construct involving posterior osteotomies</td>
<td>N = 15 (TLIF = 11; control = NR)</td>
</tr>
<tr>
<td>Buell et al.25 2021</td>
<td>Multicenter retrospective study</td>
<td>Coronal L4–S1 Cobb angle</td>
<td>L4–S1 TLIF vs. ALIF</td>
<td>Index operation that involved TLIF or ALIF at L4–5 and/or L5–S1. Minimum 2-year postoperative follow-up</td>
<td>Any patient with active infection, malignancy, diagnosis of scoliosis other than adult degenerative</td>
<td>N = 106 (TLIF = 47, ALIF = 59)</td>
</tr>
<tr>
<td>Geddes et al.26 2021</td>
<td>Single-institution retrospective</td>
<td>Coronal Cobb method</td>
<td>ALIF+PSF+S2AI screws versus PSF+S2AI screws for thoracolumbar fusion</td>
<td>1. Posterior lumbar fusion to the pelvis using S2AI screws</td>
<td>1. Patients who had posterior 3-column osteotomies</td>
<td>N = 59</td>
</tr>
<tr>
<td>Hofer et al.45 2022</td>
<td>Single-institution retrospective</td>
<td>Cobb angle method for lumbar fractional curve. The magnitude of the major lumbar coronal curve and fractional lumbar coronal curve caudal to it was measured on preoperative and follow-up anteroposterior imaging</td>
<td>T3-ilum fusion +/- kickstand placement</td>
<td>1. Deformity correction with fusion from upper thoracic spine to pelvis</td>
<td>NR</td>
<td>N = 15 (kickstand = 7, nonkickstand = 8)</td>
</tr>
<tr>
<td>Zuckerman et al.27 2023</td>
<td>Single-institution retrospective</td>
<td>Cobb angle between the sacrum and most tilted lower lumbar vertebra (either L3/4/5)</td>
<td>Instrumentation to pelvis/fusion to sacrum and TLIF</td>
<td>1. ≥6-level fusion</td>
<td>NR</td>
<td>N = 243</td>
</tr>
</tbody>
</table>

PSIF, open posterior spinal instrument fusion; TLIF, transforaminal lumbar interbody fusion; ALIF, anterior lumbar interbody fusion; LLIF, lateral lumbar interbody fusion; FC, fractional curve; LT, lower thoracic; UT, upper thoracic; cMIS, circumferential minimally invasive surgery; SVA, sagittal vertical axis; PT, pelvic tilt; PI-LL, pelvic incidence-lumbar lordosis; NR, not reported; MIS, minimally invasive surgery; CT, computed tomography; PSF, posterior spinal fusion; S2AI, S2alar iliac screw; APLCR, anteroposterior long cassette radiograph; CVA, coronal vertical axis; TK, thoracic kyphosis.
### Table 2. Results of 12 studies assessing preoperative and/or postoperative fractional curve of adult degenerative scoliosis patients

<table>
<thead>
<tr>
<th>Study</th>
<th>Preoperative FC</th>
<th>Postoperative FC</th>
<th>FC correction</th>
<th>FC radiographic predictors</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al.(^{21}) 2021</td>
<td>13.6° ± 8.2°</td>
<td>5.9 ± 5.1°</td>
<td>p &lt; 0.001</td>
<td>Preoperative FC with L4 coronal tilt toward C7 plumbline is associated with postoperative coronal imbalance</td>
<td>Directionality of preoperative FC toward C7 plumbline increasing risk of postoperative coronal imbalance</td>
</tr>
<tr>
<td>Amara et al.(^{23}) 2020</td>
<td>Level = 15.3° ± 8.2°, 2 levels = 117.9°, 3 levels = 16.3°</td>
<td>13.6° ± 8.2°</td>
<td>Group 1 vs. 2 = 0.0062; group 1 vs. 3 = 0.017; group 2 vs. 3 = 0.99</td>
<td>None</td>
<td>Additional interbody fusion levels at the FC resulted in more fractional curve correction, more major curve correction, increasing lordosis without increasing morbidity</td>
</tr>
<tr>
<td>Amara et al.(^{30}) 2019</td>
<td>FC = 15.7°, LT = 16.7°, UT = 16.9°</td>
<td>NR</td>
<td>NR</td>
<td>None</td>
<td>Treatment of only the FC was associated with lower complication rates, shorter hospital LOS and reduced blood loss than fusion to UT or LT levels; FC group had higher rates of re-extension UT or LT levels</td>
</tr>
<tr>
<td>Chou et al.(^{22}) 2018</td>
<td>FC &gt; 10° Matched cohort: preop FC–cMIS 18 and open: 18</td>
<td>Unmatched cohort: postop FC–cMIS: 17 and open: 19.6</td>
<td>cMIS = 6.9°; Open = 8.5°</td>
<td>None</td>
<td>cMIS achieved similar reduction in leg pain and correction of fractional curve as traditional open surgery, despite significantly fewer cMIS patients undergoing direct decompression</td>
</tr>
<tr>
<td>Brown et al.(^{17}) 2004</td>
<td>21°</td>
<td>10.6°</td>
<td>NR</td>
<td>Less postoperative FC decreased risk of L5–S1 degeneration</td>
<td>Patients with good postop FC achieved better outcomes with posterior fusion to L5, avoiding sacral fusion</td>
</tr>
<tr>
<td>Yagi et al.(^{21}) 2014</td>
<td>Hybrid = 23° ± 9°, control = 24° ± 10°</td>
<td>Hybrid = 7 ± 4°, control = 15 ± 8°</td>
<td>Percent correction of lumbosacral curve significantly better in hybrid versus control (p &lt; 0.001)</td>
<td>None</td>
<td>Hybrid patients had improved curve correction, fewer levels fused, decreased blood loss and fewer revision procedure when compared to control</td>
</tr>
<tr>
<td>Manwaring et al.(^{24}) 2014</td>
<td>TLIF = 9.2°, control = NR</td>
<td>TLIF = 4.1°, control = NR</td>
<td>NA</td>
<td>None</td>
<td>Significant fractional curve correction in staged cMIS is achieved through 2 stage TLIF treatment of L5–S1</td>
</tr>
<tr>
<td>Pugely et al.(^{18}) 2017</td>
<td>Group B = 19.4°; group F = 25.5°; group S = 17.7°</td>
<td>NA</td>
<td>NA</td>
<td>None</td>
<td>Sciatic nerve pain in setting of lumbar structural curves is associated with foraminal stenosis at the concavity of the caudal fractional curve; femoral nerve pain likely caused by stenosis at concavity of main structural curve (L3 or below)</td>
</tr>
</tbody>
</table>

(Continued)
Table 2. Results of 12 studies assessing preoperative and/or postoperative fractional curve of adult degenerative scoliosis patients (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Preoperative FC</th>
<th>Postoperative FC</th>
<th>FC correction</th>
<th>FC radiographic predictors</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buell et al. 25</td>
<td>All = 20.2° ± 7.0°, TLIF = 19.4° ± 7.2°, ALIF = 20.8° ± 6.9°</td>
<td>All = 6.9° ± 5.2°, TLIF = 7.1° ± 5.4°, ALIF = 6.8° ± 5.1°</td>
<td>Multiple regression demonstrated 1-mm increase in L4–5 TLIF cage height led to 2.2° reduction in L4 coronal tilt (p = 0.011), and 1° increase in L5–S1 ALIF cage lordosis led to 0.4° increase in L5–S1 segmental lordosis (p = 0.045). Matched analysis demonstrated comparable fractional correction (TLIF = -13.6° ± 6.7° vs. ALIF = -13.6° ± 8.1°, p = 0.982).</td>
<td>None</td>
<td>Results demonstrate comparable fractional curve correction (66.7% for TLIF patients versus 64.8% for ALIF patients), despite the use of significantly larger and more lordotic cages in ALIF</td>
</tr>
<tr>
<td>Geddes et al. 26</td>
<td>PSF = 13.4° ± 7.1°, ALIF+PSF = 18.3° ± 9.3°</td>
<td>PSF = 8.6° ± 4.4°, ALIF+PSF = 6.1° ± 5.3°</td>
<td>PSF = 4.8° ± 4.5° (27% curve correction), ALIF+PSF = 12.1° ± 6.0° (68% correction), p = 0.053</td>
<td>NR</td>
<td>ALIF+PSF achieves greater correction of the fractional curve than PSF alone. Though not the primary indication of ALIF, this may help facilitate overall deformity correction and pelvic balance</td>
</tr>
<tr>
<td>Hofler et al. 29</td>
<td>Kickstand = 4.3-cm coronal deviation, 43° major lumbar curve, 23° fractional lumbar curve Nonkickstand = 2.2-cm coronal deviation, 35° major lumbar curve, 14° fractional lumbar curve</td>
<td>Kickstand group = 4.3-cm intraoperative coronal deviation, 1.8-cm postoperative coronal deviation Nonkickstand group = 0.6-cm intraoperative coronal deviation, 2.1-cm postoperative coronal deviation</td>
<td>Preoperative lumbar FC was greater in patients requiring a kickstand (23° vs. 35°, p = 0.02)</td>
<td>NR</td>
<td>Intraoperative kickstand rod placement guided by intraoperative APLCR allows for satisfactory reduction in cases where the fractional coronal curve persists without loss of sagittal plane correction</td>
</tr>
<tr>
<td>Zuckerman et al. 27, 28</td>
<td>Qiu type A = 11.1°, Qiu type B = 12.7°, Qiu type C = 15.6°</td>
<td>Qiu type A = 5.3°, Qiu type B = 7.6°, Qiu type C = 6.4°</td>
<td>Type C patients had the most LSF curve correction (p = 0.023 for change in LSF curve by 9.2°)</td>
<td>NR</td>
<td>Greater correction of LSF curve was seen in Qiu type C patients compared to type A and type B. More TLIFs were associated with greater amount of LSF curve correction. No clear trends seen regarding LSF curve change and postoperative outcomes</td>
</tr>
</tbody>
</table>

PSIF, open posterior spinal instrument fusion; TLIF, transforaminal lumbar interbody fusion; ALIF, anterior lumbar interbody fusion; LLIF, lateral lumbar interbody fusion; FC, fractional curve; LOS, length of stay; LT, lower thoracic; UT, upper thoracic; cMIS, circumferential minimally invasive surgery; NR, not reported; PSF, posterior spinal fusion; APLCR, anteroposterior long cassette radiograph.
primary complaint of radiculopathy localizing to the level of the FC. However, these data suggest it is not necessary to incorporate FC correction in the surgical plan for all ASD patients. Although this appeared to be the case, the authors did acknowledge one caveat: patients who underwent FC correction alone exhibited higher rates of reoperation.

The same group subsequently described a subset of 78 patients who underwent surgery for FCs (L3–S1) > 10° involving ≥ 1 interbody fusion within each patient's FC. All patients included in the study presented with primary complaints of lumbar radiculopathy that could be localized to the level of the FC. With respect to the specifics of FC correction, interbody placement at 2–3 levels (versus 1-level interbody placement) was associated with significantly reduced fractional and major curves without a commensurate increase in complications.

More recently, in an expanded multicenter cohort, Chou et al. examined outcomes in 118 ADS patients who underwent correction of their FCs (79 underwent open correction, while 39 underwent circumferential minimally invasive surgery [cMIS] correction). Despite the fact that direct decompression was performed less frequently in the cMIS group, pain outcomes—including postoperative visual analogue scale (VAS) leg pain scores—were similar between the cMIS and open groups. This suggests foraminal stenosis on the FC concavity drives the radicular leg pain and the indirect decompression afforded by FC correction may be sufficient to offer symptomatic improvement. To clarify, neural element decompression appears to be the key process driving pain relief. While for patients with an exclusively radicular pain picture, direct decompression (such as hemilaminectomy or foraminotomy) may be sufficient. However, for those with concurrent axial pain attributable to the deformity, indirect decompression with interbody placement via either open or cMIS techniques may offer effective symptom relief without the need for concurrent direct decompression. Said strategy helps to maintain fusion surfaces that would be removed with direct decompression, and so the-

Fig. 2. ROB (Results of Risk of Bias) assessment conducted using the ROBINS-I (Risk Of Bias In Non-randomized Studies-of Interventions) tool.

https://doi.org/10.14245/ns.2347202.601
Theoretically improve the odds of successful radiographic fusion. Finally, one advantage of concurrent FC correction is that the placement of instrumentation helps to maintain the nerve root decompression and so may allow for better long-term symptom relief, though this remains a point of ongoing investigation.

Data from Brown et al. suggested that correction of the FC was associated with more favorable patient-reported outcomes in a series of 16 ADS patients who underwent posterior fusion to L5. Average FC correction was 10.4° and postoperative L5–S1 degeneration correlated significantly with the magnitude of residual FC. For example, patients requiring revision had significantly larger residual FCs (21° vs. 8.7°, p < 0.05) and those with subsequent L5–S1 disc degeneration had significantly larger residual FCs than those without such degeneration (15° vs. 8.7°, p < 0.05). As distal segment degeneration is a driver of surgical revision, these data suggest FC correction may reduce the odds of revision surgery for adjacent segment disease.

Yagi et al. compared the efficacy of a hybrid anterior-posterior surgical approach to conventional anterior column release (ACR) with posterior fusion in 33 patient pairs undergoing corrective surgery for ADS. Each pair was matched by age and coronal curve magnitude for analysis. Patients underwent either "conventional" anterior release followed by long-segment posterior instrumentation, or a "hybrid" multilevel anterior interbody fusion followed by posterior segmental instrumentation. The combined anteroposterior procedure has been previously associated with improved correction of severe deformity curves, lower likelihood for postoperative progression, and reduced rates of pseudarthrosis. In this study, the hybrid group demonstrated significantly lower rates of both postoperative coronal imbalance (38% vs. 56%, p = 0.03), complications (18% vs. 39%, p = 0.01), and rates of surgical revision (12 of 33 patients vs. 6 of 33 patients, p = 0.03), though quality of life (QoL) outcomes were similar. The major complications in the control group, which exhibited the higher overall rate of complications, included proximal junctional kyphosis (PJK) (n = 7) and deep infection. Interestingly, although one of the touted advantages of the anterior-posterior hybrid approach was a decreased risk for pseudarthrosis, the latter was not demonstrated by Yagi and colleagues. Specifically, only 3 cases of pseudarthrosis were reported, and 2 of them occurred in the hybrid group.

Both Chou et al. and Manwaring et al. compared the ability of cMIS techniques to achieve FC correction relative to open techniques. Manwaring et al. examined the impact of ACR with concomitant extreme lateral interbody fusion (XLIF) on sagittal and coronal balance in 36 patients treated for ADS. Both patients who underwent XLIF without ACR and those who underwent ACR demonstrated improvement in their coronal Cobb angle; there was significant correction in the ACR group (8.2° vs. 4.2°, p < 0.002). As mentioned previously, in their comparison of patients treated via open (n = 79) versus cMIS (n = 39) approaches, Chou et al. observed no differences in postoperative pain, change in coronal Cobb angle, change in pelvic incidence-lumbar lordosis (PI-LL) mismatch, Oswestry Disability Index (ODI) improvement or VAS back pain scores. However, cMIS techniques achieved these results with fewer direct decompression procedures.

Buell et al. have reported that correction of the FC via decompression alone or short-segment fusion both represent reasonable options, even though they may accelerate degeneration of the residual curve. Both TLIF and anterior lumbar interbody fusion (ALIF) approaches can be entertained and allow for curve correction and indirect decompression of the nerve roots ipsilateral to the FC concavity. Ultimately, a less-invasive decompression alone may be preferable in frail patients with predominately radicular symptoms and no evidence of instability on radiographic evaluation.

Buell et al. recently presented the results of a multicenter analysis of 106 patients with ≥ 30° coronal main curves and ≥ 10° lumbosacral FCs who underwent L4–5 TLIF (n = 47) versus L4–5 and/or L5–S1 (lateral) ALIF (n = 59) for FC correction. Matched analysis of 28 pairs of patients showed no significant difference in FC correction between TLIF and ALIF (55.7% vs. 64.8%, respectively). Notably, patient-reported outcomes on the ODI and 36-item Short-Form health survey were similar between groups, as were overall complication rates. It was noted that ALIF offered greater restoration of lordosis at L5–S1, suggesting that ALIF may be preferable when concomitant sagittal plane correction is desired.

Next, in their 59 patient retrospective series, Geddes et al. sought to determine whether ALIF can improve the FC in deformity surgery as compared to posterior surgery alone. The authors found the addition of ALIF led to significantly greater FC correction (12.1° vs. 4.8°, p < 0.01) and smaller postoperative FC (6.1° vs. 8.6°, p = 0.023). Major curve correction (23.5° vs. 14.9°, p = 0.006) was also greater and multivariable analysis showed the addition of ALIF to be independently predictive of FC correction even after accounting for the use of TLIFs in the posterior-only cohort.

Most recently, Zuckerman et al. re-examined the multi-institutional International Spine Study Group (ISSG) dataset previously analyzed by Buell et al. Their analysis focused on the...
relative importance of correcting the major versus FCs in ADS patients. All 243 patients included in their analysis underwent ≥ 6-level fusion for coronal Cobb angles > 30° and C7 coronal vertical axes > 3 cm. Improvement in patient-reported outcomes assessed with ODI were correlated with both lumbosacral FC and major coronal curve correction; they also correlated preoperative coronal alignment with FC magnitude. They noted that FCs were largest in patients wherein the coronal imbalance lay ipsilateral to the major curve convexity, suggesting the FC was the driver of overall coronal imbalance.27

DISCUSSION

An ongoing question pertaining to ADS is whether the FC is a driver of coronal malalignment or merely a compensatory curve for the major scoliotic curve. In general, the present literature suggests that the presence of the FC concavity ipsilateral to side of truncal shift is a risk factor for persistent coronal malalignment. It also suggests that FC correction may help to alleviate ASD-associated radiculopathy without requiring direct decompression.

An additional advantage of FC correction is a potential reduction in the odds a patient will require surgical revision, indicating that FC correction in addition to the major curve can optimize surgical outcomes. This finding coincides with the proposal recently advanced by Plais and the International Spine Study Group.28 Through investigating a multicenter cohort of 404 patients (age > 45 years) with thoracolumbar major coronal curves (> 15°; apex at T11–L3) and FCs > 5°, the authors found that among patients with global coronal malalignment, those with truncal shift ipsilateral to the FC concavity had significantly larger FCs (22.28° vs. 14.84°, p < 0.001) relative to those with truncal shift contralateral to the FC concavity. Additionally, they had greater pelvic obliquity angled towards the side of the truncal shift. This led the authors to suggest that while FCs may be compensatory in cases where truncal shift is contralateral to the side of the FC concavity, where the truncal shift is ipsilateral, the concavity likely plays a role as the primary driver of coronal malalignment. According to this paradigm, restoration of coronal balance depends upon treatment of the FC, as has been reported.

With this paradigm in mind, Obeid et al.29 released an expert’s consensus, treatment-oriented guideline for correcting coronal imbalance in ASD. Their guideline defines concave and convex coronal malalignment as types 1 and 2, respectively. For convex (type 2) coronal malalignment, Obeid et al.29 suggests avoiding correction of the main curve as doing so would risk worsening coronal malalignment. Instead, the lumbosacral curve—which may be the FC in type 2A patients—should be corrected. When the main curve is located within the lumbosacral spine and associated with a compensatory lumbar curve (type 2B), 3-column osteotomies at the apex (usually between L4–S1, most commonly at L5) of the lumbosacral curve will generally suffice for correction of the short, main curve. However, most patients will have their major curve in the lumbar or thoracolumbar spine (type 2A) and the FC will lie at the lumbosacral junction. When this is the case, FC correction should be dictated by the degree of flexibility present at the lumbosacral junction. For example, a previously fused interbody space at the lumbosacral junction merits performance of a 3-column osteotomy with L5 pedicle subtraction osteotomy to achieve FC correction. Furthermore, it is important to keep in mind that ending a construct at L4 or L5 in an ASD patient with a preexisting FC predisposes them to subsequent distal segment degeneration. The indications for excluding the sacrum from a construct include: (1) the presence of normal L5–S1 disc and facets; (2) the UIV lies at or below T10; (3) the patient has normal bone mineral density; or (4) the patient is young or relatively less active.30 Indications for extending the construct to the sacrum are prior surgical decompression at L5–S1, signs and symptoms consistent with L5–S1 radiculopathy, and/or the presence of L5–S1 spondylolisthesis.

When concave coronal malalignment (type 1) is present, on the other hand, correction of the main curve only will result in indirect correction of the FC, which is compensatory, not structural.29 If the main curve is flexible, a posterior column osteotomy at its apex is likely sufficient for correction, whereas a rigid or fused curve will require a 3-column osteotomy at the apex to achieve correction (type 1 coronal malalignment).29

1. Coronal Realignment and Kickstand Rod Technique

For patients in whom the FC appears to contribute significantly to coronal imbalance, there remains debate regarding the magnitude of major curve correction that is required. Prior work by Deviren et al.,31 among others, has suggested that the coronal curve magnitude inversely correlates with curve flexibility. The flexibility of the major curve is best assessed using preoperative lateral bending radiographs. In patients with truncal shift ipsilateral to the FC concavity (Qiu type C curves), major curve correction in the absence of FC correction will exacerbate the truncal shift and result in poorer postoperative coronal alignment. To this end, Bao et al.32 noted that patients with
type C curves are most likely to have coronal malalignment postoperatively. Consequently, for these patients, adequate FC correction is paramount; it may also be beneficial to use less aggressive correction of the major curve, especially for those with flexible major curves preoperatively.

One increasingly popular technique for coronal realignment is the kickstand rod technique, which employs a rod with distal fixation in the ilium and proximal fixation to the thoracolumbar spine. The kickstand rod can provide significant coronal correction due to the distracting force and torque it provides (Fig. 3). These forces are greater than those achieved using rod bending maneuvers (distraction/compression) alone. It is therefore a powerful tool for FC correction. Buell et al. described the successful use of this technique in 17 adult patients with thoracolumbar/lumbar degenerative scoliosis. Highlighting the strength of correction provided by the kickstand rod, they reported coronal overcorrection in one patient, though the authors indicated this is a rare complication when the technique is properly executed. Puvanesarajah et al. and Mundis et al. similarly found the kickstand rod technique to facilitate good coronal correction in their series of 20 and 21 patients treated for adult scoliosis, respectively. Interestingly, Mundis et al. from the ISSG compared coronal correction with the kickstand rod technique to correction achieved when using conventional accessory rods alone. They found that the kickstand technique can produce better overall postoperative coronal balance. However, this did not translate to improved patient-reported outcomes on the Scoliosis Research Society-22 Questionnaire or ODI assessments, suggesting that there may be a point beyond which additional coronal plane correction is no longer meaningful.

2. Radicular Pain and the FC

In approximately 90% of cases, back and/or leg pain is the primary reason for the initial hospital/clinic visit in patients with ADS. The FC appears to play a significant role in this symptomatology for thoracic and thoracolumbar major curves, as it has been demonstrated that radiculopathy at presentation localizes most commonly to the L4, L5, and S1 spinal roots on the side of the FC concavity. Ultimately, many patients will have endured chronic back pain over many years prior to presentation. As long as relative sagittal and coronal balance are maintained; however, the scoliotic deformity may not necessarily be disabling. On the other hand, the radicular pain cause by foraminal stenosis within the FC concavity often leads patients to seek surgical intervention. Furthermore, progression of the scoliotic deformity with multilevel lateral listhesis that can further augment the radicular pain. To this end, the work of Chou et al. found that correction of the FC alone (without an accompanying direct decompression procedure) was sufficient to...
alleviate lumbar radicular pain. Of note, Chou et al.\textsuperscript{22} reported that laminectomy alone is insufficient to address this pain, which results from compression of the dorsal root ganglia within the neural of the FC concavity.

In the case of lumbar structural curves, the relative contribution of the FC to patient symptomatology is less clear, though the distribution of the radicular pain can help to localize the pain to either foraminal stenosis within the structural curve or FC concavity. As suggested by the results from Pugely et al.,\textsuperscript{18} pain in the sciatic distribution (predominately L5, S1, and S2 dermatomes) is most likely attributable to foraminal stenosis at the concavity of the FC, whereas radicular pain in the femoral distribution (predominately L2–4 dermatomes) is more easily attributed to foraminal stenosis within the structural curve concavity. Accurate assessments of: (1) the concordance of the FC convexity and the directionality of the global coronal malalignment, and (2) the degree to which the patient’s symptomatology is attributable to the FC are key to designing the optimal surgical plan.\textsuperscript{18} Finally, these findings mark a paradigm shift from previous beliefs that the FC was always the source of radicular pain in ADS patients.\textsuperscript{24}

### 3. Coronal and Sagittal Balance

Achieving coronal and sagittal balance are primary goals of ADS surgery. A recent investigation by Zhang et al.\textsuperscript{26} suggests the concordance of FC concavity and global coronal malalignment predicts the odds of residual malalignment following surgical correction. The authors examined the relation of FC orientation and preoperative coronal imbalance to predict restoration of coronal balance in 101 patients (74 instrumented to pelvis).\textsuperscript{18} Those 27 patients who achieved postoperative coronal balance were more likely to have a FC concavity opposite to the preoperative net coronal imbalance (66\% vs. 19.4\%, \(p < 0.001\)).

They found on logistic regression that the best predictors of postoperative coronal imbalance were consistency pattern (preoperative coronal imbalance and FC concavity on the same side) and preoperative coronal C7 plumbline > 30 mm towards the convex side of the major curve.\textsuperscript{18} This led the authors to argue that greater attention must be paid to the directionality and treatment of the FC.

These results echoed earlier findings by Bao et al.,\textsuperscript{27} who examined the prevalence and impact of preoperative coronal imbalance on outcomes in 284 patients who underwent surgery for degenerative lumbar scoliosis. They found 34.8\% of patients presented with coronal imbalance; those with preoperative coronal imbalance > 30 mm shifted towards the convexity of the major curve were significantly more likely to have persistent coronal imbalance postoperatively. However, preoperative coronal imbalance did not impact patient-reported outcomes such as ODI or VAS for back pain. They, like Zhang et al.,\textsuperscript{18} consequently argued that asymmetric osteotomies intended to reduce the major curve had also exacerbated truncal shift and coronal instability. Consequently, the authors argued in favor of a TLIF within the FC concavity to restore neutral alignment to the L4 and L5 vertebrae, thereby restoring overall coronal alignment.\textsuperscript{28} Similarly, Bao et al.\textsuperscript{27} concur as they argued for performing a TLIF within the FC to restore neutral alignment to the L4 and L5 vertebrae and thus the overall coronal alignment.

Additional surgical measures may be warranted in these patients to ensure postoperative sagittal balance, which data suggests to have the greatest influence on long-term patient outcomes.\textsuperscript{5,38} However, significant alignment correction often involves the use of long-segment posterior constructs, which are high risk for mechanical failure in elderly osteoporotic/osteopenic patients.\textsuperscript{36}

### 4. Surgical Approaches

At present, data does not support a single optimal approach for correction of coronal plane deformities.\textsuperscript{38} In general, approaches can be divided into purely posterior approaches versus combined anterior-posterior approaches, and into open versus MIS approaches. None of the included studies directly compared MIS and open approaches for ADS alone. However, a recent article by Chou et al.\textsuperscript{29} using the ISSG database reported a propensity-matched cohort study of 154 patients (77 pairs) who underwent cMIS with anterior interbody fusion and posterior percutaneous instrumentation or open posterior fusion for ASD. Patients were matched on multiple metrics, including construct length, age, body mass index, and baseline spinopelvic parameters. Those treated with cMIS surgery had grossly similar QoL outcomes, radiographic outcomes, and surgical revision rates but lower intraoperative blood loss. These same authors also described results within a subset of patients from the ISSG cohort with adult scoliosis, characterized as those > 18 years old with a FC > 10°.\textsuperscript{22} As described in the results section, they found cMIS and open posterior fusion achieved similar changes in coronal Cobb angle, PI-LL mismatch, ODI, and VAS back pain. However, this was achieved with fewer decompressive procedures.

Further expounding on the advantages of ACR, Yagi et al.\textsuperscript{21} showed in their small experience that multilevel anterior interbody placement afforded superior coronal plane correction in

---

https://doi.org/10.14245/ns.2347202.601
patients with thoracolumbar ($61° \pm 21°$ vs. $45° \pm 25°$, $p < 0.01$) or lumbosacral curves ($67° \pm 21°$ vs. $41° \pm 19°$, $p < 0.001$) while reducing intraoperative blood loss (mean reduction 1.5 L, $p < 0.001$) and the average number of fused segments ($6.7 \pm 1.2$ vs. $14.6 \pm 1.3$, $p < 0.001$).

The final extant question concerns the necessity of including the FC within the fusion construct.\textsuperscript{30,41} In the adolescent idiopathic scoliosis population, construct termination at L3 or L4 is considered acceptable for most curve types as there appears to be limited potential for distal segment degeneration and it spares additional motion segments.\textsuperscript{10-12,14,42} These motion segments may allow for pelvic mobility to compensate for reduced motion within the stiff long-segment construct immediately cephalad to the lumbosacral junction.\textsuperscript{42} However, the biomechanics of deformity in the adolescent idiopathic scoliosis population is quite different than the biomechanics of adult deformity. In the ADS population, the lumbar FC is a common source of radicular pain and thus requires treatment at the time of surgery. When it comes to correcting the lumbar FC, multiple studies have shown that ending the construct at S1 as opposed to terminating at L4 or L5 decreases the rate of adjacent segment degeneration.\textsuperscript{9,10,41} Based upon the small series of Brown and colleagues,\textsuperscript{17} distal junctional kyphosis (DJK) may occur in over one-third of patients treated with constructs terminating at L5.\textsuperscript{17} Using the ISSG multicenter database, Yao et al.\textsuperscript{42} showed that terminating ASD constructs distally at L4–5 versus the sacrum was associated with poorer sagittal alignment restoration at 6-week follow-up, although pelvic fixation was associated with higher rates of PJK. Coronal balance restoration was similar in both groups both at 6-week and 2-year follow-up. QoL outcomes were also similar between groups, and DJK requiring surgical revision was only noted in one of 28 included patients in the matched groups analysis. However, as illustrated by Brown et al.,\textsuperscript{17} failure to correct the FC can lead to L5–S1 segment breakdown, ultimately requiring surgical revision. Therefore, optimum FC correction is critical to a good overall outcome irrespective of the distal instrumented segment; patients with inadequate FC correction are at increased risk of adjacent segment disease and coronal decompensation at the L5–S1 level.\textsuperscript{10,23} However, treatment of the FC alone appears insufficient to produce an optimal outcome. As shown by Amara et al.,\textsuperscript{20} who compared outcomes between patients receiving long-construct fusion to the

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig4.png}
\caption{Mindmap demonstrating algorithmic approach to adult degenerative scoliosis diagnosis, characterization, and correction. cMIS, circumferential minimally invasive surgery; FC, fractional curve; MC, major curve; BMD, bone mineral density; EMG, electromyography.}
\end{figure}
upper or lower thoracic spine to those undergoing treatment of the FC alone, long-segment constructs invite higher risk profiles, but lower the risk of surgical revision.

In summation, we used the findings of the present study to synthesize a summary “mindmap” diagram that illustrates an algorithmic approach to surgical management of the FC in patients with ADS (Fig. 4).

5. Limitations

There are several limitations to acknowledge regarding the present study. In general, there is currently a paucity of studies comparing cMIS and open approaches for FC management, in part due to the recent advent of cMIS techniques. Consequently, it is unclear whether cMIS techniques offer similar levels of coronal plane realignment, pain relief, or functional improvement relative to traditional open techniques. Additionally, the present review is retrospective by design and utilizes data reported by previously published studies. As we did not directly collect or report the data within each study it is difficult to determine the extent to which particular studies may be biased. The present review is therefore limited by the quality of the included studies and is subject to reporting and sampling biases at baseline. Furthermore, the degree of heterogeneity presents across studies included in this review made quantitative synthesis challenging; nonetheless, we were deliberate and meticulous in our screening and enforced strict application of inclusion and exclusion criteria to ensure that, despite differences in study designs and potential differences in participating patient populations, the focus of each study was the corrective ability and role of the FC in the operative management of ADS.

CONCLUSION

Current literature suggests that the FC—defined as the compensatory lumbosacral distal to the major curve in ADS—may be a key structural contributor to malalignment and symptoms in patients with ADS. When coronal plane malalignment is ipsilateral to the FC concavity, the FC likely contributes to deformity progression and therefore must be corrected in the final construct. By contrast, when truncal shift is contralateral to the FC concavity, the FC is likely compensatory and ending constructs proximal to the sacrum may be reasonable, allowing for preservation of an additional motion segment, albeit at the cost of increased risk of adjacent segment disease. For curves requiring significant coronal plane realignment, kickstand rods appear to be the most effective technique when combined with conventional rod bending maneuvers and osteotomy work. Preliminary data suggests that cMIS techniques targeting ACR for deformity correction may help to preserve motion segments by enabling similar degrees of coronal plane correction as open posterior-only constructs. Nevertheless, the utility of cMIS techniques can be limited, especially for severe, flexible curves. Additional investigation into the strengths and weaknesses of cMIS techniques relative to open approaches is necessary, as is an improved understanding of the extent to which the FC is a precipitant of pathology in ADS as opposed to a compensatory mechanism for the underlying degenerative major curve.

NOTES

Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: SCR, ZP, NJB, ALM, NL, LDDA, JAO, BDE, MHP; Data curation: SCR, ZP, NJB, SS, JR, AS, JCH; Formal analysis: SCR, NJB, SS, JR, JCH; Methodology: SCR, NJB, JCH, ALM, NL, LDDA, JAO, MHP; Project administration: LDDA, JAO, BDE, MHP; Visualization: SCR, ZP, NJB, ALM, NL, LDDA, JAO, BDE, MHP; Writing – original draft: SCR, ZP, NJB, SS, JR, AS; Writing – review & editing: SCR, ZP, NJB, JR, JCH, ALM, NL, LDDA, JAO, BDE, MHP.

ORCID
Seth C. Ransom: 0000-0001-5949-1703
Zach Pennington: 0000-0001-8012-860X
Nolan J. Brown: 0000-0002-6025-346X
Shane Shahrestani: 0000-0001-7561-4590
Jessica Ryvlin: 0000-0002-4964-7740
Anthony L. Mikula: 0000-0002-5143-3733
Nikita Lakomkin: 0000-0002-3987-0084
Benjamin D. Elder: 0000-0002-7782-7829
Joseph A. Osorio: 0000-0003-4057-2161
Martin H. Pham: 0000-0002-7733-8902

REFERENCES


Artificial Intelligence in Spinal Imaging and Patient Care: A Review of Recent Advances

Sungwon Lee1,2, Joon-Yong Jung1,2, Akaworn Mahatthanatrakul3, Jin-Sung Kim4
1Department of Radiology, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea
2Visual Analysis and Learning for Improved Diagnostics (VALID) Lab, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea
3Department of Orthopaedics, Faculty of Medicine, Naresuan University Hospital, Phitsanulok, Thailand
4Spine Center, Department of Neurosurgery, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea

Artificial intelligence (AI) is transforming spinal imaging and patient care through automated analysis and enhanced decision-making. This review presents a clinical task-based evaluation, highlighting the specific impact of AI techniques on different aspects of spinal imaging and patient care. We first discuss how AI can potentially improve image quality through techniques like denoising or artifact reduction. We then explore how AI enables efficient quantification of anatomical measurements, spinal curvature parameters, vertebral segmentation, and disc grading. This facilitates objective, accurate interpretation and diagnosis. AI models now reliably detect key spinal pathologies, achieving expert-level performance in tasks like identifying fractures, stenosis, infections, and tumors. Beyond diagnosis, AI also assists surgical planning via synthetic computed tomography generation, augmented reality systems, and robotic guidance. Furthermore, AI image analysis combined with clinical data enables personalized predictions to guide treatment decisions, such as forecasting spine surgery outcomes. However, challenges still need to be addressed in implementing AI clinically, including model interpretability, generalizability, and data limitations. Multicenter collaboration using large, diverse datasets is critical to advance the field further. While adoption barriers persist, AI presents a transformative opportunity to revolutionize spinal imaging workflows, empowering clinicians to translate data into actionable insights for improved patient care.

Keywords: Artificial intelligence, Machine learning, Deep learning, Spine, Patient care, Clinical decision-making

INTRODUCTION

Providing optimal spine patient care is becoming increasingly complex due to the rapid growth of patient data, the rising number of spine patients, and expanding treatment options. The large amount of information from medical imaging and electronic health records (EMRs), combined with growing patient volumes driven by an aging population, presents a major challenge for efficient data processing and analysis. At the same time, the increase in treatment methods, from minimally invasive procedures to personalized medicine, demands careful consideration of risks, benefits, and suitability for each individual. Traditional methods may struggle in this situation, highlighting artificial intelligence (AI)’s potential to process large amounts of data, identify patterns, and support evidence-based decision-making for more efficient, personalized, and effective spine care.1,2

Medical research is undergoing a revolution. AI, powered by
machine learning (ML) and deep learning (DL), is unlocking new discoveries with the potential to improve diagnoses, treatments, and patient outcomes. While spine image research faced early challenges due to complex anatomical structures, data scarcity, and nonstandardized protocols, the field has witnessed a remarkable surge in research and commercially available solutions, particularly in the past 5 or 6 years. This rapid advancement is readily apparent in a PubMed search, where a staggering 86% of all AI spine research papers have been published since 2017.

While traditional research methods like regression and correlation models remain valuable tools, the field is experiencing a paradigm shift towards AI algorithms, particularly ML, DL, and generative models. Compared to regression models, which are limited to predicting simple linear relationships, ML can learn complex, nonlinear interactions and hidden patterns within data. DL builds on this by utilizing numerous layers of computations to extract more complex features from high-dimensional data, such as image data. While both ML and DL can be used to build clinical decision support systems or prognosis prediction models, DL also excels at extracting meaningful features from images, enabling applications like landmark detection and disease classification. Generative models, such as generative adversarial networks and large language models (LLMs), go even further. They learn the underlying patterns and relationships within data and then use this knowledge to create entirely new data. This opens up possibilities like synthetic computed tomography (CT) images generated from magnetic resonance imaging (MRI) images, radiology reports written by AI, and chatbots to answer frequently asked questions for patients.

This review focuses on recent advancements in AI-powered spine image analysis and patient care. However, instead of focusing on the role of a specific AI algorithm, our aim is to cover this from the perspective of a clinical task-based manner (Fig. 1). To do this, we searched PubMed and Google Scholar for relevant articles published between January 2017 and March 2024. Our search query combined terms related to AI, ML, and DL with those related to the spine (including “spine” or “vertebra”) and spinal imaging modalities (including “radiograph,” “CT,” and “MR”). We then assessed the retrieved studies for their methodologies, evaluation metrics, and contributions to spine surgery or research. Additionally, we hand-searched the reference lists of included articles to identify relevant studies not captured by the initial database search. This process resulted in a final selection of 134 articles for review.

CURRENT STATUS AND CLINICAL APPLICATIONS ACROSS PATIENT CARE PROCESS

1. Image Improvement

The improvement of image quality through AI is a well-established field already implemented in clinical practice, particularly in CT and MRI. Traditional interpolation techniques based on simple mathematics artificially increase the resolution but often result in blurring. DL has taken this process to the next level, generating more realistic and sharper images. This AI-based interpolation scheme is incorporated as a part of DL-based image improvement protocols in the latest scanners. In CT scans, the radiation dose follows the “As low as reason-
ably achievable” principle. Lower doses typically lead to noisier images. Fortunately, AI algorithms now effectively reduce this noise postacquisition, enabling high-quality images with reduced radiation doses.

In MRI, scan time serves as a trade-off for image quality, analogous to the radiation dose in CT: Achieving high image quality often requires longer scan time, leading to patient discomfort and potential movement artifacts. Traditional techniques like parallel imaging and compressed sensing address scan time reduction but still struggle with noises and artifacts. Recent advancements in DL-based reconstruction have overcome these challenges. By combining traditional schemes with DL-based optimization, DL reconstruction significantly improves signal-to-noise ratio with significantly reduced scan times. In spine MRI, DL-reconstructed images have been found to be interchangeable with standard MRI for detecting various abnormalities, offering excellent image quality with a remarkable 70% reduction in scan time. Major magnetic resonance (MR) vendors have already incorporated these features, and vendor-neutral solutions for spine MR are also commercially available.

However, some concerns remain regarding the potential for DL-based improvements to alter lesion details or introduce artifacts. Ongoing research and validation are crucial to ensure the reliability and safety of these AI-powered tools in clinical practice.

2. Assistance at the Initial Diagnostic Stage

Spinal imaging plays a critical role in the initial step of patient care, with measurements and landmark locations being one of the fundamental information. This information guides treatment planning and risk assessment, ensuring precise and objective decision-making for effective patient management. For AI-based image analysis, accurate segmentation of the vertebral body and disc spaces is a critical first step. Segmentation assists in spine numbering and data preparation for other spine analysis algorithms. This is why structure segmentation models have been a significant focus of spine AI research, with numerous public datasets available to train them. Additionally, recent breakthroughs in natural image segmentation techniques have nudged developments of automated segmentation tools for spine surgical hardware, achieving near-perfect Dice scores (Fig. 2).

Assessment of kyphosis, lordosis, and scoliosis is another primary indication of spinal imaging. However, accurate assessment of spinal curvature relies on hand-measured parameters, such as Cobb angle, pelvic incidence (PI), sacral slope, pelvic tilt, lumbar lordosis (LL), and sagittal vertical axis. This is a laborious task, prone to high interobserver agreements, and thus has been one of the early targets of spine AI research. Like human analysis, supervised DL models can be trained to identify or segment specific landmarks on anteroposterior (AP) and lateral spine radiographs, automatically connecting them to calculate spinal parameters. Advanced models can now detect up to 78 landmarks and 18 spinopelvic parameters in whole spine lateral radiographs, demonstrating excellent agreement with human measurements. In AP radiographs, automated Cobb angles exhibit lower mean error compared to human intra-/interobservers (2°–6.32° vs. ± 9.6°/± 11.8°), showing potential for scoliosis screening in children (sensitivity, 95.7%; specificity, 88.1%), or progression monitoring in postoperative patients. Several commercially available software solutions have already received U.S. Food and Drug Administration (FDA) or Korean Ministry of Food and Drug Safety approval for landmark and curvature detection.

Vertebral segmentation and numbering within spine CT and MR images have become a well-established area of research, offering significant advancements in medical imaging analysis. This technology automatically identifies and delineates specific anatomical structures, most commonly the vertebral body, intervertebral disc, and spinal canal. Reportedly, these automated...
algorithms achieve impressive accuracy, with Dice scores ranging from 89% to 95% for these key structures. Several medical technology vendors have already received FDA approval for spine labeling software, and their incorporation into PACS (Picture Archiving and Communication Systems) is already underway. The true power of automatic segmentation lies in its ability to facilitate quantitative analysis and automated diagnoses. For instance, automated vertebral segmentation enables (1) the detection of abnormal vertebral heights, (2) the identification of abnormal spinal curvatures, and (3) the planning of surgical procedures and radiotherapy. With disc segmentation, we can automatically grade disc degeneration in sagittal MRIs. Furthermore, research with neural foramen segmentation has shown that the cross-sectional area of the neural foramen directly correlates with patient height and inversely correlates with age. These segmentation models have opened doors for automatic diagnoses of spinal stenosis and neural foraminal stenosis.

3. Image Interpretation and Diagnosis

1) Disc herniation and degeneration

While disc and spinal canal segmentation models play a crucial role in identifying anatomical structures, their potential extends far beyond that. In MRI, AI models are making significant strides in detecting and grading lumbar spinal stenosis. These models are trained to identify stenosis in the lumbar central canal, lateral recess, and neural foramina using axial or sagittal images. One study reported remarkable agreement between a trained model and subspecialized radiologists in classifying stenosis severity (normal/mild vs. moderate/severe) in an external test set, achieving kappa values ranging from 0.95 to 0.96. AI-powered MRI analysis can significantly benefit radiologists by reducing interpretation time (124–274 seconds vs. 47–71 seconds, p < 0.001) and improving interobserver agreement (kappa values = 0.71 and 0.70 with DL vs. 0.39 and 0.39 without DL, both p < 0.001). FDA-approved solutions are already emerging. For example, a lumbar spine report generation software boasts impressive sensitivity and specificity for detecting central canal stenosis, as demonstrated in 2 separate studies (92.70% and 99.04% or 77.14% and 98.95% for central canal stenosis). Beyond MRI, AI-powered spinal stenosis diagnosis is also making progress in CT scans and radiographs. AI models have achieved diagnostic accuracies of 83%–88% for the spinal canal and 71%–75% for the lateral recess on axial CT scans. This opens doors for evaluating disc herniation in CT scans or even diagnosing opportunistic disc herniation during abdomen and chest CT scans. For cervical spine radiographs, AI approaches have demonstrated promising results in detecting ossification of the posterior longitudinal ligament (accuracy 0.88, area under the receiver operating characteristics curve, area under the curve [AUC] 0.94, surpassing the performance of spine physicians) and spondylotic myelopathy (accuracy, 71.1%; AUC = 0.864). Another study reports AUC values up to 90% for spinal stenosis in lumbar radiographs, suggesting the potential of AI as a triage tool for further imaging.

2) Fracture

Compression fracture, the most common type of spinal fracture, has been one of the first targets for fracture detection in radiographs. Current DL models achieve impressive accuracy, reaching around 90% in compression fracture detections and even in differentiating between old and new compression fractures (AUC = 0.80). One model reports the performance was comparable to human readers (accuracy of 93%, p < 0.001) with lung markings as the primary source of false positives. Further advancing the field, an FDA-approved universal fracture detector is already in widespread use. It efficiently detects fractures in radiographs of all limbs, spine, and ribs. For spine fractures specifically, a recent study compared human, AI-only, and AI-assisted human detection. The results revealed a clear advantage of AI-assisted human interpretation, with sensitivity/specificity of 94.5%/100% compared to 92.4%/98.4% for humans alone and 89.1%/62.2% for AI alone. More recently, advancements in fracture detection have extended to CT scans. In 2022, the winners of the cervical spine fracture detection challenge achieved an AUC of 0.96 (95% confidence interval [CI], 0.95–0.96), sensitivity of 88% (95% CI, 86%–90%), and specificity of 94% (95% CI, 93%–96%). This same year also saw the introduction of an FDA-approved cervical spine fracture detector. Research further extends to thoracolumbar CT scans, with some studies demonstrating the ability to categorize fractures into 4 types (no injury, compression, burst, translational/rotational, and distraction) with accuracies ranging from 68.6% (burst) to 89.3% (distraction).

3) Inflammation and tumors

Identifying and classifying inflammatory diseases, infections, and tumors are actively researched fields within spine imaging, each too vast to explore fully here. However, AI models are already demonstrating impressive capabilities, achieving expert-level performance in several tasks. For example, in pelvic radiographs, DL models can diagnose sacroiliitis with accuracy that is on par with experts (Cohen kappa = 0.79). Similarly, they...
can accurately quantify inflammatory sacroiliitis in MRI\textsuperscript{51,52} and differentiate between tuberculous and pyogenic spondylitis with an AUC of 0.802 (compared to 0.729 for human experts, p = 0.079).\textsuperscript{53}

Spinal oncology has also seen exciting advancements in radiomics and DL. Algorithms can now detect metastatic lesions in CT scans (sensitivity, 75\%–90\%\textsuperscript{54,55}) and be applied to reduce interrater variability.\textsuperscript{56} In MRI, DL models outperformed fourth-year residents in differentiating malignant vertebral fractures, achieving 90\% sensitivity and 79\% specificity.\textsuperscript{57} Furthermore, ML and DL can discriminate between normal and pathologic bone marrow patterns in MRI\textsuperscript{58,59} and distinguish spine metastases from lung cancer versus other primary origins.\textsuperscript{60}

4. Surgical Planning and Intraoperative Use

CT aids surgical planning by detailing bone anatomy and pathology and guiding surgical approaches. Moreover, navigation systems using preoperative CT images improve screw placement accuracy. Synthetic CT scans are computer-generated images resembling conventional CT scans but are created using other imaging modalities like MR images through generative models or other DL algorithms.\textsuperscript{61} This technology has several advantages over CT scans, such as elimination of radiation exposure (radiation dose of an average lumbar spine CT scan: 3.5 mSv),\textsuperscript{62} improved visualization of metal structures and the peripheral field of view, and high-resolution depiction of soft tissue structures like intraosseous hemangiomas.\textsuperscript{62} Studies comparing visualization of body structures, artifacts, and geometrical measurements between synthetic and traditional CT scans have found the synthetic versions to be non-inferior.\textsuperscript{65}

However, synthetic CT is still a relatively young field with limited clinical use. While some applications like presurgical and radiotherapy planning are emerging,\textsuperscript{63,64,65} caution is advised when using measurements from synthetic spine CT scans. Studies have shown inaccuracies in pedicle measurements performed in the axial plane, with relative errors reaching up to 34\%.\textsuperscript{66}

Augmented reality (AR) and virtual reality (VR) are emerging technologies demonstrating promising benefits in various healthcare fields, including robotic surgery, laparoscopic surgery, and, notably, orthopedic surgery for the spine.\textsuperscript{66} One of their main applications is as a navigation tool in the operating room. AR and VR systems use computer vision techniques to process preoperative or intraoperative images (radiographs, CT scans, or MRIs) and overlay relevant anatomical structures, potential screw trajectories, or ideal screw locations onto the surgical field, guiding surgeons with real-time visualization.\textsuperscript{67}

The integration of DL into AR and VR systems further enhances their capabilities, particularly in object and landmark detection within images. For example, DL has been used to track a specific vertebra of interest in fluoroscopic images with high accuracy (mean error of 2.27\%)\textsuperscript{68} and identify 7 anatomical landmarks on intraoperative lumbar spine CT scans with minimal error.\textsuperscript{69} Additionally, DL has shown promise in improving robotic screw placement\textsuperscript{70} and identifying bone drill breakthroughs during surgery.\textsuperscript{71} Classification models powered by DL can even differentiate various types of pedicle screws\textsuperscript{72} and anterior cervical fusion systems\textsuperscript{73} in radiographs.

Screw navigation systems are getting a boost from neural networks. These AI tools automate screw planning, including screw size and trajectory. One study reports a dramatic 90\% reduction in workflow time, with just 3 out of 130 screws requiring manual adjustment.\textsuperscript{74} Neural networks can also personalize screw placement for each patient’s bone structure. This customization helps maximize pull-out force and reduce screw failure, which will be an essential benefit for patients with osteoporosis.\textsuperscript{75}

We anticipate that AI will play a crucial role in addressing some current limitations of AR and VR, such as low image resolution and steep learning curves. AI-powered AR and VR solutions hold potential, offering features like determining ideal spine alignment and implant size, compensating for motion, and even facilitating 3D printing.\textsuperscript{66,67}

5. Opportunistic Diagnosis

Opportunistic screening refers to usage of incidental information from existing medical images acquired for a different purpose.\textsuperscript{76} For example, an abdominal radiographs and CT scans intended for other diagnoses may incidentally reveal compression fractures.\textsuperscript{77,78} One of the most promising areas for this approach lies in body composition imaging, particularly for the assessment of bone mineral density (BMD).

While dual-energy x-ray absorptiometry (DXA) remains the standard for BMD measurement, CT scans offer valuable insights, especially for individuals with obesity, severe degenerative spine disease, or postoperative spine conditions where DXA accuracy can be compromised. Quantitative CT using reference phantoms or dual-energy CT has been used for these cases, but their accessibility remains limited. However, advancements in ML are now enabling the extraction of reliable BMD information from various CT scans, including the abdomen, chest, and lumbar spine, transforming them into valuable opportunistic screening tools.\textsuperscript{79,81}
Opportunistic CT scans are also revealing valuable insights into sarcopenia, a condition linked to vertebral compression fractures and increased mortality across various diseases.\textsuperscript{82,83} Sarcopenia is strongly associated with the cross-sectional area and mean density of muscles like the psoas or abdominal wall in CT scans.\textsuperscript{76} AI algorithms can now automatically segment muscles, subcutaneous and visceral fat, and vertebral bodies for sarcopenia and BMD measurement.\textsuperscript{84,85} Body composition measurements from opportunistic imaging could even potentially aid in the risk stratification of patients undergoing spinal surgery or predict future fractures.\textsuperscript{86,87}


ML and DL algorithms are revolutionizing the way we predict patient outcomes in spine surgery. While still in their early stages of clinical implementation, these powerful tools hold immense promise for personalized care and improved decision-making.

While the traditional studies with well-designed large cohorts or linear regression models have achieved success,\textsuperscript{88} ML algorithms can capture both linear and nonlinear relationships between diverse factors and outcomes, requiring less human intervention in model development.\textsuperscript{89} This leads to superior predictive performance compared to traditional methods.\textsuperscript{90,91} Furthermore, ML enables clustering of patients with similar data patterns. This paves the way for prognostication and treatment optimization tailored to specific groups of patients.\textsuperscript{92,93}

However, it's important to note that successful outcome prediction using ML requires high-quality, well-structured datasets with minimal missing data and clear outcome measures.

1) Herniated disc disease surgery

Herniated disc disease is a common spine condition where ML has shown utility in surgical decision support and postoperative prognosis. One key area of focus involves predicting recurrent herniation after surgery. Several ML studies have identified high-risk patients based on patient demographics, clinical parameters, and pre- and postoperative pain scores.\textsuperscript{90,94} Significant factors reported to be associated with reherniation included pain scores, Oswestry Disability Index (ODI), PI–LL mismatch, body mass index, coronal angulation, duration of symptoms, and age.\textsuperscript{94} Additionally, incorporating radiographic features such as facet orientation, herniation type, Modic changes, and disc calcification has further enhanced prediction accuracy for recurrent lumbar disc herniation.\textsuperscript{95}

The decision to proceed with surgery for herniated disc disease involves complex considerations. Models trained on patient demographics, questionnaire data, and MRI results exhibit promising potential in surgical triaging, predicting surgical referrals with AUCs ranging from 0.68 to 0.88.\textsuperscript{96,97} Similar models have also been used to predict improvements in quality of life after surgery (AUCs up to 0.78),\textsuperscript{98} or after conservative therapy (100% accuracy for 12 scale ODI).\textsuperscript{99} Mourad et al.\textsuperscript{99} built a surgical recommendation model based on clinical symptoms, MRI findings, and patient demographic factors. The root mean square error between model predictions and ground truth was 0.0964, with agreement being higher than agreement between individual doctors.

Moreover, ML models trained with demographic information, comorbidities, and preoperative/intraoperative findings have shown efficacy in predicting hospital length of stay or readmission after spine surgeries like anterior cervical discectomy and fusion.\textsuperscript{100-102} or lumbar single-level laminectomy.\textsuperscript{103} This could aid in determining the appropriate care setting (outpatient vs. inpatient) or optimize hospital resource allocations.

2) Spinal deformity

Adult spinal deformity surgery is complex and carries a high risk of complications. To improve outcomes and reduce these complications, researchers have developed algorithms to aid clinical decision-making.\textsuperscript{104} For instance, a study by the International Spine Study Group used ML-based clustering and found 4 prognostic phenotypes.\textsuperscript{105} This approach identified that younger, more resilient patients with good mental health were less likely to need repeat surgery compared to older, frail patients with poorer mental health. Another study by Scheer et al. employed decision tree ensembles to predict proximal junctional kyphosis and pseudoarthrosis with an AUC of 0.896\textsuperscript{106} and 0.947.\textsuperscript{107}

3) Postoperative complication

ML models are demonstrating remarkable potential in predicting major postoperative complications following spine surgery, such as wound infections, thromboembolism, and mortality.\textsuperscript{108,109} These models leverage diverse demographic and clinical parameters, often exceeding the performance of traditional logistic regression models and established risk scores like the American Society of Anesthesiologists (ASA) physical status classification score.\textsuperscript{110}

For instance, Hopkins et al. developed an ML model that predicts postoperative surgical site infection after spinal fusion with a median AUC of 0.787.\textsuperscript{111} Their analysis revealed factors like...
congestive heart failure, chronic pulmonary failure, hemiplegia/paraplegia, and multilevel fusion as the most influential variables, providing valuable insights for risk stratification. Additionally, another study identified 10 key predictors, including age, gender, ASA physical status classification grade, surgical approach, and preoperative laboratory values, demonstrating the broad range of factors that ML models can incorporate for comprehensive risk assessment.\(^{112}\)

Hardware failure is another concern, especially for patients with known risk factors such as osteoporosis, long fixation length, and certain fixation end points.\(^{113}\) You only look once v5 (YOLOv5), a type of convolutional neural network model specialized for both detection and classification, can help detect hardware failure in postoperative radiographs (Fig. 3).\(^{114}\)

**DISCUSSION**

1. Limitations and Challenges

While AI holds immense promise for spinal imaging, several limitations require careful consideration. The “black box” nature of certain models and the lack of interpretability in decision-making raise concerns for clinical adoption. Fortunately, researchers are developing techniques to make AI models more interpretable. These techniques, like Gradient-weighted class activation mapping or Shapley additive explanations, can create heatmaps or graphs that highlight the focus of the model when making predictions. This can help explain, for instance, how the model identified spinal stenosis\(^{115}\) or reherniation.\(^{94}\) Even more recent advancements involve using LLMs to explain AI model predictions in a more comprehensive way, further improving interpretability.\(^{116}\)

Another crucial hurdle is ensuring generalizability.\(^{117}\) Recent studies suggest external validation may not adequately assess true model performance. They recommend that researchers prioritize recurrent local validation across diverse datasets to ensure real-world applicability.\(^{118}\) The focus should shift from chasing perfect metrics to demonstrating tangible clinical value, such as reducing inter-reader variability or mitigating human errors.

Furthermore, the scarcity of high-quality labeled datasets poses a significant challenge. Compared to other radiology fields, large public benchmarks for spinal imaging remain inadequate. Existing public datasets (Table 1) tend to be small in size (<1,000 images), fragmented across institutions, and heterogeneous in acquisition and populations, leading to overfitting and limited generalizability. While current models perform well for normal

---

**Fig. 3.** Hardware failure predictions on spine radiographs. An off-the-shelf model was trained to predict 8 types of hardware failures: adjacent segment disease, loosening, migration, pseudoarthrosis, protrusion, rod fracture, screw fracture, and subsidence. The arrows were embedded in the original images to indicate screw fracture (black arrow) and adjacent segment disease (white arrow).

---

**Table 1.** Public spine radiology image datasets

<table>
<thead>
<tr>
<th>Source</th>
<th>Modality</th>
<th>Region of interest</th>
<th>Segmentation/labels</th>
<th>Data size</th>
<th>Released date</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPIDER(^{131})</td>
<td>MR</td>
<td>Lumbar</td>
<td>Segmentations of the vertebrae, disc, and spinal canal</td>
<td>447 Sagittal series (218 patients)</td>
<td>2023</td>
</tr>
<tr>
<td>Mendeley data(^{132})</td>
<td>MR</td>
<td>Lower Lumbar</td>
<td>Radiology report</td>
<td>48,345 Axial slices (515 patients)</td>
<td>2019</td>
</tr>
<tr>
<td>RSNA cervical fracture(^{47})</td>
<td>CT</td>
<td>Cervical</td>
<td>Label and segmentations of vertebrae and fractures</td>
<td>3,112 Patients</td>
<td>2023</td>
</tr>
<tr>
<td>CTSpine1K(^{133})</td>
<td>CT</td>
<td>Whole spine</td>
<td>Segmentation of vertebra</td>
<td>1,005 CT</td>
<td>2021</td>
</tr>
<tr>
<td>VerSe(^{134})</td>
<td>CT</td>
<td>Whole spine</td>
<td>Segmentations of the vertebrae</td>
<td>374 Sagittal series (355 patients)</td>
<td>2020</td>
</tr>
<tr>
<td>BUU’ spine(^{135})</td>
<td>XR</td>
<td>Lumbar</td>
<td>Labels of 4 disease entities</td>
<td>400</td>
<td>2023</td>
</tr>
<tr>
<td>VinDr-SpineXR(^{136})</td>
<td>XR</td>
<td>Whole spine</td>
<td>Labels of 13 disease entities</td>
<td>10,466</td>
<td>2021</td>
</tr>
</tbody>
</table>

MR, magnetic resonance; CT, computed tomography; XR, x-ray.
spinal anatomy, there are very few models trained on datasets with normal variants, fractures, models predicting various disease entities.

2. Future Directions

Despite the challenges, the future of AI in spinal imaging is bubbling with potential. Researchers are increasingly leveraging prospectively gathered data from clinical trials and multi-centered datasets, yielding promising results. For instance, studies utilizing the American college of surgeons-national surgical quality improvement program database have successfully generated ML models for predicting 30-day readmissions or discharge to nonhome facility after lumbar fusion using pre-discharge information.

Another exciting approach involves directly incorporating EMR or imaging data into DL models. Natural language processing shows promise in predicting intraoperative vascular injuries with superior accuracy compared to traditional methods. Similarly, DL models trained directly on images can identify hidden features invisible to the human eye. For example, DL models utilizing preoperative sagittal MRIs of the cervical spine as inputs have demonstrated higher accuracy in predicting early onset adjacent segment disease in cervical fusion patients compared to models using only preoperative clinical data. Additionally, AI models trained on spinal radiographs and CT scans are outperforming established risk assessment tools in predicting fractures. Finally, encouraging results have been achieved in predicting the early progression of scoliosis using spinal radiographs.

CONCLUSION

In conclusion, this review highlights the remarkable progress and potential of AI in advancing spinal imaging and patient care. From automated measurements to surgical planning, AI is transforming workflow efficiency, accuracy, and reliability across the spectrum of spine imaging. However, thoughtfulness is required to ensure real-world validity, utility, and adoption of AI tools.

REFERENCES

10. Yeoh H, Hong SH, Ahn C, et al. Deep learning algorithm for simultaneous noise reduction and edge sharpening in spinal imaging, there are very few models trained on datasets with normal variants, fractures, or models predicting various disease entities.


sualization of the cervical spine with deep learning-based synthetic CT compared to conventional CT: a single-center noninferiority study on image quality. Eur J Radiol 2022;154:1-10.


Comparative Review of the Socioeconomic Burden of Lower Back Pain in the United States and Globally

Diana Chang¹, Austin Lui², Alisa Matsoyan², Michael M. Safaee³, Henry Aryan², Christopher Ames²

¹Department of Neurosurgery, University of California, Los Angeles, CA, USA
²Department of Neurosurgery, University of California, San Francisco, University of California, San Francisco, CA, USA
³Department of Neurosurgery, University of Southern California, Los Angeles, CA, USA

Internationally, the United States (U.S.) cites the highest cost burden of low back pain (LBP). The cost continues to rise, faster than the rate of inflation and overall growth of health expenditures. We performed a comprehensive literature review of peer-reviewed and non-peer-reviewed literature from PubMed, Scopus, and Google Scholar for contemporary data on prevalence, cost, and projected future costs. Policymakers in the U.S. have long attempted to address the high-cost burden of LBP through limiting low-value services and early imaging. Despite these efforts, costs (~$40 billion; ~$2,000/patient/yr) continue to rise with increasing rates of unindicated imaging, high rates of surgery, and subsequent revision surgery without proper trial of non-pharmacologic measures and no corresponding reduction in LBP prevalence. Globally, the overall prevalence of LBP continues to rise largely secondary to a growing aging population. Cost containment methods should focus on careful and comprehensive clinical assessment of patients to better understand when more resource-intensive interventions are indicated.

Keywords: Low back pain, Chronic low back pain, Health care economics and organizations, Medical economics, Global health, Cost of illness

INTRODUCTION

Approximately 5% of the United States (U.S.) population consumes 50% of all healthcare spending. A portion of this high-cost minority are individuals with chronic low back pain (LBP).¹

Internationally, the U.S. cites having the highest cost burden of LBP is consistent with having the highest national health expenditure overall.² In 2013, spinal conditions consisting of both neck and back pain, accounted for the third highest national health spending after only diabetes and ischemic heart disease.³ The cost continues to rise along with a growing prevalence. In 2016, nearly 10% (31.6 million people) of the U.S. population reported suffering from chronic LBP.⁴ Globally, the age-standardized point prevalence is similar, around 8.2%.⁵ Between 1997 and 2005, total national expenditure of adults with spine problems increased 65% after adjusting for inflation and faster than the rate of growth of overall health expenditures.⁶

1. Theoretical Framework for Understanding the Cost of Care

This review comprises of 3 sections: (1) an overview of the socioeconomic burden of LBP in the U.S., (2) cost burden globally divided geographically and by level of development, and (3) cost burden of LBP compared to other chronic diseases in the U.S. For domestic data, relevant literature and statistics for each section were gathered first through government- or government-affiliated databases (Healthcare Cost and Utilization Project, Centers for Medicare & Medicaid Services, Centers for Disease Control and Prevention), then a focused literature search of academic and nonacademic databases (MEDLINE, Embase, Web
of Science, Google Scholar), including nonprofit domestic economic thinktanks for more updated statistics. Search terms for cost statistics included, "socioeconomic [AND] LBP," "cost [AND] LBP," "spending [AND] LBP," "economic* [AND] LBP," and "healthcare costs [AND] LBP," for disease burden statistics, search terms included "prevalence [AND] LBP," "incidence [AND] LBP," and "burden [AND] LBP." For statistics on international practices in the management of LBP, search terms were a combination of the prior terms with the country's name. For example, for the cost of LBP in Korea, the following search term was used: "Korea" [AND] 'LBP' [AND] 'cost.' The same formula was used for cost and prevalence statistics for every country and region. Authors selected the literature by first using publicly accessible domestic and international government databases with updated data then using nongovernment-based sources to supplement missing data. Inconsistencies were resolved to default on data from government-based databases.

1) Framework

To compare the cost of LBP care per capita across countries of different sizes. The following equation was used.

\[
P(i,t) = \text{TruePrevalance}(i,t) \times \text{Diagnosis}(i,t) \times C(i,t) = \text{TruePrevalance}(i,t) \times \text{Diagnosis}(i,t) \times \text{Prob of Surgery} \times \text{Cost of Surgery} + \text{Prob of Imaging} \times \text{Cost of Imaging} + \text{Prob of PT} \times \text{Cost of PT}
\]

where,

- \( P(i,t) \) is the per capita cost of LBP treatment in country \( i \) in year \( t \)
- TruePrevalance captures true prevalence in the population
- Diagnosis is the probability to be diagnosed given back pain (not all sick people seek care or are being diagnosed)
- \( C(i,t) \) is the cost per diagnosed patient
- Other variables are self-explanatory and are conditional on the positive diagnosis

All costs are reported in U$S2021. Domestic costs were converted to U$S2021 using the Consumer Price Index for All Urban Consumers as reported by the US Bureau of Labor Statistics for the July of the respective year to July of 2021. The mid-year (July 1st) Consumer Price Index index was chosen as the period closest reflecting global pre-COVID-19 pandemic real Gross Domestic Product trends in attempt to attenuate the drastic economic effects experienced due to the pandemic.

Per capita costs for international studies were calculated by dividing the total national annual cost by the national annual population size as reported by the World Bank for that year. This amount was then converted to U$S2021 using the same methodology described above.

All ensuing costs are given in U$S2021.

RESULTS

1. The Socioeconomic Burden of LBP in the U.S.

1) Sources of high spending: frequent ambulatory visits, surgery, imaging

A major driver of the high-cost burden of LBP in the U.S. is higher rates of surgery and frequent and often initial visits to medical specialists (and the associated interventions) instead of primary providers. Sixty-one percent of the $22.9 billion of total medical spending to address LBP in 2016 was spent on ambulatory visits. LBP accounted for 2% of all (or 2.63 million) emergency department visits in 2006. Nearly 67% of these patients were admitted and 10% receiving computed tomography (CT) or magnetic resonance imaging (MRI), 3 times higher than imaging rates in 2002. One in 4 patients who received primary care for LBP received imaging while 1 in 3 patients in the Emergency Department received imaging. Within the first 90 days of beginning sick-leave, on average 32% of patients with LBP undergo surgery in the U.S. compared to 6% in other highly developed countries like Sweden. The same trend rings true over time with 92% of U.S. patients receiving surgery within the first year compared to 75% in countries like Germany.

Not only is surgery an earlier therapeutic option in the U.S., but the rate of surgical intervention also continues to rise particularly for degenerative spinal diseases. Between 2004 and 2015, the volume of elective lumbar fusions in the U.S. increased 62.3% (from 60.4 to 79.8 fusions per 100,000 U.S. adults). Amongst those older than 65 years old, the volume increased more drastically, from 98.3 to 170.3 per 100,000 U.S. adults. The market for lumbar fusions continues to grow 18%–20% annually with fusion as the standard for treating common lumbar pathologies which do not typically involve instability, like lumbar stenosis despite few studies demonstrating definitive clinical superiority of fusion over nonfusion decompression. One proven driver of this increase in surgery is imaging overuse which may lead to faulty attribution of pain to an imaging abnormality, particularly as most imaging abnormalities are incidental findings in asymptomatic patients. Although surgery is not the most widespread intervention it is the costliest, averaging $51,500 per admission and exceeding $10 billion in ag-
aggregate in 2015.\(^5\) In another study, while only 1.2% of patients with newly diagnosed LBP received surgery, surgery accounted for almost 30% of the total 12-month costs of the entire cohort.\(^3\)

The American College of Physicians (ACP) has developed clinical guidelines for primary care physicians and Emergency Department physicians seeing a patient for the first time with LBP.\(^7\) ACP guidelines urge against imaging within 30 days of diagnosis and before trying nonsurgical treatments. Deviations from these guidelines are common and costly—responsible for an additional $373 million annually.\(^9\) In a recent study, patients who received imaging (lumbar CT, MRI, or radiograph) within 30 days of diagnosis had double the 12-month costs of those treated under guidelines even after stratifying by imaging modalities.\(^1\) Use of MRIs as the first intervention within 30 days of the LBP diagnosis led to an 8-fold increased risk of spine surgery.\(^2\) Furthermore, Lurie et al. found that 22% of the regional variation in spine surgery rates can be explained by variation in the rate of advanced spine imaging (CT and MRI), a trend that has been true over time and across regions.\(^19,20\) Notably, advanced imaging is twice as predictive of surgery than the regional density of spine surgeons, hospitals in which spine surgery is performed, or socioeconomic or insurance status.\(^21\) Thankfully, the rate of guideline deviation has decreased over the past decade.\(^4\)

One of the most common nonsurgical treatments is prescription medications, which contributed 15% of the direct medical costs in 2007.\(^22\) In 2008, there was a 50% increase in narcotic prescriptions concomitant with a 50% decrease in acetaminophen prescriptions.\(^23\) The prescription of opioids has been linked with worse pain, functioning including higher doses being directly associated with prolonged work disability, catastrophizing and depression.\(^18,24,25\) This has also contributed to the increase in substance abuse disorders and deaths due to overdose. While clinicians are increasingly wary of prescribing narcotics for chronic LBP, opioids remain the most frequently prescribed medication for LBP.\(^26\)

2) **Indirect cost burden twice as large due to lost productivity**

However high the direct cost, the indirect cost of LBP is at least twice as high.\(^1\) An estimated 149 million workdays are lost annually due to LBP, accounting for 5% of lost workdays from any cause. A major cause of LBP and thus lost workdays is occupation-related LBP from high-risk industries such as lumber retailing, gas extraction, and nursing. Occupation-related LBP was responsible for 101.8 million (68%) of lost workdays due to LBP.\(^24\) Occupations with the highest prevalence of LBP include health care providers, farmers, fishers, and forestry workers.\(^27\)

2. **The Socioeconomic Burden of LBP Globally**

   1) **LBP: leading source of global disability since 1990**

   Since the global burden of disease study was published in 1990, LBP has been a leading cause of years lived in disability (YLD).\(^28-30\) In 2017, about 580 million people worldwide reported having LBP, with an incidence of 250 million responsible for 64 million YLD annually. In the Western population, 70%–85% will develop LBP at one point in life, 60% will continue to report LBP 1 year later, and 10%–15% will have chronic LBP.\(^31-33\) Due to a growing global population, the age-standardized point-prevalence of LBP has decreased in most countries (-2.1%). However, overall number of cases has increased nearly 20% between 2007 and 2017.\(^5,28\) Prevalence is highest among those 80+ years old. Nevertheless, YLD is highest among those 45–49 years old because of the significance of disability on quality of life at a working age. Causes of this absolute rise in prevalence include increased longevity, obesity and psychiatric illness in developed countries. In emerging economies, additional causes include rapid industrialization with a growing working population and increasingly sedentary lifestyles.\(^5,29,34-36\)

   Accurate cost comparisons across international studies remains elusive as studies adopt varying methodologies for calculating costs. Regardless, a few systematic reviews have attempted to explore geographic differences.\(^5,37-38\) The findings from these reviews and more recent studies are summarized by their respective geographic region in Table 1 and Figs. 1 and 2.

2) **LBP in high-income North America and Australia**

   The age-standardized mean prevalence of LBP in the US and Canada is 10.71% (95% confidence interval, 10.06–11.39), the 7th highest prevalence globally since 1990.\(^3\) In Canada, emergency department visits for LBP has increased to 3.2% of all visits, only 9% of which are truly attributable to nerve impingement.\(^39\) Disease and treatment regimen has also evolved with increasing incidences of LBP due to “sequelae of previous back surgery” which was claimed 26 times more in 2007 than 2000.\(^40\) Utilization of instrumented lumbar surgeries more than doubled between 1993 and 2012 with the annual procedure rate among those older than 80-year old increasing 7.6-fold.\(^40\) The same is true in Australia. Between 2003 and 2013 the rate of 3+ level or 2+ approach spinal fusion grew 400%, simple fusion grew 115%, while decompression grew 16% for the treatment of spinal stenosis despite minimal evidence of their marginal benefit over decompression alone.\(^41\)

   Despite a notably distinct healthcare financing system, Canada, like the U.S., has dramatically increased the rate of surgical

https://doi.org/10.14245/ns.2448372.186
intervention to address LBP.46,47 This cost burden is only worsened by postsurgical complications, suggesting potential over-medicalization of a multifactorial pain syndrome.

3) LBP in Europe: Netherlands, Belgium, Sweden

The highest per capita direct cost was reported in 2002 from the Netherlands at 8,533€ ($410)—US$2021 calculated from foreign currency using the historical exchange rate average in the year reported. Then the historic US equivalent is covered to US$2021 using the US Bureau of Labor’s Consumer Price Index published annual inflation rate—per capita.43 Costs were highly variable depending on level of care and referral patterns. Specialist care for back pain cost significantly more than those treated at a primary care setting (4,875€ vs. 2,365€ or $6,045 vs. $2,932, p < 0.001). General practitioner referrals also accumulated lower costs relative to those referred by a specialist (1,569€ vs. 3,018€ or $1,946 vs. $3,742, p < 0.05).44 Among all patients, LBP was managed as follows: 88% treated with exercise therapy, 53% with opioids, and only 26% patients treated with surgical intervention.44 LBP alone was responsible for 25% of all drug costs for musculoskeletal pathology with an average of 2 medications per patient with chronic lower back pain.45

In Sweden, surgery and specialist care are responsible for a quarter each of all direct costs for LBP in Sweden, similar to the US where surgery is responsible for 22% of all direct costs.46,47 The landscape of medical costs for LBP has changed as treatment regimens evolve. In a recent Belgian study, LBP was responsible for 55% of all transcutaneous electrical nerve stimulation units and 60% of intrathecal pumps.48 Across the continent, the indirect cost of absenteeism varies widely, ranging from 38%–85% of all costs.49,50 In Sweden, each episode of LBP results in on average 51 days of absenteeism, equivalent to 2,753€ ($3,436).49 Cost of illness comparisons between the U.S. and European countries is difficult as most U.S. studies take the perspective of private insurance. Any cost reduction over time may be due to true cost containment efforts or costs merely shifted to another payer.

Lumbar fusion surgery popularity grows not only in the U.S. but around the world though at a different pace. In Finland, as the rate of lumbar decompressions doubled between 1997 and 2017 from 33 to 77 per 100,000 person-years, the rate of lumbar fusions tripled from 9 to 30 per 100,000 person-years.50 Norway likewise experienced a faster growth in the rate of complex lumbar surgeries, the majority of which were fusions (13.6 to 21 per 100,000 inhabitants), than simple lumbar surgeries (64.3 to 88.9 per 100,000) between 1999 and 2013. Females and adults
between 60 and 74 years old made up the most frequent and fastest growing complex lumbar surgery demographic. So while the U.S. reports the fastest growth in the total number of lumbar fusions performed year over year, rate of growth per 100,000 inhabitants in Europe currently outpaces the U.S.

4) LBP in Asia: Korea, Japan, China

In Asia, high-income countries had the highest burden of LBP (age-standardized point prevalence: 13.16 [11.74–14.73]) whereas lower-income East Asia had the lowest of all regions globally (3.92 [3.46–4.37]) likely due to high population density.

Fig. 1. Prevalence, direct cost and indirect cost of low back pain for select high-income countries. The relative prevalence, direct and indirect cost (US$2021), and direct and indirect cost per person (US$2021) are depicted for the United States (U.S.), Japan, Korea, Netherlands, Sweden, and Canada by colored ribbons with their associated rank. For example, the U.S. has the highest prevalence, annual direct cost, direct cost/person, and annual indirect cost and therefore the orange ribbon has both the widest and topmost ribbon for those categories. However, Sweden has the highest indirect cost per person and likewise, the purple ribbon is the widest and surpasses the orange for indirect cost per person. The referenced article is listed under each country. For countries with more than one referenced article, asterisks help differentiate from which article the prevalence and/or cost data originated.
In Korea, duration of pain was the major direct cost determinant consistent with findings from other high-income countries. Fifty-one percent of insurance claims for back pain was for pain lasting less than 6 months accounting for 10% of total costs due to LBP compared to the 6% of claims for pain lasting longer than 2 years which was responsible for 30% of costs.\textsuperscript{52}

In Japan, chronic LBP affects an estimated 1.5 million people, accounting for nearly one-third of patients with chronic pain.\textsuperscript{53} While no difference was found in costs per ER visits or hospitalization, chronic LBP patients sought their provider seven

---

(Continued)
times more than age, sex, body mass index, Charlson Comorbidity Index, and smoking status-matched controls leading to an incremental cost of $1,230 per person and a national direct cost burden of $24.4 billion.\textsuperscript{53} Unlike in other developed countries, loss of productivity in Japan due to chronic LBP is largely due to presenteeism, or decreased productivity while being present at work. Like in other high-income countries, the socioeconomic burden of LBP is significantly worsened by psychiatric comorbidities.\textsuperscript{54–56} Depression and anxiety is associated with higher pain, lower quality of life, increased productivity

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{Global age-standardized prevalence of low back pain (LBP) by region in 1990 and 2017. Average age-standardized prevalence percentage of LBP by geographic region and gender in 1990 and 2017 represented as black lines and colored bars, respectively, based on the global burden of disease studies. For 2017 estimates, the male average is marked by blue bars; female average by pink; male and female averages depicted in purple. The lower and upper uncertainty level markers are marked by vertical dashes. Regional percent changes between 1990 to 2017 and the male and female average regional rankings in 1990 and 2017 are listed in the right three columns, respectively. (Continued)}
\end{figure}
loss, and increased healthcare utilization in patients with chronic LBP.\textsuperscript{57,58} This recognition has shifted Japanese health policy makers towards addressing chronic LBP as a holistic pain to be treated with cognitive behavioral treatment, exercise, and sleep hygiene.\textsuperscript{59}

Since 1990, China has seen a gradual decline in point prevalence of LBP nationwide attributed to an improving primary care system as the YLD due to LBP increased 23% between 1990 and 2016 due to the population growth and increased longevity.\textsuperscript{60}

5) LBP in emerging economies: Sub-Saharan Africa, India, Brazil

Known risk factors like height and fat distribution in high-income populations have no relationship to LBP in lower-income populations.\textsuperscript{61}

In Sub-Saharan Africa, after degenerative spine disease, spinal infections are the second leading cause of LBP with tuberculosis responsible for nearly 80% of symptomatic infections.\textsuperscript{62} Human immunodeficiency virus is the cause cited for 84% of lower back spondyloarthropathies and the third leading cause of LBP.\textsuperscript{63}

Multiple studies have examined the occupational hazards affecting men and women of lower economic status in urban India.\textsuperscript{64-66} For men in Southern India, lack of educational attainment is a significant risk factor for LBP.\textsuperscript{61} For working women, the high incidence of LBP (70%-80%) has been attributed to a combination of prolonged hours in suboptimal working positions, occupational monotony, and inadequate income—highlighting the complex biopsychosocial model underpinnings of chronic pain.\textsuperscript{65} Prevalence of LBP among rural housewives in India is likewise high (83%) though the economic burden to society is significantly lower due to reduced access to healthcare and lower wages.\textsuperscript{67}

In countries with rapidly expanding economies, like Brazil, the epidemiology of LBP looks increasingly like those of higher-income countries. In 2016, two-thirds of government spending on spinal disorders was spent on LBP. The direct cost impact is growing secondary to high utilization of healthcare services, procedural interventions, and imaging.\textsuperscript{64}

Treatment and prevention of chronic LBP in lower-income regions varies significantly from those in higher-income countries due to the prevalence of preventable communicable diseases and occupational hazards while countries with rapidly growing economies are beginning to demonstrate the same over-medicalization seen in high-income countries.

| Table 2. Cost of major chronic diseases in the United States (US$2023) |
|---------------------------------|-------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Low back pain                  |                  |                  |                  | $31.92           | $142.91          | $174.83         | $524.54          |
| Stroke                         | Shmagel et al.\textsuperscript{a} 2016 | 2016 | 27,000,000 | $31.92           | $142.91          | $174.83         | $524.54          |
| Dementia                       |                  |                  |                  | $31.92           | $142.91          | $174.83         | $524.54          |
| Cardiovascular disease         |                  |                  |                  | $31.92           | $142.91          | $174.83         | $524.54          |

\textsuperscript{a}Prevalence includes only those diagnosed with Alzheimer disease.

NA, not available.

### References

3. Socioeconomic Burdens of Major Chronic Diseases in the U.S.

While neither the costliest nor the most prevalent chronic disease, chronic back pain has one of the highest cost per person. An aging population and increased longevity will only exacerbate the socioeconomic burden of these chronic diseases, in particular LBP, in the next few decades. Findings are summarized in Table 2 and Fig. 3.

4. Relative Cost of Major Chronic Diseases in Other High-Income Countries

Across high-income countries, LBP is one of, if not the costliest chronic disease per case. Between cardiovascular disease, dementia, stroke and diabetes, LBP is responsible for the highest per capita cost in Sweden. In the Netherlands and the U.S., the per capita cost of LBP ranks second only to diabetes. Both diabetes and LBP carry growing cost burdens associated with sedentary lifestyles and rising obesity rates. The cost per capita of LBP and major chronic diseases is presented in Table 3.

DISCUSSION

1. Why Does LBP Cost so Much in the U.S.? A Healthcare Pricing Issue

The high cost of LBP in the U.S. is in proportion to its high national healthcare costs, not the population's health status. While the U.S. boasts the highest obesity rates, the prevalence of LBP in the U.S. is similar to other high-income countries.68

Table 3. Relative per capita cost of major chronic diseases in select high-income countries (US$2023)

<table>
<thead>
<tr>
<th>Country</th>
<th>LBP</th>
<th>CVD</th>
<th>CVD: LBP</th>
<th>Dementia</th>
<th>Dementia: LBP</th>
<th>Stroke</th>
<th>Stroke: LBP</th>
<th>Diabetes</th>
<th>Diabetes: LBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>$1,830.25 (Waters &amp; Graf6 2018)</td>
<td>$1,407.78 (Virani et al.85 2020)</td>
<td>$1,117.84 (Waters &amp; Graf6 2018)</td>
<td>0.769</td>
<td>0.611</td>
<td>$441.93 (Waters &amp; Graf6 2018)</td>
<td>0.241</td>
<td>$2,114.94 (Waters &amp; Graf6 2018)</td>
<td>1.168</td>
</tr>
<tr>
<td>Canada</td>
<td>$0.81* (Beaudet40 2013)</td>
<td>$1,059.57 (Tarride et al.89 2019)</td>
<td>NA†</td>
<td>$251.18 (Østbye et al.90 1994)</td>
<td>NA†</td>
<td>$199.94 (Goeree et al.91 2005)</td>
<td>NA†</td>
<td>$296.44 (Dawson et al.92 2002)</td>
<td>NA†</td>
</tr>
<tr>
<td>Netherlands</td>
<td>$330.71 (Lambeek et al.93 2011)</td>
<td>$0.61 (Wilkins et al.94 2017)</td>
<td>1.63 × 10⁻³</td>
<td>$307.72 (Koopmanschap et al.95 1998)</td>
<td>0.824</td>
<td>$180.57 (Struijs et al.96 2006)</td>
<td>0.484</td>
<td>$509.00 (Peters et al.97 2017)</td>
<td>1.390</td>
</tr>
<tr>
<td>Sweden</td>
<td>$379.79 (Ekman et al.95 2005)</td>
<td>$0.26 (Wilkins et al.94 2017)</td>
<td>6.06 × 10⁻⁴</td>
<td>$363.02 (Wimo et al.96 1997)</td>
<td>0.847</td>
<td>$341.32 (Terént et al.97 1994)</td>
<td>0.796</td>
<td>$157.38 (Henriksson et al.98 1998)</td>
<td>0.367</td>
</tr>
<tr>
<td>Korea</td>
<td>$147.28 (Lee et al.76 2019)</td>
<td>$157.12 (Chang et al.102 2012)</td>
<td>0.945</td>
<td>$85.79 (Suh et al.102 2006)</td>
<td>0.516</td>
<td>$172.70 (Cha103 2018)</td>
<td>1.040</td>
<td>$422.95 (Oh et al.104 2021)</td>
<td>2.54</td>
</tr>
<tr>
<td>Japan</td>
<td>$318.50 (Montgomery et al.53 2017)</td>
<td>$157.54* (Gochi et al.105 2018)</td>
<td>0.438</td>
<td>$1,336.91 (Sado et al.106 2018)</td>
<td>3.720</td>
<td>$0.12 (Urakami et al.107 2019)</td>
<td>3.45 × 10⁻⁴</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

LBP, low back pain; CVD, cardiovascular disease.

*Direct cost per person. No total cost of LBP in Canada. †Not reported as there was no available total of LBP in Canada. ¶Includes only the total cost of ischemic heart disease.

https://doi.org/10.14245/ns.2448372.186
Furthermore, Papanicolas et al.\textsuperscript{69} found that population health factors at large (smoking, drinking, obesity) were not responsible for the substantially higher healthcare costs of the U.S. Instead, the high price of healthcare, in particular physician and hospital services, pharmaceuticals, and diagnostic testing in the U.S. drives the high cost of chronic disease.\textsuperscript{70} Across surgical specialties ranging from obstetrics, general surgery, and orthopedic surgery, surgeries in the U.S. are more expensive and thus more lucrative than other comparable countries. For example, the cost of a knee replacement is 53% more expensive in the U.S. than Switzerland and 77% more expensive than Australia.\textsuperscript{71} The higher rate and revenue of performing spine surgery and other high-margin procedures like caesarean deliveries and angioplasties account for a fifth of the difference in healthcare cost per capita between the U.S. and other high-income countries.\textsuperscript{72}

Another possible cause of the U.S.’ disproportionate spending on LBP is its well-known litigious nature, which may predispose to overutilization of indisputable clinical evidence such as imaging. The U.S. performs many more CT scans (278.5 per 1,000 people, in 2019) than any other country. Iceland, which ranks second, performs 234 CT scans per 1,000 people and Korea, third, 228 per 1,000 people.\textsuperscript{73} The price of scans is also higher in the U.S. with a the nearly 10-fold difference in CT per capita cost between the U.S. and the Netherlands ($220 vs. $23, respectively).\textsuperscript{72} Emanuel et al. notes that 7% of the cost difference between the U.S. and Netherlands is due to imaging.\textsuperscript{72}

### 2. The Future of LBP in the U.S.

The fastest growing segment of the U.S. population are people aged 60 years and older, from 962 million 2017 to 2.1 billion

#### Table 4. International evidence-based guidelines on the management of LBP\textsuperscript{74-82}

<table>
<thead>
<tr>
<th>Level of treatment</th>
<th>Acute LBP</th>
<th>Chronic LBP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-line</strong></td>
<td>Remaining active</td>
<td>Remaining active</td>
</tr>
<tr>
<td></td>
<td>Education/reassurance</td>
<td>Education/reassurance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive behavioral therapy</td>
</tr>
<tr>
<td><strong>Second-line</strong></td>
<td>Spinal manipulation</td>
<td>Spinal manipulation</td>
</tr>
<tr>
<td></td>
<td>Massage</td>
<td>Massage</td>
</tr>
<tr>
<td></td>
<td>Acupuncture</td>
<td>Acupuncture</td>
</tr>
<tr>
<td></td>
<td>NSAIDs</td>
<td>NSAIDs</td>
</tr>
<tr>
<td></td>
<td>Superficial heat</td>
<td>Acupuncture</td>
</tr>
<tr>
<td><strong>Limited use where indicated</strong></td>
<td>Exercise therapy</td>
<td>Yoga</td>
</tr>
<tr>
<td></td>
<td>Cognitive behavioral therapy</td>
<td>Midfulness-based exercises</td>
</tr>
<tr>
<td></td>
<td>Skeleton muscle relaxants</td>
<td>Interdisciplinary rehabilitation</td>
</tr>
<tr>
<td></td>
<td>Opioids*</td>
<td>NSAIDs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SNRI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Surgery</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Discectomy for herniated disc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Laminectomy for spinal stenosis</td>
</tr>
<tr>
<td><strong>Insufficient evidence</strong></td>
<td>Yoga</td>
<td>Superficial heat</td>
</tr>
<tr>
<td></td>
<td>Mindfulness-based exercises</td>
<td>Skeletal muscle relaxants</td>
</tr>
<tr>
<td></td>
<td>Interdisciplinary rehabilitation</td>
<td>Antiseizure medications</td>
</tr>
<tr>
<td></td>
<td>SNRI</td>
<td>Surgery</td>
</tr>
<tr>
<td></td>
<td>Antiseizure medications</td>
<td>(spinal fusion for nonradicular LBP with degenerative disc findings)</td>
</tr>
<tr>
<td></td>
<td>Surgery</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Discectomy for herniated disc</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Laminectomy for spinal stenosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Spinal fusion for non-radicular LBP with degenerative disc findings</td>
<td></td>
</tr>
<tr>
<td><strong>Not recommended</strong></td>
<td>Epidural glucocorticoid injection for herniated disc</td>
<td>Paracetamol</td>
</tr>
<tr>
<td></td>
<td>Systemic glucocorticoids</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paracetamol</td>
<td></td>
</tr>
</tbody>
</table>

LBP, low back pain; NSAID, nonsteroidal anti-inflammatory drugs; SNRI, selective norepinephrine reuptake inhibitor.

*Use with caution.
in 2050. In the U.S., 2 out of every 3 adult male over 60 years old reports having LBP in the past year.\textsuperscript{74} A critical subdivision of the elderly population are those older than 65, a population particularly prone to the complications of LBP (depression, falls, etc.). Superaging populations like Japan where those over 65 years old outnumber those under 18 face the economic crisis of a simultaneously decreasing labor force and increasing public sector demands on health care.\textsuperscript{75,76} By 2034, the U.S. too is projected to become a superaging population.\textsuperscript{77}

Considering these impending demographic challenges, adherence to evidence-based management of LBP can help safeguard from wasteful healthcare spending. A 2018 Lancet series highlights global recommendations on the management of acute and chronic LBP summarized in Table 4.\textsuperscript{17,28,30,78-80} A seriously underutilized tool—patient education and reassurance—is the first line therapy for both acute and chronic LBP.

**CONCLUSION**

The cost of LBP will continue to rise in the U.S. and other high-income countries largely due to an aging population becoming an ever-greater public budget strain. This urges discernment of the cost-contributors and inefficiencies in the clinical and health system-wide management of chronic LBP. Respecting guidelines for imaging and surgical management and cautious referrals to specialists for the first visit would be reasonable initial approaches to managing a complex biopsychosocial issue like chronic pain.

**NOTES**

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Author Contribution:** Conceptualization: MMS, CA; Formal analysis: DC, AL; Investigation: DC, AL, AM; Methodology: DC, MMS, CA; Project administration: DC, MMS; Writing – original draft: DC, AL; Writing – review & editing: DC, AL, MMS, CA, HA.

**ORCID**

Diana Chang: 0000-0002-0942-1675  
Austin Lui: 0000-0003-1347-8815  
Alisa Matsoyan: 0009-0007-7171-6072  
Michael M. Safaee: 0000-0003-2618-3098  
Christopher Ames: 0000-0001-8754-2512

REFERENCES

13. Hansson TH, Hansson EK. The effects of common medical interventions on pain, back function, and work resump-
68. ProCon.org. Global obesity levels [Internet]. Britannica;
Cost of LBP in US and Globally

Chang D, et al.


103. Cha YJ. The economic burden of stroke based on South Korea’s National Health Insurance Claims Database. Int J Heal Policy Manag 2018;7:904.


Radiographic and Clinical Outcomes of Transverse Process Hook Placement at the Proximal Thoracic Upper Instrumented Vertebra in Adult Spinal Deformity Surgery

Sang Hun Lee, Micheal Raad, Andrew H. Kim, David B. Cohen, Khaled M. Kebaish
Department of Orthopaedic Surgery, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

Objective: Few studies have reported radiographic and clinical outcomes of transverse process hook (TPH) placement at the proximal thoracic upper instrumented vertebra (UIV) in adult spinal deformity (ASD) surgery. This study aims to investigate radiographic and clinical outcomes of TPH placement at the UIV for ASD surgery.

Methods: This is a retrospective cohort of 56 patients with ASD (age, 59 ± 13 years; follow-up, 44 ± 19 months) from Johns Hopkins Hospital, who underwent long posterior spinal fusion to the proximal thoracic spine (T2–5). Visual analogue scale (VAS) for back pain, Oswestry Disability Index (ODI), 36-item Short Form health survey scores, thoracic kyphosis (TK), lumbar lordosis, sacral slope, pelvic tilt, pelvic incidence, proximal junctional kyphosis (PJK) angle, PJK incidence, pattern of PJK, grades of TPH dislodgement, revision surgery, and factors associated with high-grade TPH dislodgement were analyzed.

Results: VAS for back pain and ODI values improved significantly from preoperatively to final follow-up. Mean change in PJK angle was 12° (range, 0.5°–43°). Twenty patients (36%) developed PJK, of whom 13 had compression fractures at 1 vertebra distal to the UIV (UIV–1). Final TPH position was stable in 42 patients (75%). In most patients (86%), TPH dislodgement did not progress after 6-month postoperative follow-up. Three patients (5.3%) underwent revision surgery to extend the fusion because of symptomatic PJK. Unstable TPH position was associated only with revision surgery and TK.

Conclusion: TPH placement at the proximal thoracic UIV for long fusion showed favorable clinical and radiographic outcomes in terms of the incidence of PJK and mean PJK angle at mean 44-month follow-up. TPHs placed in the proximal thoracic UIV were in stable position in 75% of patients. Compression fracture at UIV–1 was the most common pattern of PJK. PJK angle progression was greater in revision cases and in patients with greater preoperative thoracic kyphosis.

Keywords: Adult spinal deformity, Kyphosis, Proximal junctional kyphosis, Proximal thoracic spine, Sagittal imbalance, Scoliosis, Transverse process hook

INTRODUCTION

Proximal junctional kyphosis (PJK) is one of the most common complications after adult spinal deformity (ASD) surgery. The reported incidence ranges from 25% to 66%, varying according to diagnostic criteria, fusion levels, and duration of follow-up, among other factors.1-6 PJK has adverse effects on clinical outcomes, and revision surgery to extend the fusion is often needed for patients with symptomatic PJK.

Numerous studies have investigated risk factors for PJK.1-9
Radiographic risk factors include greater preoperative thoracic kyphosis, excessive correction of sagittal vertical axis (SVA) and lumbar lordosis, greater preoperative pelvic incidence, and inadequate restoration of sagittal balance (e.g., pelvic incidence – lumbar lordosis mismatch). Research suggests that several factors may affect the risk of developing PJK, including UIV fixation method, the number of spinal levels fused, the use of pelvic fixation, patient age, bone mineral density, smoking status, and others.\(^1\)\(^-\)\(^9\)

Several studies comparing pedicle screws and transverse process hooks (TPHs) as the UIV fixation method noted a significantly lower incidence of PJK in adult and pediatric spinal deformity patients with TPHs.\(^3\)\(^,\)\(^6\)\(^,\)\(^10\) Biomechanical cadaver studies have also shown that TPHs reduce stress at the junction between the UIV and proximal segments and allow a more gradual transition of segmental motion compared with pedicle screws in the UIV.\(^1\)\(^1\)\(^-\)\(^1\)\(^3\)

Although the use of TPHs at the UIV is becoming more common, little is known about the morphological features, clinical outcomes, and prognostic factors associated with this technique. Therefore, we investigated radiographic and clinical outcomes of patients with ASD after TPH placement at the proximal thoracic UIV for long posterior spinal fusion.

**MATERIALS AND METHODS**

**1. Patient Selection**

We retrospectively analyzed data from patients at Johns Hopkins Hospital who underwent long posterior spinal fusion of the proximal thoracic spine that used TPHs at the UIV and pedicle screws distal to the UIV. All surgical procedures were performed by 2 coauthors from 2008 to 2014.

Inclusion criteria were as follows: aged ≥ 20 years at the time of surgery; primary diagnosis of spinal deformity (scoliosis, kyphosis, or kyphoscoliosis) or PJK treated with instrumentation involving a proximal fusion level between T2 and T5; and minimum 2-year clinical and radiographic follow-up. Of 86 patients, we excluded 30 patients for the following reasons: insufficient follow-up (n = 24), Scheuermann kyphosis (n = 3), ankylosing spondylitis (n = 1), traumatic condition (n = 1), and neoplastic condition (n = 1).

This study was approved by the Institutional Review Board (IRB) of Johns Hopkins Hospital (IRB No. 00135145).

**2. Surgical Technique**

The spine was exposed down to the distal end of the spinous processes, while preserving the interspinous ligaments at the proximal and distal segments, with minimal dissection of the paraspinal muscles. A blunt lamina finder device (DePuy Synthes Spine, Inc., Raynham, MA, USA) was used to prepare the insertion point on the transverse process, ensuring that the TPH blade was immediately lateral to the lateral edge of the pedicle. Care was taken to ensure proper sizing of the TPH so that it could latch to the entire transverse process without weakening or fracturing it.

**3. Clinical Outcome Data**

We assessed the following clinical outcome measures from patient medical records: visual analogue scale for back pain, Oswestry Disability Index (ODI), and 36-item Short Form health survey (SF-36) physical composite score and mental composite score. We compared changes in scores from the preoperative visit to latest follow-up.

**4. Radiographic Analyses**

We analyzed 36-inch standing scoliosis radiographs taken at the preoperative visit, 6-week follow-up, and latest follow-up. In addition to PJK angle, sagittal measurements were thoracic kyphosis (T4–12 angle), lumbar lordosis (L1–S1 angle), sacral slope, pelvic tilt, and pelvic incidence. To assess global sagittal spinal alignment, we measured SVA (C7–S1).

We used the following criteria from the study by Glattes et al.\(^2\) to define PJK: (1) presence of a PJK angle (proximal junction sagittal angle of ≥ 10° between the lower endplate of the UIV and the upper endplate of the vertebrae 2 levels proximal to the UIV); and (2) progression of ≥ 10° in the PJK angle from the baseline measurement.

Using lateral radiographs, we assessed the pattern of PJK and the TPH position according to our proposed novel grading system (Fig. 1). Grade 0 indicates no dislodgement or unilateral, incomplete dislodgement of the TPH. Grade 1 indicates bilateral, incomplete dislodgement of the TPH. Grade 2 indicates unilateral or bilateral complete dislodgement of the TPH. Grades 0 and 1 were considered stable positions. We graded TPH position at postoperative follow-up of 6 weeks, 6 months, 12 months, and each subsequent year.

**5. Statistical Analyses**

We analyzed associations between TPH position and patient demographic characteristics and radiographic and clinical outcomes. We stratified patients by TPH position and compared age, sex, body mass index (BMI) value, preoperative diagnosis,
UIV level (T2 or T3 vs. T4 or T5), pelvic incidence, pelvic tilt, lumbar lordosis, thoracic kyphosis, and SVA. We used $\chi^2$ tests to compare categorical variables and analysis of variance to compare continuous variables. Twenty patients were selected randomly for agreement analysis and were reviewed by 2 independent reviewers regarding final position of the TPH. Interrater and intrarater agreement were assessed using the kappa statistic for agreement and the percentage agreement. Statistical analyses were performed using IBM SPSS Statistics ver. 22.0 (IBM Co., Armonk, NY, USA). A p-value < 0.05 was considered statistically significant.

RESULTS

1. Patient Characteristics

Fifty-six patients (43 women) were enrolled. Mean ± standard deviation patient age was 59 ± 13 years, and mean follow-up duration was 44 ± 19 months. The mean preoperative patient BMI value was 29 ± 6.4 kg/m$^2$. With regards to bone mineral density, the mean lowest preoperative dual-energy X-ray absorptiometry T score was -1.36 ± 1.22. Preoperative diagnoses were primary thoracolumbar kyphoscoliosis in 37 patients and revision for thoracolumbar PJK in 19. The mean number of spinal segments fused was 14 (range, 12–16). The UIV was T2 in 10 patients, T3 in 29 patients, T4 in 15 patients, and T5 in 2 patients (Table 1).

2. Surgical Procedures

Spinal osteotomies were performed in 45 patients (80%), consisting of multilevel posterior column osteotomies (n = 31), pedicle subtraction osteotomies (n = 4), vertebral column resections (n = 3), posterior column osteotomies with pedicle subtraction osteotomies (n = 2 patients), posterior column osteotomies with vertebral column resections (n = 4), and pedicle subtraction osteotomy with vertebral column resection (n = 1). Rod material was titanium in a majority of cases (n = 53), with 3 cases using cobalt chrome rods. Additional surgical procedures during the follow-up period were performed for 9 patients (16%): extension of fusion for new-onset PJK (n = 3) (Table 2), cervical spine surgery (n = 4), and revision surgery for rod fracture (n = 2).

3. Radiographic Outcomes

Thoracic kyphosis and C2–S1 SVA improved significantly from preoperatively to latest follow-up (Table 3). The mean change in PJK angle was 12° (range, 0.5°–43°) at latest follow-up, and radiographic PJK was found in 20 patients (36%). When assessing patterns of PJK, we found that 13 of 20 patients had compression fractures at 1 vertebra distal to the UIV (UIV–1).
Other patterns were multiple compression fractures (≥ 2 vertebrae) (n = 3), screw pullout at UIV–1 (n = 2), compression fracture at the UIV (n = 1), and compression fracture at 2 levels distal to the UIV (UIV–2) (n = 1) (Table 2).

Final TPH position was stable in 42 patients (75%) (grade 0 in 27 patients, grade 1 in 15 patients) and unstable (grade 2) in 14 patients. TPH position did not change after the 6-month postoperative assessment in 48 patients (86%), after the 1-year assessment in 7 patients (12%), and after the 2-year assessment in 1 patient.

### 4. Clinical Outcomes

VAS for back pain and ODI values improved significantly from preoperatively to final follow-up (Table 3). However, SF-36 scores did not change significantly between preoperatively and final follow-up. Nine patients required revision surgery following TPH fixation. Indications for revision included surgery for additional PJK (n = 3), surgery for additional cervical spine problems (n = 4), and revision for rod fracture (n = 2).

### 5. Associations Between TPH Position and Other Parameters

Only 2 parameters—revision surgery and greater preoperative thoracic kyphosis—were associated with unstable final TPH position (both, p < 0.001) (Table 4). We found no significant associations between final TPH position and patient sex, age, or BMI; level of the UIV; pelvic incidence; preoperative pelvic tilt, lumbar lordosis, or SVA; or correction of pelvic tilt, lumbar lordosis, thoracic kyphosis, or SVA.

### 6. Intrarater and Interrater Agreement

Based on data from 20 randomly selected patients, intrarater agreement was 80% (kappa = 0.70) and intrarater agreement was 90% (kappa = 0.85) regarding final position of TPH.
DISCUSSION

In addition to the incidence of PJK, we investigated radiographic features, clinical outcomes, and factors related to TPH placement at the proximal thoracic UIV for long posterior spinal fusion. TPHs placed at the UIV were in stable position (grade 0 or 1) in 75% of patients, and TPH position did not change after 6 months postoperatively in most patients. Unstable TPH position was more common in patients who underwent surgery for PJK after previous thoracolumbar fusion than among patients who underwent primary deformity surgery, as well as in patients who had greater preoperative thoracic kyphosis. The incidence of PJK was 36%, and most cases of PJK consisted of compression fracture at UIV–1. Revision surgery for new-onset PJK was performed in 3 patients (5.3%) during a mean follow-up period of 44 months. Interrater and intrarater agreement when determining the final position of the TPH were 80% and 90%, respectively.

The effect of the load distribution for UIV fixation on the risk of PJK has been investigated in clinical and biomechanical studies.8-13 Kim et al.8 reported a lower incidence of PJK when using TPHs versus pedicle screws in pediatric patients with scoliosis. Those results were similar to findings of a subsequent study by Helgeson et al.,9 demonstrating a 5.6° change in the screw group compared with a 1.4° change in the TPH group. In 2013, Hassanzadeh et al.10 compared the use of TPHs versus pedicle screws for UIV fixation in 47 patients with ASD. The authors reported a significantly lower incidence of PJK in the TPH group (0 of 20) compared with the screw group (8 of 27) (p = 0.01) at a mean follow-up of 2.8 years.

Biomechanical research supports clinical findings of the superiority of TPH to pedicle screws.11-13 The stiffness of constructs using TPHs was significantly lower than that of constructs using pedicle screws in porcine and cadaveric spines.11-13 Moreover, TPH constructs maintained a pattern of monotonic increase in mean range of motion from distal to proximal and showed lower supra-adjacent hypermobility compared with UIV pedicle screw constructs, which had the greatest mean range of motion at the first uninstrumented segment.11-13

We are aware of no previous studies that have assessed TPH position and its change over time. Although nearly 50% of our cohort had stable, grade 1 TPH position, the remaining patients had a variable degree of dislodgement. This dislodgement may be similar to the loosening of pedicle screws at the UIV, which is typically associated with pseudarthrosis at the affected level. However, unlike pedicle screw loosening, TPH dislodgement may represent natural repositioning during follow-up and may provide the construct with a transitional level of motion from the proximal unfused segment to the distal fused segments. In most patients, TPH position did not change after 6-month follow-up, which may suggest that the adaptation period of TPHs is approximately 6 months. The low incidence of revision sur-

Table 3. Radiographic and clinical parameters for 56 patients who underwent long posterior spinal fusion for adult spinal deformity, 2008–2014

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preoperative</th>
<th>Immediate postoperative</th>
<th>Follow-up*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiographic parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7–S1 SVA (mm)</td>
<td>40 ± 69</td>
<td>15 ± 31</td>
<td>19 ± 43</td>
<td>0.038</td>
</tr>
<tr>
<td>Lumbar lordosis (°)</td>
<td>41 ± 24</td>
<td>48 ± 11</td>
<td>45 ± 14</td>
<td>0.271</td>
</tr>
<tr>
<td>Pelvic incidence (°)</td>
<td>57 ± 15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelvic tilt (°)</td>
<td>28 ± 14</td>
<td>26 ± 12</td>
<td>26 ± 14</td>
<td>0.512</td>
</tr>
<tr>
<td>Sacral slope (°)</td>
<td>30 ± 16</td>
<td>32 ± 11</td>
<td>31 ± 11</td>
<td>0.644</td>
</tr>
<tr>
<td>Thoracic kyphosis (°)</td>
<td>52 ± 19</td>
<td>46 ± 16</td>
<td>47 ± 16</td>
<td>0.040</td>
</tr>
<tr>
<td>Clinical parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODI</td>
<td>52 ± 23</td>
<td></td>
<td>27 ± 21</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SF-36-MCS</td>
<td>41 ± 14</td>
<td></td>
<td>44 ± 15</td>
<td>0.550</td>
</tr>
<tr>
<td>SF-36-PCS</td>
<td>36 ± 7.7</td>
<td></td>
<td>41 ± 15</td>
<td>0.140</td>
</tr>
<tr>
<td>VAS for back pain</td>
<td>5.5 ± 2.5</td>
<td></td>
<td>3.7 ± 2.2</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.
SVA, sagittal vertical axis; ODI, Oswestry Disability Index; SF-36, 36-item Short Form health survey; MCS, mental composite score; PCS, physical composite score; VAS, visual analogue scale.
*Mean ± standard deviation follow-up was 44 ± 19 months.
Table 4. Characteristics of 56 patients with adult spinal deformity by final* proximal transverse hook position

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Position of proximal transverse hook</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grade 0 (n = 26)</td>
<td>Grade 1 (n = 16)</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>57 ± 15</td>
<td>60 ± 13</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td>13</td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26 ± 4.8</td>
<td>29 ± 8.6</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary deformity</td>
<td>23</td>
<td>8</td>
</tr>
<tr>
<td>Revision for PJK</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>UIV level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2 or T3</td>
<td>19</td>
<td>13</td>
</tr>
<tr>
<td>T4 or T5</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Follow-up duration (mo)</td>
<td>44 ± 16</td>
<td>48 ± 24</td>
</tr>
<tr>
<td>Radiographic parameters (°)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumbar lordosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>38 ± 28</td>
<td>46 ± 19</td>
</tr>
<tr>
<td>Change</td>
<td>11 ± 26</td>
<td>1.9 ± 17</td>
</tr>
<tr>
<td>Pelvic incidence</td>
<td>61 ± 12</td>
<td>54 ± 19</td>
</tr>
<tr>
<td>Pelvic tilt</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>31 ± 12</td>
<td>25 ± 20</td>
</tr>
<tr>
<td>Change</td>
<td>4.1 ± 13</td>
<td>0.4 ± 15</td>
</tr>
<tr>
<td>Sagittal vertical axis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>35 ± 62</td>
<td>41 ± 59</td>
</tr>
<tr>
<td>Change</td>
<td>24 ± 48</td>
<td>25 ± 57</td>
</tr>
<tr>
<td>Thoracic kyphosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>46 ± 20</td>
<td>56 ± 16</td>
</tr>
<tr>
<td>Change</td>
<td>3.1 ± 17</td>
<td>8.7 ± 14</td>
</tr>
</tbody>
</table>

Grade 0, no dislodgement or unilateral, incomplete dislodgement of the transverse process hook (TPH); grade 1, bilateral, incomplete dislodgement of the TPH; grade 2, unilateral or bilateral complete dislodgement of the TPH. Grades 0 and 1 were considered stable position. SD, standard deviation; BMI, body mass index; PJK, proximal junctional kyphosis; UIV, upper instrumented vertebra.

*Mean ± standard deviation follow-up was 44 ± 19 months.

Surgery for new-onset PJK (5.3%) and a mean change in PJK angle of 12° may further support this theory.

The 36% incidence of PJK in our cohort is similar to that reported in previous case series. However, if we were to use a different definition of PJK, based on a threshold of 20° rather than 10°, the incidence in our cohort would be 18% (10 of 56 patients), which is comparable to the findings of previous studies using that criterion.

The indication for surgery may be associated with the progression of PJK angle and final TPH position. Patients who underwent surgery for PJK after a previous fusion had a higher incidence of unstable TPH position and PJK angle progression than patients who underwent primary thoracolumbar fusion. Various factors, including bone and soft tissue conditions and tendencies of some patients to develop a stooping posture, may contribute to this finding.

When analyzing the patterns of PJK after TPH UIV fixation, we found that 65% of patients had a compression fracture at UIV-1. This finding could be explained by the fact that TPHs do not stabilize the anterior spinal column. If further mechani-
cal augmentation, such as preventive vertebroplasty, were to be provided at UIV–1, a substantial proportion of cases of PJK after this procedure may be potentially prevented.

This study has several limitations. First, we lacked information regarding fusion status from flexion-extension radiographs or computed tomography. Second, 24 of 80 otherwise eligible patients had less than 2-year follow-up and were excluded from analysis. Third, this is a retrospective case series that is subject to all limitations inherent in such a design, as well as selection bias. Lastly, this study was performed at a tertiary care academic center with a unique patient population. Such findings should be interpreted with caution when extrapolating to different clinical setting. However, we believe this is the first study to describe the detailed morphological features of TPH placement for UIV fixation at the proximal thoracic spine in patients with ASD. Our study underscores the need for future comparative studies to examine differences in clinical outcomes by TPH position.

CONCLUSION

TPH placement at the proximal thoracic UIV for long fusion showed favorable clinical and radiographic outcomes in terms of the incidence of PJK and mean PJK angle. In 75% of patients treated with long posterior spinal fusion for ASD, TPHs placed in the proximal thoracic spine at the UIV were in a stable position at minimum 2-year follow-up. In most patients, TPH position did not change after the first 6 months postoperatively. Compression fracture at UIV–1 was the most common pattern of PJK. PJK angle progressed significantly more in patients with greater preoperative thoracic kyphosis and in those with PJK after previous thoracolumbar fusion.

NOTES

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Author Contribution:** Conceptualization: SHL, MR; Data curation: SHL; Formal analysis: SHL, MR, AHK; Methodology: SHL, MR; Project administration: MR; Visualization: MR; Writing – original draft: SHL, MR; Writing – review & editing: SHL, MR, AHK, DBC, KMK.

**ORCID**

Sang Hun Lee: 0000-0001-5942-8309

Micheal Raad: 0000-0001-8167-5808

Andrew H. Kim: 0000-0001-8888-8590

David B. Cohen: 0000-0002-9793-0217

Khaled M. Kebaish: 0000-0002-1235-6202

REFERENCES

10. Hassanzadeh H, Gupta S, Jain A, et al. Type of anchor at the proximal fusion level has a significant effect on the incidence of proximal junctional kyphosis and outcome in adults after...
Magnetic Resonance Imaging-Related Anatomic and Functional Parameters for the Diagnosis and Prognosis of Chiari Malformation Type I: A Systematic Review and Meta-analysis

Zairan Wang¹, Zhimin Li¹, Shiyuan Han¹, Xianghui Hu¹, Siyuan Pang¹, Yongning Li¹,², Jun Gao¹

¹Department of Neurosurgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China
²Department of International Medical Services, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China

Objective: Imaging parameters of Chiari malformation type I (CMI) development are not well established. This study aimed to collect evidence of general or specific imaging measurements in patients with CMI, analyze indicators that may assist in determining the severity of CMI, and guide its diagnosis and treatment.

Methods: A comprehensive search was conducted across various databases including the Cochrane Library, PubMed, MEDLINE, Scopus, and Embase, covering the period from January 2002 to October 2023, following predefined inclusion criteria. Meta-analyses were performed using RevMan (ver. 5.4). We performed a quantitative summary and systematic analysis of the included studies. This study was registered in the PROSPERO (International Prospective Register of Systematic Reviews) prior to initiation (CRD42023415454).

Results: Thirty-three studies met our inclusion criteria. The findings indicated that out of the 14 parameters examined, 6 (clivus length, basal angle, Boogard's angle, supraocciput lengths, posterior cranial fossa [PCF] height, and volume) exhibited significant differences between the CMI group and the control group. Furthermore, apart from certain anatomic parameters that hold prognostic value for CMI, functional parameters like tonsillar movement, obex displacement, and cerebrospinal fluid dynamics serve as valuable indicators for guiding the clinical management of the disease.

Conclusion: We collated and established a set of linear, angular, and area measurements deemed essential for diagnosing CMI. However, more indicators can only be analyzed descriptively for various reasons, particularly in prognostic prediction. We posit that the systematic assessment of patients’ PCF morphology, volume, and other parameters at a 3-dimensional level holds promising clinical application prospects.

Keywords: Chiari malformation, MRI-related parameters, Posterior cranial fossa, Meta-analysis, Systematic review

INTRODUCTION

Chiari malformation type I (CMI) is a neuroanatomical abnormality, in which the cerebellar tonsils extend up to 5 mm below the foramen magnum (FM). Patients with CMI suffer from various pain, sensory or motor deficits, and cognitive dysfunction. The primary diagnostic tool for this disorder is magnetic resonance imaging (MRI). With the development of med-
MRI Parameters Related to CMI


1. Materials and Methods

The meta-analysis and systematic review were registered on the PROSPERO (International Prospective Register of Systematic Reviews) with the registration number CRD42023415454.

1. Data Sources and Search Strategy

We systematically searched published and unpublished literature between 2002 (January) and 2023 (October) from Cochrane Library, PubMed, MEDLINE, Scopus, and Embase databases. We concatenated terms and phrases using appropriate Boolean operators. Retrieval included (“CM” OR “Chiari malformation” OR “Arnold-Chiari Malformation”) and (“MRI” OR “Magnetic Resonance Imaging” OR “Magnetic Resonance”). Language restrictions were not applied. We manually searched reference lists of all articles and gray literature to identify potentially eligible studies.

2. Inclusion and Exclusion Criteria

Studies considered for meta-analysis and systematic evaluation met the following inclusion criteria: (1) simple CMI without congenital cranio-cervical junction malformations; (2) studies that examined children or adults who had a clinical diagnosis of CMI; (3) complete data records through prospective or retrospective studies; and (4) studies published in peer-reviewed journals. Adult participants were those aged ≥ 18 years. The exclusion criteria were studies without effect sizes, unpublished duplicate publications and conference abstracts, and studies with incomplete data.

3. Risk of Bias and Analysis Plan

Since most of the included studies were nonrandom comparative studies, the risk of study bias was assessed using the Newcastle-Ottawa Scale, and the preferred Cochrane tool for non-randomized studies (Risk Of Bias In Non-randomized Studies - of Interventions, ROBINS-I). Each study was rated based on the scores of the Newcastle-Ottawa Scale obtained. A maximum score of 9 was set for 8 items. If a study scored >6, it was considered to have a good quality. Using the ROBINS-I tool, a study’s risk of bias is evaluated by considering confounding, participants’ selection, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of reported results. The risk assessment for each bias category can be classified as low, moderate, serious, or critical. The overall bias risk of a study was determined by the highest risk identified within each category. Two reviewers separately evaluated study quality; any differences will be resolved by discussion or consensus with the third reviewer.

4. Study Selection and Data Collection Process

All titles and abstracts identified from the search process were independently assessed by 2 reviewers (ZW and SH). Instances of disagreement were resolved by discussion between the 2 reviewers or by introduction of a third reviewer (ZL). All articles were summarized in detail, and data from full-text articles that met the study benchmarks were incorporated, including subjects’ demographics and narrative summaries of outcomes and methods. Summaries were performed using predesigned standardized tables. If there was a lack of relevant data in the included studies, we request data by contacting the corresponding author of the paper.

5. Quality Assessment

This study included reports with a control group or pre- and postoperative comparisons; there were no systematic differences between the groups. All studies were approved by the ethics committee and conducted with patients’ informed consent. Only studies that met the quality criteria were included.

6. Quality of the Evidence

A modified Recommendations Assessment, Development, and Evaluation grading (GRADE) method was used to categorized the quality of evidence as high, moderate, low, or very low. Because publication bias is difficult to assess in observational studies, only 4 factors were assessed in our meta-analysis: risk of bias, inconsistency, indirectness, and imprecision. When there
was a risk for a factor results, the quality of evidence for the related factor correspondingly decreased by 1 or 2 grades. The level of evidence and the strength of the recommendation were determined through discussion by all members of the research group.

7. Data Analysis and Systematic Review

Meta-analyses were performed using RevMan (ver. 5.4, Cochrane Collaboration, Oxford, UK) and IBM SPSS Statistics ver. 19.0 (IBM Co., Armonk, NY, USA). The primary outcome was the mean (M) and standard deviation of the objective assessment measures for subjects. Meta-analyses of mean differences (MDs) were expressed as 95% confidence intervals (CIs). The heterogeneity of results was estimated using I², Z, and chi-square tests (p < 0.05). Risk of publication bias was assessed by examining funnel plots for symmetry and summarizing and analyzing data that could not be combined. A leave-one-out sensitivity analysis was used to test the robustness of the results. Studies in which data could not be combined adopted a systematic narrative approach.

RESULTS

1. Search Results

The initial search identified 3,411 studies. A total of 1,832 studies remained after removing duplicate studies. After selecting titles and abstracts and browsing the complete text, a total of 33 studies which met the requirements and had complete data were finally determined. A total of 17 studies (2,097 patients with CMI and 1,055 controls) and 16 studies (involving 1,270 patients with CMI) were included to evaluate the role of imaging parameters in disease diagnosis and explore the correlation between imaging parameters and postoperative prognosis, respectively. Two of these studies were by the same author. Therefore, in order to reduce the possibility of duplication of data, only their most recent study was included in the meta-analysis to reduce bias. Among 1,055 individuals in the control group, 377 were recruited healthy volunteers and 678 were patients with no obvious posterior fossa pathology or medical problems; therefore, they were inferred to reflect PCF morphology of the normal population. A flow chart of the search is shown in Fig. 1.

2. Quality Assessment of the Studies

The authors used meta-analysis for the role of imaging parameters in diagnosing disease. According to the Newcastle-Ottawa Scale, 7 studies scored 8 points, while 10 studies scored 7 points, indicating relatively high quality of each study. Table 1 showed the characteristics and quality scores of the included meta-analysis studies. Supplementary Table 1 showed the risk of bias assessments of the included studies using the ROBINS-I tool. One study had a critical risk of bias due to missing data. Seven studies scored moderately biased mainly due to bias in confounding or measurement of outcomes. The quality of evi-

---

Fig. 1. Flow chart of study through different phases of the meta-analysis and systematic review.
dence based on GRADE was downgraded to moderate (imprecision or inconsistency) or low (imprecision and inconsistency) for these findings of the basal angle, Boogard’s angle, supraoc-ciput lengths, PCF height, and PCF volume, respectively. The detailed data for GRADE scales were presented in Table 2. In contrast, discussing the correlation between imaging parameters and postoperative prognosis used a systematic narrative approach rather than meta-analysis. The reasons for the systematic narrative approach were (1) methodological differences between studies were too great to allow for straightforward comparisons, (2) inconsistencies in the data used to calculate effect sizes, (3) inconsistencies in the data for potential confounders, and (4) heterogeneity of the data used in previous studies, which did not allow for the calculation of moderating effects.

3. Meta-Analysis Results of MRI-Related Parameters in Patients With CMI

Numerous MRI-related parameters are used to assess pa-

Table 1. Key characteristics and quality scores of the included studies regarding the value of MRI-related parameters to diagnose CMI

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>No. patients with CMI/controls</th>
<th>Mean age (yr)</th>
<th>Included quantitative parameters</th>
<th>Quality score (NOS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karagöz et al. 2002</td>
<td>Turkey</td>
<td>22/21</td>
<td>32.9 ± 12.0</td>
<td>Basal angle, Boogard’s angle, tentorium angle, the slope of the tentorium, PCF area</td>
<td>7</td>
</tr>
<tr>
<td>Milhorat et al. 2010</td>
<td>USA</td>
<td>388/80</td>
<td>33.6 ± 10.1</td>
<td>Tentorium angle, clivus length, FMAP, PCFV and PCFBV</td>
<td>7</td>
</tr>
<tr>
<td>Alperin et al. 2014</td>
<td>USA</td>
<td>36/37</td>
<td>37.0 ± 11.0</td>
<td>Clivus length, supraoc-ciput lengths, PCFV</td>
<td>8</td>
</tr>
<tr>
<td>Aydin et al. 2005</td>
<td>Turkey</td>
<td>60/30</td>
<td>35.1 ± 12.7</td>
<td>Clivus length, supraoc-ciput lengths, FMAP, AP diameter of PCF, PCF height</td>
<td>7</td>
</tr>
<tr>
<td>Urbizu et al. 2014</td>
<td>Spain</td>
<td>100/50</td>
<td>45.5 ± 12.2</td>
<td>Basal angle, Wackenheim angle, the slope of the tentorium, clivus length, supraoc-ciput lengths, FMAP, AP diameter of PCF, PCF height</td>
<td>7</td>
</tr>
<tr>
<td>Krishna et al. 2016</td>
<td>Canada</td>
<td>8/16</td>
<td>42.6 ± 10.4</td>
<td>Tentorium angle, PCFV</td>
<td>7</td>
</tr>
<tr>
<td>Milhorat et al. 2009</td>
<td>USA</td>
<td>280/75</td>
<td>33.7 ± 10.4</td>
<td>Tentorium angle, clivus length, FMAP, PCFV, and PCFBV</td>
<td>8</td>
</tr>
<tr>
<td>Dunton et al. 2011</td>
<td>Canada</td>
<td>81/107</td>
<td>42.6 ± 13.0</td>
<td>Boogard’s angle, clivus length</td>
<td>7</td>
</tr>
<tr>
<td>Heiss et al. 2012</td>
<td>USA</td>
<td>48/18</td>
<td>36.8 ± 11.2</td>
<td>Clivus length, supraoc-ciput lengths</td>
<td>7</td>
</tr>
<tr>
<td>Houston et al. 2018</td>
<td>USA</td>
<td>162/140</td>
<td>38.3 ± 10.0</td>
<td>Basal angle, Wackenheim angle, Boogard’s angle, odontoid angle, clivus length, supraoc-ciput lengths, AP diameter of PCF, PCF height</td>
<td>7</td>
</tr>
<tr>
<td>Nair and Rajshekhar 2022</td>
<td>India</td>
<td>27/10</td>
<td>&lt; 18</td>
<td>Boogard’s angle, tentorium angle, the slope of the tentorium, clivus length, supraoc-ciput lengths, FMAP, AP diameter of PCF, PCF height</td>
<td>7</td>
</tr>
<tr>
<td>Nishikawa et al. 2022</td>
<td>Japan</td>
<td>50/20</td>
<td>4–7†</td>
<td>Clivus length, supraoc-ciput lengths, PCFV, and PCFBV</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65/24</td>
<td>8–11†</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>70/23</td>
<td>12–15†</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>32/25</td>
<td>16–19†</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>230/58</td>
<td>20–49†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang et al. 2014</td>
<td>China</td>
<td>52/17</td>
<td>37.0 (22–59)†</td>
<td>Tentorium angle, clivus length</td>
<td>8</td>
</tr>
<tr>
<td>Yan et al. 2016</td>
<td>China</td>
<td>67/40</td>
<td>15.3 ± 3.9</td>
<td>Boogard’s angle, clivus length, supraoc-ciput lengths, FMAP, AP diameter of PCF</td>
<td>8</td>
</tr>
<tr>
<td>Yüksel et al. 2022</td>
<td>Turkey</td>
<td>70/69</td>
<td>38.5 (17–70)†</td>
<td>PCF area</td>
<td>8</td>
</tr>
<tr>
<td>Tubbs et al. 2003</td>
<td>USA</td>
<td>100/50</td>
<td>9.0 (&lt; 18)</td>
<td>Odontoid angle</td>
<td>8</td>
</tr>
<tr>
<td>Besacho et al. 2015</td>
<td>USA</td>
<td>55/125</td>
<td>34.0 ± 10.2</td>
<td>Wackenheim angle, odontoid angle</td>
<td>7</td>
</tr>
</tbody>
</table>

MRI, magnetic resonance imaging; CMI, Chiari malformation type I; NOS, Newcastle-Ottawa Scale; PCF, posterior cranial fossa; FMAP, anteroposterior diameter of the foramen magnum; PCFV, posterior cranial fossa volume; PCFBV, posterior cranial fossa brain volume; AP, anteroposterior.

†Range. †Mean (range).
tients with CMI. After pooling, a total of 14 indicators were included in the meta-analysis (Fig. 2, Supplementary Fig. 1).

Two studies\textsuperscript{17,18} grouped different age populations, and we included each age group as a separate sample for analysis. One study\textsuperscript{18} classified CMI into 3 subtypes (type A, normal PCF volume [PCFV] and occipital bone size; type B, normal PCFV and small volume of the area surrounding the FM [VAFM] and occipital bone size; and type C, small VAFM, PCFV, and occipital bone size). We combined these 3 datasets to exclude bias caused by screening data from this study. The results showed that 6 (clivus length, basal angle, Boogard's angle, supraocciput lengths, PCF height, and PCFV) of the 14 parameters were significantly different in the CMI group compared with the control group.

1) Clivus length

A combined analysis of this parameter included data from 16 groups,\textsuperscript{8,11,14-20} 1,543 patients with CMI and 699 controls, and showed a significant difference in clivus length between the groups (MD = -1.14; 95% CI, -1.51 to -0.77; I\textsuperscript{2} = 92%; Z = 5.97; p < 0.00001). Clivus lengths were smaller in patients with CMI than in the controls. However, we found that the p-value after the Q-test was too small. Sensitivity analyses and regression studies on the results did not show significant bias factors. By looking at the funnel plot, we found that 4 datasets\textsuperscript{8,9,14,20} of the included datasets were heavily biased. Exclusion of these 4 datasets resulted in significantly lower heterogeneity (χ\textsuperscript{2} = 20.21, p = 0.04; I\textsuperscript{2} = 46%), thus suggesting them to be the primary source of higher heterogeneity. After excluding these 4 datasets, the results still showed a significant difference in the CMI group compared with the control group (MD = -0.90; 95% CI, -1.02 to -0.78; Z = 14.53; p < 0.00001) (Fig. 3A).

2) Basal angle

The basal angle (cranial base flexion angle) is the inclination angle measured from the nasion, top of the dorsum sellae, and the basal slopes of the occipital and sphenoid bones. The combined analysis of this parameter included 3 studies\textsuperscript{7,11,16} with 283 patients with CMI and 211 controls. The studies showed that patients with CMI had a wider basal angle than controls (MD = 5.01; 95% CI, 3.81–6.21; heterogeneity: χ\textsuperscript{2} = 2.00, p = 0.37; I\textsuperscript{2} = 0%; Z = 8.15; p < 0.00001) (Fig. 3B).

3) Boogard's angle

Boogard's angle is defined by the angle measured between the top of the dorsum sellae, basion, and opisthion. The 5 data-
sets for this parameter were from 4 studies,\(^7,14,16,17\) including 385 patients with CMI and 298 controls. One study\(^17\) included data of children and adults and did not produce significant bias in the combined analysis. The results showed a greater Boogard's angle in patients with CMI than in the controls (MD = 0.61; 95% CI, 0.45–0.77; heterogeneity: \(\chi^2 = 7.43; p = 0.11; I^2 = 46\%\); \(Z = 7.42; p < 0.00001\)) (Fig. 3C).

4) Supraocciput lengths
Supraocciput lengths were defined as the length of the occipital bone medial to the PCF in the midsagittal plane. Thirteen datasets from 8 studies\(^9-11,15-18,20\) were included in evaluation of this parameter. Supraocciput lengths were shorter in patients with CMI than in the controls. However, the heterogeneity was slightly higher after the combined study (\(\chi^2 = 46.25, I^2 = 74\%\)). We analyzed the included datasets and found that 2 datasets\(^9,20\) of them were derived from subgroup studies of patients with CMI. Exclusion of these 2 datasets resulted in significantly lower heterogeneity (\(\chi^2 = 20.35, p = 0.03, I^2 = 51\%\)), thus suggesting them to be the primary source of higher heterogeneity. After excluding these 2 groups, the results still showed a significant difference between the 2 groups (MD = -0.33; 95% CI, -0.45 to -0.21; \(Z = 5.47; p < 0.00001\)) (Fig. 3D).

5) Height of the PCF
The height of the PCF was measured with a line drawn from the most anterior portion of the tentorium, perpendicular to the McRae line. Five datasets from 4 studies\(^10,11,16,17\) were included in assessment of this parameter, and the results showed significant heterogeneity (\(\chi^2 = 23.82, p < 0.0001, I^2 = 83\%\)). We performed a sensitivity analysis using a leave-one-out method. The analysis results (Supplementary Fig. 2) showed the source of heterogeneity was the study of Houston et al.,\(^16\) which had only female subjects; sex may have been a factor in determination of the height of the PCF. Because of a lack of studies on sexual differences in the included literature, a subgroup analysis could not be performed. After removing the data from this group, the heterogeneity was significantly reduced (\(\chi^2 = 4.24, p = 0.24, I^2 = 29\%\)). The height of the PCF was significantly reduced in patients with CMI (n = 281) compared with that in controls (n = 110) (MD = -1.25; 95% CI, -1.49 to -1.01; \(Z = 10.17; p < 0.00001\)) (Fig. 3E).

6) Volume of the PCF
A total of 8 datasets (879 patients with CMI and 283 controls) from 4 studies\(^8,9,12,18\) were included in the analysis of PCFV. The results showed a significant difference in PCFV between the 2 groups, but with high heterogeneity (\(\chi^2 = 242.48, p < 0.00001, I^2 = 97\%\)). We analyzed the data across the groups and found that age could be the reason for the high heterogeneity. We divided the data into adult and child groups, while excluding the study of Nishikawa et al.,\(^18\) which, unlike other studies, analyzed 16- to 19-year-olds as a subgroup. In most studies, 18 years of age was the threshold to distinguish between adults and children. When data from this age group were removed and tested for subgroups, heterogeneity was significantly reduced (adults:...
Fig. 3. Comparison of magnetic resonance imaging-related parameters in patients with Chiari malformation type I (CMI) and controls. Effect sizes are presented as mean difference with 95% confidence intervals (CIs). The heterogeneity of results was estimated using $I^2$, Z, and chi-square tests ($p < 0.05$). Risk of publication bias was assessed by examining funnel plots for symmetry.

(A) The change in clivus length. (B) The change in basal angle. (C) The change in Boogard’s angle. (D) The change in supraoccipital lengths. (E) The change in the height of posterior cranial fossa. (F) The change in the volume of posterior cranial fossa. SE, standard error; SMD, standard mean difference; SD, standard deviation; df, degrees of freedom.

Continued...
Fig. 3. Comparison of magnetic resonance imaging-related parameters in patients with Chiari malformation type I (CMI) and controls. Effect sizes are presented as mean difference with 95% confidence intervals (CIs). The heterogeneity of results was estimated using $\chi^2$, Z, and chi-square tests ($p < 0.05$). Risk of publication bias was assessed by examining funnel plots for symmetry. (A) The change in clivus length. (B) The change in basal angle. (C) The change in Boogard’s angle. (D) The change in supraocciput lengths. (E) The change in the height of posterior cranial fossa. (F) The change in the volume of posterior cranial fossa. SE, standard error; SMD, standard mean difference; SD, standard deviation; df, degrees of freedom. (Continued)

\[\chi^2 = 2.46, p = 0.29, \text{I}^2 = 19\%; \text{children: } \chi^2 = 1.14, p = 0.57, \text{I}^2 = 0\%\].

PCFV was significantly reduced in patients with CMI compared with controls, in both adults (MD = -24.25; 95% CI, -26.09 to -22.42; Z = 25.95; $p < 0.00001$) and children (MD = -10.55; 95% CI, -13.49 to -7.60; Z = 7.02; $p < 0.00001$) (Fig. 3F).

4. MRI-Related Parameters in Predicting Prognosis of Patients With CMI

MRI-related anatomical parameters such as McRae line, pBC2 line, clivoaxial angle, FM-C2 cistern, condylar-C2 sagittal vertical alignment may be associated with the prognosis of CMI. The introduction of functional parameters, such as tonsillar movement, obex displacement, and cerebrospinal fluid dynamics provide adequate guidance for clinical management of the disease. Table 3 summarizes the main characteristics and outcomes of the collected studies conducted on MRI-related imaging parameters for prediction of prognosis in patients with CMI (Supplementary Fig. 1).

DISCUSSION

This study showed that patients with CMI had significant changes in clivus and supraoccipital length and PCFV compared with controls, except for tonsillar herniation. For postoperative predictors of CMI, traditional anatomical parameters are not satisfactory predictors, whereas cerebrospinal fluid dynamics and tonsillar motion parameters may be more convincing.

1. The Value of MRI-Related Parameters in the Diagnosis of CMI

1) Measurement and assessment of PCF structures

Some studies did not reveal a difference in the anteroposterior (AP) diameter of the PCF compared with normal controls with CMI, whereas others have found AP diameters of the PCF in patients with CMI to be shorter. Herein, the AP diameter of the PCF in patients with CMI was not significantly different from that of the controls. In contrast, the height of the PCF showed some differences. The study of Houston et al. had only female subjects; application of the meta-analysis to evaluate the change in PCF height in patients with CMI showed significant heterogeneity. However, the other PCF parameters in that study did not show substantial heterogeneity when subjected to the meta-analysis, suggesting that the variation in PCF height may be related to sex. The reduction in PCF height may account for the smaller PCF in women. However, very few studies have focused on the effect of sex differences on the structure of the PCF to allow further analysis.

Volumetric measurements give a more realistic picture of the...
Table 3. The main characteristics and outcomes of the included studies on the value of MRI-related parameters for the prognosis of patients with CMI

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>No. of cases (M/F)</th>
<th>Mean age (yr)</th>
<th>Intervening method</th>
<th>Parameters</th>
<th>Effect evaluation</th>
<th>Follow-up period</th>
<th>MRI sequences</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collins et al.²⁴ 2022</td>
<td>USA</td>
<td>44 (20/24)</td>
<td>9.73 ± 4.56</td>
<td>PFDD/nonsurgery</td>
<td>CSF flow, tonsillar motion</td>
<td>Clinical outcomes</td>
<td>&gt;2 yr</td>
<td>PC-MRI and 2D fast imaging employing steady-state acquisition</td>
<td>Tonsillar motion was positively correlated with the exacerbation of clinical symptoms in patients with CMI. Tonsillar motion was associated with reduced CSF flow, manifested as aggravated clinical symptoms.</td>
</tr>
<tr>
<td>Furtado et al.²⁶ 2011</td>
<td>India</td>
<td>20 (9/11)</td>
<td>13.08 ± 4.74</td>
<td>FMD, shrinkage of tonsils, and duraplasty</td>
<td>Dimensions of the syrinx and cord, FM, morphometry, dimensions of the PCF, ICV, PCFV</td>
<td>Modified Asgari scoring system</td>
<td>24.6 (3–84) mo</td>
<td>Cranio cervical MRI</td>
<td>The age at presentation, duration, type, morphological measurements of the cranial and FM, and syrinx-related changes were not associated with outcomes at short-term postoperative follow-up.</td>
</tr>
<tr>
<td>Ladner et al.²⁸ 2015</td>
<td>USA</td>
<td>119 (64/55)</td>
<td>8.5 ± 4.2</td>
<td>PFDD</td>
<td>pB–C2 line</td>
<td>Clinical outcomes</td>
<td>2.4 ± 2.9 yr</td>
<td>Brain/cervical spine MRI</td>
<td>Better outcomes after PFDD treatment in patients with grade 1 pB–C2 lines with increased ventral canal obstruction than in those with grade 0 pB–C2 lines.</td>
</tr>
<tr>
<td>Liu et al.²⁷ 2019</td>
<td>China</td>
<td>39 (10/29)</td>
<td>48.3 ± 9.7</td>
<td>PFD</td>
<td>Morphological indicators of PCF (13 linear, 8 angular, 4 areal, and 4 ratios)</td>
<td>CCOS</td>
<td>27 (2–82) mo</td>
<td>Cranio cervical MRI</td>
<td>The morphometric measurements of the PCF did not predict the response to PFD in patients with CMI.</td>
</tr>
<tr>
<td>Mantha et al.²² 2021</td>
<td>Australia</td>
<td>59 (20/39)</td>
<td>3.9/12.7</td>
<td>Surgical decompression with suboccipital craniectomy, C1 laminectomy, Y-shaped durotomy and expansile duraplasty</td>
<td>FM–C2 cistern, pB–C2 distance, CXA, clival length, diaval angle and BoA</td>
<td>CCOS</td>
<td>4–12 mo</td>
<td>Brain/cervical spine MRI</td>
<td>In younger patients with CMI, the preoperative volume of the FM–C2 cistern is a particular indicator that may play a good role in predicting postoperative outcomes.</td>
</tr>
<tr>
<td>Marianayagam et al.²⁹ 2021</td>
<td>USA</td>
<td>35 (20/15)</td>
<td>11.7 ± 5.1</td>
<td>OCF with or without endoscopic endonasal odontoidectomy</td>
<td>CXA, pB–C2, canal diameter, atlantodental, basion-dens, and basion-axial interval</td>
<td>Clinical outcomes</td>
<td>21.3 ± 13.5 mo</td>
<td>CT/MRI</td>
<td>CXA may be the most critical morphological index for predicting clinical outcomes in pediatric patients with CMI after OCF.</td>
</tr>
</tbody>
</table>

(Continued)
Table 3. The main characteristics and outcomes of the included studies on the value of MRI-related parameters for the prognosis of patients with CMI (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>No. of cases (M/F)</th>
<th>Mean age (yr)</th>
<th>Intervening method</th>
<th>Parameters</th>
<th>Effect evaluation</th>
<th>Follow-up period</th>
<th>MRI sequences</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGirt et al.</td>
<td>USA</td>
<td>130 (59/71)</td>
<td>16 ± 13</td>
<td>PFD</td>
<td>Hindbrain CSF flow</td>
<td>Clinical outcomes</td>
<td>19 ± 17 mo</td>
<td>Cine PC-MRI</td>
<td>Preoperatively normal hindbrain CSF flow was an independent risk factor for intervention failure after decompression of patients with CMI.</td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McGirt et al.</td>
<td>USA</td>
<td>44 (24/20)</td>
<td>8 ± 6</td>
<td>PFD</td>
<td>Ventral or dorsal CSF flow</td>
<td>Clinical outcomes</td>
<td>27 ± 16 mo</td>
<td>Cine PC-MRI</td>
<td>Combined ventral and dorsal CSF flow analysis of the hindbrain could better predict patient response to PFD.</td>
</tr>
<tr>
<td>2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ravindra et al.</td>
<td>USA</td>
<td>206 (77/129)</td>
<td>11.5 ± 4.96</td>
<td>PFD, VBD/OCF</td>
<td>C–C2SVA, pB–C2 line, CXA</td>
<td>Sensitivity and specificity</td>
<td>2.28 ± 1.36/2.62 ± 1.43 yr</td>
<td>Brain/cervical spine MRI</td>
<td>A C–C2SVA ≥ 5 mm was a high predictor of requiring OCF/VBD surgery in patients with CMI.</td>
</tr>
<tr>
<td>2021</td>
<td></td>
<td></td>
<td>11.4 ± 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thakar et al.</td>
<td>India</td>
<td>57 (30/27)</td>
<td>38.29 ± 14.32</td>
<td>FMD, C-1 laminectomy, and duraplasty</td>
<td>M-line–FVV distance, caudal displacement of the obex</td>
<td>CCOS</td>
<td>40.29 ± 10.36 mo</td>
<td>Cranio cervical MRI</td>
<td>Obex caudal displacement and shorter distances of the M-line–FVV were associated with good CCOS scores, suggesting that patients with higher hindbrain pathology responded better to surgery.</td>
</tr>
<tr>
<td>2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yuksel et al.</td>
<td>Turkey</td>
<td>70 (17/53)</td>
<td>38.5 (17–70)</td>
<td>Conservative treatment and/or follow-up, surgery</td>
<td>Tonsillar herniation, Chamberlain line, McRae line, and odontoid process-McRae line angle</td>
<td>Annual MRI and periodic examinations</td>
<td>NA</td>
<td>Cranio cervical MRI</td>
<td>McRae line value and symptom severity can be used as predictors of surgical intervention decisions. A new CHIASURG scale can effectively and reliably predict the risk of surgical intervention.</td>
</tr>
<tr>
<td>2022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MRI, magnetic resonance imaging; CMI, Chiari malformation type I; PFD, posterior fossa decompression; CSF, cerebrospinal fluid; PC, phase contrast; 2D, 2-dimensional; FM, foramen magnum; PCF, posterior cranial fossa; ICV, intracranial volume; PCFV, posterior cranial fossa volume; PFDD, posterior fossa decompression with duraplasty; CCOS, Chicago Chiari Outcome Scale; FM–C2, foramen magnum–C2; pB–C2, perpendicular to the basion to C2 line; CXA, clivo-axial angle; BoA, Boogard’s angle; OCF, occipitocervical fusion; CT, computed tomography; VBD, ventral brainstem decompression; C–C2SVA, condylar–C2 sagittal vertical alignment; FMD, foramen magnum decompression; FVV, fourth ventricle vertex; NA, not applicable.
size of the PCF than linear and area measurements. We showed a significant difference in PCFV between patients with CMI and normal subjects, thus confirming that a smaller PCF is a characteristic feature of patients with CMI. It is also clear from the subgroup analysis that PCFV is associated with demographic characteristics. Age, race, sex, and body mass index have statistically significant effects on intracranial measurements and must be considered.\(^1\) Herein, age was found to be a significant factor affecting PCFV. Subgroup analysis showed that grouping by age significantly reduced statistical heterogeneity and demonstrated that adult and pediatric patients with CMI had a substantially smaller PCFV than the corresponding control group. The PCFV is significantly smaller or the PCF is underdeveloped in patients with CMI,\(^2,12,42\) which may be the leading cause of the caudal downward protrusion of the normally developing hindbrain.\(^3\) One study has explored the extent of the association between linear and volumetric measures. In addition to occipital bone length being mildly correlated with PCF and fourth ventricle volume in the CMI cohort, the other expected linear measures were not correlated in each cohort.\(^4\) These findings suggest that linear measures only complement volumetric measures.

Despite the potential diagnostic and prognostic value of PCF morphology, volumetric assessment of PCF size is not commonly used in clinical practice. It is very time-consuming to manually outline the PCF size on multiple images. In contrast, manual measurement of the length of different markers of the PCF is more time efficient and more commonly used as an alternative measurement of PCFV. With the development of 3-dimensional (3D) imaging technology and artificial intelligence, the results of the 3D evaluation are likely to be more accurate and representative, which is a good direction for future research.

2) Measurement and assessment of clivus lengths and associated angles

We found a significant difference in clivus length between patients with CMI and controls. Although heterogeneity of the data included in the meta-analysis was high, the overall effect showed that the clivus was shorter in patients with CMI. We also found that patients with CMI exhibit a wider basal angle and Boogard’s angle than those without. In the author’s opinion, angle combined with clivus length is strong evidence for diagnosing CMI, and separate indices often do not provide a comprehensive characterization.

3) Measurement and evaluation of the occipital structure

Shorter occipital bone length\(^5,11,15,18,20,42\) and wider tentorium angulation\(^17,19\) are typical features of CMI. This study used a meta-analysis to reveal that the patients with CMI and controls showed significant differences between occipital bone length measurements; however, it did not show significant differences in either the tentorium angle or slope of the tentorium. Shorter occipital bone significantly affects the overall layout of the PCF. There is an abnormality in the length of the occipital bone in patients with adult Chiari malformation due to hypoplasia of the occipital body originating from the paraxial mesoderm.

Further analysis showed that the brain tissue volume in the PCF did not differ in patients with CMI compared to controls.\(^8,18\) This allows for overcrowding of the PCF in patients with CMI, which affects the flow of cerebrospinal fluid and alters the form of cerebrospinal fluid flow in the posterior part of the PCF. As symptoms worsen, they will eventually affect the neural structures of the PCF. We believe that the rear aspect of the PCF is a better indicator of the degree of crowding than the structures on the anterior part of the PCF. Shorter occipital bone length makes the PCF morphology shallower and more likely to cause overcrowding of the PCF structures. Changes in tentorium angle make the PCF a better indicator of the degree of crowding than the structures on the anterior part of the PCF. Shorter occipital bone length may not be a characteristic change in CMI. To compensate for the small PCF, the cerebellar tentorium shifts upwards, resulting in an abnormal tentorium angle, which may be an alternation secondary to CMI.\(^13,44\) In the future, exploring the correlation between changes in tentorium angle and CMI symptoms may be done, using the angle change as an adjunct to CMI staging rather than a basis for diagnosis. Considering the PCF to be a container, shorter clivus and occipital lengths make the area of the mouth of the container much larger than the area of the bottom. The area of PCF in the midsagittal plane showed no difference, but the height had decreased. If the PCF of an ordinary person is like a bowl, then the PCF of a patient with CMI is more like a plate. Thus, it is important to measure the shape, volume, and other parameters of PCF in patients with CMI at the 3D level.

Many studies have been conducted on measuring FM in patients with CMI; however, the conclusions differ. We found no significant difference in the diameter of the FM in the midsagittal plane between patients with CMI and controls. We speculate one of the 2 possibilities: firstly, because of the small sample size of the study, the measurement of the FM area was not perfect; secondly, there may have been a lack of relationship between the FM and tonsillar herniation and obstruction of cere-
brosplinal fluid flow. One study found that cerebrospinal fluid reflux at FM may not be due to bone structure.\textsuperscript{45}

2. Value of MRI-Related Parameters in Predicting Prognosis of Patients With CMI

1) Anatomical structure parameters

The degree of cerebellar tonsillar ectopia is the basis of CMI diagnosis; however, is not a prognostic factor.\textsuperscript{24,25} An earlier study collected the following data from pediatric patients with CMI (with or without improvement): age at presentation, duration, type, morphological measurements of the cranial and FM, and presence of syringomyelia and found that these variables were not associated with outcomes at short-term postoperative follow-up.\textsuperscript{26} CMI is associated with smaller PCFV\textsuperscript{34-49} which is also an effective parameter for diagnosis. However, a study based on characteristic PCF parameters found that morphometric measurements of the PCF did not predict the response to posterior fossa decompression in patients with CMI.\textsuperscript{27} This study encompassed 13 linear, 8 angular, and 4 area parameters associated with PCF features and 4 ratios associated with these linear and area parameters to assess.

McRae line value and symptom severity can be used as predictors of surgical intervention decisions.\textsuperscript{21} Authors of this study also developed a CHIASURG scale including depth of tonsillar herniation, Chamberlain line, and McRae line to predict surgical interventions in patients with CMI.\textsuperscript{21} The scale can effectively and reliably predict the risk of surgical intervention.

The pB-C2 line is drawn perpendicular to the line from the C-2 body and the basion at the posterior extent of the odontoid process. The results of a large retrospective study showed better outcomes after posterior fossa decompression with duraplasty treatment in patients with grade I pB-C2 lines (pB-C2 line $\geq 3$ mm) with increased ventral canal obstruction than those with grade 0 pB-C2 lines (pB-C2 line $< 3$ mm).\textsuperscript{28} A pB-C2 line of $> 9$ mm may indicate brainstem compression and is an indicator for anterior decompression surgery.\textsuperscript{50} Another study found that the clivoaxial angle increased significantly only in the postoperative improvement group by analyzing clivoaxial angle, pB-C2 line, atlantoaxial interval, basion-dens interval, basion-axial interval, and canal diameter at the level of C1.\textsuperscript{29} Additionally, other studies suggested that traditional anatomical parameters do not predict prognosis,\textsuperscript{40-42} which may be related to the age and sex of the patients involved in these studies.

The preoperative volume of the FM-C2 cistern is a prevalent indicator. It is a novel cerebrospinal fluid space of the upper cervical canal extending from the FM to the inferior cortex of the C2 body. In younger patients with CMI, the preoperative volume of the FM-C2 cistern is an indicator that may play a good role in predicting postoperative outcomes.\textsuperscript{32} The condylar-C2 sagittal vertical alignment provides a more accurate description of the anatomical loading relationship between the atlantooccipital joint and the cervical segment of the upper spine. A single-center study found that children with CMI had a higher condylar-C2 sagittal vertical alignment than controls.\textsuperscript{30} A multicenter cohort study further validated the predictive value of condylar-C2 sagittal vertical alignment with a sensitivity of 100%, specificity of 86%, and misclassification rate of 12.6% in identifying high-risk patients.\textsuperscript{31} A value of condylar-C2 sagittal vertical alignment $> 5$ mm was a major predictor of the requirement of occipitocervical fusion or ventral brainstem decompression surgery in patients with CMI.\textsuperscript{30,31} The obex caudal displacement and shorter distances of the M-line–fourth ventricle vertex were associated with good Chicago Chiari Outcome Scale scores, suggesting that patients with higher hindbrain pathology responded better to surgery.\textsuperscript{33}

Although these studies have yielded optimistic results, further research is needed to demonstrate the predictive value of these indicators.

2) Cerebrospinal fluid dynamic parameters

Qualitative analysis of phase contrast-MRI can provide additional information to help clinicians decide whether to operate.\textsuperscript{24,25,34,35} The results of Fan et al.\textsuperscript{35} revealed that the subarachnoid manipulation procedure was more feasible for CMI patients with type III cerebrospinal fluid kinetic abnormalities (cerebrospinal fluid flow blockage found in the posterior fossa of the cerebellum and tonsil, the IV ventricle and the central aqueduct, and the ventral space between the clivus and the brainstem). However, the subdural decompression procedure was more suitable for patients with CMI with type I cerebrospinal fluid kinetic abnormalities (cerebrospinal fluid flow blockage found in the posterior fossa space behind the cerebellum and tonsil).\textsuperscript{35} Moreover, preoperative normal hindbrain cerebrospinal fluid flow is an independent risk factor for intervention failure after decompression of patients with CMI.\textsuperscript{24} In other words, abnormal cerebrospinal fluid flow in the hindbrain region preoperatively may predict a better surgical outcome. One study found that combined ventral and dorsal cerebrospinal fluid flow analysis of the hindbrain could better predict patient response to PCF decompression.\textsuperscript{36} Aqueductal stroke volume was crucial for determining the need of surgical treatment.\textsuperscript{37} An aqueductal stroke volume of $\leq 12$ $\mu$L is an essential factor for considering
surgical intervention. Conservative treatment is indicated for adult patients with CMI who are symptomatic and have an aqueductal stroke volume > 15 µL. This study also noted that clinical improvement was positively correlated with an increase in aqueductal stroke volume after treatment.

3) Tonsillar motion
Slight movement of the mesencephalon and brainstem that occurs with the arterial pulsations in the cardiac cycle due to the descending tonsils is a nonnegligible factor in causing abnormal cerebrospinal fluid flow as well as altered neuromotor function. The tonsillar motion was positively correlated with the exacerbation of clinical symptoms in patients with CMI. Slight movement of the mesencephalon and brainstem that occurs with the arterial pulsations in the cardiac cycle due to the descending tonsils is a nonnegligible factor in causing abnormal cerebrospinal fluid flow as well as altered neuromotor function. The tonsillar motion was positively correlated with the exacerbation of clinical symptoms in patients with CMI.34 Tonsillar motion is associated with reduced cerebrospinal fluid flow that manifests as aggravated clinical symptoms.34 Similar results were obtained in another study, which compared morphological and physiological parameters in patients with CMI with excellent and poor prognosis after decompression surgery and found that maximal spinal cord displacement during the cardiac cycle was a better predictor of prognosis than morphological indicators.38 One study even found normalization of the cerebellar tonsils and brainstem in patients, 6 months after cranio cervical decompression by follow-up observation. Thus, the authors inferred that CMI is not a congenital disorder but an acquired malformation caused by cerebellar tonsillar pulsation embedded into the greater occipital foramen.14

3. Limitations
There were some limitations to this study. Primarily, we collected and established some MRI-related morphological and functional parameters in patients with CMI, and obtained some good results. However, most of these parameters are based on traditional measurement methods and are too affected by human factors. Second, most studies did not clearly differentiate between symptomatic and asymptomatic CMI subjects, therefore, imaging-based features specific to the onset and nononset of symptoms are not available. Third, there is very little data in the literature on female subjects. Due to the lack of studies on gender differences, subgroup analysis was not possible. Finally, due to the lack of studies on the prognostic value of MRI-related parameters in patients with conservative follow-up, this aspect has not been evaluated.

CONCLUSION
With the development of MRI technology, diagnosing CMI should become more accurate and comprehensive. The traditional measurement method is affected by human factors and has too many parameters. The complexity of the cranial structure also makes it more challenging to select valid parameters. We collated and established a set of linear, angular, and area measurements deemed essential for diagnosing CMI. However, more indicators can only be analyzed descriptively for various reasons, particularly in prognostic prediction. It has become a major clinical challenge to determine and manage the disease faster and more accurately through multi-dimensional analysis. We posit that the systematic assessment of patients’ PCF morphology, volume, and other parameters at a 3D level holds promising clinical application prospects. Alternatively, additional criteria could be introduced to evaluate the occipital bone, slope, and basal angle as references for confirming the disease diagnosis. Multimodal MRI can be used for determination of disease and prediction of prognosis by introducing parameters to provide adequate guidance for clinical management.

NOTES

Supplementary Material: Supplementary Table 1 and Figs. 1-2 can be found via https://doi.org/10.14245/ns.2347150.575.
Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This work was supported by the National High Level Hospital Clinical Research Funding (2022-PUMCH-B-112); the CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-C&T-B-016); and the Yinhua Public Welfare Foundation.
Author Contribution: Conceptualization: ZW, YL, JG; Formal analysis: ZW, ZL, SH, SP; Investigation: ZW, ZL, SH, XH, SP; Methodology: ZW, ZL, XH; Project administration: YL, JG; Writing – original draft: ZW, ZL, SH; Writing – review & editing: ZW, ZL, SH, XH, SP, YL, JG.

ORCID
Zairan Wang: 0000-0001-8205-7016
Yongning Li: 0000-0002-1305-2286
Jun Gao: 0000-0003-1520-0235

REFERENCES
2. Sekula RF, Jannetta PJ, Casey KF, et al. Dimensions of the posterior fossa in patients symptomatic for Chiari I malfor-


28. Ladner TR, Dewan MC, Day MA, et al. Evaluating the rela-


524 www.e-neurospine.org
**Supplementary Table 1. Risk of bias assessments of the included studies using the ROBINS-I tool**

<table>
<thead>
<tr>
<th>Study</th>
<th>Bias due to confounding</th>
<th>Bias in selection of participants into the study</th>
<th>Bias in classification of interventions</th>
<th>Bias due to deviations from intended interventions</th>
<th>Bias due to missing data</th>
<th>Bias in measurement of outcomes</th>
<th>Bias in selection of the reported results</th>
<th>Overall bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karagöz et al.’ 2002</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Milhorat et al.ª 2010</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Alperin et al.’ 2014</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Aydın et al.ª 2005</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Urbizu et al.ª 2014</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Krishna et al.ª 2016</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Critical</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Critical</td>
</tr>
<tr>
<td>Milhorat et al.ª 2009</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Duftron et al.ª 2011</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Heiss et al.ª 2012</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Houston et al.ª 2018</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
</tr>
<tr>
<td>Nair and Rajshperhar  º 2022</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Nishikawa et al.ª 2022</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Wang et al.ª 2014</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Yan et al. 2016</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Yuksel et al.ª 2022</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Tubbs et al.ª 2003</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
<tr>
<td>Besachio et al.ª 2015</td>
<td>Moderate</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

ROBINS-I, Risk Of Bias In Non-randomised Studies - of Interventions.
Supplementary Fig. 1. MRI-related anatomic and functional parameters for evaluating the diagnosis and prognosis of Chiari malformation type I. MRI, magnetic resonance imaging; PCF, posterior cranial fossa; CSF, cerebrospinal fluid; FM, foramen magnum; C-C2SVA, condylar-C2 sagittal vertical alignment; pB-C2, perpendicular to the basion to C2 line; FM-C2, foramen magnum-C2; FVV, fourth ventricle vertex.
Supplementary Fig. 2. Sensitivity analysis of heterogeneity about the result of posterior cranial fossa height. CI, confidence interval.
Are There Advantages in Cervical Intrafacetal Fusion With Minimal Posterolateral Fusion (PLF) Compared to Conventional PLF in Posterior Cervical Fusion?

Sun Woo Jang¹, Sang Hyub Lee², Jeong Kyun Joo³, Hong Kyung Shin¹, Jin Hoon Park¹, Sung Woo Roh¹, Sang Ryong Jeon¹

1Department of Neurological Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea
2Department of Neurosurgery, Spine Center, The Leon Wiltse Memorial Hospital, Suwon, Korea
3St. Mary Central Madichuk Clinic, Goyang, Korea

Objective: We propose that cervical intrafacetal fusion (cIFF) using bone chip insertion into the facetal joint space additional to minimal PLF is a supplementary fusion method to conventional posterolateral fusion (PLF).

Methods: Patients who underwent posterior cervical fixation accompanied by cIFF with minimal PLF or conventional PLF for cervical myelopathy from 2012 to 2023 were investigated retrospectively. Radiological parameters including Cobb angle and C2–7 sagittal vertical axis (SVA) were compared between the 2 groups. In cIFF with minimal PLF group, cIFF location and PLF location were carefully divided, and the fusion rates of each location were analyzed by computed tomography scan.

Results: Among enrolled 46 patients, 31 patients were in cIFF group, 15 in PLF group. The postoperative change of Cobb angle in 1-year follow-up in cIFF with minimal PLF group and conventional PLF group were 0.1° ± 4.0° and -9.7° ± 8.4° respectively which was statistically lower in cIFF with minimal PLF group (p = 0.022). Regarding the fusion rate in cIFF with minimal PLF group in postoperative 6 months, the rates was achieved in 267 facets (98.1%) in cIFF location, and 244 facets (89.7%) in PLF location (p < 0.001).

Conclusion: Postoperative sagittal alignment was more preserved in cIFF with minimal PLF group compared with conventional PLF group. Additionally, in cIFF with minimal PLF group, the bone fusion rate of cIFF location was higher than PLF location. Considering the concerns of bone chip migration onto the spinal cord and relatively low fusion rate in PLF method, applying cIFF method using minimized PLF might be a beneficial alternative for posterior cervical decompression and fixation.

Keywords: Intrafacetal fusion, Posterolateral fusion, Posterior cervical fusion

INTRODUCTION

Posterior cervical fusion (PCF) is an important method for degenerative cervical spine disease, cervical ossification of posterior longitudinal ligament (OPLL), cervical trauma and tumor disease. Historically, PCF was developed at the beginning of the 20th century from the methods of Hibbs¹ and Albee² who used only an autologous bone graft. Following the description of the wiring technique by Rogers³ in 1942, new methods such as translaminar screws, lateral mass screws (LMS), and pedicle screws...
were introduced. These instrumentations provide immediate stability, and patients benefit from earlier mobilization and rehabilitation while waiting for bone fusion. However, long-term stabilization through instrumental fixation should achieve final bone fusion such as posterolateral fusion (PLF) and proper preparation of the fusion bed is inevitable. However, conventional PLF has some disadvantages including bone chip migration to the spinal cord, which is exposed for decompression, posterior neck pain due to far lateral muscle detachment to make a sufficient fusion bed, or poor fusion rate due to bone chip resorption which has only been reported to be approximately 60%–80%. Another procedure for PCF is cervical intrafacetal fusion (cIFF) using bone chip insertion into well-dissected and prepared facet joints. Therefore, cIFF technique added with minimal PLF could prevent disadvantage of conventional PLF and increase bone fusion rate; however, this is not currently common in cervical spine surgery. In this study, we hypothesized that cIFF achieves a higher rate of bone fusion compared to cervical PLF and demonstrated that adding cIFF to minimal PLF helps maintain cervical alignment and promotes bone fusion.

MATERIALS AND METHODS

This study was approved by the Institutional Review Board (IRB) of Asan Medical Center (IRB No. 2023-1087), ensuring compliance with ethical standards. Given the retrospective nature of the study, the requirement for informed consent was waived by the IRB.

1. Study Design

In this single-center study, patients who underwent posterior cervical laminectomy and screw fixation accompanied by cIFF with minimal PLF and conventional PLF for cervical myelopathy from March 2012 to December 2023 were investigated retrospectively. Among 179 patients, 46 patients were finally selected for the study: 15 patients had posterior cervical fusion accompanied by conventional PLF and 31 patients had cIFF with minimal PLF. Inclusion criteria were a diagnosis of degenerative cervical myelopathy or OPLL. Exclusion criteria consisted of (1) fusion level equal to or less than 2, (2) fusion including occipito-cervical or C1 level, (3) cervical spine tumor, (4) infectious disease, (5) congenital disease, (6) traumatic spine disease, (7) patients followed-up for less than 6 months.

2. Surgical Procedure of cIFF With Minimal PLF

The patient was placed in a prone position with the head secured in a 3-point Mayfield skull clamp (Integra Life Science Corporation, Cincinnati, OH, USA). A midline skin incision was conducted. For visualization of the LMS insertion point, paravertebral muscles were dissected to the lateral margin of the lateral mass. Bilateral gutters were made to remove the lamina for cord decompression using high-speed drilling (Fig. 1A). Then, the lamina was removed en bloc to minimize spinal cord damage (Fig. 1B–D). Drilling with 2- and 3-mm diamond burs was performed to widen the facetal cavity (Fig. 1E, F). Next, curettage was conducted for the meticulous decortication of the intrafacetal space (Fig. 1G, H). In this step, to avoid ventral penetration beyond the facetal space, drilling was kept within a 10-mm depth. Then, bone chips harvested from the removed lamina were inserted into the intrafacetal space using a small impactor (Fig. 1I, J). Afterwards, 2 different screw insertion techniques were used according to the targeted surgical areas: LMS was applied for C3 to C6, and a cervical pedicle screw was used for C2, C7, and T1. Alternatively, for small lateral masses, a pedicle screw was inserted. After cervical screws were inserted, rods were placed and tightened. Remaining bone chips mixed with demineralized bone materials were applied along the lateral side of the lateral mass, which was decorticated for PLF. However, the volume of bone chips for PLF was applied limitedly to avoid bone chip migration onto the spinal cord, and lateral surface decortication in lateral mass was not performed to avoid postoperative neck pain which we named minimal PLF. Finally, muscles and skin were closed tightly. The patient was instructed to wear a Philadelphia collar (Ossur Orthopedics, Reykjavik, Iceland) for 5 months after the surgery for solid bone fusion and a computed tomography (CT) scan was conducted to assess whether bone fusion had been achieved at the 6 months after surgery.

3. Clinical and Radiological Analysis

The patients were divided into 2 groups, conventional PLF group and cIFF with minimal PLF group, and patients’ medical and radiographical records were reviewed. Clinical outcomes were assessed by examining the changes in Nurick grades before and after the surgery. Radiological parameters including C2–7 Cobb angle and C2–7 sagittal vertical axis (SVA) were measured preoperatively, immediately postoperatively, and at the 1-year follow-up.

In cIFF with minimal PLF group, cIFF location and PLF location were divided, and the fusion rates of each location were
compared in postoperative 6 months. The achievement of fusion was decided by the presence of trabecular bone bridging on a CT scan. To analyze the successful fusion, the fusion bridge was carefully evaluated at each facetal level. The first step was examination of the CT sagittal image to determine whether IFF or PLF had occurred. This evaluation was performed by drawing the facet outline (Fig. 2A, B) and identifying a direct bone bridge inside the facet joint was decided to intrafacetal fusion, whereas a circumferential fusion bridge outside the facet joint was considered as evidence of PLF (Fig. 2C, D). If a direct bone bridge was not observed at a specific level, the CT coronal image at that same level was examined (Fig. 2E, F). The levels in

Fig. 1. Representative images of the surgical procedure of cervical intrafacetal fusion. (A) Bilateral gutters were made by high-speed drilling. (B–D) Lamina was removed en bloc to avoid spinal cord damage. (E, F) Drilling using 2- and 3-mm diamond burs was conducted to widen the facetal cavity. (G, H) Curettage was performed for the meticulous decortication of the intrafacetal space. (I, J) Bone chips harvested from removed lamina were inserted into the intrafacetal space using a small size impactor.
which the CT sagittal and coronal images showed no direct bone bridge were classified as nonfusion.

4. Statistics

The simple t-test and paired t-test were conducted using IBM SPSS Statistics ver. 20.0 (IBM Co., Armonk, NY, USA) to assess preoperative and postoperative parameters and fusion mass. Statistical significance was set at \( p < 0.05 \).

RESULTS

1. Demographics

The demographic data of the enrolled patients is presented in Table 1. The 46 patients included 31 males (67%) and 15 females (33%), with a mean age at surgery of 63.4 years (range, 38–85 years). The average clinical follow-up was 19.5 months (range, 6–106 months). In cIFF with minimal PLF group, 6 patients had pedicle screws instead of LMS in the segment of the small lateral mass between C3 and C6 levels. Two patients experienced wound dehiscence from superficial infection. Among them, 1 patient underwent a revision procedure and 1 patient received hyperbaric oxygen therapy without revision. During the screw insertion procedure, 1 patient experienced a vertebral artery injury when trying to insert the pedicle screw, and then, this was converted to LMS. Two patients developed C5 palsy: in 1 patient, this was caused by the ventral extrafacetal impaction of a bone chip into the neural foramen, and they underwent revision surgery to remove the bone chip, resulting in a
Table 1. Summary of demographic data obtained in enrolled patients

<table>
<thead>
<tr>
<th>No.</th>
<th>Age (yr)</th>
<th>Sex</th>
<th>Diagnosis</th>
<th>Operation</th>
<th>Nurick score</th>
<th>Follow-up duration (mo)</th>
<th>BMD</th>
<th>Fusion level</th>
<th>Screw other than routine</th>
<th>Complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>M</td>
<td>OPLL</td>
<td>cIFF with minimal PLF</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>2.1</td>
<td>C3–6</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>66</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>3</td>
<td>2</td>
<td>10</td>
<td>1.5</td>
<td>C3–6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>53</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>-0.2</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>-0.2</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>57</td>
<td>F</td>
<td>CSM</td>
<td></td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>-2.4</td>
<td>C2–T1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>70</td>
<td>F</td>
<td>CSM</td>
<td></td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>-0.9</td>
<td>C3–T1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>64</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>2</td>
<td>1</td>
<td>8</td>
<td>0.8</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>41</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>-0.3</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>38</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>-2.2</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>78</td>
<td>F</td>
<td>CSM</td>
<td></td>
<td>4</td>
<td>3</td>
<td>12</td>
<td>-0.6</td>
<td>C4–7</td>
<td>CPS at left C4</td>
</tr>
<tr>
<td>11</td>
<td>60</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>0.5</td>
<td>C2–6</td>
<td>CPS at C6 bilateral</td>
</tr>
<tr>
<td>12</td>
<td>61</td>
<td>F</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>-0.5</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>61</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>2</td>
<td>-</td>
<td>6</td>
<td>-0.4</td>
<td>C3–6</td>
<td>CPS at right C3, 4, 5</td>
</tr>
<tr>
<td>14</td>
<td>69</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>1.1</td>
<td>C2–6</td>
<td>CPS at right C6</td>
</tr>
<tr>
<td>15</td>
<td>68</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>3.5</td>
<td>C2–T1</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>66</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>0.5</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>58</td>
<td>F</td>
<td>OPLL</td>
<td></td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>-1.3</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>82</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>-2.2</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>53</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>-1.3</td>
<td>C3–6</td>
<td>CPS at right C3, 4, 5, 6</td>
</tr>
<tr>
<td>20</td>
<td>55</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>-1</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>62</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>25</td>
<td>-1.1</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>48</td>
<td>F</td>
<td>OPLL</td>
<td></td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>1</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>57</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>-</td>
<td>C3–7</td>
<td>CPS at left C3, 4</td>
</tr>
<tr>
<td>24</td>
<td>68</td>
<td>F</td>
<td>CSM</td>
<td></td>
<td>3</td>
<td>3</td>
<td>20</td>
<td>-2.2</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>77</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>3</td>
<td>-</td>
<td>6</td>
<td>-2.9</td>
<td>C3–7</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>55</td>
<td>F</td>
<td>CSM</td>
<td></td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>-0.3</td>
<td>C3–T4</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>75</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>0.8</td>
<td>C2–7</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>85</td>
<td>M</td>
<td>CSM</td>
<td></td>
<td>3</td>
<td>2</td>
<td>16</td>
<td>-0.9</td>
<td>C2–6</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>66</td>
<td>F</td>
<td>CSM</td>
<td></td>
<td>1</td>
<td>-</td>
<td>49</td>
<td>-1.6</td>
<td>C3–T1</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>53</td>
<td>M</td>
<td>OPLL</td>
<td></td>
<td>1</td>
<td>1</td>
<td>42</td>
<td>-0.7</td>
<td>C3–7</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### Table 1. Summary of demographic data obtained in enrolled patients (Continued)

<table>
<thead>
<tr>
<th>No.</th>
<th>Age (yr)</th>
<th>Sex</th>
<th>Diagnosis</th>
<th>Operation</th>
<th>Nurick score</th>
<th>Follow-up duration (mo)</th>
<th>BMD</th>
<th>Fusion level</th>
<th>Screw other than routine</th>
<th>Complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>53</td>
<td>F</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>4</td>
<td>3</td>
<td>-0.9</td>
<td>C3–6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>64</td>
<td>F</td>
<td>OPLL</td>
<td>Conventional PLF</td>
<td>2</td>
<td>1</td>
<td>36</td>
<td>C2–6</td>
<td>CPS at right C3, 4, 5, 6</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>60</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>2</td>
<td>2</td>
<td>40</td>
<td>0.9</td>
<td>C2–5</td>
<td>CPS at right C3, 4 and bilat, C5</td>
</tr>
<tr>
<td>34</td>
<td>61</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>4</td>
<td>2</td>
<td>13</td>
<td>-</td>
<td>C3–7</td>
<td>CPS at bilat C3 and right C5</td>
</tr>
<tr>
<td>35</td>
<td>58</td>
<td>F</td>
<td>OPLL</td>
<td>Conventional PLF</td>
<td>2</td>
<td>3</td>
<td>42</td>
<td>-0.6</td>
<td>C2–T2</td>
<td>CPS at right 4, 5</td>
</tr>
<tr>
<td>36</td>
<td>69</td>
<td>F</td>
<td>OPLL</td>
<td>Conventional PLF</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>2.1</td>
<td>C2–5</td>
<td>CPS at bilat C5</td>
</tr>
<tr>
<td>37</td>
<td>78</td>
<td>F</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>-3.4</td>
<td>C3–7</td>
<td>CPS at bilat C3, right C4, left C5, 6</td>
</tr>
<tr>
<td>38</td>
<td>69</td>
<td>M</td>
<td>OPLL</td>
<td>Conventional PLF</td>
<td>2</td>
<td>2</td>
<td>35</td>
<td>0.5</td>
<td>C2–5</td>
<td>CPS at left C5</td>
</tr>
<tr>
<td>39</td>
<td>66</td>
<td>M</td>
<td>OPLL</td>
<td>Conventional PLF</td>
<td>4</td>
<td>3</td>
<td>36</td>
<td>-</td>
<td>C4–7</td>
<td>CPS at right C4, 5, 6</td>
</tr>
<tr>
<td>40</td>
<td>60</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>-0.1</td>
<td>C4–7</td>
<td>CPS at bilat C4, 6, right C5</td>
</tr>
<tr>
<td>41</td>
<td>71</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>-1</td>
<td>C2–5</td>
<td>CPS at right C4, bilat C5</td>
</tr>
<tr>
<td>42</td>
<td>71</td>
<td>F</td>
<td>OPLL</td>
<td>Conventional PLF</td>
<td>2</td>
<td>1</td>
<td>12</td>
<td>-2</td>
<td>C2–7</td>
<td>CPS at right C3, 4, 6 and translaminar screw, left C2</td>
</tr>
<tr>
<td>43</td>
<td>78</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>3</td>
<td>2</td>
<td>12</td>
<td>-1</td>
<td>C3–7</td>
<td>CPS at bilat C5, 6</td>
</tr>
<tr>
<td>44</td>
<td>73</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>1</td>
<td>1</td>
<td>59</td>
<td>-0.3</td>
<td>C4–7</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>58</td>
<td>M</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>2</td>
<td>2</td>
<td>106</td>
<td>-</td>
<td>C4–T2</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>67</td>
<td>F</td>
<td>CSM</td>
<td>Conventional PLF</td>
<td>2</td>
<td>1</td>
<td>36</td>
<td>-2.5</td>
<td>C4–T1</td>
<td>CPS at bilat C4, 5, 6</td>
</tr>
</tbody>
</table>

BMD, bone marrow density; OPLL, ossification of posterior longitudinal ligament; CSM, cervical spondylotic myelopathy; cIFF, cervical intrafacetal fusion; PLF, posterolateral fusion; CPS, cervical pedicle screw; bilat, bilateral.
gradual recovery of motor function (Fig. 3A, B). The other C5 palsy was of unknown origin and recovered spontaneously. In conventional PLF group, 13 patients had pedicle screws instead of LMS in the selected segment dependent on surgeon’s decision. One patient underwent a revision surgery due to postoperative hematoma.

The comparison of demographic data between cIFF with minimal PLF group and conventional PLF group is presented in Table 2. In cIFF with minimal PLF group, 15 patients were diagnosed with CSM, and 16 patients were diagnosed with OPLL. In conventional PLF group, 9 patients were diagnosed with CSM, and 6 patients were diagnosed with OPLL. The Nurick score improvement after operation was 0.55 ± 0.49 in cIFF with minimal PLF group which was significantly lower than conventional PLF group (0.53 ± 0.71, p = 0.040). In cIFF with minimal PLF group, the mean bone marrow density (BMD) was -0.41 ± 1.39 and in conventional PLF group, BMD was -0.74 ± 1.45. In all patients, 7 patients had osteopenia, and 3 had osteoporosis. However, all osteopenia and osteoporosis patients did not receive any additional medications for osteoporosis.

2. Comparison of Cervical Parameters Between Selected cIFF With Minimal PLF Group and Conventional PLF Group

Comparison of cervical parameters of preoperative, postoperative, 1-year follow-up, and postoperative change (1-year follow-up – immediate postoperative) between cIFF with minimal PLF group and conventional PLF group is presented in Table 3. Among 31 patients of cIFF with minimal PLF group, 16 patients whose follow-up period was above 1-year were selected and sagittal parameter was calculated. The average Cobb angle be-

---

**Fig. 3.** The ventral extrafacetal impaction of a bone chip into the neural foramen (indicated by a red arrow), causing compression of the C5 nerve root is shown on sagittal (A) and axial computed tomography images (B).

**Table 2.** Comparison of demographic data between cIFF with minimal PLF group and conventional PLF group

<table>
<thead>
<tr>
<th>Variable</th>
<th>cIFF with minimal PLF group (n = 31)</th>
<th>Conventional PLF group (n = 15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>61.8 ± 10.7</td>
<td>66.8 ± 6.4</td>
<td>0.117</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>21 (67)</td>
<td>10 (66)</td>
<td>0.888</td>
</tr>
<tr>
<td>Female</td>
<td>10 (33)</td>
<td>5 (34)</td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSM</td>
<td>15 (48)</td>
<td>9 (60)</td>
<td>0.298</td>
</tr>
<tr>
<td>OPLL</td>
<td>16 (52)</td>
<td>6 (40)</td>
<td></td>
</tr>
<tr>
<td>Nurick score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (pre)</td>
<td>2.03 ± 0.97</td>
<td>2.26 ± 0.77</td>
<td>0.301</td>
</tr>
<tr>
<td>Postoperative (post)</td>
<td>1.43 ± 0.73</td>
<td>1.73 ± 0.68</td>
<td>0.803</td>
</tr>
<tr>
<td>∆Post–pre</td>
<td>-0.55 ± 0.21</td>
<td>-0.53 ± 0.71</td>
<td>0.040*</td>
</tr>
<tr>
<td>Follow-up (mo)</td>
<td>13.6 ± 10.2</td>
<td>31.6 ± 24.7</td>
<td>0.004*</td>
</tr>
<tr>
<td>BMD</td>
<td>-0.41 ± 1.39</td>
<td>-0.74 ± 1.45</td>
<td>0.975</td>
</tr>
<tr>
<td>Fusion level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6 (19)</td>
<td>7 (47)</td>
<td>0.900</td>
</tr>
<tr>
<td>4</td>
<td>12 (39)</td>
<td>5 (33)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>10 (32)</td>
<td>2 (13)</td>
<td></td>
</tr>
<tr>
<td>≥ 6</td>
<td>3 (10)</td>
<td>1 (7)</td>
<td></td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%). cIFF, cervical intrafacetal fusion; PLF, posterolateral fusion; CSM, cervical spondylotic myelopathy; OPLL, ossification of posterior longitudinal ligament; BMD, bone mineral density.

*p < 0.05, statistically significant differences.

---

fore surgery for the patients was 10.5° ± 10.0° and 8.6° ± 8.2° in immediate postoperative and 8.8° ± 7.6° in 1-year follow-up. In conventional PLF group, the average Cobb angle before surgery
Table 3. Comparison of cervical parameters of preoperative, immediate postoperative, 1-year follow-up, and Δ (1-year follow-up – immediate postoperative) between cIFF with minimal PLF group and conventional PLF group

<table>
<thead>
<tr>
<th>Variable</th>
<th>cIFF with minimal PLF group (n = 16)</th>
<th>Conventional PLF group (n = 15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cobb angle (°)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>10.5 ± 10.0</td>
<td>16.3 ± 7.2</td>
<td>0.080</td>
</tr>
<tr>
<td>Postoperative</td>
<td>8.6 ± 8.2</td>
<td>13.3 ± 8.4</td>
<td>0.784</td>
</tr>
<tr>
<td>1-Year follow-up</td>
<td>8.8 ± 7.6</td>
<td>3.7 ± 8.8</td>
<td>0.677</td>
</tr>
<tr>
<td>Δ (1-Year F/U–postoperative)</td>
<td>0.1 ± 4.0</td>
<td>-9.7 ± 8.4</td>
<td>0.022*</td>
</tr>
<tr>
<td>C2–7 SVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>34.8 ± 18.9</td>
<td>29.8 ± 13.8</td>
<td>0.369</td>
</tr>
<tr>
<td>Postoperative</td>
<td>36.2 ± 12.3</td>
<td>32.5 ± 13.2</td>
<td>0.951</td>
</tr>
<tr>
<td>1-Year follow-up</td>
<td>34.5 ± 15.3</td>
<td>42.6 ± 20.3</td>
<td>0.466</td>
</tr>
<tr>
<td>Δ (1-Year F/U–postoperative)</td>
<td>-1.7 ± 10.5</td>
<td>10.0 ± 16.7</td>
<td>0.430</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. cIFF, cervical intrafacetal fusion; PLF, posterolateral fusion; F/U, follow-up; SVA, sagittal vertical axis. *p < 0.05, statistically significant differences.

was 16.3° ± 7.2° and 13.3° ± 8.4° in immediate postoperative and 3.7° ± 8.8° in 1-year follow-up. The change (1-year follow-up – immediate postoperative) in cIFF with minimal PLF group was 0.1° ± 4.0° which was significantly lower than conventional PLF group (-9.7° ± 8.4°, p = 0.022). In the cIFF with minimal PLF group, the average C2–7 SVA before surgery for the patients was 34.8 ± 18.9 and 36.2 ± 12.3 in immediate postoperative and 34.5 ± 15.3 in 1-year follow-up. In conventional PLF group, the average C2–7 SVA before surgery was 29.8 ± 13.8 and 32.5 ± 13.2 in immediate postoperative and 42.6 ± 20.3 in 1-year follow-up. The change (1-year follow-up – immediate postoperative) in cIFF with minimal PLF group was -1.7 ± 10.5 which was not significant but lower than conventional PLF group (10.0 ± 16.7, p = 0.430).

3. Comparison of Fusion Rate According to cIFF and PLF Locations in cIFF With Minimal PLF Group

In cIFF with minimal PLF group, cIFF location and PLF location were divided and the fusion rates of each location in postoperative 6 months were carefully compared (n = 31). The total number of facets among the cIFF with minimal PLF group was 272. Among them, fusion success in cIFF location was observed in 267 facets (98.1%), and in PLF location, 244 facets (89.7%) were successful, which was significantly different (p < 0.001).

Table 4. Comparison of the fusion rate between cIFF location and PLF location in cIFF with minimal PLF group after 6 months

<table>
<thead>
<tr>
<th>Fusion level</th>
<th>cIFF</th>
<th>PLF</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2–3 (n = 30)</td>
<td>28 (93.3)</td>
<td>22 (73.3)</td>
<td>0.038*</td>
</tr>
<tr>
<td>C3–4 (n = 58)</td>
<td>57 (98.2)</td>
<td>49 (84.4)</td>
<td>0.008*</td>
</tr>
<tr>
<td>C4–5 (n = 60)</td>
<td>60 (100)</td>
<td>59 (98.3)</td>
<td>0.315</td>
</tr>
<tr>
<td>C5–6 (n = 60)</td>
<td>60 (100)</td>
<td>58 (96.6)</td>
<td>0.154</td>
</tr>
<tr>
<td>C6–7 (n = 46)</td>
<td>45 (97.8)</td>
<td>40 (86.9)</td>
<td>0.049*</td>
</tr>
<tr>
<td>C7–T1 (n = 12)</td>
<td>11 (91.6)</td>
<td>10 (83.3)</td>
<td>0.537</td>
</tr>
<tr>
<td>T1–2 (n = 2)</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>1.000</td>
</tr>
<tr>
<td>T2–3 (n = 2)</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>1.000</td>
</tr>
<tr>
<td>T3–4 (n = 2)</td>
<td>2 (100)</td>
<td>2 (100)</td>
<td>1.000</td>
</tr>
<tr>
<td>Total (n = 272)</td>
<td>267 (98.1)</td>
<td>244 (89.7)</td>
<td>&lt; 0.001*</td>
</tr>
</tbody>
</table>

Values are presented as number (%). cIFF, cervical intrafacetal fusion; PLF, posterolateral fusion. *p < 0.05, statistically significant differences.

When analyzed by level, the fusion rate of cIFF in C2–3, 3–4, and 6–7 was significantly higher than that of PLF (Table 4).

DISCUSSION

In the situation where wide cervical laminectomy for spinal cord decompression is performed, placing bone chips around a decorticated lateral mass for PLF can potentially induce the migration of bone chips onto the spinal cord, resulting in unfortunate spinal cord compression. The concerns of bone chip migration make it insufficient for the application of bone chips in the PLF bed.\(^{13,14}\) Moreover, the bone fusion rate at the area treated with PLF was reported to be approximately 60%–90% and tended to be poorly fused where preoperative instability was found.\(^{16–22}\) Additionally, extensive lateral muscle dissection to create a PLF bed provokes uncomfortable postoperative neck pain.\(^{13,14}\) Therefore, we utilized a technique that involves the direct placing of bone chips into the intrafacetal space to induce cIFF, which applies as additional methodology to limited PLF to overcome the disadvantages of PLF alone. When we conduct cIFF combined with minimal PLF, it is possible to perform less invasive PLF as an adjuvant procedure by applying a small volume of bone chips as well as less lateral muscle detachment and avoiding lateral surface decortication of lateral mass. After removing the soft tissue within the intrafacetal space and utilizing the autologous bone chips obtained during laminectomy, we can safely and conveniently impact bone chips without concerns of their migration towards the spinal cord. Because the
bone chips are fixed within the intrafacetal tight space, they are expected to promote effective bone fusion. Intrafacetal fusion, unlike onlay grafts that performed in PLF, is expected to allow a higher rate of direct fusion considering the transmission of axial loading and Wolff’s law.23

There have been a few reports of the intrafacetal technique in PCF. Cofano et al.24 reported the insertion of cervical interfacetal spacer allograft during PCF in patients who had developed pseudarthrosis after anterior cervical disectomy and fusion. The paper suggested that the fusion rate could be improved, and foraminal stenosis could be addressed by increasing the foraminal height and width. However, their study targeted patients who had undergone the anterior approach. Although Kansal et al.25 used the terminology of intrafacetal fusion in bilateral C1–2 transarticular screw insertion, this was approached by anterior high cervical retropharyngeal approach. Therefore, this present study is the first report to address intrafacetal fusion in the posterior approach of a general degenerative cervical spine and OPLL. In the previous reports, the terminology “intrafacetal” and “interfacetal” were intermingled, but we considered “intrafacetal” might be more appropriate terminology to describe a fusion method that inserts bone chips into the cervical facetal space.

Based on the data, when comparing the follow-up results on the sagittal alignment of cIFF with minimal PLF group and conventional PLF group, it was noted that the Cobb angle and C2–7 SVA were observed as having better maintenance in lordotic curve during the follow-up in cIFF with minimal PLF group rather than conventional PLF group. In comparison with previous study using LMS fixation with PLF only, this result of present study showed better outcome. Lee et al.26 reported the difference of Cobb angle between final follow-up and immediate postoperative was -0.70° after LMS with PLF only, which is higher in postoperative loss of lordotic curve than our results (0.1°±4.0°). And, in C2–7 SVA, the difference between final follow-up and immediate postoperative was 2.06 which higher than our results (-1.7 ± 10.5). Inose et al.27 documented alterations in cervical lordosis after posterior cervical decompression and PLF between preoperative and postoperative 1-year results which were 9.9° ± 11.3° and 4.6° ± 10.4°. Additionally, the C2–7 SVA from 21.8 ± 20.2 to 29.3 ± 20.2. Comparing to our preoperative and postoperative 1 year data of cIFF with minimal PLF group, in which the preoperative Cobb angle was 10.5° ± 10.0° and reduced to 8.8° ± 7.6° 1 year postoperative, and the C2–7 SVA measured 34.8 ± 18.9 preoperatively and 34.5 ± 15.3 1 year postoperatively, the present novel technique demonstrated a better preservation of cervical sagittal alignment.

The previous literatures which studied about PCF rates used dynamic x-ray and the absence of instability was used as the indication of fusion.29,30,31 Highsmith et al.32 reported a fusion rate of 92% in 24 months and Heller et al.33 61.5% in 25.5 months in which they used absence of pseudoarthrosis in simple x-ray as a criteria of fusion. However, Hong et al.34 reported 100% of fusion in 18.9 months. Therefore, the evaluation of fusion by simple x-ray has a variable range of fusion rate. In our study, the CT scan observing bone bridge or trabecular bone formation enabled a more objective comparison of fusion rates by directly evaluating the cIFF location and PLF location in the fusion bed in patients and we evaluated them in relatively early follow-up period than previous studies. In cIFF with minimal PLF group, when counting the number of fused facets in cIFF location, the fusion rate 98.1% in 6 months was significantly higher when compared to PLF location (89.7%, p < 0.001). Because of the minimal PLF procedure leading to a relatively small fusion bed space due to limited muscle dissection and the use of fewer bone chips, the lower fusion rate observed may be attributed to these factors. However, at just 6 months postoperative state, the cIFF demonstrated a significantly higher fusion rate of 98.1%, suggesting it a promising option to improve the fusion success in PCF surgery. The locations of the worse facet fusion rate in PLF compared to cIFF were C2–3, C3–4, and C6–7 levels, which is supposed to be the results of micromotion in the uppermost and lowermost portion of the cervical region. Therefore, the need of cIFF is emphasized in these levels.

In cIFF with minimal PLF group, C5 palsy was observed in 2 cases postoperatively, and in one of them, we considered it was caused by bone chips which were deeply inserted and penetrated beyond ventral facetal surface resulting in nerve root compression. The penetrated bone chips were confirmed on postoperative CT. Therefore, we performed a revision procedure to remove the penetrated bone chips. The patient recovered gradually from C5 palsy during the 19-month follow-up. It is important to note that during cIFF procedures, there might be a risk of nerve root or vertebral artery compression from the penetration of bone chips beyond the ventral facetal margin. This complication could be prevented by avoiding drilling more than 10-mm depth from the dorsal facet surface and aggressive bone chip impaction during cIFF.

This study had several limitations. First, the retrospective design of the present study might have the potential selection bias and limited data; therefore, we need to perform a more precise study design such as prospective study including patient-re-
ported outcomes with Japanese Orthopaedic Association score or visual analogue scale score to demonstrate the advantage of cIFF compared to PLF. Second, the data of fusion rate did not directly compare cIFF and PLF using double-arm study. Instead, we analyzed fusion results based on patients who underwent both procedures simultaneously, making an independent comparison difficult. Third, the comparison of fusion rates at the 6 months postoperatively may not have shown sufficient fusion. Comparing the long-term follow-up CT-based fusion rates with the conventional PLF group, would provide a more accurate analysis. Fourth, factors that can influence fusion outcomes, such as BMD and types of screwfixation were not taken into consideration: in cIFF with minimal PLF group, pedicle screws were inserted in 6 patients (19%) because of a small lateral mass between the C3 and C6 levels, and in conventional PLF group, pedicle screws were inserted in 13 patients (86%) because of surgeon’s preference. Finally, the cases enrolled in this study were chosen with a specific focus on degenerative spine conditions and OPLL. Further validation in extended indications, for example trauma or spine metastasis, is required to determine the effectiveness of cIFF.

**CONCLUSION**

In this study, to overcome the limitations of PLF, which has a relatively lower fusion rate, risks of bone chip migration to the spinal cord, and can induce neck pain due to wide muscle dissection and extended decortication, we introduced an alternative technique that directly inserts bone chips into the intrafacetal space combined with minimal PLF. Postoperative sagittal alignment was more preserved in cIFF with minimal PLF group rather than conventional PLF group. The fusion rate according to location, intracetal space (98.1%) was significantly higher than PLF bed (89.7%) in cIFF with minimal PLF technique. Considering the concerns of disadvantages in the PLF procedure as well as the convenience of the intrafacetal fusion technique, the additional cIFF combined with minimal PLF might be a beneficial alternative for posterior cervical decompression and fixation.

**NOTES**

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This research was supported by the Korean Fund for Regenerative Medicine (KFRM) grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Health & Welfare) (No. 23C0120L1).

**Author Contribution:** Conceptualization: SRJ; Formal analysis: SWJ; Investigation: SWJ; Methodology: SWJ, SRJ; Project administration: SRJ; Writing – original draft: SWJ; Writing – review & editing: SWJ, SHL, JKJ, HKS, JHP, SRJ, SWR.

**ORCID**

Sun Woo Jang: 0000-0001-8425-4961
Sang Hyub Lee: 0000-0002-6642-5285
Jeong Kyun Joo: 0000-0002-5722-1213
Hong Kyung Shin: 0000-0001-8182-3321
Jin Hoon Park: 0000-0002-0903-3146
Sung Woo Roh: 0000-0001-6562-4154
Sang Ryong Jeon: 0000-0002-8340-7978

**REFERENCES**

Additional Intrafacetal Fusion to Minimal PLF for Posterior Cervical Fusion

Jang SW, et al.


Finite Element Analysis of Stress Distribution and Range of Motion in Discogenic Back Pain

Pyung Goo Cho¹*, Seon-Jin Yoon²*, Dong Ah Shin³, Min Cheol Chang³

¹Department of Neurosurgery, Ajou University Medical Center, Suwon, Korea
²Department of Neurosurgery, Yonsei University College of Medicine, Seoul, Korea
³Department of Physical Medicine & Rehabilitation, Yeungnam University College of Medicine, Daegu, Korea

Objective: Precise knowledge regarding the mechanical stress applied to the intervertebral disc following each individual spine motion enables physicians and patients to understand how people with discogenic back pain should be guided in their exercises and which spine motions to specifically avoid. We created an intervertebral disc degeneration model and conducted a finite element (FE) analysis of loaded stresses following each spinal posture or motion.

Methods: A 3-dimensional FE model of intervertebral disc degeneration at L4–5 was constructed. The intervertebral disc degeneration model was created according to the modified Dallas discogram scale. The von Mises stress and range of motion (ROM) regarding the intervertebral discs and the endplates were analyzed.

Results: We observed that mechanical stresses loaded onto the intervertebral discs were similar during flexion, extension, and lateral bending, which were greater than those occurring during torsion. Based on the comparison among the grades divided by the modified Dallas discogram scale, the mechanical stress during extension was greater in grades 3–5 than it was during the others. During extension, the mechanical stress loaded onto the intervertebral disc and endplate was greatest in the posterior portion. Mechanical stresses loaded onto the intervertebral disc were greater in grades 3–5 compared to those in grades 0–2.

Conclusion: Our findings suggest that it might be beneficial for patients experiencing discogenic back pain to maintain a neutral posture in their lumbar spine when engaging in daily activities and exercises, especially those suffering from significant intravertebral disc degeneration.

Keywords: Finite element, Intervertebral disc, Posture, Discogenic back pain, Lumbar spine

INTRODUCTION

Lower back pain is a widespread issue affecting numerous individuals, and it can cause a wide range of disabilities and often imposes a significant socioeconomic burden,¹ with a reported lifetime prevalence of 60%–80%.[²] Lower back pain is a complex condition that is influenced by various physiological and psychological factors and changes in the brain.[³,⁴] Intervertebral disc degeneration plays a significant role in patients with lower back pain.[⁵] Discogenic back pain refers to lower back pain associated with intervertebral disc degeneration that occurs without disc herniation, anatomical deformity, or other clearly identifiable causes of pain and disability.[⁶]

Discogenic back pain is often refractory to oral medication or various procedures, and even surgical treatment does not exhibit a high success rate.[⁶,⁷] For patients with discogenic back pain, regular exercise and postural education are crucial.[⁸,⁹] Numerous previous studies have focused on proper posture in patients with discogenic back pain.[¹⁰,¹¹] However, it is not clearly understood which posture causes the most significant mechan-
tical stresses on the intervertebral disc. Precise knowledge regarding the mechanical stress applied to the intervertebral disc following each individual spine motion enables physicians and patients to understand how individuals should be guided in their exercises and which spinal motions to specifically avoid. Previous in vivo studies have been conducted to evaluate stress loading on the intervertebral disc resulting from various spinal postures or motions. However, these studies were conducted on healthy subjects with no lower back pain, or did not measure intradiscal pressure under various postures. In a previous in vivo study measuring intradiscal pressure, a needle or pressure sensor was inserted into the intervertebral disc, and this can cause damage to the intervertebral disc and accelerate its degeneration. Therefore, due to the ethical issues, in vivo studies are currently limited in their ability to evaluate stress loading on the intervertebral disc following various spinal postures or motions.

Recently, in an effort to overcome the limitations of in vivo studies, finite element (FE) modeling that is typically used in industrial fields has been utilized for spinal research. FE modeling measures mechanical stresses on each spinal structure without inserting an invasive device into the structure. Several studies have evaluated pressures on the intervertebral discs of the lumbar spine according to different postures or motions using FE modeling. However, these studies were not conducted using an intervertebral disc degeneration model. Normal and degenerated intervertebral discs experience different intradiscal pressures and possess different anatomical characteristics. To obtain clinically relevant research results that are applicable to patients with discogenic back pain, an FE analysis of the intradiscal stress occurring during each spinal posture or motion should be conducted using an intervertebral disc degeneration model.

In the current study, we created an intervertebral disc degeneration model and conducted an FE analysis of loaded stresses following each individual spinal posture or motion.

**MATERIALS AND METHODS**

1. **Model Development**

After obtaining approval from the Institutional Review Board of the Severance Hospital (2013-0515-001), a computed tomography (CT) scan of the lumbar spine was obtained from one patient (a 45-year-old man with lower back pain with no structural abnormality in the lumbar spine magnetic resonance imaging scan. It was a high-resolution CT scan encompassing a 1.0-mm section. Based on the CT images, a 3-dimensional (3D) model of the L4–5 lumbar spine was created using the Mimics (Materialise, Leuven, Belgium) software. The obtained 3D geometry was transformed into a hexahedral mesh using IA-FEMesh (The University of Iowa, Iowa City, IA, USA). The final FE model possessed 2 vertebrae (L4 and L5), one intervertebral disc, cartilage endplates, and spinal ligaments. The element type, number of nodes, and number of elements of each component are indicated in Table 1. The facet joint gap was modeled based on the original CT data. Surface-to-Surface contact and frictionless sliding between facet joints were applied. The ligaments were modeled as linear elastic in compression-free conditions. The intervertebral disc was made hyperelastic using the Mooney-Rivlin model. The material properties of each component are listed in Table 2. The final FE model was exported to Abaqus (Dassault Systemes, Paris, France) for analysis (Fig. 1). To validate the intact FE model, the same load and boundary conditions were used as those described in Yamamoto et al.

<table>
<thead>
<tr>
<th>Component</th>
<th>Element type</th>
<th>No. of nodes</th>
<th>No. of elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortical bone</td>
<td>Hexahedral C3D8RH</td>
<td>5,388</td>
<td>2,652</td>
</tr>
<tr>
<td>Cancellous bone</td>
<td>Hexahedral C3D8RH</td>
<td>8,590</td>
<td>7,197</td>
</tr>
<tr>
<td>Posterior bone</td>
<td>Hexahedral C3D8RH</td>
<td>5,707</td>
<td>3,830</td>
</tr>
<tr>
<td>Nucleus pulposus</td>
<td>Hexahedral C3D8RH</td>
<td>1,440</td>
<td>1,044</td>
</tr>
<tr>
<td>Annulus fibrosus</td>
<td>Hexahedral C3D8RH</td>
<td>1,300</td>
<td>832</td>
</tr>
<tr>
<td>Cartilage endplate</td>
<td>Hexahedral C3D8RH</td>
<td>1,984</td>
<td>938</td>
</tr>
<tr>
<td>Facet surface</td>
<td>Hexahedral C3D8RH</td>
<td>204</td>
<td>65</td>
</tr>
<tr>
<td>Ligaments</td>
<td>Line T3D2H</td>
<td>62</td>
<td>31</td>
</tr>
<tr>
<td>Annulus fibers</td>
<td>Line T3D2H</td>
<td>260</td>
<td>416</td>
</tr>
<tr>
<td>Total components</td>
<td></td>
<td>24,935</td>
<td>17,005</td>
</tr>
</tbody>
</table>

2. **Intervertebral Disc Degeneration Model**

The intervertebral disc degeneration model was established according to the modified Dallas discogram scale: grade 0: normal disc; grade 1: radial tears confined to the inner third of the annulus fibrosis; grade 2: radial tears extending to the middle third of the annulus fibrosis; grade 3: a radial tear extending to the outer third of the annulus fibrosis; grade 4: a grade 3 tear with dissection into the outer third of the annulus that involves greater than 30° of the disc circumference; and grade 5: full-thickness tear. As a boundary condition, the inferior surface of the lower vertebra was constrained in all directions. The inter-

https://doi.org/10.14245/ns.2347216.608
Table 2. Material properties used in the finite element model

<table>
<thead>
<tr>
<th>Component</th>
<th>Young’s modulus (MPa)</th>
<th>Poisson ratio</th>
<th>Cross section area (mm²)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortical bone</td>
<td>12,000</td>
<td>0.3</td>
<td></td>
<td>Shirazi-Adl et al. 2019</td>
</tr>
<tr>
<td>Cancellous bone</td>
<td>100</td>
<td>0.2</td>
<td></td>
<td>Wang et al. 2016</td>
</tr>
<tr>
<td>Posterior bone</td>
<td>3,500</td>
<td>0.25</td>
<td></td>
<td>Polikeit et al. 2003</td>
</tr>
<tr>
<td>Nucleus pulposus</td>
<td>Hyperelastic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C10: -0.219197</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C01: 0.43494</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D1: 0.000927066</td>
<td>0.499</td>
<td></td>
<td>Shirazi-Adl et al. 2019</td>
</tr>
<tr>
<td>Annulus fibrosus</td>
<td>Hyperelastic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C10: -0.117485</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C01: 0.273737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D1: 0.66206</td>
<td>0.45</td>
<td></td>
<td>Lavaste et al. 1991</td>
</tr>
<tr>
<td>Annulus fibers</td>
<td>500</td>
<td>0.3</td>
<td></td>
<td>Little et al. 2008</td>
</tr>
<tr>
<td>Cartilage endplate</td>
<td>24</td>
<td>0.4</td>
<td></td>
<td>Goel et al. 1995</td>
</tr>
<tr>
<td>Facet cartilage</td>
<td>24</td>
<td>0.4</td>
<td></td>
<td>Wang et al. 2016</td>
</tr>
<tr>
<td>Ligament</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL</td>
<td>20</td>
<td>63.7</td>
<td></td>
<td>Zhong et al. 2006</td>
</tr>
<tr>
<td>PLL</td>
<td>20</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>32.9</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITL</td>
<td>58.7</td>
<td>3.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISL</td>
<td>11.6</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSL</td>
<td>15</td>
<td>30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ALL, anterior longitudinal ligament; PLL, posterior longitudinal ligament; CL, capsular ligament; ITL, intertransverse ligament; ISL, interspinous ligament; SSL, supraspinous ligament.

Fig. 1. A finite element model of the L4–5 functional spinal unit implemented in Abaqus software. (A) A section cut in the midsagittal plane. The cortical bone, cancellous bone, endplates, and intervertebral disc are implemented. (B) The annulus fibrosus is composed of fibers at an angle of 45° to the ground substance. (C) The intervertebral disc consists of the nucleus pulposus and 4 layers of the annulus fibrosus.
vertebral disc was composed of the nucleus pulposus and 4 layers of annulus fibrosus. The annulus fibrosus was reinforced with fibers in the ground material. Elements were removed in the middle of the posterior portion of the annulus according to grade. The superior surface of the upper vertebra was coupled to a reference point. Loading was applied at that point and included a 7.5-Nm moment and 280-N compression. The compression force was the weight of the upper body on the spine. The compressive loading was applied in the form of a follower load as suggested by Patwardhan et al.\textsuperscript{29} The follower load was constructed by coupling the reference points at the center of gravity of each vertebral body and connecting the 2 points with a connector element. The von Mises stress and range of motion (ROM) regarding the intervertebral discs and endplates were analyzed.

**RESULTS**

The von Mises stress values during flexion, extension, and lateral bending were not significantly different, and all were greater than those that occurred during torsion (Fig. 2). Based on the comparison among the grades during flexion, lateral bending, and torsion, there was no significant difference in the peak von Mises stress values. However, during extension, peak von Mises stresses were greater in grades 3–5 than those in the others. Grades 3 and 4 were 22.7% and 25.7% higher than grade 0, respectively; grade 5 was 17.8% larger than grade 0 (Fig. 2). ROM was measured during flexion, extension, lateral bending, and torsion by applying a moment of 6 Nm. The ROM values were within 10% of the results reported by Yamamoto et al.\textsuperscript{20} ROM for each motion is indicated in Fig. 3. The ROM values during flexion and extension were greater than those during lateral bending and torsion. From the comparison among the grades, during lateral bending and torsion there was no significant difference in peak von Mises stress. However, during flexion, ROM in grade 5 was greater than that in others, and grade 5 exhibited 5.2% greater ROM than did grade 0. During extension, ROM in grade 4 was greater than that in others, and grade 4 exhibited 8.2% greater ROM than did grade 0 (Fig. 3). The von Mises stress was greatest in the posterior portions of the intervertebral disc (Fig. 4) and endplates (Fig. 5) in all the grades. Peak stresses loaded onto the intervertebral discs were greater in grades 3–5 than those in grades 0–2.

![Fig. 2.](image-url) The peak von Mises stress value acting on the end plate during extension. During extension, the peak von Mises stress values are greater in grades 3–5 than those in the others.

![Fig. 3.](image-url) Range of motion in the L4–5 functional spinal unit according to each spinal motion. During flexion, the value is greatest in grade 5; during extension, it is greatest in grade 4.

![Fig. 4.](image-url) Contours of von Mises stress acting on the intervertebral disc during extension. All peak stresses are loaded onto the posterior portion of the annulus fibrosus. Peak stresses in grades 3–5 are greater than those in grades 0–2.
DISCUSSION

In our study, we observed that mechanical stresses loaded onto the intervertebral disc were similar during flexion, extension, and lateral bending, and all of these were greater than those observed during torsion. From the comparison among the grades, the mechanical stresses during extension were greater in grades 3–5 than that in others. During extension, mechanical stress loaded onto the intervertebral disc and endplate was greatest in the posterior portion. Mechanical stresses loaded onto the intervertebral disc were greater in grades 3–5 compared to those in grades 0–2. Additionally, ROMs during flexion and extension were greater than those observed during lateral bending and torsion. From the comparison among the grades, ROM during flexion was highest in grade 5; during extension, it was highest in grade 4.

We demonstrated that flexion, extension, and lateral bending directs the stress or pressure to the L4–5 intervertebral disc to a similar degree in patients with discogenic back pain. Therefore, medical staff should emphasize to patients with discogenic back pain the importance of maintaining a neutral posture in the lumbar spine during daily activities and exercise. Neutral posture in the lumbar spine refers to a natural and relaxed alignment with slight lumbar lordosis to produce the least amount of pressure on the spinal column, discs, and nerves. In clinical practice, the importance of maintaining a neutral position for patients with lower back pain has been emphasized. We scientifically confirmed the importance of a neutral position of the lumbar spine using FE analysis. Also, stress loaded onto intervertebral discs was particularly high during extension for patients with severe degeneration of intervertebral discs (grades 3–5). This indicates that patients with severe degeneration of the intervertebral disc should particularly avoid extension of the lumbar spine.

The extent of ROM values at specific positions indicates the potential to generate stress or pressure in the intervertebral disc, and these values were observed to be significantly greater during flexion and extension. Furthermore, ROM was even greater in severe disc degeneration (grade 4 or 5). These findings highlight the importance of the neutral position of the lumbar spine, particularly in patients with severe intervertebral disc degeneration.

Additionally, during extension, mechanical stress loaded onto the intervertebral disc and endplate was greatest in the posterior portion. The sinuvertebral nerve (a branch of the spinal nerve root), is considered to be a primary contributor to discogenic back pain in response to mechanical and chemical irritation. The sinuvertebral nerves are distributed at the posterior portion of the disc, along the outer layer of the annulus fibrosus. Furthermore, when the annulus of the intervertebral disc is torn, the sinuvertebral nerves grow inward along the tear. When extending the lumbar spine, there is a high possibility of mechanical irritation to sinuvertebral nerves in the posterior area of intervertebral disc. The extension position can also lead to an annulus tear in the posterior portion of the intervertebral disc and cause the sinuvertebral nerve to grow inward toward the nucleus pulposus. The extension posture of the lumbar spine can trigger or exacerbate discogenic back pain. This negative influence was particularly pronounced when disc degeneration was severe.

Several studies have evaluated the influence of posture and motion on the lumbar spine through FE analysis. Kuo et al. investigated increments in intradiscal pressure during standing, flexion, extension, and rotation. Intradiscal pressure was increased in all of the postures, most markedly during flexion. They suggested that lumbar flexion is the posture or motion that should be most avoided to prevent disc degeneration. Cho et al. evaluated how pressures on lumbar spine change during different postures such as standing, erect sitting on a chair, slumped sitting on a chair, and sitting on the floor using FE analysis. The pressures on the nucleus pulposus, annulus fibrosus, and cortical bone during standing and erect sitting postures were not...
significantly different. However, during slumped sitting in a chair and sitting on the floor, there was significantly increased pressure on the nucleus pulposus, annulus fibrosus, and cortical bone. In particular, sitting on the floor induced even greater pressure on the nucleus pulposus and annulus fibrosus than did slumped sitting in a chair. They concluded that maintaining a neutral posture is important to reduce intradiscal pressure and cortical bone stress associated with degenerative disc disease or spinal deformities. These previous studies emphasized the importance of adopting or maintaining a neutral posture based on the results of their FE analyses.\textsuperscript{11,12} However, the FE analyses conducted in the previous studies did not utilize the intervertebral disc degeneration model. Our study is first to conduct FE analysis to investigate stresses loaded onto various spinal postures or motions using a model of intervertebral disc degeneration. Furthermore, we analyzed and compared the stress loaded onto the lumbar spine based on the severity of intervertebral disc degeneration. However, our study has some limitations. First, we did not evaluate the distribution of stress loaded onto the intervertebral disc or endplate during postures or motions other than extension. Second, we conducted our research targeting only the L4–5 lumbar region, where discogenic back pain occurs most commonly, rather than targeting the entire lumbar spine. Third, Young’s modulus or the Poisson ratio, both of which can vary for each grade, were not taken into account in this experiment. Our study solely acquired morphology from the patient’s CT images, failing to fully represent the physical properties of the degenerative disc. There are various methods available to ascertain the physical properties of degenerative discs,\textsuperscript{33-35} and it is recommended to integrate these into future research. Fourth, FE analysis simplifies complex spinal structures and cannot reflect all the factors that can occur \textit{in vivo}. Accordingly, additional clinical research based on the results of our study should be conducted.

CONCLUSION

In conclusion, we demonstrated using FE analysis that mechanical stresses loaded onto intervertebral discs are significantly increased in flexion, extension, and lateral bending in patients with discogenic back pain. In patients with severe disc degeneration, mechanical stress was observed to be greater (particularly during extension) compared to that observed in mild disc degeneration. Moreover, during extension mechanical stress loaded onto the intervertebral disc and endplate was focused in the posterior portion and was greater in patients with severe disc degeneration than that in patients with mild disc degeneration. Our findings suggest that it might be beneficial for patients experiencing discogenic back pain to maintain a neutral posture in their lumbar spine when engaging in daily activities and exercises, especially those suffering from significant intravertebral disc degeneration.

NOTES

\textbf{Conflict of Interest:} The authors have nothing to disclose.
\textbf{Funding/Support:} This study was supported by CGBio.
\textbf{Author Contribution:} Conceptualization: PGC, SJY, DAS, MCC; Data curation: PGC, SJY, DAS, MCC; Formal analysis: PGC, SJY, DAS, MCC; Methodology: PGC, SJY, DAS, MCC; Visualization: PGC, SJY, DAS, MCC; Writing – original draft: PGC, SJY, DAS, MCC; Writing – review & editing: PGC, SJY, DAS, MCC.
\textbf{ORCID}
Pyung Goo Cho: 0000-0001-7087-8597
Seon-Jin Yoon: 0000-0002-3255-5081
Dong Ah Shin: 0000-0002-5225-4083
Min Cheol Chang: 0000-0002-7629-7213

\textbf{REFERENCES}
3. Chang MC. Mild cognitive impairment and major depressive disorder as confounders in the study on association between chronic low back pain and brain atrophy. Pain 2023;164:e237.

www.e-neurospine.org


Comparison of Transoral Anterior Jefferson-Fracture Reduction Plate and Posterior Screw-Rod Fixation in C1-Ring Osteosynthesis for Unstable Atlas Fractures

Mandi Cai1,2,*, Yifeng Wu1,2,*, Rencai Ma2, Junlin Chen1,2, Zexing Chen1,2, Chenfu Deng1,2, Xinzhao Huang1,2, Xiangyang Ma1,2, Xiaobao Zou2

1 The First School of Clinical Medicine, Southern Medical University, Guangzhou, China
2 Department of Orthopedics, General Hospital of Southern Theatre Command of PLA, Guangzhou, China

Objective: To compare the clinical outcomes of transoral anterior Jefferson-fracture reduction plate (JeRP) and posterior screw rod (PSR) surgery for unstable atlas fractures via C1-ring osteosynthesis.

Methods: From June 2009 to June 2022, 49 consecutive patients with unstable atlas fractures were treated by transoral anterior JeRP fixation (JeRP group) or PSR fixation (PSR group) and followed up at General Hospital of Southern Theatre Command of PLA; 30 males and 19 females were included. The visual analogue scale (VAS) score, Neck Disability Index (NDI), distance to anterior arch fracture (DAAF), distance to posterior arch fracture (DPAF), lateral mass displacement (LMD), Redlund-Johnell value, postoperative complications, and fracture healing rate were retrospectively collected and statistically analyzed.

Results: Compared with that in the PSR group, the bleeding volume in the JeRP group was lower, and the length of hospital stay was longer. The VAS scores and NDIs of both groups were significantly improved after surgery. The postoperative DAAF and DPAF were significantly smaller after surgery in both groups. Compared with the significantly shorter DPAF in the PSR group, the JeRP group had a smaller DAAF, shorter LMDs and larger Redlund-Johnell value postoperatively and at the final follow-up. The fracture healing rate at 3 months after surgery was significantly greater in the JeRP group (p < 0.05).

Conclusion: Both C1-ring osteosynthesis procedures for treating unstable atlas fractures yield satisfactory clinical outcomes. Transoral anterior JeRP fixation is more effective than PSR fixation for holistic fracture reduction and short-term fracture healing, but the hospital stay is longer.

Keywords: Atlas fracture, Unstable fractures, Transoral anterior approach, Posterior approach, C1-ring osteosynthesis

INTRODUCTION

Atlas fractures account for approximately 2%–13% of cervical fractures and approximately 1%–2% of total spinal fractures. Currently, conservative treatment is favored for stable atlas fractures. However, how to handle unstable atlas fractures, which include all fractures except anterior arch single fractures without transverse ligament rupture or posterior arch fractures, is debatable. Conventional atlantoaxial or upper cervical fusion can result in loss of cervical motor function and lower quality of life, while nonsurgical treatment methods are associated with a high nonunion rate of atlas fracture. In 2004, Ruf et al.
first proposed C1-ring osteosynthesis via a transoral approach to treating unstable Jefferson fractures. Subsequently, C1-ring osteosynthesis has gradually become one of the ideal procedures for treating unstable atlas fractures because it can instantly reduce and fix fractures while preserving the motor function of the upper cervical spine.5,8–10

Currently, C1-ring osteosynthesis is most commonly performed via transoral anterior plate fixation9,11 and posterior screw-rod fixation.12,11 Previous studies have aimed to investigate the effectiveness of each C1-ring osteosynthesis procedure. However, to the best of our knowledge, no comparative studies of these 2 procedures have been reported. In this study, we retrospectively analyzed the clinical data of 49 patients with unstable atlas fractures treated with C1-ring osteosynthesis via the use of the transoral anterior Jefferson-fracture reduction plate (JeRP)9,15 or the conventional posterior screw rod (PSR). We compared the clinical efficacy of these 2 procedures, which is important for the selection of clinical procedures.

MATERIALS AND METHODS

1. Patient Selection

From June 2009 to June 2022, a total of 49 consecutive patients who met the inclusion and exclusion criteria and were treated by C1-ring osteosynthesis at General Hospital of Southern Theatre Command of PLA were recruited and followed up (Table 1). The detailed screening criteria were (1) diagnosis of traumatic atlas fracture by clinical and imaging examination, (2) Landells type II or III fracture, (3) Dickman type I or II injury, (4) no previous cervical disease or cervical trauma or operation, and (5) signed informed consent forms. The exclusion criteria were (1) other cervical vertebral fractures or mixed fractures, (2) chronic or nonunion atlas fractures, (3) neurological dysfunction of spinal cord injury, (4) inability to tolerate surgery, and (5) incomplete follow-up data.

The indications for anterior JeRP surgery were (1) diagnosis of unstable atlas fracture, (2) obvious symptoms that cannot be treated conservatively, (3) no contraindications to anterior surgery, such as oral inflammation or previous history of oral surgery affecting the anterior approach, and (4) after the preoperative conversation, patients independently selected the approach and signed the informed consent form. Indications for the PSR surgery included (1) diagnosis of unstable atlas fracture, (2) obvious symptoms that cannot be treated conservatively, (3) no contraindications to posterior surgery such as infection or inflammation of the posterior cervical tissue space, (4) unsuitable for anterior surgery such as patients with oral infections, oral deformities, low immunity, small intraoral volume, high atlantoaxial position, long fracture line extending lateral mass, or when the posterior pharyngeal wall tissue is too thin to cover the plate, etc., and (5) after the preoperative conversation, patients independently selected the approach and signed the informed consent form. All surgeries were performed by one skilled senior spine surgeon and his team.

Thirty-one patients were treated with transoral anterior JeRP fixation (JeRP group), including 19 males and 12 females, with a mean age of 43.9 ± 14.0 years. The mean follow-up time was 20.1 ± 10.6 months (range, 12–51 months); the other 18 patients were treated with PSR fixation (PSR group), which included 11 males and 7 females, with a mean age of 42.4 ± 12.4 years. The mean follow-up time was 21.5 ± 1.9 days (range, 12–36 months). All patients had a history of trauma, with a mean injury time of 4.9 ± 3.3 days in the JeRP group and 4.7 ± 3.5 days in the PSR group. The types of trauma included falling, motor vehicle accidents and crashing. All patients had symptoms of neck and occipital pain but no neurological symptoms. The

The indications for anterior JeRP surgery were (1) diagnosis of unstable atlas fracture, (2) obvious symptoms that cannot be treated conservatively, (3) no contraindications to anterior surgery, such as oral inflammation or previous history of oral surgery affecting the anterior approach, and (4) after the preoperative conversation, patients independently selected the approach and signed the informed consent form. Indications for the PSR surgery included (1) diagnosis of unstable atlas fracture, (2) obvious symptoms that cannot be treated conservatively, (3) no contraindications to posterior surgery such as infection or inflammation of the posterior cervical tissue space, (4) unsuitable

### Table 1. Comparison of 2 groups of baseline information

<table>
<thead>
<tr>
<th>Variable</th>
<th>JeRP group (n = 31)</th>
<th>PSR group (n = 18)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>43.9 ± 14.0</td>
<td>42.4 ± 12.4</td>
<td>0.722</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>19 (61.3)</td>
<td>11 (61.1)</td>
<td>0.812</td>
</tr>
<tr>
<td>Female</td>
<td>12 (38.7)</td>
<td>7 (38.9)</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>21.5 ± 1.9</td>
<td>21.2 ± 1.6</td>
<td>0.577</td>
</tr>
<tr>
<td>Symptoms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occipital neck pain</td>
<td>31 (100)</td>
<td>18 (100)</td>
<td>&gt; 0.999</td>
</tr>
<tr>
<td>AIS grade</td>
<td>2.1 ± 0.3</td>
<td>2.3 ± 0.5</td>
<td>0.091</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>5 (16.1)</td>
<td>3 (16.7)</td>
<td>0.961</td>
</tr>
<tr>
<td>Diabetes</td>
<td>6 (19.4)</td>
<td>3 (16.7)</td>
<td>0.815</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
<td>9 (29.0)</td>
<td>5 (27.7)</td>
<td>0.925</td>
</tr>
<tr>
<td>Combined other trauma</td>
<td>7 (22.5)</td>
<td>3 (16.7)</td>
<td>0.726</td>
</tr>
<tr>
<td>Injury time (day)</td>
<td>4.9 ± 3.3</td>
<td>4.7 ± 3.5</td>
<td>0.882</td>
</tr>
<tr>
<td>Type of injury</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Falling</td>
<td>10 (32.3)</td>
<td>7 (38.9)</td>
<td>0.638</td>
</tr>
<tr>
<td>MVA</td>
<td>15 (48.4)</td>
<td>9 (50.0)</td>
<td>0.913</td>
</tr>
<tr>
<td>Crushing</td>
<td>6 (19.3)</td>
<td>2 (11.1)</td>
<td>0.693</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard error or number (%). JeRP, Jefferson-fracture reduction plate; PSR, posterior screw-rod; BMI, body mass index; AIS, American Spinal Injury Association Impairment Scale; MVA, motor vehicle accident.
American Spinal Injury Association Impairment Scale grade was used to assess the severity of cervical spine injury. Common preoperative comorbidities in patients included osteoporosis, diabetes, cardiovascular disease, and combined other trauma. All patients had anterior and posterior atlas arch fractures as shown by cervical spine x-ray, computed tomography, and magnetic resonance imaging examinations without significant spinal cord compression (Table 1). This study was approved by the Ethical Review Committee of General Hospital of Southern Theatre Command of PLA (2023011).

2. Surgical Procedure

1) Transoral anterior JeRP fixation

Preoperative preparation: All patients were instructed to gargle 3 to 6 times daily with 0.02% chlorhexidine acetate before surgery. A professional dental cleaning procedure was also performed. A nasogastric feeding tube was placed 1 hour before surgery, and prophylactic broad-spectrum antibiotics were applied conventionally 30 minutes before surgery.

Surgical procedures: While under general anesthesia with nasotracheal intubation, the patient was placed in the supine position, and the neck was situated slightly hyperextended with skull traction. After routine oral cleaning and disinfection, the oral cavity was opened by a Codman retractor. A longitudinal incision of 3–4 cm was then made in the median posterior pharyngeal wall to incise the mucosa and split the longitudinal muscles. After the subperiosteal layer of the muscle and the prevertebral fascia were separated to reveal the anterior C1 arch and the lateral mass, the fracture was exposed. An appropriately sized plate was selected and placed transversely in front of the atlas. The wider side of the JeRP was placed on the lateral mass near the fracture line and fixed by 2 screws. Then, one temporary reduction screw was inserted through the slide hole of the plate onto the anterior arch of the other fracture side. After the compressing reduction forceps was installed, one arm hooked the reduction hole of the JeRP and the other arm clasped the reduction screw and then closed the handle of the forceps to achieve fracture end closure. (C, D) Coronal and axial pictures present the fixation: after the fracture was satisfactorily reduced, the remaining anterior arch screws and lateral mass screws on the other side were fixed.

Postoperative management: After surgery, the tracheal airway cannula was removed after 24–48 hours, and the nasal feeding tube was left for 1 week. The 0.02% chlorhexidine ace-
both sides, and only one side was tightened. The fracture was repositioned by lateral compression with a pair of reduction forceps, after which the nut was locked on the other side. After the placement of the plate and screws was verified to be satisfactory by C-arm fluoroscopy, the incision was closed (Fig. 3).

Postoperative management: The drainage tube was removed when the postoperative drainage volume was less than 50 mL/24 hr. The sutures were removed 12–14 days after surgery. Cervical spine x-ray and CT examination were performed 1 week after surgery. The cervical brace was fixed and protected for 3 months, and regular follow-up was performed (Fig. 4).

3. Observed Indexes
The surgical time, bleeding volume, hospital stay and postop-
Comparison of 2 C1-Ring Osteosynthesis
Cai M, et al.

https://doi.org/10.14245/ns.2347230.615

www.e-neurospine.org

operative complications were recorded. The visual analogue scale (VAS) was used to assess the degree of occipital neck pain; the Neck Disability Index (NDI) was used to evaluate the cervical spine function. The distance of anterior arch fracture (DAAF), distance of posterior arch fracture (DPAF), lateral mass displacement (LMD), and the Redlund-Johnell value were measured before and after surgery to assess reduction (Fig. 5). All imaging indicators were measured and evaluated by 2 independent observers who were not involved with the study. Postoperative cervical spine x-ray and CT were reviewed to evaluate the internal fixation and fracture healing. We generally recommend our patients have follow-up imaging at 3, 6, and 12 months after surgery, and then every 12 months or at times of discomfort. All cases were followed up with imaging for at least 1 year.

4. Statistical Analysis
Statistical analysis was performed using IBM SPSS Statistics ver. 21.0 (IBM Co., Armonk, NY, USA). The measurement data were expressed as mean ± standard deviation. The pre- and post-surgery comparisons between the 2 groups were performed using the paired t-test, and the independent sample t-test was used for intergroup comparisons. The enumeration data were expressed as number (%), and the χ² test was used for intergroup comparisons. The level of significance was set at p < 0.05. Intraobserver and interobserver reliability was assessed by kappa (κ) statistic and intraclass correlation coefficient (ICC).

RESULTS

1. Patient Characteristics
All 49 patients completed the surgery successfully. The mean surgical times were 104.5 ± 14.7 minutes (JeRP group) and 110.6 ± 18.5 minutes (PSR group), with no significant difference between the 2 groups (p = 0.214). The mean blood loss in the JeRP group was lower (69.4 ± 20.2 mL vs. 103.3 ± 25.4 mL, p < 0.001). The mean length of stay in the PSR group was 8.1 ± 1.6 days, which was significantly shorter than the 11.6 ± 1.6 days in the JeRP group (p < 0.001).

2. Clinical Symptom Parameters
There was no significant difference in the preoperative VAS scores (5.5 ± 1.1 vs. 5.3 ± 1.0, p = 0.569) or NDI (58.1% ± 4.1% vs. 57.6% ± 4.2%, p = 0.640) between the 2 groups. The postoperative (JeRP: 5.5 ± 1.1 vs. 0.8 ± 0.8, p < 0.001; PSR: 5.3 ± 1.0 vs. 1.1 ± 0.8, p < 0.001) and final follow-up (JeRP: 5.5 ± 1.1 vs. 0.4 ± 0.5, p < 0.001; PSR: 5.3 ± 1.0 vs. 0.3 ± 0.5, p < 0.001) VAS scores of both groups were significantly greater than the preoperative scores, with no difference between groups (0.8 ± 0.8 vs. 1.1 ± 0.8, p = 0.312; 0.4 ± 0.5 vs. 0.3 ± 0.5, p = 0.714). The NDI was significantly lower in both groups after surgery (JeRP: 58.1% ± 4.1% vs. 26.8% ± 4.0%, p < 0.001; PSR: 57.6% ± 4.2% vs. 26.3% ± 4.4%, p < 0.001) and at final follow-up (JeRP: 58.1% ± 4.1% vs. 1.7% ± 2.4%, p < 0.001; PSR: 57.6% ± 4.2% vs. 2.2% ± 2.4%, p < 0.001), with no differences between the groups (26.8% ± 4.0% vs. 26.3% ± 4.4%, p = 0.721; 1.7% ± 2.4% vs. 2.2% ± 2.4%, p = 0.501).

3. Radiographical Parameters
There were no differences in preoperative parameters, such as DAAF, DPAF, LMD, or the Redlund-Johnell value, between the 2 groups. Postoperative DAAF and DPAF were significantly smaller in the PSR group (7.1 ± 2.0 mm vs. 4.6 ± 4.9 mm, p = 0.006;
Fig. 4. A 61-year-old male with combined fractures of the anterior and posterior atlantoaxial arches was treated by posterior C1-ring osteosynthesis using the posterior screw rod. (A) Preoperative open-mouth x-ray imaging showed displacement of the lateral masses. (B, C) The axial images of computed tomography (CT) scan and 3-dimensional reconstruction revealed fractures of the anterior and posterior arches of the atlas with displacement of the lateral mass. (D) Preoperative magnetic resonance imaging showed no spinal cord compression. Red arrows showed the fracture breaks. (E, F) Postoperative open-mouth and lateral x-ray imaging showed good C1–2 alignment with adequate placement of PSR. (G, H) Postoperative CT image after surgery revealed reduction of fracture. (I, J) Open-mouth and lateral x-ray images at 9 months after surgery showed no loosening of the rod and screws. (K, L) CT images at 9 months after surgery revealed solid bone fusion.

2.9 ± 1.9 mm vs. 1.1 ± 1.4 mm, p < 0.001). In the JeRP group, the postoperative DAAF was also significantly smaller (7.2 ± 3.1 mm vs. 1.5 ± 1.6 mm, p < 0.001), whereas there was no significant difference in the DPAF reduction (2.2 ± 1.4 mm vs. 2.0 ± 1.4 mm, p = 0.408). However, the postoperative DAAF in the JeRP group was much smaller than that in the PSR group (1.5 ± 1.6 mm vs. 4.6 ± 4.9 mm, p = 0.002), while the postoperative DPAF in the JeRP group was greater than that in the PSR group (2.0 ± 1.4 mm vs. 1.1 ± 1.4 mm, p = 0.028). The LMDs of the JeRP group were 5.6 ± 2.6 mm, 0.9 ± 1.4 mm, and 0.5 ± 1.1 mm before surgery,
Comparison of 2 C1-Ring Osteosynthesis
Cai M, et al.

Comparison of 2 C1-Ring Osteosynthesis
Cai M, et al.

https://doi.org/10.14245/ns.2347230.615

www.e-neurospine.org

550

after surgery, and at the last follow-up, respectively; and those of the PSR group were 5.5 ± 3.3 mm, 2.2 ± 2.8 mm, and 1.9 ± 3.0 mm, respectively. The LMDs in both groups were significantly smaller after surgery (JeRP: p < 0.001; PSR: p < 0.001), and the postoperative and final follow-up LMDs in the JeRP group were significantly smaller than those in the PSR group (0.9 ± 1.4 mm vs. 2.2 ± 2.8 mm, p = 0.042; 0.5 ± 1.1 mm vs. 1.9 ± 3.0 mm, p = 0.025). The Redlund-Johnell values in the JeRP group were 37.9 ± 4.9 mm, 40.9 ± 4.3 mm, and 40.7 ± 4.2 mm before surgery, after surgery, and at the last follow-up, respectively; and these values equal 37.3 ± 2.7 mm, 38.5 ± 2.1 mm, and 38.6 ± 2.1 mm, respectively, in the PSR group. After surgery (JeRP: 37.9 ± 4.9 mm vs. 40.9 ± 4.3 mm, p < 0.001; PSR: 37.3 ± 2.7 mm vs. 38.5 ± 2.1 mm, p < 0.001) and at the last follow-up (JeRP: 37.9 ± 4.9 mm vs. 40.7 ± 4.2 mm, p < 0.001; PSR: 37.3 ± 2.7 mm vs. 38.6 ± 2.1 mm, p = 0.002), the Redlund-Johnell value significantly improved in both groups, and the Redlund-Johnell value was greater in the JeRP group than in the PSR group (40.9 ± 4.3 mm vs. 38.5 ± 2.1 mm, p = 0.014; 40.7 ± 4.2 mm vs. 38.6 ± 2.1 mm, p = 0.029). Both observers’ intraobserver and interobserver reliability showed well in all radiographic parameters measured, including preoperative DAAF (ICC, 0.91; ICC, 0.86), postoperative DAAF (ICC, 0.89; ICC, 0.83), preoperative DPAF (ICC, 0.90; ICC, 0.84), postoperative DPAF (ICC, 0.91; ICC, 0.81), preoperative LMDs (ICC, 0.84; ICC, 0.80), postoperative LMDs (ICC, 0.85; ICC, 0.80), last followed-up LMDs (ICC, 0.88; ICC, 0.81), preoperative R-J value (ICC, 0.90; ICC, 0.81), postoperative R-J value (ICC, 0.91; ICC, 0.83), and last followed-up R-J value (ICC, 0.87; ICC, 0.84).

4. Complications and Healing

There was no difference in the incidence of complications after surgery between the 2 groups (atlantoaxial instability: p = 0.9; implants loosening: p = 0.526; screw misplacement: p = 0.13). Bone fusion was confirmed by continuous bone bridge formation without a visible fracture line at the fracture site on x-ray or thin-layer CT (0.9 mm) images.17,18 Fracture healing was independently diagnosed by 2 orthopedic surgeons based on imaging. The fracture healing rates at 3, 6, and 12 months after surgery were 61.3%, 83.9%, and 90.3%, respectively, in the JeRP group. Similarly, the rates were 22.2%, 66.7%, and 83.3% in the PSR group, respectively. The fracture healing rate at 3 months after surgery was greater in the JeRP group (61.3% vs. 22.2%, p = 0.008), but no differences were found between the 2 groups at 6 and 12 months after surgery (83.9% vs. 66.7%, p = 0.286; 90.3% vs. 83.3%, p = 0.656). Two patients in the JeRP group and one patient in the PSR group exhibited atlantoaxial instability after surgery, which were revised by posterior atlantoaxial fixation and fusion; 2 patients with osteoporosis in the JeRP group developed implant loosening and were revised by posterior atlantoaxial fixation fusion after the anterior implants were removed; 2 cases of screw misplacement occurred in the PSR group after surgery and was revised by adjusting the screw position (Table 2). Well intraobserver and interobserver agreements were found for bone fusion rate at 3-month (κ = 0.86; κ = 0.79), 6-month (κ = 0.84; κ = 0.81) and 12-month (κ = 0.82; κ = 0.78) follow-up.

DISCUSSION

In this study, we compared imaging and clinical indices between 2 types of C1-ring osteosynthesis. We found that there was no difference in surgical time between the 2 groups. The PSR group had more intraoperative blood loss but a shorter length of stay. The posterior pharyngeal wall is adjacent to the prevertebral fascia without much neuromuscular tissue. Compared with the posterior approach, the anterior approach does not require more dissection of paravertebral tissue for expo-

Fig. 5. Schematic diagram of measurement indicators in computed tomography (CT) scan. (A) Axial CT image shows the distance of anterior arch fracture equals “a”; the distance of posterior arch fracture equals “b”. (B) Coronal CT image shows the lateral mass displacement equals “c1+c2”. (C) Sagittal CT image shows the Redlund-Johnell value equals “d”. 
Table 2. Comparisons of clinical data before and after surgery between the 2 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>JeRP group (n = 31)</th>
<th>PSR group (n = 18)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical time (min)</td>
<td>104.5 ± 14.7</td>
<td>110.6 ± 18.5</td>
<td>0.214</td>
</tr>
<tr>
<td>Bleeding (mL)</td>
<td>69.4 ± 20.2</td>
<td>103.3 ± 25.4</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Length of stay (day)</td>
<td>11.6 ± 1.6</td>
<td>8.1 ± 1.6</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Follow-up (mo)</td>
<td>20.1 ± 10.6</td>
<td>19.4 ± 8.0</td>
<td>0.813</td>
</tr>
<tr>
<td>DAAF (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before operation</td>
<td>7.2 ± 3.1</td>
<td>7.1 ± 2.0</td>
<td>0.865</td>
</tr>
<tr>
<td>At discharge</td>
<td>1.5 ± 1.6</td>
<td>4.6 ± 4.9</td>
<td>0.002*</td>
</tr>
<tr>
<td>DPAF (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before operation</td>
<td>2.2 ± 1.4</td>
<td>2.9 ± 1.9</td>
<td>0.370</td>
</tr>
<tr>
<td>At discharge</td>
<td>2.0 ± 1.4</td>
<td>1.1 ± 1.4</td>
<td>0.028*</td>
</tr>
<tr>
<td>LMD (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before operation</td>
<td>5.6 ± 2.6</td>
<td>5.5 ± 3.3</td>
<td>0.900</td>
</tr>
<tr>
<td>At discharge</td>
<td>0.9 ± 1.4</td>
<td>2.2 ± 2.8</td>
<td>0.042*</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>0.5 ± 1.1</td>
<td>1.9 ± 3.0</td>
<td>0.025*</td>
</tr>
<tr>
<td>R-J value (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before operation</td>
<td>37.9 ± 4.9</td>
<td>37.3 ± 2.7</td>
<td>0.638</td>
</tr>
<tr>
<td>At discharge</td>
<td>40.9 ± 4.3</td>
<td>38.5 ± 2.1</td>
<td>0.014*</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>40.7 ± 4.2</td>
<td>38.6 ± 2.1</td>
<td>0.029</td>
</tr>
<tr>
<td>VAS score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before operation</td>
<td>5.5 ± 1.1</td>
<td>5.3 ± 1.0</td>
<td>0.569</td>
</tr>
<tr>
<td>At discharge</td>
<td>0.8 ± 0.8</td>
<td>1.1 ± 0.8</td>
<td>0.312</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>0.4 ± 0.5</td>
<td>0.3 ± 0.5</td>
<td>0.714</td>
</tr>
<tr>
<td>NDI (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before operation</td>
<td>58.1 ± 4.1</td>
<td>57.6 ± 4.2</td>
<td>0.640</td>
</tr>
<tr>
<td>At discharge</td>
<td>26.8 ± 4.0</td>
<td>26.3 ± 4.4</td>
<td>0.721</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>1.7 ± 2.4</td>
<td>2.2 ± 2.4</td>
<td>0.501</td>
</tr>
<tr>
<td>Complications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlantoaxial instability</td>
<td>2 (6.5)</td>
<td>1 (5.6)</td>
<td>0.900</td>
</tr>
<tr>
<td>Implants loosing</td>
<td>2 (6.5)</td>
<td>0 (0)</td>
<td>0.526</td>
</tr>
<tr>
<td>Screw misplacement</td>
<td>0 (0)</td>
<td>2 (11.1)</td>
<td>0.130</td>
</tr>
<tr>
<td>Fracture healing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-Month follow-up</td>
<td>19 (61.3)</td>
<td>4 (22.2)</td>
<td>0.008*</td>
</tr>
<tr>
<td>6-Month follow-up</td>
<td>26 (83.9)</td>
<td>12 (66.7)</td>
<td>0.286</td>
</tr>
<tr>
<td>12-Month follow-up</td>
<td>28 (90.3)</td>
<td>15 (83.3)</td>
<td>0.656</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard error or number (%). JeRP, Jefferson-fracture reduction plate; PSR, posterior screw-rod; DAAF, distance of anterior arch fracture; DPAF, distance of posterior arch fracture; LMD, lateral mass displacement; R-J value, Redlund-Johnell value; VAS, visual analogue scale; NDI, Neck Disability Index.

*p < 0.05, statistically significant differences.
groups showed significant postoperative improvements in VAS scores and NDI, there was no statistically significant difference between the 2 groups. The effects of the 2 procedures on improving clinical symptoms were similar.

There was no difference in the overall complication rate between the 2 groups. Two of the 5 patients with osteoporosis in the JeRP group experienced internal fixation loosening at the long-term follow-up. The diagnostic criterion for osteoporosis was a DXA test with the T value less than -2.5. The anterior lateral mass screw fixation was performed in transoral anterior JeRP fixation, whose fixation strength is weaker than that of posterior pedicle screw fixation, especially in osteoporosis patients. Three cases of postoperative atlantoaxial instability were observed in both groups. This finding is similar to that of Tu et al., which suggests that C1-ring osteosynthesis may not be infallible. Because injuries of the transverse ligament itself are difficult to heal over time. At the same time, patients may also have injuries of muscles and other ligaments while the bone-muscle-ligament complex plays a crucial role in joint stabilization. Therefore, we suggest that preoperative 3-dimensional CT reconstruction of the ligaments around the atlantoaxial joint should be performed to more definitively judge ligament injuries, including the transverse ligament and others. It is also recommended that dynamic cervical x-rays be taken to judge the potential risk of atlantoaxial instability at postoperative follow-up. No postoperative complications such as dysphagia, hoarseness, and any disturbance in drinking and eating occurred in the JeRP group. We consider possible reasons including the small size of the steel plates used in JeRP and less postoperative stimulation of the esophagus and trachea. Besides, the indications for JeRP do not involve lateral mass fractures, with small surgical exposure and low risk of peripheral nerve injury. All surgeries were performed by an experienced senior surgeon, and surgical proficiency is also an important factor in minimizing complications such as peripheral nerve injuries.

In 1919, Jefferson\(^1\) first proposed one classification method for atlas fractures based on the mechanism of injury and anatomical site. In 1988, Landells and Van Peteghem\(^2\) categorized atlas fractures into 3 types: anterior or posterior arch fractures, anterior-posterior arch fractures, and lateral mass fractures. Subsequently, in 1991, Levine and Edwards\(^3\) also divided atlas fractures into 3 types, namely, posterior arch fractures, lateral mass fractures, and both anterior and posterior arch fractures. Both Landells and Levine-Edwards classifications take fracture morphology into consideration and are widely used.

According to the guidelines of the American Congress of Neurological Surgeons (CNS), the integrity of the transverse ligament is the main basis for assessing the stability of atlas fractures, which means that atlas fractures with the intact transverse ligament are considered stable fractures; otherwise, fractures with the ruptured transverse ligament are considered unstable.\(^4\) This method is now widely accepted. Then Dickman et al.\(^5\) also described atlantoaxial transverse ligament injuries, where type I refers to rupture in the middle of the transverse ligament and type II refers to avulsion fracture of the transverse ligament at the attachment point of the lateral mass. However, through retrospective analysis of a large number of cases, Lee and Woodring\(^6\) concluded that only anterior arch single fractures without combined transverse ligament rupture or simple posterior arch fractures should be considered stable fractures, while all other types of fractures are unstable fractures. The reason is that when multiple fractures of the anterior arch exist, even if the transverse ligament is intact, the anterior arch is too weak to restrain the odontoid from moving forward, thus leading to posterior dislocation. Additionally, when the anterior and posterior arches are both fractured, although the intact transverse ligament may prevent the separation of the lateral mass, there is a possibility of rotational displacement of the fractured mass using the attachment point of the transverse ligament as the fulcrum. Radcliff et al.\(^7\) pointed out that the traditional Spence Rule which states that a displacement of the C1 lateral masses by > 6.9–8.1 mm suggests the loss of transverse ligament integrity, can be inaccurate. Based on the above, we believe that the method of assessing whether a C1 fracture is stable based on the integrity of the transverse ligament may not be accurate. We are inclined to Lee et al., that is, all fractures except for anterior arch single fracture without transverse ligament rupture and posterior arch fracture are unstable fractures, which is one of the exclusion criteria of this study.

Usually, separation and displacement of the atlas increase the space of the spinal canal, and patients with neurological dysfunction are rare. Therefore, reduction and stabilization are the most important aspect in the treatment of unstable atlas fractures. Nonoperative treatment is still recommended for patients with stable atlas fractures in the first stage. However, this therapy may not be appropriate for unstable fractures. One multicenter study indicated that surgical treatment was associated with a higher fusion rate, shorter fracture healing time, more favorable clinical outcomes, and better fracture reduction for unstable atlas fractures.\(^8\)

For surgical methods, atlantoaxial fusion has been used mostly in the past and has been tested in a timely manner to
obtain satisfactory results. However, this surgery sacrifices a significant portion of the motor function of the upper cervical spine, especially axial rotational motion, thus reduces the life quality of patients after surgery. In 2004, Ruf et al. first treated 6 Jefferson-fracture patients with transverse ligament rupture via transoral anterior screw-rod fixation as C1-ring osteosynthesis. With no postoperative atlantoaxial instability, all patients achieved good outcomes with preserved cervical motion. Subsequently, reports of C1-ring osteosynthesis for unstable atlas fractures, a physiological surgical fixation, have increased. However, whether transverse ligament rupture leads to late atlantoaxial instability has become a point of controversy for this surgery. Li et al. proposed the “buoy phenomenon,” suggesting that C1-ring osteosynthesis can restore the height of the occipital-atlantoaxial complex, which tightens the loose longitudinal ligaments to maintain the stability of the atlantoaxial joints. Koller et al. and Li-Jun et al. also reported in their biomechanical studies that C1-ring osteosynthesis can restore the stability of the atlantoaxial joints when combined with atlas fracture and transverse ligament rupture. A study showed that posterior C1-ring osteosynthesis is superior to atlantoaxial fusion in terms of preserving the physiological function of the cervical spine and long-term relief of neck pain.

Now, the mainstream approaches for C1-ring osteosynthesis include the posterior and transoral anterior approaches. Posterior C1-ring osteosynthesis is typically represented by horizontal screw-rod fixation. Ruf et al. used a transoral anterior screw rod system to fix the atlas with satisfactory results, but this method did not fit the anatomical features of the posterior atlas structures. Transoral anterior C1-ring osteosynthesis is now most often performed with plate fixation. The JeRP fixation system based on the anatomical parameters of the anterior atlantoaxial spine was designed by General Hospital of Southern Theatre Command of PLA for transoral anterior C1-ring osteosynthesis and has showed satisfactory results in clinical application. It is important to mention that for the JeRP technique, when exposing C1, especially the lateral side of the lateral mass, some important anatomical structures around C1, such as the internal carotid artery (IC) and hypoglossal nerve and so on should receive more attentions. In cases of poor preoperative conditions, such as elderly patients with atherosclerosis of IC or in cases where the fracture extends far enough to make exposure difficult, intraoperative maneuvers should be performed with greater caution.

Our study presents some limitations. First, as our study was retrospective in nature, selection bias was inevitable in the case and procedure selection process. Particularly, for different procedures we did not use more sophisticated algorithms to screen for, and these make our results potentially biased and need to be confirmed by more prospective studies in the future. In addition, the sample size in this study was small, and a larger number of patients is needed to validate the differences between the 2 techniques. Finally, the follow-up period was short, and additional attention should be given to distant complications in the future.

CONCLUSION

Both anterior transoral JeRP and PSR fixation in C1-ring osteosynthesis for unstable atlas fractures have satisfactory clinical efficacy. Transoral anterior JeRP fixation provides better comprehensive fracture reduction and facilitates short-term fracture healing but has a longer hospital stay and a greater risk of potential infection. For patients with combined osteoporosis, internal fixation is more prone to loosening. PSR fixation has a low risk of internal fixation loosening but poorly reduces anterior arch fractures. Thus, increased fracture healing time is needed. Each type of surgery has advantages, and surgeons should design a reasonable surgical program based on the patient’s individual situation and the surgical techniques they have mastered.

NOTES

Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: XM, XZ; Formal analysis: MC, YW, JC, CD; Investigation: MC, YW, RM, ZC, XH; Methodology: MC, YW, RM, ZC, CD, XH, XZ; Project administration: XM, XZ; Writing – original draft: MC; Writing – review & editing: MC, YW, XM, XZ.

ORCID
Mandi Cai: 0000-0003-2581-4353
Chenfu Deng: 0000-0003-0253-5650
Xiaobao Zou: 0000-0002-2101-1900

REFERENCES


Knockdown of *best1* Gene in Zebrafish Caused Abnormal Neuronal and Skeletal Development - A Subtype of Craniovertebral Junction Malformation?

Zhenlei Liu1,2,*, Kang Li1,2,*, Kai Wang1,2, Lei Zhang1,2, Shanhang Jia1,2, He Wang1,2, Fengzeng Jian1,2, Hao Wu1,2

1Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Beijing, China
2Spine Center, China International Neuroscience Institute (China-INI), Beijing, China

**Objective:** To investigate the developmental defects caused by knockdown of *best1* gene in zebrafish as a model for a subtype of craniovertebral junction (CVJ) malformation.

**Methods:** Two antisense morpholinos (MOs) were designed targeting zebrafish *best1* to block translation (ATG-MO) or to disrupt splicing (I3E4-MO). MOs were microinjected into fertilized one-cell embryos. Efficacy of splicing MO was confirmed by reverse transcription-polymerase chain reaction. Phenotypes were analyzed and quantified by microscopy at multiple developmental stages. Neuronal outgrowth was assed in transgenic zebrafish expressing green fluorescent protein in neurons. Skeletal ossification was visualized by Calcein staining.

**Results:** Knockdown of *best1* resulted in zebrafish embryos with shorter body length, curved axis, low survival rate, microcephaly, reduced eye size, smaller head and brain, impaired neuronal outgrowth, and reduced ossification of craniofacial and vertebral bone.

**Conclusion:** *Best1* gene plays critical roles in ophthalmologic, neurological and skeletal development in zebrafish. A patient with a premature stop codon in *BEST1* gene exhibited similar phenotypes, implying a subtype of CVJ malformation.

**Keywords:** Craniovertebral junction malformation, Genetic variations, *BEST1* gene, Zebrafish, Skeletal ossification

**INTRODUCTION**

The craniovertebral junction (CVJ) is a highly complex functional unit consisting of the base of the occipital bone, the atlas (C1), the axis (C2) and the soft tissues around these skeletal structures.1,2 It provides complex motion of upper cervical region and protects the crucial area of the brainstem and upper cervical spinal cord. CVJ skeletal malformation can result in structural instability causing spectrum of neurological deficits including sensory and motion dysfunction, or even life-threatening dyspnea and cardiac arrest.3 The spectrum of CVJ malformation includes atlantooccipital fusion, atlantoaxial dislocation, basilar invagination, platybasia, and vertebral artery variations.4 It is intuitive to attribute the etiology of CVJ malformation to disturbance of embryonic development. Both genetic and environmental factors can be involved. Homeobox (HOX)
genes are pivotal genetic regulators of embryonic axial development. For instance, inactivating Hoxd3 in mice causes abnormal fusion of the atlas vertebra and the occiput, which is homologous to atlantooccipital fusion in human. However, the genetic causes for the CVJ malformation were largely unclear. Several studies suggested genetic predisposition underlying CVJ malformation. CVJ malformation has been reported to be much more common in Down syndrome (MIM: 190685). Mucopolysaccharidosis type IVA (MIM: 253000), and Wiedemann-Steiner syndrome (MIM: 605130). Sadler et al. found that chromodomain (CHD) genes were associated with Chiari I malformation. Chauhan et al. reported mutations in the FBN1 gene may play a part in CVJ malformation. In the previous study, we performed exome sequencing on 92 sporadic CVJ malformation patients and found damaging mutations of the BEST1 gene were significantly enriched in the patient group (p = 5.8 × 10⁻⁶). Variants in BEST1 are associated with autosomal dominant vitelliform macular dystrophy-2 (VMD2 [MIM: 153700]). Inhibition of BEST1 resulted in an enhanced neuronal excitability in the cerebellum. These studies indicated a disease-causing role of BEST1 gene mutations for eyeball, central nervous system, and maybe skeletal system disorders.

Given that the BEST1 gene is highly conserved in zebrafish and human, and zebrafish has been proven to be suitable for the study of human skeletal disease, we designed best1 mor- pholinos (MOs) and performed knockdown studies to investigate the role of best1 gene in early zebrafish development.

MATERIALS AND METHODS

1. Zebrafish Care and Maintenance

Adult wild-type AB strain zebrafish were maintained at 28.5°C on a 14-hour light/10-hour dark cycle. Five to six pairs of zebrafish were set up for natural mating every time. On average, 200–300 embryos were generated. Embryos were maintained at 28.5°C in fish water (0.2% Instant Ocean Salt in deionized water). The embryos were washed and staged according to Kimmel et al. The establishment and characterization of HuC-EGFP transgenic lines has been described elsewhere. The zebrafish facility at SMOC (Shanghai Model Organisms Center, Inc., Shanghai, China) is accredited by the AAALAC (Association for Assessment and Accreditation of Laboratory Animal Care) International.

2. Zebrafish Microinjections and Phenotyping

It is imperative to identify a sequence within the best1 gene that is pivotal for its physiological function. We also avoid regions with high sequence homology to other genes, to reduce the chance of off-target effects. Gene Tools, LLC (http://www.gene-tools.com/) designed the MOs. Antisense MOs (GeneTools) were microinjected into fertilized one-cell stage embryos according to standard protocols. The sequences of the best1 translation-blocking and splice-blocking MOs were 5'-CTGCGGGAGAGTCAAACACCTCA-3' (ATG-MO) and 5'-GAGTACGTACCCGTCATCCTAGG-3' (I3E4-MO), respectively. The sequence for the standard control-MO was 5'-CCTCTTACCTCAGTTACATAAAT-3' (Gene Tools). The amount of the MOs used for injection was as follows: Control-MO and I3E4-MO, 4 ng per embryo; ATG-MO, 4 ng per embryo. Primers spanning best1 exon 3 (forward primer: 5'-GGAGATTGGAGGAGGCCGAGACT-3') and exon 6 (reverse primer: 5'-AAACCTGCGCAGCAATCAGAC-3') were used for reverse transcription-polymerase chain reaction (RT-PCR) analysis for confirmation of the efficacy of the I3E4-MO. The primer efla sequences used as the internal control were 5'-GGAAATTGGAGCAGACGCAAATAC-3' (forward) and 5'-GATACTGACCCGTCATAAC-3' (reverse). At 2-dpf and 5-dpf, embryos were anesthetized with 0.016% MS-222 (tricaine methanesulfonate, Sigma-Aldrich, St. Louis, MO, USA). Zebrafish were then oriented on lateral side (anterior, left; posterior, right; dorsal, top), and mounted with 3% methylcellulose in a depression slide for observation by fluorescence microscopy. The phenotypes of gross morphology were quantified analyzed.

3. Quantitative Real-Time PCR

Total RNA was extracted from 20 to 30 embryos per group with Trizol (Roche, Basel, Switzerland) according to the manufacturer’s instructions. RNA was reverse transcribed using the the PrimeScript RT reagent Kit with gDNA Eraser (Takara, Tokyo, Japan). Quantification of gene expression was performed in triplicates using Bio-rad iQ SYBR Green Supermix (Bio-Rad Laboratories Inc., Hercules, CO, USA) with detection on the Realplex system (Eppendorf Co., Enfield, CT, USA). Relative gene expression quantification was based on the comparative threshold cycle method (2^ΔΔCt) using efla as endogenous control gene. Primer sequences are listed in Supplementary Table 1.

4. Calcein Staining

To evaluate craniofacial structure formation in zebrafish, fertilized one-cell wild-type AB embryos were injected with best1-MO and control-MO. At 6-dpf, zebrafish were washed with fish water 3 times and immersed in 0.2% Calcein solution.
for 10 minutes. Then, zebrafish were rinsed thoroughly in fish water 3 times (5 minutes for each time) and anaesthetized with 0.016% MS-222 (tricaine methanesulfonate, Sigma-Aldrich). Zebrafish were then oriented on lateral side and ventral side, and mounted with 3% methylcellulose (Sigma-Aldrich) in a depression slide for observation by fluorescence microscopy.

5. Image Acquisition

Embryos and larvae were analyzed with Nikon SMZ 18 Fluorescence microscope and subsequently photographed with digital cameras. A subset of images was adjusted for levels, brightness, contrast, hue and saturation with Adobe Photoshop 7.0 software (Adobe, San Jose, CA, USA) to optimally visualize the expression patterns. Quantitative image analyses were processed using image based morphometric analysis (NIS-Elements D4.6, Nikon, Tokyo, Japan) and ImageJ software (U.S. National Institutes of Health, Bethesda, MD, USA; http://rsbweb.nih.gov/ij/). 10 animals were including in group and average signal per animal was calculated.

6. Statistical Analysis

All data are presented as mean ± standard error of the mean (SEM). Statistical analysis and graphical representation of the data were performed using GraphPad Prism 5.0 (GraphPad Software, San Diego, CA, USA). Statistical significance was performed using a Student t-test, analysis of variance or χ² test as appropriate.

RESULTS

1. Effectiveness of best1 I3E4-MO

Customized MOs were microinjected into one-cell stage embryos (Supplementary Fig. 1A). Two specific antisense MOs targeting zebrafish best1 gene were designed to prevent either the translation of best1 mRNA (ATG-MO) or proper splicing of intron3 and exon4 (I3E4-MO) of best1 pre-mRNA (Supplementary Fig. 1B). We confirmed that I3E4-MO prevented proper splicing of intron3 at 30-hpf by RT-PCR. The additional band indicated that the splicing between intron 3 and exon 4 was altered with inserting part of intron3 (Supplementary Fig. 1C, lane 2). In addition, the expression of house-keeping genes ef1α was not influenced by both control-MO and I3E4-MO (Supplementary Fig. 1C, lanes 5 and 6). The expression level of wild-type best1 mRNA in I3E4-MO injected embryos was significantly lower than that of the control (p < 0.0001; Supplementary Fig. 1D) with quantitative RT-PCR. This demonstrated the effi-

![Fig. 1. Body length, body axis and survival rate of best1 zebrafish morphants. (A–F) Gross morphology of transgenic (HuC:EGFP) zebrafish embryos at 2-dpf. Compared with control morpholino (MO), knock down best1 causes shorter body length (C–F), curved body axis (C–F). (G) A time-course plot of percent survival in control versus best1 morphants for 3 days. (H) The percentage of embryos with development defects. (I, J) Quantification of body length (I) and curvature angles (J) of embryos. Columns, mean; standard error of the mean (n = 10; analysis of variance), ***p < 0.0001. Scale bar, 100 μm. dpf, days post fertilization.](https://doi.org/10.14245/ns.2347238.619)
2. Knockdown of best1 Causes Potent Developmental Defects in Zebrafish

We used transgenic (Tg) zebrafish (HuC: EGFP transgenic line) to observe the phenotypes of zebrafish embryos after best1 knockdown by injection of ATG-MO and I3E4-MO. To increase the objectivity of the measurements, we performed the quantification process in a blinded manner. This approach concealed the experimental groups during the measurement procedure, avoiding potential bias in the evaluation of phenotypic changes. Compared to control-MO injected embryos (Fig. 1A and B), both I3E4-MO injected (Fig. 1C and D) and ATG-MO injected (Fig. 1E and F) embryos exhibited significantly reduced body length (p < 0.0001) (Fig. 1G) and apparent curved body axis (p < 0.0001) (Fig. 1D, F, and H; Supplementary Fig. 2). The sur-

---

**Fig. 2.** Nervous system developmental abnormalities caused by best1 knockdown. (A–L) Gross morphology of transgenic (HuC:EGFP) zebrafish embryos at 2-dpf. Compared with control morpholino (MO), knockdown of best1 gene causes head and brain patterning defects (E, I, blue arrowheads), microcephaly (E, I, yellow dotted line), small eyes (E, I, black arrowheads), brain-size reductions (F, J, H, and L), and abnormal neuronal outgrowth (F, J, asterisks). (M–O) Quantification of the head area (M), brain area (N) and average motor axon length (O) shows significantly decrease in best1 morphants. Columns, mean; standard error of the mean (n = 10; analysis of variance), ***p < 0.0001. Scale bar, 100 μm. dpf, days post fertilization.
Knockdown of best1 Gene in Zebrafish

Fig. 3. Loss of best1 causes abnormal craniofacial structure and interorbital distance. (A–N) Representative bright field and fluorescent images of zebrafish head skeleton at 6-dpf. In vivo visualization of the skeleton is achieved by the administration of a fluorescent dye (Calcein) directly to the fish water. Dyes that bind to calcified matrix can be used to label the entire skeleton. Lateral view (A–F) and ventral view (G–L) of the head skeleton of day-6 embryos labeled with Calcein. (B, C, E, H, I, K, and L) When embryos were injected with best1-morpholino (MO) at the one-cell stage, the amount of stained mineralized tissue was significantly reduced compared to fish injected with control-MO. Panel N shows increased intraocular distance (yellow arrows) in best1 morphants. Panel M and N show measurements of the distance between the eyes, and panel O shows the distances depicted graphically as the mean for 10 embryos of each type. Columns, mean; standard error of the mean (n = 10; analysis of variance), ***p < 0.0001. Scale bar, 100 μm. 5ba, fifth branchial arch; op, opercular bone; ec, ectopterygoid; e, ethmoid plate; pq, palatoquadrate; m, Meckel cartilage; dpf, days postfertilization.
vival rate of best1 knockdown embryos was decreased (Fig. 1I) and the percentage of live embryos with aforementioned gross defects was significantly higher in the best1 knockdown group versus the control group (p < 0.0001) (Fig. 1J).

Knockdown of best1 caused several developmental abnormalities in the nervous system, including defects in head and brain patterning (Fig. 2A, E, and I, denoted by dotted lines and blue arrowheads) and defective neuronal outgrowth (Fig. 2B, F, and J, arrows in control embryo indicate normal motor axons; asterisks in I3E4-MO and ATG-MO embryos showed no outgrowth of axons). Quantitative morphometric analysis revealed substantially decreased head and brain sizes (p < 0.0001) (Fig. 2M and N) and markedly shortened motor neuron length (p < 0.0001) (Fig. 2O) in I3E4-MO embryos compared to controls. Knockdown of best1 also resulted in reduced eye size of zebrafish (Fig. 2E and I, marked by black arrowheads).

Taken together, these results demonstrate that the knockdown of best1 in zebrafish embryos can lead to shorter body length, curved body axis, low survival rate, microcephaly, small eyes, reduced head and brain size, and impaired neuronal outgrowth.

3. Knockdown of best1 Causes Craniofacial and Vertebral Skeletal Ossification Defect

By application of the fluorescent dye Calcein to label the calcified matrix of the skeleton, a significant reduction of craniofacial skeletal ossification was found in best1 knockdown zebrafish (6 days post fertilization, 6 dpf) compared to that of control (Fig. 3). In the control embryos, 5 pairs of craniofacial bone structures were nicely labeled in both lateral (Fig. 3A and D) and ventral (Fig. 3G and J). However, only the fifth branchial arch (5ba) was faintly labeled in the I3E4-MO embryos (Fig. 3B, E, H, and K), and no calcified bone structures were identified in the ATG-MO embryos (Fig. 3C, F, I, and L). Impaired craniofacial development also led to increased interorbital distance in best1 knockdown embryos (p < 0.0001) (Fig. 3M–O).

Furthermore, we administrated Calcein at 8.5 dpf to visualize calcified vertebrae of zebrafish embryos. On average, 14 vertebral segments could be clearly identified in the control group.

![Fig. 4. Loss of best1 causes abnormal craniofacial structure and vertebral development.](image-url)

(A–H) Representative bright field and fluorescent images of zebrafish skeleton at 8.5-dpf. In vivo visualization of the skeleton is achieved by the administration of a fluorescent dye (Calcein) directly to the fish water. Dyes that bind to calcified matrix can be used to label the entire skeleton. Lateral view (B, C) of the skeleton and mineralized vertebrae of day-8.5 embryos labeled with Calcein. Vertebrae 1, 2, 3, 4, 5, and 6 are indicated. (E, F, and H) When embryos injected with best1-morpholino (MO) at one-cell stage, the amount of stained mineralized tissue is markedly reduced compared to control-MO-injected fish. Vertebral development is significantly delayed in best1 morphants. (I) Quantification of the number of mineralized vertebrae at 8.5-dpf. Columns, mean; standard error of the mean (n = 10; analysis of variance), ***p < 0.0001. Scale bar, 100 μm. dpf, days postfertilization.

560 www.e-neurospine.org  https://doi.org/10.14245/ns.2347238.619
(Fig. 4A–C). Whereas zero vertebrae could be distinguished in best1 knockdown embryos (Fig. 4D–H). There is a significant difference among the groups (p < 0.0001) (Fig. 4I).

These findings indicated that best1 might play a critical role in the early skeletal development of zebrafish.

**DISCUSSION**

Congenital CVJ skeletal malformation is a rare disease that can lead to disastrous neurological deficits in patients. The etiology of congenital CVJ malformation is still undefined. In our previous study, 5 loss-of-function or damaging missense were detected out of 92 congenital CVJ skeletal malformation patients (carrier rate = 5.43%), which is significantly enriched in the patient group compared with in-house control (p = 5.77 × 10⁻⁶; odds ratio, 27.74; 95% confidence interval, 7.15–94.25). In this study, we designed MOs targeting zebrafish best1 to knockdown best1 expression and investigated the corresponding phenotypes.

First, the knockdown of best1 in zebrafish embryos caused smaller eyes than control. Microphthalmia (axial length ≤ 20 mm) has been reported in multiple VMD2 patients. It has been reported that BEST1 played an essential role in ocular patterning, maintenance of retinal pigment epithelium (RPE) cells, and normal growth of the eye. Bestrophin is a 585 amino acid transmembrane protein acting as an oligomeric chloride channel that regulates calcium channels for calcium and transmembrane potential homeostasis. Bestrophin mutations alter chloride ion-related conductance across the RPE cell membrane, and abnormal channel function may account for the abnormal electro-oculogram seen in patients with Best disease. Our results in zebrafish agreed with that in VMD2 patients.

Second, the knockdown of best1 in zebrafish embryos resulted in multiple skeletal defects, including shorter body length, curved body axis, microcephaly, and impaired ossification of craniofacial and vertebral bone. Early zebrafish skeleton patterning is regulated by complex signaling pathways, including hox, fgf, slh, tbx, hmx, runx2, bmp, and others. Intriguingly, skeleton malformations can be associated with eye and ear disorders. In human Oculoauricular Syndrome (MIM 612109), patients manifest with complex ocular anomalies, including microphthalmia, dysplastic ears with abnormal external ear cartilage, as well as spina bifida occulta and platybasia. In mice, mutations in Mitf gene can cause microphthalmia, deafness and osteopetrosis. BEST1 is regulated by MITF; these results imply that BEST1 gene plays an important role in eye development and skeletal ossification. In addition, BEST1 interacts with ANO6, another calcium-dependent transmembrane channel, which is crucial for apatite deposition in bone mineralization. ANO6 activates NCX1, a calcium channel in osteoblasts, to transport calcium to the extracellular matrix for bone formation. ANO6 mice embryos (E15.5) also show delayed bone mineralization, which is highly consistent with our zebrafish results.

Third, best1 knockdown caused microcephaly, reduced brain size, and impaired neuronal outgrowth in zebrafish embryos. Woo et al. prepared Best1 knockout mice to investigate tonic inhibition in the brain and found that inhibition of Best1 resulted in enhanced neuronal excitability in the cerebellum. Apparently Bestrophins distribute and function as ion channels in the central nervous system. However, no studies have yet reported these neurological developmental defects.

Intriguingly, the patient in our cohort with a premature stop codon of BEST1 gene (p.S79F*) did present with ophthalmic, spinal and neurological developmental defects (Fig. 5). This is a 27-year-old female patient suffering with decreased sensation on all limbs and weakness of bilateral lower limbs for one year. Physical examination revealed increased tendon reflexes and positive Hoffmann and Babinski signs. Imaging workup found Chiari malformation with syringomyelia and hydrocephalus, multiple spinal anomalies (including assimilation of atlas, basilar invagination, fusion of C5 and C6 vertebrae, butterfly vertebra at T12 and scoliosis) and microphthalmia. She denied decreased visual acuity, which is a later-onset symptom of VMD2, and refused a fundus examination. The specific genetic mutation identified in the patient was a premature stop codon, which could result in truncation of the Bestrophin-1 protein, affecting its normal structure and function. Interestingly, this mutation’s impact in humans is similar to that of MO, which knocks down the expression of the best1 gene in zebrafish. Notably, this premature stop codon mutation was identified in a CVJ malformation patient who also presented comorbidities such as microphthalmia, neurological defects, and multiple skeletal malformations, particularly along the spine. Given the homology between zebrafish and humans, it is hypothesized that decreased BEST1 gene function, whether due to a premature stop codon in humans or through MO knockdown in zebrafish, could disrupt critical molecular pathways related to eye, neurological, and skeletal development, leading to similar phenotypes. Further research is required to investigate these mechanisms.

Our study has several limitations. First, the VMD2 manifests with later-onset macular/retinal dystrophy. Thus, we did not...
Knockdown of *best1* Gene in Zebrafish


Knockdown of *best1* Gene in Zebrafish

examine the retina of zebrafish embryos for macular dystrophy. Second, this is an *in vivo* knockdown study that validated the essential role of the *best1* gene in the early development of zebrafish. Further knockout studies and *in vitro* experiments are warranted to delineate upstream and downstream targets and pathways involved. Third, the CVJ structures of zebrafish are not as analogical to humans compared with that of mice, which may be more suitable for our future study.

**CONCLUSION**

Damaging mutation of *BEST1* gene may cause a subtype of CVJ malformation with comorbidities such as microphthalmia, VMD2, spinal anomalies, scoliosis, and neurological defects. Further studies are required to investigate the mechanism and possible preventive measures for the subtype of CVJ malformation.

**NOTES**

Supplementary Material: Supplementary Table 1 and Figs. 1-2 can be found via https://doi.org/10.14245/ns.2347238.619.

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This research was funded by the National High Level Hospital Clinical Research Funding (2022-PUMCH-D-004).

Author Contribution: Conceptualization: ZL, FJ, HWu; Data curation: ZL, KL, KW, LZ, SJ, HWang, FJ, HWu; Formal analy-
sis: ZL, KL, KW, LZ; Funding acquisition: HWu; Methodology: ZL, KL, KW, LZ, SJ, HWang, FJ, HWu; Project administration: ZL, FJ, HWu; Visualization: SJ, HWang; Writing – original draft: ZL, KL; Writing – review & editing: ZL, KL, FJ, HWu.

**ORCID**

Zhenlei Liu: 0000-0003-0827-4420  
Kang Li: 0009-0008-3146-7179  
Kai Wang: 0009-0000-1971-647X  
Lei Zhang: 0000-0001-8142-403X  
Shanhang Jia: 0000-0003-1397-1773  
He Wang: 0000-0001-6937-0665  
Fengzeng Jian: 0000-0001-7860-278X  
Hao Wu: 0000-0002-8118-9989

**REFERENCES**


Supplementary Table 1. Primer sequences quantitative real-time polymerase chain reaction

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward primer 5′-3′</th>
<th>Reverse primer 5′-3′</th>
</tr>
</thead>
<tbody>
<tr>
<td>best1</td>
<td>TCAGCACGGCTGTCTACAAG</td>
<td>TGCGACGTCGTTATTGATTC</td>
</tr>
<tr>
<td>efla</td>
<td>GGAAATTCGAGACCAGAAATAC</td>
<td>GATACCAGCCTCAAACCTCACC</td>
</tr>
</tbody>
</table>
Supplementary Fig. 1. Morpholino (MO)-targeted down-regulation of best1 was confirmed by reverse transcription-polymerase chain reaction (RT-PCR) and quantitative RT-PCR (qRT-PCR). (A) Four nonogram of each type of MOs were microinjected into fertilized one-cell stage embryos according to standard protocols. (B) The zebrafish best1 gene was targeted by specific morpholino antisense to prevent start of expression (ATG-MO) or proper splicing of exon 4 (I3E4-MO). Primers 3F and 6R interrogate the presence of wild-type (nonmutant) transcripts or those in which intron 3 has been inserted. (C) RT-PCR of best1 transcript from control-MO and I3E4-MO morpholino-injected embryos 30-hpf, demonstrating insertion of intron3. Injection of 4 ng of best1 morpholino alters the splicing between intron 3 and exon 4, as revealed by shift in PCR bands between control and best1 morpholino-injected embryos. (D) Quantitative measurements of best1 expression levels measured by qRT-PCR (**p < 0.0001). Samples were collected at 30-hpf after introduction of 4 ng of MO at the one-cell stage (N = 30). hpf, hours post fertilization.
Supplementary Fig. 2. Reduced body length and apparent curved body axis in I3E4-MO (B, E) and ATG-MO (C, F) injected embryos compared with control-MO injected embryos (A, D).
Clinical and Radiological Outcomes in C2 Recapping Laminoplasty for the Pathologies in the Upper Cervical Spine

Dong Hun Kim¹, Jae Taek Hong¹, Jung Woo Hur¹, Il Sup Kim², Ho Jin Lee³, Jong Beom Lee⁴

¹Department of Neurosurgery, Eunpyeong St. Mary’s Hospital, The Catholic University of Korea, Seoul, Korea
²Department of Neurosurgery, St. Vincent Hospital, The Catholic University of Korea, Suwon, Korea
³Department of Neurosurgery, Chungbuk National University Hospital, Cheongju, Korea

Objective: To evaluate C2 muscle preservation effect and the radiological and clinical outcomes after C2 recapping laminoplasty.

Methods: Fourteen consecutive patients who underwent C2 recapping laminoplasty around C1–2 level were enrolled. To evaluate muscle preservation effect, the authors conducted a morphological measurement of extensor muscles between the operated and nonoperated side. Two surgeons measured the cross-sectional area (CSA) of obliquus capitis inferior (OCI) and semispinalis cervicis (SSC) muscle before and after surgery to determine atrophy rates (ARs). Additionally, we examined range of motion (ROM), sagittal vertical axis (SVA), neck visual analogue scale (VAS), Neck Disability Index (NDI), and Japanese Orthopaedic Association (JOA) score to assess potential changes in alignment and consequent clinical outcomes following posterior cervical surgery.

Results: We measured the CSA of OCI and SSC before surgery, and at 6 and 12 months postoperatively. Based on these measurements, the AR of the nonoperated SSC was 0.1% ± 8.5%, the AR of the operated OCI was 2.0% ± 7.2%, and the AR of the nonoperated OCI was -0.7% ± 5.1% at the 12 months after surgery. However, the AR of the operated side’s SSC was 11.2% ± 12.5%, which is a relatively higher value than other measurements. Despite the atrophic change of SSC on the operated side, there were no prominent changes observed in SVA, C0–2 ROM, and C2–7 ROM between preoperative and 12 months postoperative measurements, which were 11.8 ± 10.9 mm, 16.3° ± 5.9°, and 48.7° ± 7.7° preoperatively, and 14.1 ± 11.6 mm, 16.1° ± 7.2°, and 44.0° ± 10.3° at 12 months postoperative, respectively. Improvement was also noted in VAS, NDI, and JOA scores after surgery with JOA recovery rate of 77.3% ± 29.6%.

Conclusion: C2 recapping laminoplasty could be a useful tool for addressing pathologies around the upper cervical spine, potentially mitigating muscle atrophy and reducing postoperative neck pain, while maintaining sagittal alignment and ROM.

Keywords: Laminectomy, Laminoplasty, Postoperative complications, Spinal cord neoplasms, Cervical vertebrae, Muscular atrophy

INTRODUCTION

The surgery of craniovertebral junction (CVJ) could be complicated due to discreet relationships in the surrounding neuro-vascular structures, complex biomechanical issues, and intricate muscular structures.¹–¹⁰ The C2 spinous process gives attachment to obliquus capitis inferior (OCI), rectus capitis posterior major (RCPM), bulky portions of the semispinalis cervicis (SSC),
spinalis cervicis, interspinalis and multifidus muscles. These C2 muscles form the long arms with the critical function of neck extension and rotation.\textsuperscript{11-14} The detachment of C2 muscles not only hampers muscle function but also causes considerable postoperative pain and postoperative kyphosis.\textsuperscript{15,16}

The preservation of C2 muscle attachments is essential to prevent postoperative cervical kyphosis.\textsuperscript{17} Sparing C2 spinous process importantly preserves anchor points for extensor muscles and ligament structures and avoids the risk of cervical kyphosis after laminoplasty and spinal cord tumor surgery.\textsuperscript{15} However, it is common practice to detach the C2 muscles in conventional upper cervical spine surgeries.

Recently, muscle-preserving technique was introduced for exposure of the posterior cervical spine.\textsuperscript{15,18} However, the C2 recapping technique has been rare in the literature, and no study has investigated this new technique regarding long-term clinical and radiological outcomes.

Therefore, the purpose of this study is to quantitatively analyze muscle atrophy around the CVJ by measuring muscle volume before and after C2 recapping laminoplasty. Additionally, we will assess whether there is aggravation of neck pain or changes in radiologic parameters related to cervical alignment and clinical outcome indicators, thus confirming the clinically beneficial potential of this surgical methods.

MATERIALS AND METHODS

1. Study Design

This is a retrospective study of consecutive patients who had undergone C2 recapping laminoplasty surgeries around CVJ between January 2010 and January 2022. After obtaining approval from the Institutional Review Board (IRB) of Catholic Medical Center of The Catholic University of Korea (IRB No. PC23DASS0107), which waived the need for informed consent, patient data were retrospectively reviewed.

In this study, criteria for inclusion were patients who were older than 18 years old, had pre- or postoperative computed tomography (CT) images and C-spine x-rays including flexion and extension. And all patients also had patients-reported outcome measures such as neck visual analogue scale (VAS) score, Neck Disability Index (NDI) score, and Japanese Orthopaedic Association (JOA) score. Based on this data, C2 muscle atrophy of OCI and SSC muscles, change of cervical alignment, and postoperative complications were evaluated. In addition, we examined the volume of C2 muscles on CT scan, C2–7 SVA in neutral x-ray, C0–2 angle, and C2–7 angle in dynamic x-ray before and after surgery.

The patients were followed-up for at least one year following the surgery. The use of soft cervical collar was recommended for a month. Regular outpatient follow-up was conducted at 1, 3, 6, and 12 months after surgery to analyze the clinical and radiological evaluation.

To exclude factors that could potentially impact precise result analysis, the exclusion criteria were set as follows: cases with missing data in clinical and radiological evaluation within a year after surgery, cases where recapping laminoplasty was performed bilaterally, and cases with a history of previous cervical spine surgery.

2. Surgical Techniques (Figs. 1, 2; Supplementary video clip 1)

After general anesthesia, the patient was placed in the prone position, and the head was held in a slightly flexed position using a Mayfield head holder.

The surgical technique started with a midline skin incision from the inion to the C3 vertebra. The dissection continued through the subcutaneous tissues. The trapezius, splenius capitis, and the semispinalis capitis muscles were reflected from the midline and then the muscle plane of the suboccipital muscles was identified. The plane between the suboccipital and SSC muscles were developed deeply, and the semispinalis capitis muscles superficially.

C2 spinous process is split longitudinally with a threaded surgical wire or ultrasonic bone scalpel, leaving all muscular attachments (RCPM, OCI, and SSC muscles). The C2 lamina, pedicle, C1–2 joint, and C2–3 joints can be exposed by blunt dissection through the intermuscular plane between the SSC and the OCI muscle without damaging muscles. Next, a bone scalpel was used to make a lateral gutter on the lateral aspect of the C2 lamina.

If the tumor is skewed to one side, only one side C2 lamina can be opened and operated on. When wider exposure is needed because of the tumor size, the bilateral C2 lamina could be expanded. In this case, while the separation of the spinous process may pose a lack of physical support, it can still be beneficial in minimizing muscle injury.

Either C1 laminectomy or C1 laminoplasty could be possible to expand the surgical exposure. C1 laminectomy and C1 laminoplasty require resection of the RCPM muscle, of which function is trivial in humans.

The dura was suspended with tenting sutures and was cut along the midline for treatment of intradural tumors. Then the tumor or pathologic lesion was carefully separated from the...
nerve root and spinal cord. Finally, the blood supply to the lesion was blocked by coagulation and repeatedly washed with physiological saline; then, it was removed completely. To remove intramedullary tumors, the spinal cord was longitudinally cut from the most prominent and nearest area without blood vessels. The tumor was separated and removed along its border. When the tumor was separated, we only pulled the tumor, not the spinal cord. After tumor resection, the separated C2 laminae is brought back to its counterpart with stitches using non-absorbable suture composed of ethylene terephthalate, passed through drill holes in each split half of the C2 spinous process without damaging the C2 muscles. Finally, working layer by layer, the surgeon will close the incision using absorbable sutures.

Fig. 1. Schematic representation of the C2 muscle preservation procedure. (A) The rectus capitis posterior major (RCPM), obliquus capitis inferior (OCI), and semispinalis cervicis (SSC) muscles are attached to C2 spinous process. The rectus capitis posterior minor is attached to C1 posterior arch. The OCI is the only suboccipital muscle that does not have an attachment to the cranium. (B) Division of the plane between the OCI muscle and the SSC muscles. (C) C2 spinous process is split longitudinally, and lateral gutter are made between the SSC muscles and the OCI muscles. Dotted line indicates the resection margin in the midline spinous process and lateral gutter. The bilateral SSC muscles and OCI muscles were completely preserved. (D) Unilateral C2 spinous process and lamina was retracted to expose spinal canal. While the C2 attachments of the SSC, OCI, and RCPM muscles are left intact, the C2 laminal flaps are elevated to swing open. Either C1 laminectomy or C1 laminoplasty could be possible to expand the surgical exposure. (E) Main surgical procedure and tumor removal can be performed after the opening of C2 lamina. (F) Reconstructing the C2 spinous process/lamina and the muscle attachments. Expanded half of the C2 spinous process then reattached to counterpart with stitch passed through drill-hole in each split half of spinous process.

Fig. 2. Intraoperative photographs of C2 muscle preservation procedure. (A) C2 spinous process is split longitudinally with a surgical threaded wire (white arrows), leaving all muscular attachments (RCPM, OCI, and SSC muscles). (B) Dividing the plane between the OCI muscle (black arrow) and the SSC muscles (white arrow). A scalpel and bipolar forceps are used in the sharp dissection process to avoid heat damage to the muscles. (C) Making lateral gutter on the C2 lamina using bone scalpel for muscle-preserving C2 laminoplasty. (D) Unilateral C2 spinous process and lamina is retracted to expose spinal canal. While the attachments of the SSC, OCI, and RCPM muscles (at C2) are left intact, the C2 laminal flaps are elevated to swing open. (E) Spinal cord is exposed. Main surgical procedure and tumor removal can be performed after the opening of C2 lamina. (F) Reconstructing the C2 spinous process/lamina and the muscle attachments. Expanded half of the C2 spinous process then reattached to counterpart with stitch (white arrows) passed through drill-hole in each split half of spinous process. The bilateral SSC muscles and OCI muscles are completely preserved. RCPM, rectus capitis posterior major; OCI, obliquus capitis inferior; SSC, semispinalis cervicis.
3. Radiological Evaluation
We compared the radiological outcomes, such as C2 muscle atrophy, change of cervical alignment and segmental motion before and after the surgery. Standing lateral, flexion, and extension radiographs of the cervical spine were performed both preoperatively and postoperatively. Radiography and CT imaging examination were performed before surgery and at 6, 12 months after surgery. We examined the C2–7 sagittal vertical axis (SVA) in the neutral position, C0–2 angle, and C2–7 angle in extension and flexion before and after surgery. The angles were measured based on the following lines: McGregor line, the lines of lower endplates of C2, and the lines of the upper endplate of C7. The C2–7 SVA was defined as the horizontal offset of a plumb line dropped from the center of the C2 vertebral body to the posterior superior corner of the C7 vertebra.

All enrolled patients underwent CT preoperatively, 6 months, and 1 year after the surgery. Axial CT images were aligned parallel to the inferior endplate of the vertebral body. Axial CT images were used to measure the cross-sectional area (CSA) of the muscles around the C2 spinous process (Fig. 3). We measured the CSA of the OCI muscle at the middle of the C2 spinous process and the CSA of the SSC at the C2–3 intervertebral disc level (Fig. 3). The muscle atrophy rate (AR) was measured as follows: muscle AR = [(preoperative CSA − postoperative CSA)/preoperative CSA] × 100. Two spine surgeons independently reviewed the imaging studies. All CSA measurements were performed twice by the same person to minimize the potential for error in constructing the polygons around the muscles’ margins, and the average values were analyzed.

4. Clinical Evaluation
The clinical features of each patient, including age, sex, and diagnosis were recorded. Clinical outcome was assessed using JOA score for cervical myelopathy preoperatively and postoperatively. The recovery rate (RR) was calculated according to the following formula (Hirabayashi method): RR (%) = (postoperative JOA – preoperative JOA)/(17 [full score] – preoperative JOA) × 100.19

We used the VAS score to evaluate the neck pain and the NDI to assess the functional status of the patient’s neck. Postoperative neck pain was assessed immediately after surgery, 1 month, 3 months, 6 months, and 1 year after the surgery.

We also evaluated the following postoperative wound complications: cerebrospinal fluid (CSF) leakage and postoperative infection.

RESULTS
1. Patients Demographics
During the specified period, a total of 18 patients satisfying the inclusion criteria were identified. Two patients were not followed-up for a year and consequently excluded due to incomplete evaluation. And a patient who underwent bilateral laminoplasty and another patient who had history of previous cervical spine surgery were excluded.

A total of 14 patients were enrolled in this study in the end (Table 1). Overall, mean age of patients was 48.1 ± 15.9 years (range, 23–77 years), and 6 patients (42.3%) were male, and 8 patients (57.7%) were female.

The types of pathologies were schwannoma (n = 5, 35.7%), meningioma (n = 5, 35.7%), extradural ossified mass (n = 2, 14.2%), intramedullary tumor (n = 1, 7.1%), and syringomyelia (n = 1, 7.1%). The average follow-up period was 27.4 months (range, 12–74 months). There was 1 case of multiple schwann-
nomas who needed extended cervical laminectomy in the subaxial cervical spine. C1 laminectomy was required in 5 cases, accounting for 35.7% of the total surgeries. C1 laminoplasty was carried out in 5 cases (35.7%), while in 4 cases (28.6%), no additional procedures were performed.

2. Radiographic Evaluation
Inter- and intraobserver variability analyses showed intra-class correlation coefficient values were excellent (0.835–0.994) for measured cervical angles and that values measured by the 2 observers were well correlated.

The postoperative changes of each radiographical parameter are shown in Tables 2 and 3. The average postoperative AR of the OCI and the SSC muscles at 6 months were 0.8% ± 3.2% and 5.0% ± 8.8%, respectively, in the operated side. Meanwhile, the average postoperative AR of the OCI and SSC muscles at 6 months were -0.9% ± 2.8% and -0.9% ± 5.2%, respectively, in the nonoperated side. The trend persists at 12 months postoperatively, where the average postoperative AR of the OCI and the SSC muscles on the operated side were 2.0% ± 7.2% and 11.2% ± 12.5%, respectively. In contrast, on the nonoperated side, the values were -0.7% ± 5.1% and 0.1% ± 8.5%, respectively.

The preoperative C2–7 ROM was 48.7° ± 7.7°, and the postoperative C2–7 ROM at 6 months and 12 months were 44.9° ± 8.4° and 44.0° ± 10.3°. C02 ROM and SVA revealed no prominent change between preoperative and postoperative periods. (16.3 ± 5.9 vs. 16.3 ± 8.0 vs. 16.1 ± 7.2, 11.8 ± 10.9 vs. 12.8 ± 10.6 vs. 14.1 ± 11.6) (Table 2). And there was also no distinct change in sagittal alignment and ROM change after the surgery.

The bony fusion between the bisected C2 spinous process was completed in all patients. However, fusion was observed in the lateral gutter of the operated side in 11 out of the total 14 cases.

Table 1. Summary of patients’ characteristics (n = 14)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>48.1 ± 15.9</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6 (42.3)</td>
</tr>
<tr>
<td>Female</td>
<td>8 (57.7)</td>
</tr>
<tr>
<td>Types of pathologies</td>
<td></td>
</tr>
<tr>
<td>Meningioma</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td>Schwannoma</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td>Extravascular ossified mass</td>
<td>2 (14.2)</td>
</tr>
<tr>
<td>Intramedullary metastasis</td>
<td>1 (7.1)</td>
</tr>
<tr>
<td>Syringomyelia</td>
<td>1 (7.1)</td>
</tr>
<tr>
<td>Follow-up period (mo)</td>
<td>27.4 ± 23.1</td>
</tr>
<tr>
<td>Removal status of C1 lamina</td>
<td></td>
</tr>
<tr>
<td>C1 laminectomy</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td>C1 laminoplasty</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td>Neither</td>
<td>4 (28.6)</td>
</tr>
<tr>
<td>Fusion rate of C2 spinous process</td>
<td>14 (100)</td>
</tr>
<tr>
<td>Fusion rate of lateral gutter on the lateral aspect of the C2 lamina</td>
<td>11 (78.6)</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%).

Table 2. Radiologic parameters between preoperative and postoperative period

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preoperative</th>
<th>Postoperative (6 mo)</th>
<th>Postoperative (12 mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCI CSA (nonoperated) (mm²)</td>
<td>261.1 ± 119.3</td>
<td>262.9 ± 119.6</td>
<td>262.0 ± 120.0</td>
</tr>
<tr>
<td>OCI CSA (operated) (mm²)</td>
<td>241.1 ± 103.7</td>
<td>239.2 ± 105.1</td>
<td>238.7 ± 111.6</td>
</tr>
<tr>
<td>SCC CSA (nonoperated) (mm²)</td>
<td>119.1 ± 42.5</td>
<td>120.1 ± 43.3</td>
<td>119.4 ± 43.5</td>
</tr>
<tr>
<td>SCC CSA (operated) (mm²)</td>
<td>116.1 ± 41.2</td>
<td>109.4 ± 36.0</td>
<td>104.1 ± 37.5</td>
</tr>
<tr>
<td>SVA (mm)</td>
<td>11.8 ± 10.9</td>
<td>12.8 ± 10.6</td>
<td>14.1 ± 11.6</td>
</tr>
<tr>
<td>C02 ROM (°)</td>
<td>16.3 ± 5.9</td>
<td>16.3 ± 8.0</td>
<td>16.1 ± 7.2</td>
</tr>
<tr>
<td>C27 ROM (°)</td>
<td>48.7 ± 7.7</td>
<td>44.9 ± 8.4</td>
<td>44.0 ± 10.3</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.

OCI, oblique capitis inferior muscle; CSA, cross-sectional area; SSC, semispinalis cervicis muscle; SVA, sagittal vertical axis.

Table 3. AR between operated and nonoperated sides

<table>
<thead>
<tr>
<th>Variable</th>
<th>Nonoperated</th>
<th>Operated</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCI AR (% , 6 mo)</td>
<td>-0.9 ± 2.8</td>
<td>0.8 ± 3.2</td>
</tr>
<tr>
<td>OCI AR (% , 12 mo)</td>
<td>-0.7 ± 5.1</td>
<td>2.0 ± 7.2</td>
</tr>
<tr>
<td>SSC AR (% , 6 mo)</td>
<td>-0.9 ± 5.2</td>
<td>5.0 ± 8.8</td>
</tr>
<tr>
<td>SSC AR (% , 12 mo)</td>
<td>0.1 ± 8.5</td>
<td>11.2 ± 12.5</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. OCI, oblique capitis inferior muscle; AR, atrophy rate; SSC, semispinalis cervicis muscle; SVA, sagittal vertical axis.
3. Clinical Outcome Evaluation

In the 1-year follow-up period, the JOA score increased from 11.9 ± 3.6 preoperatively to 15.0 ± 3.5 postoperatively. The RR of the JOA score was 77.3% ± 29.6% while the VAS and NDI scores were improved after surgery (Table 4).

4. Postoperative Complication

Regarding complications, no severe intraoperative complications occurred after the surgery. There was no postoperative CSF leakage or wound infection after C2 recapping laminoplasty.

Table 4. Postoperative clinical outcome

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preoperative</th>
<th>Postoperative (6 mo)</th>
<th>Postoperative (12 mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS neck</td>
<td>5.2 ± 1.9</td>
<td>1.3 ± 1.3</td>
<td>0.9 ± 1.1</td>
</tr>
<tr>
<td>NDI</td>
<td>28.1 ± 8.1</td>
<td>9.0 ± 4.4</td>
<td>6.9 ± 5.5</td>
</tr>
<tr>
<td>JOA score</td>
<td>11.9 ± 3.6</td>
<td>14.6 ± 3.8</td>
<td>15.0 ± 3.5</td>
</tr>
<tr>
<td>JOA score RR (%), 6 mo)</td>
<td>71.2 ± 35.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOA score RR (%), 12 mo)</td>
<td>77.3 ± 29.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.
VAS, visual analogue scale; NDI, Neck Disability Index; JOA, Japanese Orthopaedic Association; RR, recovery rate.

Fig. 4. Case 1. (A) Preoperative axial magnetic resonance image are showing that homogenously enhanced intradural mass compressed spinal cord at the C2 level. (B) Postoperative magnetic resonance imaging is showing that there is no enhanced mass or cord compression after the surgery. (C–E) Axial computed tomography images show the change of C2 spinous process and attached muscles before, 3 months after, and 1 year after surgery.

Fig. 5. (A) Preoperative enhanced magnetic resonance image showing that homogenously enhanced ventral intradural mass compressed spinal cord at the C1–2 level. The axial computed tomography images demonstrating C2 spinous process and the bilateral OCI muscles (CSA) at the C2 level before (B) and after surgery (C). The lateral radiographs demonstrating cervical alignment before (D) and after surgery (E). OCI, obliquus capitis inferior; CSA, cross-sectional area.
2) Case 2

A 47-year-old male patient presented with neck pain and both hand clumsiness. The preoperative MRI demonstrated homogeneously enhanced intradural mass to the right side at the C1 and C2 levels. C2 recapping laminoplasty and C1 laminoplasty was performed for tumor removal. The mass was completely excised, and the pathology result was meningioma. Postoperative CT showed that the volume of C2 muscles preserved one year after surgery. The lateral radiographs demonstrated that there was no prominent change in cervical alignment before and after surgery (Fig. 5).

DISCUSSION

The cervical extensor muscles play an essential role in cervical alignment.13,18,20-23 Panjabi et al.24 reported that the neck muscles provide nearly 80% of the needed mechanical stability of the cervical spine, while osteo-ligamentous structures contribute about only 20%. Previous reports have emphasized preserving muscles that attach to the C2 spinous process to prevent cervical lordosis loss after laminoplasty.17,23,25-27

Moreover, C2 spinous process is one of the key structures for the extensor muscles because the height of this spinous process increases the moment arm of the functioning muscles complex.28

The OCI muscle is a fleshy, thick muscle located in the neck. The OCI is the largest muscle of the 4 suboccipital muscles and the only suboccipital muscle that does not attach to the cranium. It instead inserts into the transverse process of the atlas on the infero-posterior part. Its origin is at the C2 spinous process. Bilateral contraction of this muscle causes head extension and unilateral contraction performs the critical role of providing rotation of the head towards the ipsilateral side.29-31

The SSC muscles originate from the transverse processes from T1 to T6 and insert to the spinous processes from C2 to C5. These muscles act as a stabilizer and one of the main extensors of the cervical spine, which are related to the cervical motion and alignment.15,16,32

The SSC inserting to C2 spinous process is the most developed among this muscle group. Preservation of the SSC insertion to C2 could prevent both the postoperative axial pain and the loss of cervical lordosis that can affect long-term outcomes after laminoplasty. There have been lots of studies about the significance of the C2 muscles so far. However, most of them are about the lower cervical spine surgeries, and studies on C2 muscle detachment for CVJ pathologies are rare.

So, the purpose of this study was to quantitatively measure volume of OCI and SSC muscle and evaluate clinical outcomes to determine the potential clinical usefulness of C2 recapping laminoplasty.

Conventional C2 laminectomy damaged the posterior muscle structures attached to the C2 spinous process and caused muscle atrophy after surgery (Fig. 3).

This study provided some interesting findings regarding the effect of C2 recapping laminoplasty.

First, when calculating the AR of each muscle, it is found that, except for the operated SSC, there was muscle atrophic change of less than 2%. Second, it seems that there were similar flexion/extension ROM both in the upper and subaxial cervical spine before and after surgery. Although atrophy of the SSC muscle on the operated side was confirmed, it did not appear to induce ROM change before and after surgery. This finding suggests that despite the loss of CSA of the ipsilateral SSC muscle, functional outcomes may not have been affected due to preserving contralateral SSC muscle and other extensor muscles may have prevented a ROM decrease after surgery. Third, it appears that the C2 recapping technique does not result in postoperative cervical malalignment. Finally, the C2 muscle preservation technique allows enough space for complete tumor excision and minimizing the amount of dead space while avoiding muscle damage encountered after conventional laminectomy. Undamaged C2 muscles with a rich blood supply could minimize the amount of dead space created and diminishes the incidence of deep wound infection and persistent CSF leakage. Moreover, the free muscle-bone fragment receives a rich blood supply through the preserved muscular attachments to the spinous process, facilitating bone fusion in the bisected spinous process.

Subaxial alignment change is not uncommon after upper cervical spine surgery.28,33 Although the little study has sought to identify the risk factors of postoperative cervical malalignment following upper cervical spine surgery, we recently reported lower cervical spine alignment might change during the first year after CVJ posterior fixation, and lower cervical alignment is related to upper cervical angle after CVJ fixation.28,34 Besides, the risk of subaxial kyphotic change increased after CVJ fixation when combined with lower cervical laminoplasty and comprehensive dissection of deep extensor muscle.28

These data showed that the C2 muscle detachment itself was not a risk factor of malalignment or postoperative neck pain. The risk of cervical malalignment and related neck disability increased only when combined dissection of deep extensor muscle down to the lower cervical spine.
In the patients with lower cervical kyphosis and sagittal malalignment, lateral radiographs show hyperlordotic angle in the upper cervical spine.\(^1,2,8,35-39\) It is because when patients have sagittal malalignment in the lower cervical spine, the C0–2 segment’s hyperextension holds up the head to compensate for distal kyphosis, sagittal imbalance, and maintaining horizontal gaze.\(^1,3,36\)

These findings suggest that reciprocal interaction may likely affect not only global balance but also regional balance. We believe that the proposed technique may minimize the extent of soft tissue dissection, muscle splitting, and postoperative dead space. Additionally, minimizing the risk of soft tissue devascularization, denervation, and biomechanical change could decrease postoperative wound complication and adjacent level disease as previously reported. Furthermore, this could diminish postoperative pain and expedite postoperative recovery.

This study has several limitations that warrant consideration. First, it is limited by its retrospective, single-surgeon design, which may have caused selection bias. Second, the sample size could be relatively small, and the case etiologies were heterogeneous to draw meaningful conclusions. Finally, we did not evaluate axial plane movement and the rotation capacity of the OCI muscles. Although this study did not include all the C2 muscles and the whole direction of cervical motion, analyzing the 2 largest C2 muscles could reflect the pattern of postoperative changes.

Nevertheless, we believe our study might have a sufficient clinical impact. We have shown the results of a relatively long-term follow-up and time-dependent change in cervical alignment and related neck disability score following new techniques handling the C2 spinous process and its attached muscles. This study included the largest number of cases that have used the C2 recapping laminoplasty technique for CVJ pathologies to date to the best of the authors’ knowledge.

However, we suggest multicenter, multiple neck movements, and larger-scale comparative studies to obtain more accurate information on the value of this C2 muscle preservation technique in the future.

**CONCLUSION**

C2 recapping laminoplasty might be effective for CVJ pathologies to preserve C2 muscle structures. Keeping the C2 musculature could help with pathologies around the upper cervical spine. It would help prevent the C2 muscles atrophy, maintain cervical ROM, and reduce postoperative neck pain and malalignment in the postoperative period.

However, it is necessary to conduct a comparative analysis with a larger sample size using statistical methods, including patients who have undergone conventional laminectomy to validate these findings.

**NOTES**

**Supplementary Material:** Supplementary video clip 1 can be found via https://doi.org/10.14245/ns.2347270.635.

**Conflict of Interest:** The authors have nothing to disclose.

**Funding/Support:** This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Author Contribution:** Conceptualization: DHK, JTH, JWH, ISK, HJL, JBL; Formal analysis: DHK, JTH; Investigation: DHK, JTH, JWH, ISK, HJL, JBL; Methodology: DHK, JTH, HJL, JBL; Project administration: JTH, ISK, HJL, JBL; Writing – original draft: DHK, JTH; Writing – review & editing: DHK, JTH.

**ORCID**

Dong Hun Kim: 0000-0002-4422-2997
Jae Taek Hong: 0000-0001-6453-0439
Jung Woo Hur: 0000-0003-2184-0044
Il Sup Kim: 0000-0001-6328-2955
Ho Jin Lee: 0000-0002-9208-5931
Jong Beom Lee: 0000-0002-6966-6033

**REFERENCES**


Original Article

Corresponding Author
Biao Wang
https://orcid.org/0000-0001-6631-9127

Department of Spine Surgery, Honghui Hospital, Xian Jiaotong University, No. 76 Nanguo Road, Nanshaomen, Xian 710054, Shaanxi Province, China
Email: wangbiaowb1987@126.com

Received: December 1, 2023
Revised: February 21, 2024
Accepted: February 22, 2024

*Jie Guo and Yesheng Bai contributed equally to this study as co-first authors.

A Comparative Study of 2 Techniques to Avoid Bone Cement Loosening and Displacement After Percutaneous Vertebroplasty Treating Unstable Kummell Disease

Jie Guo1,2,*, Yesheng Bai3,*, Liang Li3, Jiangtao Wang1,4, Yuhang Wang1, Dingjun Hao1, Biao Wang1

1Department of Spine Surgery, Honghui Hospital, Xian Jiaotong University, Xian, China
2Shaanxi University of Chinese Medicine, Xian, China
3Pain Ward, Rehabilitation Hospital of Xian Honghui Hospital, Xian Jiaotong University, Xian, China
4Medical School of Yan’an University, Yan’an, China

Objective: Percutaneous vertebroplasty (PVP) is currently the most common surgical procedure for unstable Kummell disease (KD), but cement loosening or displacement often occurs after PVP. We had been using percutaneous pediculoplasty (PPP) or a self-developed bone cement bridging screw system to avoid this severe complication. This study intends to compare these novel surgical procedures through a 2-year follow-up evaluation.

Methods: From May 2017 to May 2021, 77 patients with single-level unstable KD were included in the PPP group, and 42 patients received the PVP-bone cement bridging screw system were included in the screw group. The changes in the vertebral body index (VBI), bisegmental Cobb angle, visual analogue scale (VAS) and Oswestry Disability Index (ODI) and the cement loosening rate and displacement rate at different follow-up time points were used to evaluate the clinical efficacy.

Results: There was no significant difference in VBI or bisegmental Cobb angle between the 2 groups (p > 0.05) before operation, immediately after operation and at 6-month follow-up, while at 1-year and 2-year postoperative evaluations, the screw group had higher VBI and bisegmental Cobb angle than the PPP group (p < 0.05). Before operation, immediately after operation, at 6-month and 1-year follow-up, there was no significant difference in VAS or ODI score between the 2 groups (p > 0.05), while at 2-year follow-up, the screw group still had higher VAS and ODI scores than the PPP group (p < 0.05). No bone cement displacement occurred in both groups, but the rate of bone cement loosening was 14.29% in group PPP, and 0 in screw group (p < 0.05).

Conclusion: This 2-year follow-up study shows that the PVP-bone cement bridging screw system combined therapy had better midterm treatment efficacy than the PVP-PPP combined therapy in patients with unstable KD, and the bone cement bridging screw system is a preferred therapy with better anti-cement loosening ability.

Keywords: Pediculoplasty, Vertebroplasty, Kummell disease, Bone cement bridging screw, Bone cement loosening
INTRODUCTION

Kummell disease (KD), also known as vertebral ischemic osteonecrosis, is manifested by vertebral osteonecrosis and accompanied by vertebral instability, progressive kyphosis, and even neurological dysfunction. Almost all patients with KD have osteoporosis, so KD is also called symptomatic thoracic osteoporotic vertebral fracture.1,2 It usually involves a single vertebra, and is mostly observed in the thoracolumbar segment.3 Intra-vertebral vacuum cleft (IVC), as the signature pathological feature of KD, plays a pivotal role in the diagnosis of KD.4 The sensitivity and specificity of IVC in the diagnosis of KD is 85% and 99% respectively, meaning that KD is usually accompanied by IVC.5-7

KD is extremely rare in patients without osteoporosis, and it is often secondary to osteoporotic fractures. With today’s aging population and the increasing number of patients with osteoporotic vertebral fracture, the number of KD patients is also increasing.8,9 Conservative treatment does not have a satisfactory efficacy and may further deteriorate the symptoms, so it has gradually become an adjunctive therapy for the surgical treatment.10,11 With the progress in minimally invasive technology for the spine, percutaneous vertebroplasty (PVP) has become the mainstream therapy for unstable KD for its outstanding pain relief effect and improvement of kyphosis.12,13 Recovery of spinal stability and complete IVC filling by bone cement are the most important analgesic mechanisms in KD. However, due to the inability of bone cement in IVC to infiltrate normal bone tissues, the postoperative loosening rate of cement can reach as high as 25%, leading to recurrent and even more severe pain for patients.14 Meanwhile, bone cement loosening may further lead to the disastrous complication of cement displacement, causing patients to take out bone cement and undergo open surgery to rebuild spinal stability. Both anterior and posterior revision surgeries have the disadvantages of large trauma, high difficulty and high risk.

Any therapy that can avoid severe complications should be the best treatment. In past clinical work, our center used the percutaneous pediculoplasty (PPP) combined technique during PVP (PVP-PPP) and PVP-bone cement bridging screw system (Changzhou Geasure Medical Apparatus and Instruments Co., Ltd., Changzhou, China) combined technique to avoid bone cement loosening and displacement in unstable KD patients. Our previous studies involving 3-dimensional finite element analysis and biomechanical investigations have unequivocally demonstrated the substantial superiority of these 2 treatment methods over PVP.15,16 These 2 therapies intend to link the bone cement in IVC to normal bone tissues in the vertebrae as a “bridge” to anti bone cement loosening or displacement. It is the first time for both techniques to be used to avoid postoperative cement loosening or displacement in KD, and there were several problems need to be solved: (1) What are the differences in actual clinical efficacy, such as complication, pain relief and functional improvement, between these 2 therapies? (2) Can the bone cement bridging screw system effectively avoid bone cement loosening and displacement? (3) What are the difficulties, advantages and disadvantages of these 2 therapies? Given these, we carried out this retrospective case-control study on 119 patients who were treated with these 2 therapies and were followed up on for over 2 years during a 4-year period to define the above scientific questions.

MATERIALS AND METHODS

1. Patients

This clinical study protocol was approved by the Institutional Review Board of Honghui Hospital of Xi’an Jiaotong University (approval number: 201701007). All patients signed the written informed consent.

One hundred nineteen patients with single-segment KD in the thoracolumbar segment and without neurological symptoms were included in this study from January 2017 to May 2021. All patients met the criteria for diagnosis of unstable KD. The changes in the angles between the extension line of the upper endplate and the extension line of the lower endplate of the affected vertebra in the forward flexion position (vertebral body angle) and in the posterior extension position were compared in the thoracolumbar dynamic position x-ray films. When the change was > 11°, the patient was defined to have dynamic instability of the affected vertebra and was diagnosed with unstable KD.17 All patients had the main symptom of severe pain in the corresponding affected part or lower waist when the body position changed due to instability, and had it progressively aggravated.

Inclusion criteria: (1) patients with single-segment unstable KD; (2) patients whose bone mineral density (BMD) T-value was below -2.5; (3) patients with severe lumbago and backache but without neurological impairment; (4) patients having bilateral pedicles intact and not damaged.

Exclusion criteria: (1) patients with neurological symptoms; (2) patients with multivertebral lesions; (3) patients also with a metastatic spinal tumor, lumbar spondylolisthesis, and other diseases that caused lumbago and backache; (4) patients with-
out osteoporosis; (5) patients with congenital absence, deformity or damage of the pedicle of the affected vertebra; (6) patients who were followed up on for less than 2 years.

All patients underwent plain radiography, dynamic position x-ray, computed tomography (CT) and magnetic resonance imaging examinations. The x-ray films and CT findings before and after operation were retrospectively compared and analyzed. By surgical method, the patients were divided into 2 groups. Those who received the PVP-PPP combined therapy were grouped to the PPP group, and those who received the PVP-bone cement bridge screw system combined therapy were grouped to the screw group.

2. Surgical Methods

1) PVP-PPP combined therapy

After general anesthesia, a patient was placed in the prone position, and had soft pads under the chest and anterior superior iliac spine to completely suspend the abdomen. Under the lateral C-arm fluoroscopy, the operating bed was adjusted so that the patient was in an appropriate hyperextension position. Manual compression reduction was used for optimal vertebral height recovery. After satisfactory reduction, C-arm fluoroscopy was used to mark the body surface projection of the lateral margin of bilateral pedicle of the affected vertebra. After disinfection and draping, a unilateral puncture operation was performed on the side with a larger IVC fissure on the CT cross section. According to the distance from the body surface of the patient's affected site to the articular process, the skin was incised longitudinally 5 to 10 mm outside the body surface marker, with the incision of about 5 mm in length. A 3.0 mm-diameter bone cement trocar was used for puncture. Along the direction of the pedicle of the affected vertebra, the bone cement trocar was ensured to target and puncture into IVC inside the pedicle. C-arm fluoroscopy was used several times to verify the accuracy of the position of the bone cement trocar. With a satisfactory position of the trocar as verified, bone cement was prepared and, when it was in the wire-drawing stage, an appropriate volume was slowly injected under anteroposterior-lateral x-ray fluoroscopy, till the interspaces in the affected vertebra were fully filled. PVP was thus done. After that, when the bone cement was in the toothpaste-like stage, the slow injection of it was continued under lateral x-ray fluoroscopy, while the trocar was slowly receded toward the backside. These 2 operations should be performed in perfect union. Bone cement injection ended when the trocar was receded throughout the entire pedicle and close to the initial puncture point. PPP was thus done.

Fig. 1. Pediculoplasty combined with vertebroplasty and bone cement bridging screw system combined with vertebroplasty to treat Kummell disease surgical operation diagram. (A) Kummell disease causes vertebral body collapse and kyphosis. (B) After the collapsed vertebral body is reset and the spine hyperextension is corrected, the anterior edge of the vertebral body has a bone defect. (C) Use bone cement to fill the bone defect and the pedicle to complete the pediculoplasty combined with vertebroplasty treatment. (D) Sagittal view after pediculoplasty combined with vertebroplasty. (E) Schematic diagram of the postoperative axial position after pediculoplasty combined with vertebroplasty. (F) The vertebroplasty treatment is completed by the minimally invasive implantation of a bone cement bridging screw. (G) Sagittal view of the bone cement bridging screw system combined with vertebroplasty. (H) Schematic diagram of the postoperative axial position of the bone cement bridging screw system combined with vertebroplasty.

https://doi.org/10.14245/ns.2347274.637
After bone cement had completely set, the trocar was withdrawn. The detailed diagrams illustrating the surgical technique are presented in (Fig. 1C–E), while (Fig. 2B and D) display postoperative x-ray and CT data of a representative patient.

2) PVP-bone cement bridge screw system combined therapy

After general anesthesia, a patient was placed in the prone position. Similarly, after satisfactory reduction, C-arm fluoroscopy was used to mark the body surface projection of the lateral margin of bilateral pedicle of the affected vertebra. Through a unilateral approach, the operation was performed on the side with more severe vertebral fissures on the CT cross section. Depending on the patient’s body shape (fat or thin), the skin was incised longitudinally 5 to 10 mm outside the body surface marker, with the incision of about 10 mm in length, and then subcutaneous tissues and lumbodorsal fascia were incised layer by layer. Under the guidance of anterior-posterior x-ray fluoroscopy, the puncture needle tip was placed at the outer margin of the projection of the pedicle (The left corresponds to the 10 o’clock position, while the right corresponds to the 2 o’clock position), and then the puncture was performed at the fracture in the vertebral body of the affected vertebra with an inward tilt of 10° to 15°. Under, When the puncture needle tip reached the posterior edge of the vertebral body under the guidance of lateral x-ray fluoroscopy, the anterior-posterior x-ray fluoroscopy was used again, under which the puncture needle tip did not break through the medial margin in the projection of the pedicle (this step could avoid the screw going into the spinal canal as much as possible). After that, the targeted puncture was continued into the fissure in the vertebral body under lateral x-ray fluoroscopy, and the guide needle was placed in. Under the guidance of the guide needle, a new-type bone cement bridge screw with a bone cement outlet was placed into IVC inside the vertebral body of the affected vertebra. The screw placing handle and the core of the screw placing extension rod were then removed. Bone cement was prepared, and injected into a special push rod for the bone cement bridge screw. When bone cement was in the wire-drawing stage or clustering stage, an appropriate volume of bone cement was injected.
The bone cement was slowly injected along the core of the extension rod under anteroposterior-lateral x-ray fluoroscopy. After IVC in the vertebral body of the affected vertebra was fully filled, bone cement injection ended. After bone cement had completely set, the screw tail extension rod was removed and the incision was sewn up. The detailed diagrams illustrating the surgical technique are presented in (Fig. 1F–H), while (Fig. 2A and C) display postoperative x-ray and CT data of a representative patient.

3. Postoperative Treatment
After the patient woke up from general anesthesia, he/she was observed for symptoms of nerve damage. After the operation, patients were routinely given antibiotics for 24 hours, and asked to rest in bed until 1 day after operation, when they could get out of bed and move under the protection of braces. The rehabilitation doctors guided postoperative functional exercises and matters needing attention. The patients were discharged after having stable disease development. After discharge, the patients were asked to keep wearing the braces for protection for 12 weeks, and also to take calcium supplements, Vitamin D3 and diphosphonates to treat osteoporosis under the guidance of osteoporosis doctors.

4. Follow-up Evaluation
Follow-up evaluations were carried out at 6 months, 1 year, and 2 years after operation. X-ray or CT was used to evaluate whether bone cement was displaced and whether the bone cement bridge screw was loose, displaced or broken. The vertebral body index (VBI), which represents the ratio of anterior to posterior vertebral height multiplied by 100% for the affected vertebral body, was measured to assess postoperative improvement in vertebral body height. The bisegmental Cobb angle, defined as the angle between the upper endplate of the first vertebra above the affected one and the lower endplate of the first vertebra below it, was measured to evaluate kyphosis correction (Fig. 3). CT was used to evaluate bone cement loosening. The criterion for bone cement loosening was defined as the presence of a continuous low-density region observed between the bone tissues and bone cement during CT follow-up examinations.

The comparison results of the patients’ visual analogue scale (VAS) were used to evaluate pain relief in the patients, with the score ranging from 0 to 10 points.

The comparison results of the patients’ Oswestry Disability Index (ODI) scores were used to evaluate the recovery of thoracolumbar functions in the patients. The ODI questionnaire has 10 questions covering the intensity of pain, personal care, lifting, walking, sitting, standing, sleeping, sex life, social life and traveling. Each question has 6 options, and is scored 0 to 5 points. ODI is used to evaluate the degree of impact of low back and leg pain on the daily life of patients. The calculation formula is ODI (%) = actual score/(actual number of questions answered × 5) × 100%.

5. Statistical Analysis
This study is a retrospective cohort study. It used the IBM SPSS Statistics ver. 26.0 (IBM Co., Armonk, NY, USA) statistical software package for statistical analysis. All measurement data were expressed as (mean ± standard deviation). The results of VBI, bisegmental Cobb angle, VAS, and ODI of the same group before operation, after operation and at the last follow-up visit were compared using repeated measures analysis of variance. Pairwise comparisons of VBI, bisegmental Cobb angle, VAS, or ODI were conducted using paired t-test. Results of them of the PPP group and the screw group were compared using independent samples t-test. Postoperative bone cement loosening of the 2 groups was compared using Fisher exact test. Enumeration data of the 2 groups were compared using chi-square test. When p < 0.05, the difference was deemed significant.

Fig. 3. Schematic diagram of imaging parameter measurement. The measurement methods of bisegmental Cobb angle is shown in the figure, and the measurement method of vertebral body index is (a/b) × 100%.
RESULTS

1. Operation Results

Among the 119 patients, 77, including 19 males and 58 females, were grouped to the PPP group. They were 72.52 ± 5.44 years old on average (range, 63–85 years), their preoperative BMD T-value measured by dual x-ray absorptiometry was -3.21 ± 0.57 on average (range, -2.5 to -5.2), and 3 had segment T10 affected, 9 had segment T11 affected, 37 had segment T12 affected, 21 had segment L1 affected, and 7 had segment L2 affected. The screw group had 42 patients, including 11 males and 31 females. They were 74.36 ± 5.07 years old on average (range, 64–88 years), their preoperative BMD T-value measured by dual x-ray absorptiometry was -3.38 ± 0.51 on average (range, -2.6 to -4.5), and 2 had segment T10 affected, 5 had segment T11 affected, 19 had segment T12 affected, 13 had segment L1 affected, and 3 had segment L2 affected. Statistical analysis showed that there were no significant differences in age, sex, BMD and affected segment between the 2 groups (p > 0.05), and the baseline data of the 2 groups were comparable. All 119 patients had the operation successfully done. The average operation duration of the PPP group was 85.52 ± 10.78 minutes (range, 70–115 minutes), and its average bone cement injection volume was 4.98 ± 0.67 mL (range, 4–6 mL). The average operation duration of the screw group was 52.07 ± 9.90 minutes (range, 36–65 minutes), and its average bone cement injection volume was 4.43 ± 0.89 mL (range, 2.5–6 mL). The screw group had a shorter operation duration and less bone cement volume than the PPP group, indicating that the PVP-bone cement bridge screw system combined therapy could shorten the operation duration and reduce bone cement volume. Among the 119 patients, 111 were not observed bone cement leakage (93.28%) and 8 had bone cement leakage (6.72%). In the PPP group, 6 patients had bone cement leakage, including 2 had it leak from the anterior margin of the vertebral body, 2 had it leak from the paravertebral vein, 1 had it leak from the sidewall of the vertebral body and 1 had it leak from the superior intervertebral disc. In the screw group, 2 had bone cement leakage, including 1 had it leak from the anterior margin of the vertebral body and 1 had it leak from the sidewall of the vertebral body. None of the patients had leakage inside the spinal canal or from the intervertebral foramen. None of the 119 patients had severe complications, such as symp-

Table 1. Basic information and clinical information of the patients

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>PPP group (n = 77)</th>
<th>Screw group (n = 42)</th>
<th>t/χ²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, male:female</td>
<td>19:58</td>
<td>11:31</td>
<td>0.033</td>
<td>0.856</td>
</tr>
<tr>
<td>Age (yr)</td>
<td></td>
<td></td>
<td>-1.341</td>
<td>0.183</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>72.52 ± 5.44</td>
<td>74.36 ± 5.07</td>
<td>1.661</td>
<td>0.099</td>
</tr>
<tr>
<td>Range</td>
<td>63–85</td>
<td>64–88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMD T-value</td>
<td></td>
<td></td>
<td>0.347</td>
<td>0.987</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>-3.21 ± 0.57</td>
<td>-3.38 ± 0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>-2.5 to -5.2</td>
<td>-2.6 to -4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T10</td>
<td>3</td>
<td>2</td>
<td>16.645</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>T11</td>
<td>9</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T12</td>
<td>37</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>21</td>
<td>13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>7</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operation duration (min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>85.52 ± 10.78</td>
<td>52.07 ± 9.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>70–115</td>
<td>36–65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volum of cement (mL)</td>
<td></td>
<td></td>
<td>3.816</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>4.98 ± 0.67</td>
<td>4.43 ± 0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>4–6</td>
<td>2.5–6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone cement loosening</td>
<td>11</td>
<td>0</td>
<td></td>
<td>Fisher 0.008</td>
</tr>
</tbody>
</table>

PPP, percutaneous pediculoplasty; SD, standard deviation; BMD, bone mineral density.
2. Follow-up Visit Results

During the 2-year follow-up, no patient in the screw group had screw loosening, displacement or fracture, no had bone cement loosening or delayed displacement, and no received revision surgery. Besides, during the follow-up, no patient had skin or soft tissue wear or discomfort caused by the screw. In the PPP group, 11 patients had bone cement loosening (14.29%), and no had bone cement displacement or received revision surgery. Analysis of the bone cement loosening of the 2 study groups by Fisher exact test showed that the PPP group had 11 patients with bone cement loosening, significantly greater than 0 patients in the screw group, with the difference of significant statistical significance (p = 0.008) (Table 1).

The evaluation of the improvement of the vertebral body height of the affected vertebra showed that the VBI of the PPP group improved from 62.51 ± 6.88 before operation to 83.28 ± 2.15 after operation, with the difference of statistical significance (p < 0.001), and although VBI dropped slightly in later follow-up visits and to 82.60 ± 2.18 at the last follow-up visit, the differences between it and that immediately after operation and at 1 year after operation (83.55 ± 2.19) were not statistically significant (p = 0.752, p = 0.967). The statistical results of VBI of the 2 groups were demonstrated in (Table 2, Fig. 4). The comparison of VBI scores between the 2 groups at different follow-up time points showed that the VBI score improved more significantly after the PVP-bone cement bridge screw system combined therapy. There was no significant difference in VBI score between the PPP group and the screw group before and immediately after operation and at 6 months after operation (p > 0.05), but the follow-up visits at 1 year and 2 years after operation showed that there were significant statistical differences between the screw group (83.55 ± 2.19 at 1 year after operation, 83.53 ± 2.45 at 2 years after operation) and the PPP group (82.69 ± 2.22 at 1 year after operation, 82.60 ± 2.18 at 2 years after operation) (p < 0.05).

The evaluation of kyphosis correction showed that the bisegmental Cobb angle of the PPP group was corrected from 29.06° ± 3.89° before operation to 16.01° ± 3.53° after operation, with the difference of statistical significance (p < 0.001), and at the last follow-up visit, the patients had the thoracolumbar kyphosis angle slightly increase, to 17.06° ± 3.70° on average, but the differences between it and that immediately after operation and at 1 year after operation (16.81° ± 3.61°) were not statistically significant (p = 0.057, p = 0.697). The evaluation also showed that the bisegmental Cobb angle of the PPP group was corrected from 29.19° ± 4.12° before operation to 14.72° ± 3.52° after operation, with the difference of statistical significance (p < 0.001),

Table 2. Statistical results of vertebral body index in 2 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PPP group</th>
<th>Screw group</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative</td>
<td>62.51 ± 6.88</td>
<td>62.02 ± 6.89</td>
<td>0.365</td>
<td>0.716</td>
</tr>
<tr>
<td>Postoperative</td>
<td>83.28 ± 2.15</td>
<td>83.70 ± 2.32</td>
<td>-0.990</td>
<td>0.324</td>
</tr>
<tr>
<td>6-Month follow-up</td>
<td>82.82 ± 2.40</td>
<td>83.58 ± 2.55</td>
<td>-1.599</td>
<td>0.112</td>
</tr>
<tr>
<td>1-Year follow-up</td>
<td>82.69 ± 2.22</td>
<td>83.55 ± 2.19</td>
<td>-2.030</td>
<td>0.045</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>82.60 ± 2.18</td>
<td>83.53 ± 2.45</td>
<td>-2.134</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.
PPP, percutaneous pediculoplasty.

https://doi.org/10.14245/ns.2347274.637
and at the last follow-up visit, the patients had the thoracolumbar kyphosis angle slightly increase, to 15.56° ± 3.25° on average, but the differences between it and that immediately after operation and at 1 year after operation (15.40° ± 3.66°) were not statistically significant (p = 0.270, p = 0.858). The statistical results of the bisegmental Cobb angle of the 2 groups were demonstrated in (Table 3, Fig. 5). The comparison of the bisegmental Cobb angle between the 2 groups at different follow-up time points showed that the bisegmental Cobb angle after the PVP-bone cement bridge screw system combined therapy was significantly smaller than that of the PPP group. There was no significant difference in bisegmental Cobb angle between the PPP group and the screw group before and immediately after operation and at 6 months after operation (p > 0.05), but the follow-up visits at 1 year and 2 years after operation showed that there were significant statistical differences between the screw group (15.40° ± 3.66° at 1 year after operation, 15.56° ± 3.25° at 2 years after operation) and the PPP group (16.81° ± 3.61° at 1 year after operation, 17.06° ± 3.70° at 2 years after operation) (p < 0.05).

The VAS pain score and ODI function score of the PPP group improved from 7.55 ± 0.27 and 76.90 ± 4.77 points before operation to 2.81 ± 0.41 and 49.19 ± 3.87 points after operation, 2.00 ± 0.40 and 20.33 ± 2.64 points at 1 year after operation, and 1.87 ± 0.30 and 19.32 ± 2.60 points at the last follow-up visit. The VAS score improved by 75.23% and the ODI score improved by 74.88 from that before operation to that at the 2-year last follow-up visit. When comparing either of the 2 scores before operation, that after operation, that at 6-month follow-up visit and that at the last follow-up visit, there were statistically significant differences between either of them at 1 year after operation and that at the last follow-up visit (p = 0.091, p = 0.060). The statistical results of the VAS and ODI scores of the 2 groups were demonstrated in (Tables 4, 5; Figs. 6, 7). The comparison of the VAS pain score and ODI function score between the 2 groups at different follow-up time points showed that the VAS and ODI scores after the PVP-bone cement bridge screw system combined therapy were significantly smaller than those of the PPP group. There was no significant difference in VAS or ODI score between the PPP group and the screw group before and immediately after operation and at 6 months and 1 year after operation (p > 0.05), but at the 2-year follow-up visit, there were significant statistical differences between the screw group (1.87 ± 0.30, 19.32 ± 2.60) and the PPP group (2.09 ± 0.33, 20.47 ± 2.90) (p < 0.05).

Table 3. Statistical results of bisegmental Cobb angle (°) in 2 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PPP group (n = 77)</th>
<th>Screw group (n = 42)</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative</td>
<td>29.06 ± 3.89</td>
<td>29.19 ± 4.12</td>
<td>-0.175</td>
<td>0.861</td>
</tr>
<tr>
<td>Postoperative</td>
<td>16.01 ± 3.53</td>
<td>14.72 ± 3.52</td>
<td>1.913</td>
<td>0.058</td>
</tr>
<tr>
<td>6-Month follow-up</td>
<td>16.58 ± 3.41</td>
<td>15.25 ± 3.89</td>
<td>1.943</td>
<td>0.054</td>
</tr>
<tr>
<td>1-Year follow-up</td>
<td>16.81 ± 3.61</td>
<td>15.40 ± 3.66</td>
<td>2.034</td>
<td>0.044</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>17.06 ± 3.70</td>
<td>15.56 ± 3.25</td>
<td>2.204</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.

PPP, percutaneous pediculoplasty.
DISCUSSION

As the global population aging aggravates, the number of patients with osteoporotic fracture of the vertebral body is increasing year by year, and so is the incidence rate of KD.\(^1,9,18,19\) IVC is the characteristic imaging manifestation of KD, and it, as the characteristic pathological change, causes the loss of spinal stability in KD and neuropathic pain in patients.\(^4,20\)

Different surgical methods, such as PVP, percutaneous kyphoplasty (PKP), percutaneous fixation, open reduction and osteotomy-orthopedics have been reported in literature.\(^21-28\) Although PVP can satisfactorily relieve pain and improve kyphosis and has become one of the most important ways to treat KD, the surgical therapies for KD are still controversial in clinical practice. IVC is a dead cavity formed by pathologically confirmed necrotic bone tissues. Filling it with bone cement can rebuild spinal stability and relieve patients’ pain. However, it is difficult for bone cement to penetrate into the normal trabecular structure, and bone cement occupies fissures only and has no mechanical interlocking or biocompatibility with the surrounding bone tissues, so the two are not bonded. Therefore, even if the postoperative efficacy is satisfactory, there are still risks of bone cement loosening and displacement during follow-up visit.

In this study, we used the PVP-bone cement bridge screw system combined therapy for KD to investigate the differences in efficacy between it and the PVP-PPP combined therapy, as well as its clinical effectiveness. With the use of the bone cement bridge screw system, bone cement in IVC can join the whole vertebral body through the bone cement bridge screw and the bone cement inside the screw, thus avoiding the risks of bone cement loosening and displacement. None of the 42 patients in

### Table 4. Statistical results of visual analogue scale score in 2 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PPP group (n = 77)</th>
<th>Screw group (n = 42)</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative</td>
<td>7.47 ± 0.35</td>
<td>7.55 ± 0.27</td>
<td>-1.326</td>
<td>0.188</td>
</tr>
<tr>
<td>Postoperative</td>
<td>2.91 ± 0.30</td>
<td>2.81 ± 0.41</td>
<td>1.374</td>
<td>0.174</td>
</tr>
<tr>
<td>6-Month follow-up</td>
<td>2.64 ± 0.28</td>
<td>2.55 ± 0.28</td>
<td>1.651</td>
<td>0.101</td>
</tr>
<tr>
<td>1-Year follow-up</td>
<td>2.14 ± 0.35</td>
<td>2.00 ± 0.40</td>
<td>1.908</td>
<td>0.059</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>2.09 ± 0.33</td>
<td>1.87 ± 0.30</td>
<td>3.510</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. PPP, percutaneous pediculoplasty.

### Table 5. Statistical results of Oswestry Disability Index score in 2 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PPP group (n = 77)</th>
<th>Screw group (n = 42)</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative</td>
<td>78.08 ± 4.49</td>
<td>76.90 ± 4.77</td>
<td>1.335</td>
<td>0.184</td>
</tr>
<tr>
<td>Postoperative</td>
<td>50.39 ± 5.75</td>
<td>49.19 ± 3.87</td>
<td>1.365</td>
<td>0.175</td>
</tr>
<tr>
<td>6-Month follow-up</td>
<td>32.81 ± 2.96</td>
<td>31.87 ± 3.00</td>
<td>1.651</td>
<td>0.101</td>
</tr>
<tr>
<td>1-Year follow-up</td>
<td>21.21 ± 2.83</td>
<td>20.33 ± 2.64</td>
<td>1.660</td>
<td>0.100</td>
</tr>
<tr>
<td>Final follow-up</td>
<td>20.47 ± 2.90</td>
<td>19.32 ± 2.60</td>
<td>2.141</td>
<td>0.034</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. PPP, percutaneous pediculoplasty.

Fig. 6. Comparison of preoperative and postoperative visual analogue scale (VAS) score between percutaneous pediculoplasty (PPP) group and screw group at different time points. p < 0.05, VAS score of patients in PPP group compared with screw group.

Fig. 7. Comparison of preoperative and postoperative Oswestry Disability Index (ODI) score between percutaneous pediculoplasty (PPP) group and screw group at different time points. p < 0.05, ODI score of patients in PPP group compared with screw group.

https://doi.org/10.14245/ns.2347274.637  www.e-neurospine.org  583
the screw group had postoperatively delayed bone cement loosening or displacement, while 11 of the 77 patients in the PPP group had it (p = 0.008), which well demonstrated the effectiveness of the fixation strength of the bone cement bridge screw system. After 2 years of follow-up evaluations, all patients had imaging parameters significantly improved. The statistical analysis results showed that, although there was no significant difference in the improvement of VBI or bisegmental Cobb angle between the screw group and the PPP group immediately after operation and at 6 months after operation (p > 0.05), the screw group had significantly higher VBI and bisegmental Cobb angle than the PPP group at 1-year and 2-year midterm postoperative evaluations (p < 0.05). Moreover, in terms of pain relief and functional improvement, there was no significant difference in VAS pain score or ODI function score between the screw group and the PPP group before and immediately after operation and at 6 months and even 1 year after operation (p > 0.05). However, at 2 years after operation, the screw group had higher VAS pain and ODI function scores than the PPP group (p < 0.05). Bone cement loosening is closely related to the patient's imaging manifestations and pain degree. Bone cement loosening causes the anterior margin of the patient’s vertebral body to collapse slightly, the correction of the patient’s VBI and bisegmental Cobb angle to loosen, the patient’s pain to deteriorate, and the patient's functions to be affected. The results of this study proved that the PVP-bone cement bridge screw system combined therapy was safer and more effective for unstable KD than the PVP-PPP combined therapy. After PVP-PPP treatment, it is evident from the outcomes that bone cement loosening results in heightened pain levels and diminished lumbar function among patients, frequently observed 1 year after surgery.

The cause of bone cement loosening and displacement after KD surgery is still unknown. We summarized and analyzed the information of almost all KD patients with bone cement displacement after PVP or PKP reported in literature in recent years, and found that all these patients had bone defects at the anterior margin of the vertebral body and bone cement relocated from the anterior part of the defected vertebral body to the ventral part. Although dynamic x-ray images have not been used to measure changes in vertebral body angle, the majority of cases can be inferred as unstable KD based on the alterations in preoperative and postoperative vertebral body angles and significant anterior vertebral bone defects. During PVP, bone cement tends to be distributed at the anterior margin of the vertebral body under the action of injection pressure, and it easily penetrates into the anterior vertebra in the event of insufficient bone mass at the anterior margin of the vertebral body. Even if the correction of kyphosis is significantly improved after operation, the patient may still have mild kyphosis after operation, so in postoperative movements, the center of gravity of the affected vertebra moves forward and the stress extrudes bone cement to produce a forward force, thus causing bone cement to move forward if the anterior margin of the vertebral body lacks the bone batter. Furthermore, in unstable KD, after the filling by bone cement, some patients may still have poor stability in the affected vertebra and are likely to have bone cement loosening or even displacement under repeated flexion and extension activities after operation. Due to the presence of IVC, it is difficult for bone cement to penetrate into the normal trabecular structure, and bone cement occupies fissures only and has no mechanical interlocking or biocompatibility with the surrounding bone tissues, which is another important reason for bone cement leakage, loosening and displacement.

Given the above factors, we developed a new-type bone cement bridge screw system in order to link bone cement to the surrounding bone tissues to avoid bone cement loosening and displacement during or after operation. The key technology was that after the targeted placement of the bone cement bridge screw in IVC, bone cement would be slowly released through the lateral hole at the front of the screw so that the injection pressure of bone cement was smaller, and meanwhile, the screw remained in the vertebral body to link bone cement with normal bone tissues and the pedicle. In this way, the bone cement screw acted as a “bridge” linking bone cement to surrounding tissues, even to the toughest pedicle to enhance mechanical fusion and avoid bone cement loosening and displacement. The tailless design of the screw also allowed the bone cement screw to be fully positioned in the bone structure and less likely to wear away muscle and other soft tissues while acting as a bridge during movement. At the end of the 2-year follow-up, none of the 42 patients had complications such as loosening, displacement, or fracture of the bone cement screw, none had delayed bone cement loosening or displacement or skin or soft tissue wear or discomfort caused by the screw. In terms of the main measurement indicators, including VBI, bisegmental Cobb angle, VAS pain score, and ODI function score, the PVP-bone cement bridge screw system combined therapy had much better effects than the PVP-PPP combined therapy.

Compared with the PVP-PPP combined therapy, the PVP-bone cement bridge screw system combined therapy also had great advantages in terms of surgical difficulty. The average operation duration of the screw group was 52.07 ± 9.90 minutes.
(range, 36–65 minutes) and its average bone cement injection volume was 4.43 ± 0.89 mL (range, 2.5–6 mL), while the average operation duration of the PPP group was 85.52 ± 10.78 minutes (range, 70–115 minutes) and its average bone cement injection volume was 4.98 ± 0.67 mL (range, 4–6 mL). The statistical analysis showed that the PVP-bone cement bridge screw system combined therapy had a far shorter operation duration than the PVP-PPP combined therapy, and far less bone cement volume. Based on our treatment experience, the greatest technical difficulties in PPP surgery were the accuracy of bone cement trocar displacement and prevention of bone cement leakage, which could lead to severe complications. First of all, it is a must to ensure the trocar is accurately placed into IVC, and secondly, the intactness of the 4 walls of the pedicle must be ensured during the operation because damage to the inner or lower wall may lead to leakage from the spinal canal or nerve root foramen during the subsequent injection of bone cement, thus resulting in symptoms of nerve injury. In this case, surgeons usually have to perform open surgery to remove the bone cement that compresses the nerve. Tomasian et al. reported that bone cement tended to be distributed in the spinal canal during PPP surgery due to its low impedance. The study conducted by Liu et al. revealed that the incidence of cement leakage was higher in PVP-PPP (29.4%) compared to PVP alone (15.4%). In order to avoid the aforesaid leakages of bone cement, Wang et al. even attempted robots in PPP to ensure safety. However, such leakages of bone cement barely occur in bone cement bridge screw operation, and the operation is easy to operate and safer.

This study, as the first and early clinical retrospective study of the PVP-bone cement bridge screw system combined therapy and the PVP-PPP therapy for unstable KD, has the following deficiencies. First, all the patients included in this study had the lesion in the single vertebral body, but there were also patients with the lesion in multiple vertebral segments in clinical practice, so the results of this study cannot be extensively applied to patients with multivertebral KD for the time being. Second, although this paper has the largest sample size so far among the studies of its kind, 119 patients are still a small sample size, which may limit the universality of the results of this study. Studies of a larger sample size are necessary. Moreover, this study cannot exclude potential interference factors such as the surgeon’s habits and surgical conditions. Fully randomized controlled studies, longer follow-up and multicenter studies are still needed in the future to ultimately lay a solid clinical data basis for the promotion of the bone cement bridge screw system.

CONCLUSION

This 2-year follow-up study on 119 patients showed that compared with the PVP-PPP combined therapy, the PVP-bone cement bridge screw system combined therapy had better mid- and short-term treatment efficacy in patients with unstable KD, could more effectively avoid bone cement loosening and displacement, quickly relieve pain, restore vertebral body height and improve kyphosis, but its long-term efficacy needs further research.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: The National Natural Science Foundation of China (No.81820167); Key Research and Development Program of Shaanxi Province (No.2020GXLH-Y-003); Key Research and Development Program of Shaanxi Province (No. 2020SFY-095).

Author Contribution: Conceptualization: JG, LL, BW; Formal analysis: JG, YB, JW; Investigation: JG, LL, YB, YW, JW; Methodology: JG, YB, BW, DH; Project administration: BW; Writing – original draft: JG; Writing – review & editing: BW, DH.

ORCID

Jie Guo: 0000-0002-4513-0701
Yesheng Bai: 0009-0003-2176-8158
Liang Li: 0000-0002-3807-2291
Jiangtao Wang: 0000-0002-4625-2559
Yuhang Wang: 0009-0002-8926-2615
Dingjun Hao: 0000-0002-7768-1700
Biao Wang: 0000-0001-6631-9127

REFERENCES


https://doi.org/10.14245/ns.2347274.637
Incidence and Survival of Patients With Malignant Primary Spinal Cord Tumors: A Population-Based Analysis

Huanbing Liu, Linnan Duan, Zhibin Li, Yuanhao Liu, Yubo Wang
Department of Neurosurgery, First Hospital of Jilin University, Changchun, China

Objective: Epidemiological studies on spinal cord tumors are rare, and studies on primary intramedullary tumors are even rarer. The incidence and survival of patients with primary intramedullary spinal cord tumors have not been well documented. We aimed to study the incidence and survival of patients with primary spinal cord malignant and borderline malignant tumors based on data from the Surveillance, Epidemiology, and End Results (SEER) database and provide information for revealing the epidemiology and exploring the prognosis of patients with primary intramedullary tumors.

Methods: Patients in the SEER database with microscopically diagnosed malignant and borderline malignant primary spinal cord tumors from 2000 and 2019 were included in this study. We analyzed the distribution of patients according to the demographic and clinical characteristics. Then, we extracted the incidence rate and 5-year relative survival for the whole cohort and different subgroups of the cohort. Finally, multivariate Cox proportional hazards models were used to analyze the independent prognostic factors associated with overall survival.

Results: A total of 5,211 patients with malignant and borderline malignant primary spinal cord tumors were included in this cohort study. Ependymoma, astrocytoma (including oligodendrogliomas and glioblastoma), lymphoma and hemangioblastoma were the most common pathological types. The age-adjusted incidence rates of primary spinal cord ependymoma was 0.18 per 100,000. The incidence rate for females was significantly lower than that for males. The incidence rate was highest in Caucasian. The incidence rate of ependymoma was significantly higher than that of other pathological types. The incidence of astrocytoma was highest among people aged 0–19 years, the incidence of ependymoma was highest among people aged 40–59 years, and the incidence of lymphoma was highest among people aged 60 years or older. The 5-year observed survival and relative survival rates for the whole cohort were 82.80% and 86.00%, respectively. Patients diagnosed with ependymoma had significantly better survival than their counterparts. We also found the impact of surgery and chemotherapy on the prognosis of patients with different tumors varies a lot.

Conclusion: We conducted a population-based analysis of malignant and borderline malignant primary spinal cord tumors with the aim of revealing the epidemiology and survival of patients with primary intramedullary spinal cord tumors. Despite some shortcomings, this study provides valuable information to help us better understand the epidemiological characteristics of primary intramedullary spinal cord tumors.

Keywords: Intramedullary tumor, Ependymoma, Epidemiology, Survival, SEER Program

INTRODUCTION

The spinal cord is one of the 2 components of the central nervous system (CNS). According to the recent CBTRUS (Central Brain Tumor Registry of the United States) report on primary CNS tumors,1,3 tumors of the cranial nerves and the spinal cord/cauda equina accounted for approximately 10% of all CNS tumors. We searched PubMed with the medical subject heading...
(MeSH) terms for spinal cord tumors, which yielded less than 7% of the results obtained for a search on brain tumors. We obtained even fewer results when searching PubMed with the MeSH terms for spinal cord tumors and epidemiology. We believe that the incidence and survival of patients with primary intramedullary spinal cord tumors have not been well documented. The Surveillance, Epidemiology, and End Results (SEER) Program provides information on cancer incidence and survival in an effort to reduce the cancer burden among the U.S. population; it is one of the largest databases of clinical information worldwide. The SEER database includes approximately 48.0% of the U.S. population, including 42.0% of Caucasian, 44.7% of African American people, 66.3% of Hispanic people, 59.9% of American Indian and Alaska Native people, 70.7% of Asian people, and 70.3% of Hawaiian/Pacific Islanders.4 The number of SEER-based studies targeting spinal canal tumors has been increasing in recent years. For example, Cao et al.5 reported the epidemiology and survival of patients with spinal meningiomas. Primary spinal ependymomas,6 lymphomas7 and hemangioblastomas8 have also been reported. However, to our knowledge, no studies of primary intramedullary tumors have been conducted. Although it is difficult to isolate cases of intramedullary tumors with the variables in the SEER database, we found that if we included malignant and borderline malignant primary spinal cord tumors, the data would cover almost 99% of intramedullary tumors. Therefore, we conducted this analysis of data in the SEER database that described the incidence and survival of patients with malignant and borderline malignant primary spinal cord tumors, with the aim of helping clinicians better understand the epidemiology of patients with primary intramedullary tumors.

MATERIALS AND METHODS

Since the data from the SEER registry were de-identified and publicly available, no institutional review board approval was necessary and no informed consent was signed for this study. This study was exempt from Institutional Review Board approval because the original data were from a public database.

The latest version of the SEER database, SEER Research Plus Data 17 Registries, was adopted.9 The patients who were diagnosed between 2000 and 2019 were considered for inclusion in the study. The patients with microscopically diagnosed malignant and borderline malignant primary spinal cord tumors were indexed. Then, we selected patients with primary sites of C72.0, in the spinal cord, and C72.1, in the cauda equina. The exclusion criteria were as follows: (1) patients with unknown age, (2) surviving patients with unknown survival time, and (3) patients diagnosed by death certificate only or autopsy only. The primary site of the tumor and tumor behavior were confirmed according to the third edition of the International Classification of Diseases for Oncology (ICD-O-3).

The detailed demographic and clinical data were extracted by the case listing session. We analyzed the characteristics of the patients according to the demographic and clinical variables, including age, sex, race, pathology type and treatments. The SEER adolescents and young adults (AYA) site recode 2020 revision10 was adopted as the classification scheme to analyze the distribution of pathology types among the patients. Then, we divided the patients into different groups according to their demographic data and extracted the incidence rates of the whole cohort and different subgroups. We aimed to find the most common tumors in different demographic groups. Incidence rates are expressed as number of patients affect per 100,000 and age-adjusted to the 2000 United States Standard Population (19 age groups-Census P25-1130) standard; 95% confidence intervals (Tiwari modified11) were calculated for rates and ratios. The rate ratio was got by comparing the rate for each subsequent grouping the reference grouping’s rate.12 Statistical significance for the rate ratio was defined as p < 0.05. We extracted the 5-year observed survival (overall survival with all causes of death), relative survival (survival in the absence of other causes of death) data based on the survival session using the Kaplan-Meier method, and the Ederer II method was adopted as the cumulative expected method (which is necessary to calculate the relative survival).13,14 The incidence and survival were calculated and extracted by SEER*Stat version 8.4.0 software.

Finally, multivariate Cox proportional hazards models were used to analyze the independent prognostic factors associated with overall survival for the whole cohort and patients with different tumors separately. The patients were stratified based on demographic and clinical data and treatments. The overall survival was defined as the number of months from the diagnosis of tumor to death due to any cause; statistical significance was defined as p < 0.05. The hazard ratios and 95% confidence intervals were calculated by IBM SPSS Statistics 25.0 software (IBM Corporation, Armonk, NY, USA).

RESULTS

We indexed 8,412 patients with benign primary spinal cord
Malignant Primary Spinal Cord Tumors

Liu H, et al.

A total of 91.8% of them were diagnosed with meningiomas, neurilemmoma and neurofibroma, with only 62 cases of glioma. Ultimately, 5,211 patients with malignant and borderline malignant primary spinal cord tumors were indexed and included in this study. The demographic and clinical data and treatments of the patients are shown in Table 1. We found that female patients accounted for 45.5% of the whole cohort.

Table 1. Demographic and clinical data and treatments of patients with primary spinal cord malignant tumors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2,372 (45.5)</td>
</tr>
<tr>
<td>Male</td>
<td>2,839 (54.5)</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>42.92 ± 20.376</td>
</tr>
<tr>
<td>Median (range)</td>
<td>45 (0–95)</td>
</tr>
<tr>
<td>0–19</td>
<td>791 (15.2)</td>
</tr>
<tr>
<td>20–39</td>
<td>1,381 (26.5)</td>
</tr>
<tr>
<td>40–59</td>
<td>1,889 (36.3)</td>
</tr>
<tr>
<td>≥ 60</td>
<td>1,150 (22.1)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>4,360 (83.7)</td>
</tr>
<tr>
<td>Black</td>
<td>399 (7.7)</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>333 (6.4)</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>41 (0.8)</td>
</tr>
<tr>
<td>Unknown</td>
<td>78 (1.5)</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4,653 (89.3)</td>
</tr>
<tr>
<td>None</td>
<td>542 (10.4)</td>
</tr>
<tr>
<td>Unknown</td>
<td>16 (0.3)</td>
</tr>
<tr>
<td>Radiation</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1,355 (26)</td>
</tr>
<tr>
<td>None/unknown</td>
<td>3,856 (74)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>697 (13.4)</td>
</tr>
<tr>
<td>None/unknown</td>
<td>4,514 (86.6)</td>
</tr>
<tr>
<td>Survival months</td>
<td></td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>87.28 ± 66.062</td>
</tr>
<tr>
<td>Median (range)</td>
<td>79 (0–239)</td>
</tr>
<tr>
<td>Vital status</td>
<td></td>
</tr>
<tr>
<td>Alive</td>
<td>4,051 (77.7)</td>
</tr>
<tr>
<td>Dead</td>
<td>1,160 (22.3)</td>
</tr>
</tbody>
</table>

Values are presented as number (%) unless otherwise indicated. SD, standard deviation.

The median age at diagnosis was 45 years old. Most of the patients, 36.3% of the whole cohort, were diagnosed at 40–59 years old. Most of the patients were white (n = 4,360, 83.7%). Currently, surgery is the first-line treatment in clinical use, and 89.3% of the patients chose surgery (n = 4,653). The median survival time of the patients was 79 months. The distribution of pathological types, classified according to the SEER AYA site record 2020 revision scheme, in the cohort is shown in Table 2. We found that ependymoma, astrocytoma (including oligodendrogliomas and glioblastoma), lymphoma and hemangioblastoma were the 4 most common pathology types.

The age-adjusted incidence rates of malignant (including borderline malignant) primary spinal cord and ependymoma

Table 2. Pathology type of the tumors classified according to the SEER AYA site recode 2020 revision

<table>
<thead>
<tr>
<th>Pathology type</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Leukemias and related disorders</td>
<td>7 (0.13)</td>
</tr>
<tr>
<td>2. Lymphomas</td>
<td>388 (7.45)</td>
</tr>
<tr>
<td>3. CNS and other intracranial and intraspinal neoplasms</td>
<td>4,235 (81.27)</td>
</tr>
<tr>
<td>3.1 Astroglial and related neoplasms</td>
<td></td>
</tr>
<tr>
<td>3.1.1 Oligodendrogliomas</td>
<td>36 (0.69)</td>
</tr>
<tr>
<td>3.1.2 Glioblastoma</td>
<td>118 (2.26)</td>
</tr>
<tr>
<td>3.1.3 Ependymoma</td>
<td>3,012 (57.8)</td>
</tr>
<tr>
<td>3.1.4 Other astrocytoma/astroglial neoplasms</td>
<td>678 (13.01)</td>
</tr>
<tr>
<td>3.1.4.1 Pilocytic astrocytoma</td>
<td>265 (5.09)</td>
</tr>
<tr>
<td>Others</td>
<td>413 (7.93)</td>
</tr>
<tr>
<td>3.2 Medulloblastoma and other invasive embryonal CNS tumors</td>
<td>91 (1.75)</td>
</tr>
<tr>
<td>3.3 Neuroblastomas/ganglioneuromas</td>
<td>51 (0.98)</td>
</tr>
<tr>
<td>3.4 Neuronal and mixed neuronal-glia neoplasms</td>
<td>185 (3.55)</td>
</tr>
<tr>
<td>3.5 Meningiomas</td>
<td>4 (0.08)</td>
</tr>
<tr>
<td>3.6 Choroid plexus neoplasms</td>
<td>1 (0.02)</td>
</tr>
<tr>
<td>3.10 Other and unspecified CNS neoplasms</td>
<td>59 (1.13)</td>
</tr>
<tr>
<td>4. Sarcomas</td>
<td>95 (1.82)</td>
</tr>
<tr>
<td>4.13 Chordoma</td>
<td>69 (1.32)</td>
</tr>
<tr>
<td>Others</td>
<td>26 (0.5)</td>
</tr>
<tr>
<td>5. Blood and lymphatic vessel tumors</td>
<td>328 (6.29)</td>
</tr>
<tr>
<td>5.1.1 Hemangioblastoma and tufted hemangioma</td>
<td>276 (5.3)</td>
</tr>
<tr>
<td>Others</td>
<td>52 (1)</td>
</tr>
<tr>
<td>6. Nerve sheath tumors</td>
<td>130 (2.49)</td>
</tr>
<tr>
<td>7. Gonadal and related tumors</td>
<td>19 (0.36)</td>
</tr>
<tr>
<td>8. Melanoma-malignant</td>
<td>9 (0.17)</td>
</tr>
</tbody>
</table>

SEER, Surveillance, Epidemiology, and End Results; AYA, adolescents and young adults; CNS, central nervous system.

https://doi.org/10.14245/ns.2347300.650
Table 3. The incidence rate of malignant primary spinal cord tumors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Count</th>
<th>Rate</th>
<th>Rate ratio (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>5,211</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2,839</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2,372</td>
<td>0.28</td>
<td>0.799 (0.756–0.844)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>791</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–39</td>
<td>1,381</td>
<td>0.30</td>
<td>1.722 (1.576–1.881)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>40–59</td>
<td>1,889</td>
<td>0.43</td>
<td>2.427 (2.232–2.641)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>≥ 60</td>
<td>1,150</td>
<td>0.40</td>
<td>2.241 (2.079–2.420)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>4,360</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>399</td>
<td>0.21</td>
<td>0.595 (0.535–0.660)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>41</td>
<td>0.16</td>
<td>0.464 (0.330–0.638)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>333</td>
<td>0.19</td>
<td>0.557 (0.496–0.623)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Pathology type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ependymoma</td>
<td>3,012</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Astrocytoma</td>
<td>832</td>
<td>0.05</td>
<td>0.231 (0.212–0.251)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>388</td>
<td>0.02</td>
<td>0.128 (0.114–0.142)</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Hemangioblastoma</td>
<td>276</td>
<td>0.02</td>
<td>0.091 (0.080–0.103)</td>
<td>&lt; 0.001*</td>
</tr>
</tbody>
</table>

CI, confidence interval.
*The rate ratio indicates that the rate is significantly different than the rate for reference (p < 0.05).

Table 4. The incidence rates of primary spinal cord ependymoma, astrocytoma, lymphoma and hemangioblastoma in different populations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ependymoma</th>
<th>Astrocytoma</th>
<th>Lymphoma</th>
<th>Hemangioblastoma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rate (95% CI)</td>
<td>Rate (95% CI)</td>
<td>Rate (95% CI)</td>
<td>Rate (95% CI)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.20</td>
<td>Reference</td>
<td>0.06</td>
<td>Reference</td>
</tr>
<tr>
<td>Female</td>
<td>0.17*</td>
<td>0.861 (0.801–0.926)</td>
<td>0.04*</td>
<td>0.749 (0.651–0.862)</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>0.05</td>
<td>Reference</td>
<td>0.07</td>
<td>Reference</td>
</tr>
<tr>
<td>20–39</td>
<td>0.21*</td>
<td>4.007 (3.467–4.647)</td>
<td>0.04*</td>
<td>0.609 (0.505–0.732)</td>
</tr>
<tr>
<td>40–59</td>
<td>0.29*</td>
<td>5.561 (4.830–16.430)</td>
<td>0.05*</td>
<td>0.671 (0.558–0.805)</td>
</tr>
<tr>
<td>≥ 60</td>
<td>0.19*</td>
<td>3.777 (3.236–4.420)</td>
<td>0.05*</td>
<td>0.690 (0.558–0.850)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.21</td>
<td>Reference</td>
<td>0.05</td>
<td>Reference</td>
</tr>
<tr>
<td>Black</td>
<td>0.09*</td>
<td>0.458 (0.391–0.534)</td>
<td>0.05</td>
<td>0.876 (0.696–1.093)</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>0.09*</td>
<td>0.424 (0.259–0.660)</td>
<td>0.02*</td>
<td>0.390 (0.140–0.923)</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>0.09*</td>
<td>0.459 (0.390–0.538)</td>
<td>0.04</td>
<td>0.824 (0.635–1.053)</td>
</tr>
</tbody>
</table>

CI, confidence interval.
*The rate ratio indicates that the rate is significantly different than the rate for reference (p < 0.05).
tumors were 0.32 and 0.18 per 100,000 population, respectively. The incidence rate for females was significantly lower than that for males. The incidence rate was highest among Caucasian, and lowest among American Indian/Alaska Native people. Regarding age, the incidence rate was highest among people aged 40–59 years. The incidence rate of ependymoma was significantly higher than that of other pathology types. The incidence rate results are shown in Table 3. The incidence rates of primary spinal cord ependymoma, astrocytoma, lymphoma and hemangioblastoma in different populations are shown in Table 4. We found that females had significantly lower incidences of ependymoma, lymphoma and astrocytoma than males. Children between the ages of 0 and 19 years old had a significantly lower incidence of ependymoma but a significantly higher incidence of astrocytoma than the adult population. The highest incidence rate of ependymoma was observed in people aged 40–59 years old, while the highest incidence rate of lymphoma was observed in the ≥ 60-year-old population. We also found the incidence of ependymoma was significantly higher in Caucasian than in people of other races.

The 5-year observed survival and relative survival for the whole cohort were 82.80% and 86.00%, respectively. Female patients, white patients, patients diagnosed between 40–59 years old and patients diagnosed with ependymoma had better five-year relative survival than their counterparts (Table 5). The results of the Cox proportional hazards models are shown in Tables 6 and 7. These results reaffirmed that patients diagnosed with ependymoma had significantly better survival than their counterparts. We also found that for the patients with ependymoma and astrocytoma surgery of the primary tumor can significantly improve the survival, while chemotherapy can significantly improve the survival for the patients with lymphoma.

### DISCUSSION

Spine tumors were classified into 3 main groups: extradural, intradural extramedullary and intramedullary.\textsuperscript{15} Primary spinal tumors are primarily systemic cancer metastases.\textsuperscript{16} Primary spinal cord ependymomas have been found to be associated with a better prognosis than other primary spinal cord tumors,\textsuperscript{17} and a recent review of the literature indicated that these tumors have a better prognosis than other spinal cord tumors.\textsuperscript{18}

### Table 5. Five-year observed survival and relative survival for different populations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observed survival</th>
<th>Expected survival</th>
<th>Relative survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>82.80%</td>
<td>96.30%</td>
<td>86.00%</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>81.00%</td>
<td>95.90%</td>
<td>84.50%</td>
</tr>
<tr>
<td>Female</td>
<td>84.90%</td>
<td>96.90%</td>
<td>87.70%</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>81.70%</td>
<td>99.80%</td>
<td>81.80%</td>
</tr>
<tr>
<td>20–39</td>
<td>88.50%</td>
<td>99.30%</td>
<td>89.10%</td>
</tr>
<tr>
<td>40–59</td>
<td>88.10%</td>
<td>97.50%</td>
<td>90.30%</td>
</tr>
<tr>
<td>≥ 60</td>
<td>68.20%</td>
<td>87.50%</td>
<td>77.90%</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>83.50%</td>
<td>96.10%</td>
<td>86.80%</td>
</tr>
<tr>
<td>Black</td>
<td>78.30%</td>
<td>96.80%</td>
<td>80.80%</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>79.10%</td>
<td>95.70%</td>
<td>82.40%</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>78.20%</td>
<td>98.50%</td>
<td>79.40%</td>
</tr>
<tr>
<td>Pathology type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ependymoma</td>
<td>93.60%</td>
<td>96.70%</td>
<td>96.70%</td>
</tr>
<tr>
<td>Astrocytoma</td>
<td>60.40%</td>
<td>97.70%</td>
<td>61.80%</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>61.90%</td>
<td>91.10%</td>
<td>67.70%</td>
</tr>
<tr>
<td>Hemangioblastoma</td>
<td>91.10%</td>
<td>95.80%</td>
<td>94.40%</td>
</tr>
</tbody>
</table>

### Table 6. The hazard ratios (HR), 95% confidence intervals (CI), and p-values were calculated using multivariable Cox regression

<table>
<thead>
<tr>
<th>Variable</th>
<th>HR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.831 (0.739–0.935)</td>
<td>0.002</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>0.727 (0.596–0.888)</td>
<td>0.002</td>
</tr>
<tr>
<td>20–39</td>
<td>0.886 (0.741–1.059)</td>
<td>0.182</td>
</tr>
<tr>
<td>40–59</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>≥ 60</td>
<td>2.647 (2.292–3.056)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>1.243 (1.064–1.450)</td>
<td>0.006</td>
</tr>
<tr>
<td>Pathology type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ependymoma</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>0.369 (0.319–0.426)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.854 (0.734–0.994)</td>
<td>0.042</td>
</tr>
<tr>
<td>None/unknown</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Radiation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2.000 (1.762–2.270)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>None/unknown</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.521 (1.313–1.762)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>None/unknown</td>
<td>Reference</td>
<td></td>
</tr>
</tbody>
</table>
Table 7. The results of hazard ratio (HR), 95% confidence interval (CI), and p-value calculated through multivariable Cox regression for patients with different tumors

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ependymoma</th>
<th>Astrocytoma</th>
<th>Lymphoma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>p-value</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Reference</td>
<td>0.007</td>
<td>Reference</td>
</tr>
<tr>
<td>Female</td>
<td>0.745 (0.601–0.923)</td>
<td>0.007</td>
<td>0.928 (0.753–1.145)</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–19</td>
<td>0.684 (0.397–1.177)</td>
<td>0.170</td>
<td>0.562 (0.414–0.763)</td>
</tr>
<tr>
<td>20–39</td>
<td>0.812 (0.591–1.115)</td>
<td>0.198</td>
<td>0.843 (0.633–1.124)</td>
</tr>
<tr>
<td>40–59</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>≥ 60</td>
<td>5.024 (3.919–6.441)</td>
<td>&lt; 0.001</td>
<td>2.266 (1.706–3.009)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Reference</td>
<td>0.130</td>
<td>Reference</td>
</tr>
<tr>
<td>Others</td>
<td>1.272 (0.931–1.739)</td>
<td>0.130</td>
<td>0.950 (0.740–1.219)</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.468 (0.318–0.690)</td>
<td>&lt; 0.001</td>
<td>0.748 (0.583–0.959)</td>
</tr>
<tr>
<td>None/unknown</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Radiation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1.891 (1.473–2.428)</td>
<td>&lt; 0.001</td>
<td>2.013 (1.557–2.602)</td>
</tr>
<tr>
<td>None/unknown</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5.366 (2.686–10.719)</td>
<td>&lt; 0.001</td>
<td>2.425 (1.924–3.057)</td>
</tr>
<tr>
<td>None/unknown</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
</tbody>
</table>

Cord tumors are anatomically separable into 2 broad categories: intradural intramedullary and intradural extramedullary. At first, we aimed to choose intramedullary spinal tumors as the study objective. However, we could not define intramedullary tumors by their primary site according to ICD-O-3. At the same time, we found that the most common primary intradural extramedullary tumors are meningioma, neurofibroma, and schwannoma, and most of them are benign tumors. Gliomas (astrocytoma and ependymoma) account for 80% of all intramedullary tumors, and most gliomas are considered to be malignant or borderline malignant according to ICD-O-3 classification. Subsequently, we chose spinal cord malignant tumors as the topic of research and included borderline malignant tumors to cover almost 99% of spinal cord gliomas in the SEER database. Among the benign tumors in the database, classified according to ICD-O-3 guidelines, only 62 cases were gliomas. We believe this study design can elucidate the epidemiology and survival of intramedullary tumors. At the same time, we acknowledge that this cohort included 130 cases of nerve sheath tumors, 4 cases of meningiomas and 1 case of choroid plexus neoplasms etc., accounting for 2.6% of the whole cohort. We had not excluded these cases in this study due to the difficulty of defining the query syntax. This inclusion might lead the reported incidence rate to be slightly higher than the true incidence rate of primary intramedullary tumors.

Park et al. reported that the frequent pathologies of primary intramedullary spinal cord tumors were spinal ependymoma (45.1%), hemangioblastoma (20.0%), and astrocytic tumors (17.4%). Spinal cord hemangioblastomas are reported to account for 2% to 15% of primary intramedullary spinal cord tumors. Most of our results were consistent with the previous report. Because benign tumors were not included in our analysis, the resulting incidence of primary spinal cord hemangioblastoma should be lower than the actual incidence. It was reported that the most common primary intramedullary tumor pathology in pediatric patients is astrocytoma; however, as age increases, the most common pathology becomes ependymoma. In our cohort study, we clearly demonstrate this by comparing the incidence of different tumors in different age groups. Primary intramedullary spinal cord lymphoma has been re-
ported as a rare diagnosis with poorly understood disease progression. However, we found that the incidence of primary spinal cord lymphoma was only exceeded by the incidences of ependymoma and astrocytoma and was significantly higher in populations older than 60 years than in other age groups. Therefore, when treating elderly patients, lymphoma should be considered in addition to gliomas.

Because intramedullary spinal cord tumors comprise diverse tumor types, distinct management strategies are chosen based on histopathology; nonetheless, advances in microsurgical techniques and technological adjuncts have improved the extent of resection and outcomes. Persson et al. performed a retrospective cohort study of 95 patients who underwent surgery for intra- or juxtamedullary tumors and found that long-term progression-free survival could be achieved by gross-total resection without additional adjuvant treatment. Matthew et al. reviewed the treatment strategies for intramedullary spinal cord tumors and concluded that most evidence-based treatments involve resection. Patients who cannot undergo gross-total resection or have subtotal resection only have radiotherapy and chemotherapy as treatment options. However, these treatments are associated with the potential for significant adverse side effects and still leave patients with a poor prognosis. In this study, we also found the impact of surgery and chemotherapy on the prognosis of patients with different tumors varies a lot. We believe that more efforts should be made to design specific treatment plans for individual patients.

One limitation of our analysis was that use of the registry-based approach precluded access to some specific clinical data. For example, we could not confirm the specific location (cervical, thoracic, or lumbar spinal cord) of the tumor. We also could not confirm the pathology according to the 2021 edition of World Health Organization classification, because of lack of genetic and molecular data. There was also a lack information about specific therapeutic methods used as well as imaging manifestations of the tumors. As a registry-based population analysis, we must also acknowledge the possibility that there may be inaccurate data collection. Nevertheless, to the best of our knowledge, this study represents the first attempt to characterize the incidence and survival of patients with primary intramedullary spinal cord tumors based on a nationwide registry. Our results provide information that may be useful for further investigating the epidemiology and exploring the prognosis of patients with primary intramedullary tumors.

CONCLUSION

We conducted a population-based analysis of malignant and borderline malignant primary spinal cord tumors to reveal the epidemiology and survival of patients with primary intramedullary spinal cord tumors. Despite some shortcomings, this study still provides valuable information to help us better understand the epidemiological characteristics of primary intramedullary spinal cord tumors.

NOTES

Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: YW; Data curation: LD, ZL, YL, YW; Formal analysis: LD, ZL, YL, YW; Methodology: YW; Writing – original draft: HL, YW; Writing – review & editing: HL, LD, ZL, YL, YW.

ORCID

Huanbing Liu: 0009-0005-2511-0681
Linnan Duan: 0009-0007-1347-0799
Zhibin Li: 0009-0005-8408-4479
Yuanhao Liu: 0009-0005-4572-527X
Yubo Wang: 0000-0002-3404-0225

REFERENCES

INTRODUCTION

Most countries face serious healthcare issues, such as increased life and health expectancy and population aging. This demographic shift will continue and by 2050 approximately half of the world’s population will live in aging countries. In particular, Japan has become a super-aged society, with 28.4% of the population aged ≥ 65 years, and the mean life expectancy for Japanese women in 2019 was 87.5 years. Conversely, the Healthy Life Years of Japanese women was 75.4 years in 2019.
Thus, a 12-year time lag exists between Healthy Life Years and the mean life expectancy. To close this time lag and extend Healthy Life Years, it is important to focus on the health of patients aged ≥75 years. Concurrently, in the future, surgical treatments, such as lumbar fusion, will become the treatment of choice for older patients who require treatment for lumbar degenerative diseases to improve their quality of life.3,4

With an increasing focus on patient health and satisfaction, spine surgeons are increasingly expected to reduce postoperative leg numbness.6 Moreover, 3 previous reports have described the predictors of increased postoperative persistent leg numbness following lumbar surgery. These reports have demonstrated that increased postoperative persistent leg numbness following lumbar decompression surgery may be related to preoperative leg numbness intensity, diabetes mellitus, dural injury, preoperative symptom duration, and preoperative severity of spinal canal stenosis.7-9 However, these previous studies examined patients of all ages. To the best of our knowledge, no studies have focused on postoperative persistent leg numbness following lumbar fusion in older adults despite its importance in an aging society. Therefore, in this study, we aimed to evaluate the specific demographic, preoperative, and perioperative predictors associated with greater postoperative persistent leg numbness following lumbar fusion in patients aged ≥75 years.

MATERIALS AND METHODS

1. Patients

The study was approved by the Institutional Review Board of Hakodate Central General Hospital (No. 2023-9). We conducted this study according to the 1964 Helsinki Declaration and its later amendments. Informed consent was obtained from all individual participants included in the study.

We retrospectively reviewed the medical records of patients aged ≥75 years who underwent posterior lumbar interbody fusion (PLIF) or transforaminal lumbar interbody fusion (TLIF) for lumbar degenerative diseases at our hospital between March 2013 and October 2019. The exclusion criteria were as follows: (1) follow-up period < 2 years; (2) additional lumbar fusion for adjacent segment degeneration; (3) surgery for spinal injury, infection, or tumor; and (4) postoperative infection. Neurologists were consulted in advance if a neurologic disease such as polyneuropathy was suspected. In our hospital, preoperative and postoperative epidural nerve block were not performed. In total, 304 patients met the inclusion criteria. The 2-year follow-up rate was 80.9% (304 of 376 patients). Of the 72 dropouts, 47 did not return to the doctor, 14 required revision surgery, 7 had postoperative infections, and 4 died of systemic diseases.

2. Operative Procedures

The operative procedures were performed in an open setting. We performed PLIF/TLIF in patients with spondylolisthesis or neuroforaminal stenosis, including far-lateral disc herniation. Patients with spondylolisthesis requiring bilateral decompression underwent PLIF; whereas those with neuroforaminal stenosis or spondylolisthesis requiring unilateral decompression underwent TLIF. Pedicle screws were positioned correctly. The upper articular processes, lower articular processes, and lamina were removed to expand the central spinal canal. After removing the nucleus pulposus, the endplate cartilage was scraped using a curette and a suitable polyetheretherketone cage was inserted into the intervertebral space with an autogenous bone graft. Two box-type cages were used for PLIF and a banana-shaped cage was used for TLIF.

3. Data Collection

Patient demographics and operative factors were reviewed. These variables included age, sex, body mass index (BMI), type of surgery, pathogenesis, alcohol, smoking, presence of preexisting vertebral fractures, central spinal canal stenosis, history of lumbar spine surgery, diabetes mellitus, depression, dural injury, symptom duration, serum calcium level, opioids use, anticonvulsants (pregabalin or mirogabalin) use, number of fusion segments, operation time, and estimated blood loss (EBL). To determine the appropriate cutoff value, we have performed receiver operating characteristic curve analysis in the univariate analysis (Supplementary Fig. 1). Slimness and obesity were defined as BMI < 20 and ≥ 30 kg/m², respectively. Central spinal canal stenosis was defined as Schizae grade C or D.10 History of previous lumbar spine surgery was defined as previous decompression, including discectomy, in the fusion range. Depression was confirmed from the medical records. We evaluated the visual analogue scale (VAS) scores for low back pain (LBP), leg pain, and leg numbness preoperatively, and at 1 and 2 years postoperatively. Improvement was evaluated using minimum clinically important differences (MCIDs). We set MCIDs as a 2-point improvement of the VASs for LBP, leg pain, and numbness based on a previous study.11 Since the MCID score is 2, patients with preoperative VAS less than 2 were excluded from analysis. Spinopelvic sagittal parameters including sagittal vertical axis (SVA), lumbar lordosis (LL), pelvic tilt (PT), and pelvic incidence (PI) were measured using standing radiographs of the
Predictors of Postoperative Persistent Numbness in the Elderly

Tsujimoto T, et al.

4. Statistical Analysis

As a scale for leg numbness intensity evaluation, the presence of postoperative persistent leg numbness was defined as a VAS score for leg numbness ≥ 5 points at the 2-year postoperative follow-up evaluation. Patients with a VAS score for leg numbness ≥ 5 and < 5 points were categorized into the persistent numbness (PN) and non-PN (N-PN) groups, respectively. All statistical analyses were performed using JMP Pro version 16.0 statistical software (SAS Institute, Cary, NC, USA). Fisher exact probability test was performed for the univariate analysis to compare categorical variables between the PN and N-PN groups. Multivariate stepwise logistic regression analysis was performed to investigate the predictors of postoperative persistent leg numbness after lumbar fusion using variables with p < 0.2 in the univariate analysis. Statistical significance was set at p < 0.05.

RESULTS

In total, 304 patients who underwent primary PLIF/TLIF were included in the analysis. Table 1 shows the patients’ demographic characteristics and clinical outcomes. This study included 102 male and 202 female patients with a mean age of 79.2 ± 3.5 (75–90) years at the time of surgery. The mean BMI was 24.7 ± 11.8 kg/m², and the mean follow-up period was 46.0 ± 18.9 months. Pathologies included lumbar spondylolisthesis in 176 patients (57.9%), lumbar spinal stenosis (LSS) with foraminal stenosis in 122 patients (40.1%), and lumbar disc herniation in 6 patients (2.0%). Surgery included PLIF in 149 patients (49.0%) and TLIF in 155 patients (51.0%). The number of fusion segments was one in 194 cases (63.8%), 2 in 89 cases (29.3%), 3 in 18 cases (5.9%), and 4 in three cases (1.0%). Additionally, we have investigated a subgroup analysis based on the number of fusion segments and the result was shown in Supplementary Table 1. The mean symptom duration was 24.6 ± 44.6 months. The mean operation time and EBL were 202.3 ± 64.2 minutes and 246.0 ± 213.9 mL, respectively. The mean preoperative VAS score for LBP, leg pain, and leg numbness were 5.1 ± 3.2, 6.5 ± 3.3, and 5.4 ± 3.5 points, respectively. The

Table 1. Characteristics of patient demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>79.2 ± 3.5</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>24.7 ± 11.8</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>102 (33.6)</td>
</tr>
<tr>
<td>Female</td>
<td>202 (66.4)</td>
</tr>
<tr>
<td>Type of surgery</td>
<td></td>
</tr>
<tr>
<td>Posterior lumbar interbody fusion</td>
<td>149 (49.0)</td>
</tr>
<tr>
<td>Transforaminal lumbar interbody fusion</td>
<td>155 (51.0)</td>
</tr>
<tr>
<td>Pathogenesis</td>
<td></td>
</tr>
<tr>
<td>Spondylolisthesis</td>
<td>176 (57.9)</td>
</tr>
<tr>
<td>Lumbar spinal stenosis with foraminal stenosis</td>
<td>122 (40.1)</td>
</tr>
<tr>
<td>Lumbar disc herniation</td>
<td>6 (2.0)</td>
</tr>
<tr>
<td>No. of fusion segments</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>194 (63.8)</td>
</tr>
<tr>
<td>2</td>
<td>89 (29.3)</td>
</tr>
<tr>
<td>3</td>
<td>18 (5.9)</td>
</tr>
<tr>
<td>4</td>
<td>3 (1.0)</td>
</tr>
<tr>
<td>Duration of symptoms (mo)</td>
<td>24.6 ± 44.6</td>
</tr>
<tr>
<td>Operation time (min)</td>
<td>202.3 ± 64.2</td>
</tr>
<tr>
<td>Estimated blood loss (mL)</td>
<td>246.0 ± 213.9</td>
</tr>
<tr>
<td>Preoperative VAS for LBP</td>
<td>5.1 ± 3.2</td>
</tr>
<tr>
<td>Preoperative VAS for leg pain</td>
<td>6.5 ± 3.3</td>
</tr>
<tr>
<td>Preoperative VAS for leg numbness</td>
<td>5.4 ± 3.5</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%). VAS, visual analogue scale; LBP, low back pain.

Fig. 1. Comparison of the change of visual analogue scales (VAS) between the persistent numbness (PN) and non-PN (N-PN) groups. (A) VAS for low back pain. (B) VAS for leg pain. (C) VAS for leg numbness. Preop, preoperative; PO, postoperative. *p < 0.05, **p < 0.001.
The changes in the VAS scores for LBP, leg pain, and leg numbness are presented in Fig. 1. The preoperative VAS scores for LBP and leg numbness were significantly higher in the PN than in the N-PN group, whereas there was no significant difference in the VAS scores for leg pain between the 2 groups. The postoperative VAS scores for LBP, leg pain, and leg numbness mean operation times of PLIF and TLIF were 205.0 ± 67.1 minutes and 199.7 ± 61.3 minutes, respectively. There was no significant difference between the operation time of 2 techniques (p = 0.473). A total of 71 patients (23.4%) were categorized into the PN group. In the N-PN group, 3 patients (1.3%) had worsening numbness from the preoperative period; in the PN group, 24 patients (33.8%) had worsening numbness (p < 0.001). There was no significant difference in pathogenesis between PN and N-PN groups (Supplementary Table 2) (p = 0.244).

Table 2. Patient population (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. (%)</th>
<th>Variable</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dural injury</td>
<td></td>
<td>BMI, body mass index; PLIF, posterior lumbar interbody fusion; TLIF, transforaminal lumbar interbody fusion; VAS, visual analogue scale; LBP, low back pain.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>267 (87.8)</td>
<td>Symptom duration (mo)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>37 (12.2)</td>
<td>&lt; 16</td>
<td>196 (64.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥ 16</td>
<td>108 (35.5)</td>
</tr>
<tr>
<td>No. of fusion segments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 3</td>
<td>282 (92.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 3</td>
<td>22 (7.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operation time (min)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 254</td>
<td>240 (78.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 254</td>
<td>64 (21.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimated blood loss (mL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 220</td>
<td>177 (58.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 220</td>
<td>127 (41.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative VAS for LBP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5</td>
<td>127 (41.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 5</td>
<td>177 (58.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative VAS for leg pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5</td>
<td>79 (26)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 5</td>
<td>225 (74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative VAS for leg numbness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5</td>
<td>115 (37.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 5</td>
<td>189 (62.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Continued)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3. Univariate analysis between PN and N-PN groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PN group (%)</th>
<th>N-PN group (%)</th>
<th>Odds ratio</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (yr)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 78</td>
<td>25.15</td>
<td>74.85</td>
<td>1.24</td>
<td>0.72–2.14</td>
<td>0.423</td>
</tr>
<tr>
<td>&lt; 78</td>
<td>21.21</td>
<td>78.79</td>
<td>1.24</td>
<td>0.72–2.14</td>
<td>0.423</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>22.77</td>
<td>77.23</td>
<td>0.95</td>
<td>0.54–1.67</td>
<td>0.848</td>
</tr>
<tr>
<td>Female</td>
<td>23.76</td>
<td>76.24</td>
<td>0.95</td>
<td>0.54–1.67</td>
<td>0.848</td>
</tr>
<tr>
<td><strong>Slimness (BMI &lt; 20 kg/m²)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24.44</td>
<td>75.56</td>
<td>0.60</td>
<td>0.24–1.50</td>
<td>0.269</td>
</tr>
<tr>
<td>Yes</td>
<td>16.22</td>
<td>83.78</td>
<td>0.60</td>
<td>0.24–1.50</td>
<td>0.269</td>
</tr>
<tr>
<td><strong>Obesity (BMI ≥ 30 kg/m²)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23.34</td>
<td>76.66</td>
<td>1.09</td>
<td>0.34–3.51</td>
<td>0.879</td>
</tr>
<tr>
<td>Yes</td>
<td>25.00</td>
<td>75.00</td>
<td>1.09</td>
<td>0.34–3.51</td>
<td>0.879</td>
</tr>
<tr>
<td><strong>Type of surgery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLIF</td>
<td>24.16</td>
<td>75.84</td>
<td>0.91</td>
<td>0.54–1.56</td>
<td>0.745</td>
</tr>
<tr>
<td>TLIF</td>
<td>22.58</td>
<td>77.42</td>
<td>0.91</td>
<td>0.54–1.56</td>
<td>0.745</td>
</tr>
<tr>
<td><strong>Alcohol</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24.26</td>
<td>75.74</td>
<td>1.73</td>
<td>0.64–4.67</td>
<td>0.275</td>
</tr>
<tr>
<td>Yes</td>
<td>15.63</td>
<td>84.38</td>
<td>1.73</td>
<td>0.64–4.67</td>
<td>0.275</td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24.11</td>
<td>75.89</td>
<td>1.30</td>
<td>0.62–2.75</td>
<td>0.488</td>
</tr>
<tr>
<td>Yes</td>
<td>19.61</td>
<td>80.39</td>
<td>1.30</td>
<td>0.62–2.75</td>
<td>0.488</td>
</tr>
<tr>
<td><strong>Central spinal canal stenosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>20.00</td>
<td>80.00</td>
<td>0.77</td>
<td>0.40–1.46</td>
<td>0.419</td>
</tr>
<tr>
<td>Yes</td>
<td>24.56</td>
<td>75.44</td>
<td>0.77</td>
<td>0.40–1.46</td>
<td>0.419</td>
</tr>
<tr>
<td><strong>Foraminal stenosis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>20.51</td>
<td>79.49</td>
<td>1.40</td>
<td>0.82–2.38</td>
<td>0.217</td>
</tr>
<tr>
<td>Yes</td>
<td>26.53</td>
<td>73.47</td>
<td>1.40</td>
<td>0.82–2.38</td>
<td>0.217</td>
</tr>
<tr>
<td><strong>Vertebral fractures</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>22.27</td>
<td>77.73</td>
<td>1.23</td>
<td>0.70–2.17</td>
<td>0.471</td>
</tr>
<tr>
<td>Yes</td>
<td>26.09</td>
<td>73.91</td>
<td>1.23</td>
<td>0.70–2.17</td>
<td>0.471</td>
</tr>
<tr>
<td><strong>Previous lumbar decompression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>19.60</td>
<td>80.40</td>
<td>3.19</td>
<td>1.67–6.10</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Yes</td>
<td>43.80</td>
<td>56.20</td>
<td>3.19</td>
<td>1.67–6.10</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Diabetes mellitus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23.24</td>
<td>76.76</td>
<td>1.05</td>
<td>0.55–2.03</td>
<td>0.874</td>
</tr>
<tr>
<td>Yes</td>
<td>24.19</td>
<td>75.81</td>
<td>1.05</td>
<td>0.55–2.03</td>
<td>0.874</td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23.16</td>
<td>76.84</td>
<td>1.28</td>
<td>0.44–3.71</td>
<td>0.654</td>
</tr>
<tr>
<td>Yes</td>
<td>27.78</td>
<td>72.22</td>
<td>1.28</td>
<td>0.44–3.71</td>
<td>0.654</td>
</tr>
<tr>
<td><strong>Serum calcium level (mg/dL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 9.0</td>
<td>30.00</td>
<td>70.00</td>
<td>0.65</td>
<td>0.34–1.22</td>
<td>0.175</td>
</tr>
<tr>
<td>≥ 9.0</td>
<td>21.72</td>
<td>78.28</td>
<td>0.65</td>
<td>0.34–1.22</td>
<td>0.175</td>
</tr>
</tbody>
</table>

(Continued)
were significantly higher in the PN than in the N-PN group at 1 and 2 years postoperatively. Reaching MCIDs between PN and N-PN groups were shown in Supplementary Table 3. Rate of reaching MCIDs of LBP, leg pain, and leg numbness in PN group were 74.2%, 70.5%, and 36.5%, respectively, and those in N-PN group were 80.6%, 89.3%, and 93.5%, respectively. Although there was no significant difference in rate of reaching MCID of LBP (p = 0.289), those of leg pain and leg numbness were significantly lower in PN group than N-PN group (p < 0.001 each).

Preoperative opioid use was 19 patients in PN group (24.4%) and 52 patients in N-PN group (23.0%), and preoperative anticonvulsants use was 37 patients in PN group (52.1%) and 118 patients in N-PN group (50.6%). There were no significant differences in preoperative use of opioid and anticonvulsants between the 2 groups (opioid, p = 0.808; anticonvulsants, p = 0.828). Postoperative opioid use was 9 patients in PN group (12.7%) and 39 patients in N-PN group (16.7%), and postoperative anticonvulsants use was 23 patients in PN group (32.4%) and 18 patients in N-PN group (7.7%). Although there was no significant difference in postoperative opioid use between the 2 groups (p = 0.461), the rate of postoperative anticonvulsants use was significantly higher in PN group than N-PN group (p < 0.001).

The majority of patients in our study sample were aged ≥ 78 years (56.6%), women (66.4%), had no slimness (87.8%), not obese (94.7%), no alcohol (89.5%), no smoking (83.2%), and had a symptom duration of < 16 months (64.5%). The majority of patients underwent TLIF (51.0%). Regarding baseline comorbidities, most patients had central spinal canal stenosis (75.0%), no foraminal stenosis (51.3%), no existing vertebral fractures (69.7%), no history of lumbar decompression (84.2%), no diabetes mellitus (79.6%), no depression (94.1%), serum calcium level ≥ 9.0 mg/dL (80.3%), and no dural injury (94.4%).

<table>
<thead>
<tr>
<th>Table 3. Univariate analysis between PN and N-PN groups (Continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Dural injury</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Symptom duration (mo)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>No. of fusion segments</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Operation time (min)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Estimated blood loss (mL)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Preoperative VAS for LBP</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Preoperative VAS for leg pain</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Preoperative VAS for leg numbness</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

PN, persistent numbness; N-PN, non-PN; CI, confidence interval; BMI, body mass index; PLIF, posterior lumbar interbody fusion; TLIF, transforaminal lumbar interbody fusion; VAS, visual analogue scale; LBP, low back pain.
ing intraoperative factors, most patients had < 3 fusion segments (92.8%), operative time < 254 minutes (78.9%), and EBL < 220 mL (58.2%). Most patients had a preoperative VAS score for LBP ≥ 5 points (58.2%), leg pain (74.0%), and leg numbness ≥ 5 (62.2%). The patient demographics are presented in Table 2.

Univariate analyses were performed on the PN and N-PN groups to verify the effects of each preoperative and intraoperative factor on persistent leg numbness following lumbar fusion in older patients. Table 3 presents the results of the analysis. A history of previous lumbar decompression (p < 0.001), serum calcium level < 9.0 mg/dL (p = 0.175), symptom duration ≥ 16 months (p = 0.014), EBL ≥ 220 mL (p = 0.040), preoperative VAS score for LBP ≥ 5 points (p = 0.073), and preoperative VAS score for leg numbness ≥ 5 points (p < 0.001) were detected as variables with p-values < 0.2. However, other factors, such as age, sex, slimness, obesity, type of surgery, alcohol, smoking, central spinal canal stenosis, vertebral fracture, diabetes mellitus, depression, dural injury, symptom duration, number of fusion segments, operation time, and preoperative VAS score for leg pain ≥ 5 points, were not associated with postoperative persistent leg numbness (p ≥ 0.2 each).

The results of comparison of pre- and postoperative spinopelvic sagittal parameters between PN and N-PN groups were shown in Supplementary Table 4. There were no significant differences in spinopelvic sagittal parameters including pre- and postoperative SVA, LL, PI, PT, and PI-LL between the PN and N-PN groups.

Predictors of an increased risk of postoperative persistent leg numbness in older patients were determined using multivariate stepwise logistic regression analysis. Table 4 presents the results of this analysis. Two of the 4 factors detected as variables (p < 0.2) in the univariate analysis were significant in the multivariate analysis. A history of previous lumbar decompression (odds ratio [OR], 3.89; p < 0.001), symptom duration ≥ 16 months (OR, 2.11; p = 0.013), and a preoperative VAS score for leg numbness ≥ 5 points (OR, 3.63; p < 0.001) were associated with greater postoperative persistent leg numbness in older patients.

**DISCUSSION**

Managing leg numbness after lumbar fusion is important for postoperative patient care. A previous study showed that leg numbness had a greater impact than that of leg/back pain on patient satisfaction in patients who underwent lumbar decompression.6,12 Indeed, in clinical settings, we frequently encounter patients who complain that leg pain has disappeared, but numbness is still persistent. Given that lower extremity and trunk muscle strengths in older adults decrease with age,13-16 management of postoperative leg symptoms becomes more important for maintaining activities of daily living. Therefore, the predictors of persistent leg numbness following lumbar surgery in older patients are of great interest. Several studies have investigated the predictors of postoperative persistent leg numbness.7,9 However, previous studies have investigated the predictors of postoperative persistent leg numbness in all generations. Furthermore, all previous studies have investigated decompression surgery, and no study has focused on persistent leg numbness after lumbar fusion. Thus, the predictors of postoperative persistent leg numbness after lumbar fusion in older patients remain unclear. Here, we first investigated the predictors of postoperative persistent leg numbness following lumbar fusion in patients aged ≥ 75 years and found that a history of previous lumbar decompression, symptom duration ≥ 16 months, and preoperative VAS score for leg numbness ≥ 5 points were predictors of persistent leg numbness.

Prolonged recovery of leg numbness following lumbar spine surgery has been known,17 and alleviation of leg numbness by surgery for LSS has been reported to be more difficult than other neurological symptoms, such as muscle weakness or pain.18,19
Huang and Sengupta\textsuperscript{17} showed that although leg pain dramatically improved within 6 weeks following lumbar decompression, leg numbness improved at a much slower speed than pain and gradually improved from 3 months to 1 year of follow-up. Furthermore, Oba et al.\textsuperscript{8} reported that postoperative leg numbness following lumbar decompression improved until 2 weeks postoperatively, after which it reached a plateau and showed no improvement. Additionally, at 1 year postoperatively, leg numbness remained significantly higher than leg pain. These previous studies have shown that leg numbness following lumbar spine surgery is associated with poorer recovery than leg pain.

We defined PN in cases of VAS score $\geq$ 5 points, which is obviously of significant intensity, according to a previous study, because mild residual numbness may not significantly impact postoperative patient satisfaction.\textsuperscript{9} As a result, in the current study, the incidence of postoperative persistent leg numbness was 23.4%. Previous studies have reported that the incidence of postoperative PN following lumbar surgery was 15.5\%–74.1\%.\textsuperscript{7,8} This result suggested that the incidence of postoperative leg numbness following lumbar fusion in older adults is equal to that following lumbar decompression. The results of this study are consistent with those of previous studies. This study suggested that a history of lumbar decompression is a predictor of persistent postoperative leg numbness after lumbar fusion in older patients. The impact of previous lumbar decompression on a range of persistent postoperative symptoms, including back pain, leg pain, and leg numbness, is known.\textsuperscript{20-22} Postoperative paravertebral tissue fibrosis and disc degeneration, including recurrent disc herniation, retained disc fragment, and internal disc disruption, are caused by lumbar decompression.\textsuperscript{23-25} Therefore, persistent postoperative symptoms, including leg numbness, may be caused by multiple lumbar surgeries that affect the progression of lumbar degeneration. Older patients often have a history of surgery; thus, surgeons planning lumbar fusion should consider the impact of previous lumbar decompression on persistent postoperative leg numbness.

The use of the VAS or Numerical Rating Scale (NRS) score for leg numbness is standard in previous studies analyzing leg numbness after lumbar spine surgery.\textsuperscript{5,6,9} Oba et al. evaluated recovery from leg numbness following decompression surgery for LSS with VAS for leg numbness.\textsuperscript{8} Furthermore, Ogura et al.\textsuperscript{5,6,9} used NRS for leg numbness to assess patients’ satisfaction and risk factors following decompression surgery for LSS. Therefore, we believe that the research method of the current study, which is using VAS for leg numbness, is appropriate.

In the present study, the patients with persistent postoperative numbness had higher rates of postoperative LBP and leg pain. Furthermore, reaching MCIDs of leg pain, and leg numbness were associated with postoperative PN. Persistent back pain, leg pain, and numbness after lumbar spine surgery are known as failed back surgery syndrome.\textsuperscript{20-22,26,27} Moreover, leg pain alone can be related to the sacroiliac joint, which is not an uncommon presentation after spinal fusion.\textsuperscript{28} Thus, in this study, patients in the PN group may have failed back surgery syndrome or sacroiliitis after spinal surgery.

Several previous studies have also explored whether the symptom duration of preoperative leg numbness is related to postoperative PN.\textsuperscript{5,29,30} Their study investigated the relationship between postoperative persistent leg numbness following lumbar decompression, spinal endoscopic surgery, or postoperative lumbar epidural hematoma, resulting in showing that postoperative PN was associated with longer symptom duration. Here, the current study has demonstrated a longer preoperative symptom duration is a risk factor of persistent leg numbness following lumbar fusion in the elderly patients. A longer duration of preoperative leg numbness is associated with a longer compression of the nerve root, which may lead to irreversible damage to the nerve root and the accompanying PN.\textsuperscript{31}

The current study suggested that a preoperative VAS score for leg numbness $\geq$ 5 points can predict persistent postoperative leg numbness. Similarly, previous studies reported that greater preoperative leg numbness was a risk factor for postoperative persistent leg numbness following lumbar surgery.\textsuperscript{23} Hara et al.\textsuperscript{7} reported that 65\% of patients with severe preoperative numbness still showed residual leg numbness at 2 years postoperatively. Furthermore, Ogura et al.\textsuperscript{5} reported that patients with postoperative persistent leg numbness following lumbar decompression had greater preoperative leg numbness than patients with no PN. Altogether, the findings of previous studies were consistent with those of our work in that greater preoperative leg numbness was a risk factor for postoperative PN following lumbar surgery.

The current study had 2 strengths. First, we investigated postoperative persistent leg numbness in patients who underwent lumbar fusion 2 years postoperatively with a sufficient sample size, even in this older adult cohort. Second, to our knowledge, this is the first study to focus on predictors of PN following lumbar fusion in older patients. When a surgeon plans to perform lumbar fusion in an older patient, the results of the current study may help explain the risk of persistent postoperative leg numbness and contribute to the choice of treatment.

However, the study had several limitations. First, it had a retrospective design. Therefore, if some older patients who were pathologically eligible for lumbar fusion could not undergo sur-
gery because of systemic problems, a selection bias could exist. Moreover, because the symptoms could not be determined from medical records alone, which often lacked detailed patient complaints, we could not classify the patient’s leg numbness as cauda equina syndrome or radiculopathy. Second, we used the VAS score, a patient-reported outcome measure, to evaluate persistent leg numbness after lumbar fusion. Generally, as patients may use “numbness” to express various symptoms, such as paresthesia, dysesthesia, or loss of sensation, patient-reported leg numbness might include a variety of sensory disturbances. Furthermore, elderly patients may have bias in description of their symptoms, numbness in particular. Distinguishing the difference between a type of sensory symptoms was difficult in the current study. Third, this was a single-center study. Additionally, in our hospital, we do not routinely perform electromyography, cervical spine magnetic resonance imaging (MRI), or thoracic spine MRI for patients with lumbar disease. Therefore, we were not able to evaluate polyneuropathy or spinal cord compression in this study. Despite these limitations, the current study identified 2 independent predictors of persistent postoperative leg numbness following lumbar fusion in older patients. Additional information on our cohort from further prospective long-term follow-up studies at multiple facilities is required to make this study more meaningful.

CONCLUSION

This work showed that a history of preoperative lumbar decompression, longer symptom duration, and a greater preoperative VAS score for leg numbness were independent predictors of persistent postoperative leg numbness following lumbar fusion in older patients. When surgeons plan to perform lumbar fusion in older patients, they should pay attention to the patient’s surgical history, symptom duration, and preoperative numbness intensity, and explain in advance the potential for postoperative persistent leg numbness in patients who expect recovery of leg numbness following lumbar fusion.

NOTES

Supplementary Materials: Supplementary Tables 1-4 and Fig. 1 can be found via https://doi.org/10.14245/ns.2347312.656.

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: TT; Data curation: TT, SF; Formal analysis: TT; Methodology: TT; Project administration: TT, MK; Visualization: TT; Writing – original draft: TT; Writing – review & editing: MK, SF, FO, YS, YH, TH, KK, TY, YT, KM, NI.

ORCID

Takeru Tsujimoto: 0000-0002-1516-3107
Masahiro Kanayama: 0000-0001-6925-8673
Yukitoshi Shimamura: 0000-0002-9491-3198
Yuichi Hasegawa: 0000-0002-8976-4933
Kenichiro Kakutani: 0000-0002-4006-037X
Takashi Yurube: 0000-0002-3007-361X
Yoshiki Takeoka: 0000-0001-5360-5618
Kunihiko Miyazaki: 0000-0002-7627-3529
Norimasa Iwasaki: 0000-0002-6819-3473

REFERENCES


Supplementary Table 1. A subgroup analysis based on the number of fusion segments

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of fusion segments</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (n = 193)</td>
<td>2 (n = 90)</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>79.2 ± 3.6</td>
<td>79.0 ± 3.2</td>
</tr>
<tr>
<td>Sex, male:female</td>
<td>67:126</td>
<td>26:64</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>24.1 ± 3.6</td>
<td>23.9 ± 3.2</td>
</tr>
<tr>
<td>Type of surgery, PLIF:TLIF</td>
<td>96:97</td>
<td>42:48</td>
</tr>
<tr>
<td>Central spinal canal stenosis</td>
<td>143 (66)</td>
<td>70 (78)</td>
</tr>
<tr>
<td>Vertebral fractures</td>
<td>64 (33)</td>
<td>20 (22)</td>
</tr>
<tr>
<td>Previous lumbar decompression</td>
<td>32 (17)</td>
<td>13 (14)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>35 (18)</td>
<td>12 (7)</td>
</tr>
<tr>
<td>Depression</td>
<td>12 (6)</td>
<td>6 (7)</td>
</tr>
<tr>
<td>Dural injury</td>
<td>8 (4)</td>
<td>6 (7)</td>
</tr>
<tr>
<td>Symptom duration (mo)</td>
<td>22.6 ± 42.7</td>
<td>23.8 ± 32.8</td>
</tr>
<tr>
<td>Operation time (min)</td>
<td>178.2 ± 51.2</td>
<td>229.5 ± 54.8</td>
</tr>
<tr>
<td>EBL (mL)</td>
<td>176.5 ± 121.8</td>
<td>326.0 ± 242.6</td>
</tr>
<tr>
<td>Preop VAS for LBP</td>
<td>5.0 ± 3.3</td>
<td>5.4 ± 3.0</td>
</tr>
<tr>
<td>Preop VAS for leg pain</td>
<td>6.4 ± 3.3</td>
<td>6.6 ± 3.2</td>
</tr>
<tr>
<td>Preop VAS for leg numbness</td>
<td>5.4 ± 3.6</td>
<td>5.5 ± 3.4</td>
</tr>
<tr>
<td>PN group</td>
<td>45 (23)</td>
<td>21 (23)</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%).
BMI, body mass index; PLIF, posterior lumbar interbody fusion; TLIF, transforaminal lumbar interbody fusion; EBL, estimated blood loss; VAS, visual analogue scale; LBP, low back pain; PN, persistent numbness.
## Supplementary Table 2. A subgroup analysis based on pathogenesis

<table>
<thead>
<tr>
<th>Pathogenesis</th>
<th>PN group</th>
<th>N-PN group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spondylolisthesis</td>
<td>38 (21.6)</td>
<td>138 (78.4)</td>
<td></td>
</tr>
<tr>
<td>Lumbar spinal stenosis with foraminal stenosis</td>
<td>30 (24.8)</td>
<td>91 (75.2)</td>
<td></td>
</tr>
<tr>
<td>Lumbar disc herniation</td>
<td>3 (50.0)</td>
<td>3 (50.0)</td>
<td>0.244</td>
</tr>
</tbody>
</table>

Values are presented as number (%).
PN, persistent numbness; N-PN, non-PN.
**Supplementary Table 3.** Reaching MCIDs between PN and N-PN groups

<table>
<thead>
<tr>
<th>Reaching MCID</th>
<th>Total</th>
<th>PN group</th>
<th>N-PN group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>191 (78.9)</td>
<td>46 (74.2)</td>
<td>145 (80.6)</td>
<td>0.289</td>
</tr>
<tr>
<td>Leg pain</td>
<td>219 (84.9)</td>
<td>43 (70.5)</td>
<td>176 (89.3)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Leg numbness</td>
<td>181 (78.0)</td>
<td>23 (36.5)</td>
<td>158 (93.5)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Values are presented as number (%).

MCID, minimum clinically important difference; PN, persistent numbness; N-PN, non-PN; LBP, low back pain.
**Supplementary Table 4.** Comparison of pre- and postoperative spinopelvic sagittal parameters between PN and N-PN groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>PN group</th>
<th>N-PN group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preoperative parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVA (mm)</td>
<td>63.6 ± 54.9</td>
<td>68.6 ± 45.3</td>
<td>0.504</td>
</tr>
<tr>
<td>LL (°)</td>
<td>32.3 ± 16.9</td>
<td>31.4 ± 14.9</td>
<td>0.710</td>
</tr>
<tr>
<td>PI (°)</td>
<td>52.1 ± 11.2</td>
<td>52.4 ± 11.0</td>
<td>0.821</td>
</tr>
<tr>
<td>PT (°)</td>
<td>25.6 ± 9.3</td>
<td>25.0 ± 9.8</td>
<td>0.690</td>
</tr>
<tr>
<td>PI-LL mismatch (°)</td>
<td>19.8 ± 14.6</td>
<td>20.4 ± 19.0</td>
<td>0.839</td>
</tr>
<tr>
<td><strong>Postoperative parameters</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVA (mm)</td>
<td>66.2 ± 48.9</td>
<td>68.1 ± 51.3</td>
<td>0.795</td>
</tr>
<tr>
<td>LL (°)</td>
<td>33.9 ± 16.2</td>
<td>33.2 ± 16.7</td>
<td>0.754</td>
</tr>
<tr>
<td>PI (°)</td>
<td>52.3 ± 10.0</td>
<td>52.2 ± 12.8</td>
<td>0.943</td>
</tr>
<tr>
<td>PT (°)</td>
<td>25.2 ± 10.2</td>
<td>24.9 ± 9.8</td>
<td>0.853</td>
</tr>
<tr>
<td>PI-LL mismatch (°)</td>
<td>18.2 ± 14.5</td>
<td>19.0 ± 19.7</td>
<td>0.762</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%).
PN, persistent numbness; N-PN, non-PN; SVA, sagittal vertical axis; LL, lumbar lordosis; PI, pelvic incidence; PT, pelvic tilt.
Supplementary Fig. 1. Receiver operating characteristic curves for age (A), symptom duration (B), estimated blood loss (C), operation time (D), and serum calcium level (E). AUC, area under the curve.
Biomechanical Analysis of Hybrid Artificial Discs or Zero-Profile Devices for Treating 1-Level Adjacent Segment Degeneration in ACDF Revision Surgery

Weishi Liang1,2,3, *, Yihan Yang1,2,3, *, Bo Han1,2,3, Duan Sun1,2,3, Peng Yin1,2,3, Yong Hai1,2,3

1Department of Orthopedic Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China
2Joint Laboratory for Research and Treatment of Spinal Cord Injury in Spinal Deformity, Laboratory for Clinical Medicine, Capital Medical University, Beijing, China
3Center for Spinal Deformity, Capital Medical University, Beijing, China

Objective: Cervical hybrid surgery optimizes the use of cervical disc arthroplasty (CDA) and zero-profile (ZOP) devices in anterior cervical discectomy and fusion (ACDF) but lacks uniform combination and biomechanical standards, especially in revision surgery (RS). This study aimed to investigate the biomechanical characteristics of adjacent segments of the different hybrid RS constructs in ACDF RS.

Methods: An intact 3-dimensional finite element model generated a normal cervical spine (C2–T1). This model was modified to the primary C5–6 ACDF model. Three RS models were created to treat C4–5 adjacent segment degeneration through implanting cages plus plates (Cage-Cage), ZOP devices (ZOP-Cage), or Bryan discs (CDA-Cage). A 1.0-Nm moment was applied to the primary C5–6 ACDF model to generate total C2–T1 range of motions (ROMs). Subsequently, a displacement load was applied to all RS models to match the total C2–T1 ROMs of the primary ACDF model.

Results: The ZOP-Cage model showed lower biomechanical responses including ROM, in-tradiscal pressure, maximum von Mises stress in discs, and facet joint force in adjacent segments compared to the Cage-Cage model. The CDA-Cage model exhibited the lowest biomechanical responses and ROM ratio at adjacent segments among all RS models, closely approached or lower than those in the primary ACDF model in most motion directions. Additionally, the maximum von Mises stress on the C3–4 and C6–7 discs increased in the Cage-Cage and ZOP-Cage models but decreased in the CDA-Cage model when compared to the primary ACDF model.

Conclusion: The CDA-Cage construct had the lowest biomechanical responses with minimal kinematic change of adjacent segments. ZOP-Cage is the next best choice, especially if CDA is not suitable. This study provides a biomechanical reference for clinical hybrid RS decision-making to reduce the risk of ASD recurrence.

Keywords: Biomechanical analysis, Cervical revision surgery, Hybrid surgery, Adjacent segment degeneration, Zero-profile device, Cervical disc arthroplasty

INTRODUCTION

Adjacent segment degeneration (ASD) is a commonly observed long-term complication in middle-aged and elderly patients who have undergone cervical fusion.1 The main manifestations on imaging are disc height reduction and herniation, facet joint proliferation, and segmental instability. When a patient exhibits neurological symptoms that correspond to imag-
Biomechanical Analysis of Hybrid ACDF Revision Strategies

Liang W, et al.

Furthermore, many biomechanical factors are considered the most significant contributors to the development of ASD. ASD disrupts the biomechanical stability of the cervical spine, and in cases of severe nerve compression, surgical revision becomes necessary to address the condition.

However, for patients with secondary ASD after primary ACDF surgery, traditional titanium alloy cages plus plates are still conventionally used in revision surgery (RS). Segmental fusion in these cases leads to a greater loss of motor segments, an increase in range of motion (ROM) compensation, and an elevated risk of ASD. The novel zero-profile (ZOP) device has been gradually applied in ACDF surgery. The ZOP device is superior to cage plus plate in reducing the incidence of ASD, minimizing intraoperative blood loss, and alleviating dysphagia. Moreover, a finite element (FE) study also reported that the ZOP device effectively reduces biomechanical responses by reducing the ROM and intradiscal pressure (IDP) in the nucleus pulposus at adjacent segments.

In recent years, anterior cervical hybrid surgery developed rapidly as an innovative approach to treat multi-segmental cervical degeneration disease (CDD). This procedure combines ACDF with cervical disc arthroplasty (CDA), providing a new strategy for RS after ACDF. The implantation of an artificial disc in the cervical spine can improve the distribution ratio of intervertebral ROM to a certain extent by preserving motion in the surgical segments. Furthermore, many biomechanical studies have demonstrated that hybrid surgery surpasses ACDF surgery in restoring the normal biomechanical state of the cervical spine.

It is well recognized that the significant increase in IDP and facet joint force (FJF) caused by the compensatory increase in the ROM of the upper and lower adjacent segments after ACDF serves as a critical biomechanical mechanism underlying ASD. The RS strategy for the ASD following ACDF necessitates a reconstruction of the cervical spine sequence and a reduction in the biomechanical responses of adjacent segments. The hybrid strategies used in the treatment of multilevel CDDs are diversified, mostly based on the surgical experience of surgeons, and lack of strong biomechanical and long-term follow-up evidence. Particularly in RS, there is a deficiency in biomechanical data as a reference for making surgical decisions and forecasting the long-term degenerative prognosis of adjacent segments. Both the ZOP device and the cervical artificial disc exhibit the capability to reduce the biomechanical responses on adjacent segments. However, their effectiveness in reducing the biomechanical responses of adjacent segments and reducing the risk of ASD recurrence compared with traditional ACDF revision has not been reported, and further investigation is needed.

In this study, we established the C2–T1 intact, primary ACDF model, and 3 other RS FE models integrating previous ACDF procedures with new-implanted revision devices: the traditional cage plus plate (Cage-Cage), the ZOP device (ZOP-Cage), and the Bryan disc (CDA-Cage) models. These FE models were designed to explore the biomechanical characteristics of adjacent segments in the RS models. From the biomechanical perspective, this FE study will provide a basis for selecting suitable RS methods to reduce the risk of ASD recurrence following ACDF RS.

MATERIALS AND METHODS

1. Establishment of the Intact Cervical FE Model

The geometric model of the C2–T1 vertebrae was generated using computed tomography (CT) scan data derived from the cervical spine of a healthy 30-year-old female, as illustrated in Fig. 1. Initial processing of the CT scan data involved its importation into Mimics (Materialise Inc., Leuven, Belgium), where
it was converted into a geometric structure. The resulting geometric model was then subjected to meshing procedures using Hypermesh (Altair Engineering Inc., Troy, MI, USA). Subsequently, the meshed model underwent preprocessing and analysis using Abaqus (Dassault Systemes Simulia Corp., Johnston, RI, USA). The FE study protocol has been reviewed and approved by the Ethics Committee of Beijing Chaoyang Hospital, Capital Medical University (No. 2019-ke-212).

### 2. Material Properties and FE Modeling

The material properties utilized in this study’s FE models followed previously published literature, and the detailed information was summarized in Table 1. The vertebrae were divided into 2 regions: cortical bone (1-mm thickness) and cancellous bone, both meshed using tetrahedral elements. The cortical endplates of the discs and facet joints were meshed with hexahedral elements. Furthermore, the intervertebral disc was divided into 3 components: the nucleus pulposus, annulus fibrosus (Fig. 1), and endplates (0.5-mm thick). The nucleus pulposus and annulus fibrosus were meshed utilizing hexahedral elements, and they occupied approximately 40% and 60% of the intervertebral disc volume, respectively. Mooney-Rivlin constitutive model was used to model the nonlinear behavior of the ground substance of both the nucleus pulposus and the annulus fibrosus. The annulus fibers, consisting of 8 layers, were modeled as hypoelastic materials using truss elements and were embedded within the annulus ground substance with an inclination of approximately 30° to the transverse plane. Moreover, the main ligaments were established using nonlinear tension-only spring elements placed in their anatomically accurate positions, encompassing the anterior longitudinal ligament, posterior longitudinal ligament, ligamentum flavum, interspinous ligament, and facet capsular ligament

<table>
<thead>
<tr>
<th>Component</th>
<th>Element type</th>
<th>Constitutive model</th>
<th>Young’s modulus (MPa)</th>
<th>Poisson ratio</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortical bone</td>
<td>C3D4</td>
<td>Isotropic elastic</td>
<td>E = 10,000</td>
<td>ν = 0.3</td>
<td>18</td>
</tr>
<tr>
<td>Cancellous bone</td>
<td>C3D4</td>
<td>Neo-Hookean</td>
<td>E = 100</td>
<td>ν = 0.3</td>
<td>20</td>
</tr>
<tr>
<td>Annulus ground substance</td>
<td>C3D8H</td>
<td>Mooney-Rivlin</td>
<td>C₁₀ = 0.1333, C₀₁ = 0.0333, D₁ = 0.6</td>
<td>-</td>
<td>20</td>
</tr>
<tr>
<td>Annulus fibers</td>
<td>T3D2</td>
<td>Hypoelastic</td>
<td>350–550</td>
<td>ν = 0.3</td>
<td>17</td>
</tr>
<tr>
<td>Nucleus pulposus</td>
<td>C3D8H</td>
<td>Mooney-Rivlin</td>
<td>C₁₀ = 0.12, C₀₁ = 0.09, D₁ = 0</td>
<td>-</td>
<td>13</td>
</tr>
<tr>
<td>Spine ligaments (ALL, PLL, ISL, LF, CL)</td>
<td>Spring</td>
<td>Nonlinear elastic</td>
<td>-</td>
<td>-</td>
<td>19</td>
</tr>
<tr>
<td>PEEK cage</td>
<td>C3D4</td>
<td>Linear elastic</td>
<td>E = 3,760</td>
<td>ν = 0.38</td>
<td>17</td>
</tr>
<tr>
<td>Screws/plate (titanium alloy)</td>
<td>C3D4</td>
<td>Linear elastic</td>
<td>E = 110,000</td>
<td>ν = 0.3</td>
<td>7</td>
</tr>
<tr>
<td>PEEK ZOP device</td>
<td>C3D4</td>
<td>Linear elastic</td>
<td>E = 3,760</td>
<td>ν = 0.38</td>
<td>17</td>
</tr>
<tr>
<td>ZOP device screws (titanium alloy)</td>
<td>C3D4</td>
<td>Linear elastic</td>
<td>E = 110,000</td>
<td>ν = 0.3</td>
<td>7</td>
</tr>
<tr>
<td>Bryan disc shell</td>
<td>C3D8</td>
<td>Linear elastic</td>
<td>E = 110,000</td>
<td>ν = 0.3</td>
<td>35</td>
</tr>
<tr>
<td>Bryan disc Nucleus</td>
<td>C3D8</td>
<td>Linear elastic</td>
<td>E = 30</td>
<td>ν = 0.45</td>
<td>35</td>
</tr>
</tbody>
</table>

ALL, anterior longitudinal ligament; PLL, posterior longitudinal ligament; ISL, interspinous ligament; LF, ligamentum flavum; CL, capsular ligament; PEEK, polyetheretherketone; ZOP, zero-profile.
In the C5–6 primary ACDF segment, polyetheretherketone cages filled with center bone graft ( cancellous bone properties)\textsuperscript{7,24} were inserted, followed by the placement of anterior titanium alloy plates (length 21 mm) and screws (insertion angle approximately 0°) to achieve solid fusion. The nodes within the interface region of the devices and bone were shared to connect them to the models. The material properties of different implant devices, including cage, screw plus plate, ZOP devices, and Bryan disc, are also listed in Table 1.

5. Loading Conditions of the Surgical Models

A follower load of 73.6 N and a pure moment of 1.0 Nm was imposed on the primary ACDF model to produce ROMs in 3 planes. All RS models, including Cage-Cage, ZOP-Cage, and CDA-Cage, were subjected to displacement loads; the displacement load applied to all RS models matched the total C2–T1 ROMs of the primary ACDF model. Soft and frictionless contact properties were used to replicate the sliding contact between the cartilage endplates of the facet joints.\textsuperscript{25} Finally, the segmental ROM, IDP, and FJF in different surgical models during 3 motion planes were calculated. During extension, the mean FJFs for both left and right facet joints at the same level were recorded and calculated. The FJF during flexion was not calculated because the C2–T1 facet joints had no contact forces. The FJFs on the loaded side were recorded and then averaged for the left and right loading conditions.

RESULTS

1. Model Validation

In the present study, the segmental ROMs for C2–T1 in the intact model were compared with those from both \textit{in vitro} studies and FE studies. During flexion-extension, the ROMs observed in the present study are close to the data from the \textit{in vitro} study conducted by Wheeldon et al.\textsuperscript{22} In the present study, segmental ROM is slightly higher at C4–5 and C6–7, exceeding 3° compared to the study of Erbulut et al.,\textsuperscript{26} but similar in other segments (Fig. 3A).

During lateral bending, the ROMs of C2–3 and C7–T1 are close to the \textit{in vitro} study by Yogananda et al.\textsuperscript{23} and the FE studies by Wu et al.\textsuperscript{27} and Wang et al.\textsuperscript{28} In the C3–4, C4–5, and C5–6 segments, our study’s ROM values are notably lower than those reported by Yogananda et al.\textsuperscript{23} and Wang et al.\textsuperscript{28} However, the difference is relatively smaller, approximately 1° or less, when compared to the data presented by Wu et al.\textsuperscript{27} (Fig. 3B).
Finally, the segmental ROM for axial rotation in this study was also compared with the data of the study conducted by Yogananda et al., Wu et al., and Wang et al. Except for the C7–T1 segment, the segmental ROMs of the other segments exhibit close to at least 2 of the 3 studies referenced. Moreover, the ROM for the C7–T1 segment in the present study is higher than the 1.48° reported by Wu et al. (Fig. 3C).

2. ROMs at the Adjacent Segments

Under a follower load of 73.6 N, the total C2–T1 ROMs in

Fig. 3. Model validation of C2–T1 intact finite element (FE) model of the present study under a 1.0-Nm moment, compared with the published studies. (A–C) Comparison of segmental range of motions (ROMs) of C2–T1 during flexion-extension, lateral bending, and axial rotation.

Fig. 4. Comparison of range of motions (ROMs) in different finite element models during flexion (A), extension (B), lateral bending (C), and axial rotation (D). ACDF, anterior cervical discectomy and fusion; ZOP, zero-profile; CDA, cervical disc arthroplasty.
the primary ACDF model following a 1-Nm moment during flexion, extension, lateral bending, and axial rotation were 19.5°, 29.8°, 18.5°, and 17.6°, respectively. The total C2–T1 ROMs of all 3 RS models matched those of the primary ACDF model. Fig. 4 illustrates the segmental ROMs for the primary ACDF and 3 RS models in different motion directions. In comparison to the primary ACDF model, the ROMs of C3–4 and C6–7 adjacent segments of the Cage-Cage model and ZOP-Cage model were increased in all motion directions; and the ROMs of the ZOP-Cage model were lower than those of the Cage-Cage model in all motion directions. The CDA-Cage model has the lowest adjacent segment ROM among all RS models in all motion directions. Furthermore, the ROMs of the upper and lower adjacent segments in the CDA-Cage model were lower than those in the primary ACDF model during flexion, lateral bending, and axial rotation, except for C6/7, while during flexion was close to those in the primary ACDF model.

3. Distribution Percentage of Segmental ROMs

The ROM percentages for each C2–T1 segment are shown in Fig. 5. In comparison to the primary ACDF model, both the Cage-Cage and ZOP-Cage models exhibited an increase in the ROM ratio of the C3–4 and C6–7 segments in all motion directions. The CDA-Cage model demonstrated the lowest ROM ratio of C3–4 and C6–7 segments among all RS constructs, which is close to or lower than those of the primary ACDF model in all motion directions. Additionally, the intervertebral ROM at the C4–5 segment in the CDA-Cage model was preserved and exhibited overactivity compared to the primary ACDF model during extension, lateral bending and axial rotation.

4. Segmental IDP Analysis at Adjacent Segments

The IDP of the adjacent segments in the different FE models are shown in Fig. 6. Among the RS models, the Cage-Cage model had the highest IDP in the C3–4 and C6–7 segments, followed by the ZOP-Cage model, both surpassing the primary ACDF
model in all motion directions. The CDA-Cage model exhibited IDP values similar to the primary ACDF model. Specifically, during extension, the IDP in the C3–4 and C6–7 segments of the CDA-Cage model was lower than that of the primary ACDF model.

5. **Stress Analysis of the Discs at Adjacent Segments**

The disc stress distribution features of the C3–4 and C6–7 discs in each model are shown in Figs. 7 and 8, respectively. The maximum von Mises stresses of annulus fibrosus at the C3–4 and C6–7 discs were higher than those of the nucleus pulposus in all directions. In different motion directions, the maximum von Mises stress was concentrated at the loading side’s corresponding edge. Comparing RS models to the primary ACDF model, the Cage-Cage and ZOP-Cage models exhibited increased maximum von Mises stress in C3–4 and C6–7 intervertebral discs. Conversely, the CDA-Cage model showed reduced maximum von Mises stress levels, except for the C6–7 segment during flexion.

6. **FJF Analysis at Adjacent Segments**

The FJF values of the different FE models in the upper and lower adjacent segments are presented in Fig. 9. In comparison to the primary ACDF model, the FJF values at the C3–4 and C6–7 levels in the Cage-Cage and ZOP-Cage models were increased, with the Cage-Cage model demonstrating the highest values, except at C3–4 during lateral bending. Moreover, the CDA-Cage model exhibited a reduced FJF values at the C3–4 and C6–7 segments in comparison to the primary ACDF model during extension, lateral bending and axial rotation.

**DISCUSSION**

1. **Main Findings May Benefit the Decision-Making of ACDF RS**

Cervical hybrid surgery is an effective method to preserve the ROM in cases of CDD, and it can effectively delay the process of cervical degeneration. In the selection of internal fixators for ACDF surgery, ZOP devices exhibit clear advantages over traditional cage plus plate structures in terms of both clinical efficacy and biomechanical response reduction. This study innovatively compared the RS methods focusing on the hybrid CDA-Cage construct, as well as ACDF involving ZOP devices and traditional cage plus plate structures following primary
Fig. 7. The von Mises stress cloud map of C3–4 intervertebral disc in operation finite element (FE) models. Stress distribution characteristics of C3–4 intervertebral discs in different FE models during flexion, extension, left bending, right bending, left rotation, and right rotation were shown. ACDF, anterior cervical discectomy and fusion; ZOP, zero-profile; CDA, cervical disc arthroplasty.
Fig. 8. The von Mises stress cloud map of C6–7 intervertebral disc in operation finite element (FE) models. Stress distribution characteristics of C6–7 intervertebral discs in different FE models during flexion, extension, left bending, right bending, left rotation, and right rotation were shown. ACDF, anterior cervical discectomy and fusion; ZOP, zero-profile; CDA, cervical disc arthroplasty.
ACDF from the perspective of biomechanics. The authors established a C2–T1 intact model, in which the primary operation was conducted in the C5–6 segment that was most prone to degeneration in the clinic, while the RS addressed the C4–5 higher adjacent segment due to ASD. Using the 3 revision methods mentioned above, we systematically compared the changes in biomechanical responses of the upper and lower adjacent segments following different RS models. Our results indicated that both ZOP-Cage and CDA-Cage constructs reduced biomechanical responses, providing valuable data for guiding clinical decision-making when selecting surgical methods to minimize the risk of re-ASD after RS.

2. ROMs at Adjacent Segments and Distribution of Segmental ROMs

After ACDF, the loss of ROM in the fusion segment was compensated by other nonfusion segments. These adjacent segments are most susceptible to biomechanical changes, exhibiting a significant increase in ROM and a higher risk of ASD. When different types of internal fixators were implanted in the surgical segment, the ROM compensation of the adjacent segments was different. In this study, compared to the Cage-Cage model, the ZOP-Cage model has a larger ROM, which results in a slightly lower adjacent segmental ROM than the Cage-Cage model. When the ROMs of the surgical segments with CDA were all retained or even overactive, making CDA-Cage model has the lowest adjacent segment ROM among all RS models. These findings are in line with those of Faizan et al., which suggest that the adjacent segmental ROM, facet joint loads, and endplate stresses of CDA combined with ACDF were closer to the normal model than the 2-level ACDF model, and had less effect on the biomechanical responses of adjacent segments. Wong et al. demonstrated that in the FE models of the three-level continuous hybrid surgery model, the reduction in ROM in the lower adjacent segment was at the cost of increased ROM in the upper adjacent segments. Additionally, Wu et al. found that the location of ACDF and CDA is an important factor in the kinematics of adjacent segments, and the ROM of adjacent ACDF segments increases much more than that of adjacent CDA segments. However, our study is inconsistent with the findings of the 2 studies of Wong et al. and Wu et al., and as we did not observe significantly higher or lower ROM in the C3–4 segment compared to the C6–7 segment in the CDA model. This discrepancy may be attributed to the number of total segments or surgical segments used in the study, as well as the location limitations of the surgical segment of the CDA imposed by revision. In this study, when compared to the other 2 RS models, the CDA-Cage model demonstrates the most significant reduction in ROM in the 2 adjacent segments. However, there were no significant differences observed in the reduction of ROM between the upper and lower adjacent segments in the CDA-Cage model. The effects of the number of total segment and hybrid segments in different models and the location of CDA and ACDF on adjacent ROMs still need to be further clarified.

3. IDP and Disc Stress Analysis at Adjacent Segments

The changes in IDP after ACDF may be attributed to many...
The rise in IDP obstructs the diffusion of nutrient substances from the endplate to the intervertebral disc, leading to a deterioration in the nutritional status of the disc. This is also considered a crucial factor contributing to disc degeneration in patients following fusion surgery for ASD. An increased number of fusion segments is a high-risk factor for ASD after multilevel cervical fusion. Patients who underwent multilevel fusion experienced a more substantial decrease in overall ROM and exhibited increased compensatory ROM in the upper adjacent segments. This increased ROM was associated with a higher IDP, which in turn elevated their risk of developing degeneration.

Previous study has demonstrated that IDP in adjacent segments following 2-level continuous hybrid surgery is significantly lower than that following 2-level continuous ACDF surgery. Our study found a consistent phenomenon in the hybrid CDA-Cage model exhibiting the lowest adjacent segmental IDP among the 3 RS models. Additionally, the present study also used a stress cloud map to display the distribution of von Mises stress on the disc. The stress cloud maps showed that the maximum stress on the disc concentrated at the corresponding edge of the loading side. Among all RS models, both Cage-Cage and ZOP-Cage models showed an increased maximum stress at the C3–4 and C6–7 intervertebral discs when compared to the primary ACDF model, whereas the CDA-Cage model consistently exhibited the lowest maximum stress. In the hybrid CDA-Cage model, the changes of IDP and maximum disc stress in adjacent segments are close to the trends observed in ROM and FJF during different postures, suggesting that CDA has a significant protective effect on the disc of both adjacent segments.

4. FJF Analysis at Adjacent Segments

Hypermobile facet joints in adjacent segments following cervical fusion can lead to increased stress load, potentially resulting in cervicogenic neck pain and even headaches. ACDF induces a decrease in FJF within the fusion segment due to rigid fixation. Studies have demonstrated that the implantation of an artificial intervertebral disc keeps the operative segment from overloading the facet joints and keeps them in a relatively healthy state. The Bryan disc used in this study has been found to reduce FJFs at the surgical segments, potentially delaying facet joint degeneration and lowering the associated degenerative risk. Conversely, other clinically used discs such as Mobi-C and Prestige-LP discs have shown an increased FJF. In addition to the FJF of the operative segment, the bearing load by facet joints of adjacent segments plays a pivotal role in predicting degeneration. Our study showed that the FJF at adjacent segments of both Cage-Cage and ZOP-Cage constructs increased compared with the primary ACDF model. In contrast, FJF in adjacent segments of CDA-Cage hybrid constructs exhibited a significant decrease compared to the primary ACDF model. The changes in FJF in adjacent segments were consistent with the changes in ROM, with CDA-Cage constructs demonstrating a positive biomechanical effect in preventing facet joint degeneration. Faizan et al. found that CDA can protect adjacent facet joints by compensating for ROM, resulting in a tendency for increased facet joint load at the CDA level while not significantly increasing FJF at adjacent segments. This phenomenon was also corroborated in our model, as FJF in the adjacent segments above the CDA level was lower than that in the CDA-operated segment. Previous studies have shown that the increase in contact force and load of adjacent facet joints after fusion surgery may lead to pathological injuries of the articular surface and ultimately accelerate joint surface degeneration. In this study, CDA effectively alleviates adjacent segment FJF, thereby mitigating ASD and providing an effective means to alleviate neck pain resulting from facet joint hypermobility.

5. Summary Analysis of Biomechanical Parameters Combined With Clinical Application

From a biomechanical perspective, our study reveals that in the RS methods, the hybrid CDA-Cage construct provides superior protection for adjacent segments compared to both ZOP-Cage and Cage-Cage constructs. Among the 3 RS constructs, CDA-Cage performed the most significant reduction in ROM, IDP, maximum disc stress, and FJF at the adjacent segments. Because of the large number of patients undergoing revision due to ASD after 1-level ACDF, the results of this study can be used as a reference for surgeons when determining the optimal approach (CDA-Cage or ZOP-Cage) for single-stage ACDF revision. Due to the relatively strict surgical indications of CDA, the hybrid CDA-Cage construct is recommended as the preferred surgical approach in patients not involving severe osteoporosis, severe cervical instability, trauma, or pathological bone injury. For the clinical effect, hybrid surgery exhibited the advantages of less intraoperative blood loss, shorter return duration to work, and a lower incidence of nerve injury. For patients who do not meet CDA criteria, ZOP-Cage emerges as a superior surgical choice over Cage-Cage, and both of these op-
operations can meet the need for strong fixation of the cervical spine.4

6. Limitations
The present study has several limitations. Firstly, the cervical spine FE model was generated based on the cervical spine of a single healthy individual, and related data could not be analyzed statistically between groups. Simple biases that occur during model construction may have a small impact on biomechanical response parameter trends. Secondly, while the plate-screw system, ZOP device and artificial disc used in this study have been proven to have good representativeness and clinical effects, they may still differ from other products of the same category. For example, the biomechanically distinctive ZEVO plate-screw system (Medtronic Sofamor Danek) with short plates and high-angled screws has demonstrated superior mechanical stability and load-sharing capabilities.8 Furthermore, it has shown efficacy in effectively preventing cage subsidence and reducing adjacent-level ossification development.2 Given its proven advantages, future comparative research should incorporate this short plate-high-angled screw system, along with the different cortical-cancellous composition of allograft spacers. The generalizability of the study's conclusions to other similar products requires further validation. Thirdly, some common cervical degenerative manifestations (osteophytic hyperplasia, disc, facet joints, endplate degenerative injuries) were not simulated in this study because this study focuses more on the comparison of different revision procedures. Future studies are needed to investigate the different material properties of degeneration and plate fixation as RS for biomechanically assuming the risk of ASD.

CONCLUSION
In this FE study, the biomechanical responses on the adjacent segments of the CDA-Cage constructs were significantly lower than those of Cage-Cage and ZOP-Cage constructs with decreased ROM, IDP, maximum disc stress, and FJF. Moreover, CDA-Cage exhibited the best performance in reducing IDP and FJF at segments C3–4 and C7–T1. The CDA in hybrid CDA-Cage constructs reduces the biomechanical responses of the adjacent segments, making it the preferred revision procedure for preventing ASD. When patients do not meet the indications for CDA, the ZOP-Cage construct, with its superior biomechanical performance, can provide a more favorable alternative to traditional Cage-Cage constructs for preventing ASD. Further biomechanical and clinical studies are anticipated to investigate the effects of multilevel hybrid revision after ACDF surgery and to assess the impact of the relative positioning of CDA and ACDF on the biomechanics of adjacent segments. This study offers a reference for surgical decision-making based on biomechanical evidence regarding the use of hybrid surgery for the treatment of ASD following primary ACDF. It is necessary for future research to validate these conclusions.

NOTES

Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This study is supported by the Beijing Hospitals Authority Youth Programme (No. QML20230315), Beijing Natural Science Foundation (No. 7242061), and Clinical Research Incubation Project, Beijing Chaoyang Hospital, Capital Medical University (No. CYFH202322).

Author Contribution: Conceptualization: WL, BH, YY, DS, PY, YH; Data curation: WL, BH, YY, DS, PY, YH; Formal analysis: WL, BH, YY, DS, PY, YH; Funding acquisition: PY, YH; Methodology: WL, YY, DS, PY, YH; Project administration: PY, YH; Visualization: WL, BH, PY, YH; Writing – original draft: WL, BH, YY; Writing – review & editing: WL, YY, PY, YH.

ORCID
Weishi Liang: 0000-0003-3565-6504
Bo Han: 0000-0002-3618-5700
Yihan Yang: 0009-0005-8254-8567
Duan Sun: 0009-0009-2374-2756
Peng Yin: 0000-0001-9984-7663
Yong Hai: 0000-0002-7206-325X

REFERENCES
changes after one- or two-level anterior cervical discectomy and fusion using either a zero-profile device or cage plus plate: a finite element analysis. Comput Biol Med 2020;120:103760.


Using Machine Learning Models to Identify Factors Associated With 30-Day Readmissions After Posterior Cervical Fusions: A Longitudinal Cohort Study

Aneysis D. Gonzalez-Suarez¹, Paymon G. Rezaii¹, Daniel Herrick¹, Seth Stravers Tigchelaar¹, John K. Ratliff¹, Mirabela Rusu², David Scheinker¹, Ikchan Jeon¹, Atman M. Desai¹

¹Department of Neurosurgery, Stanford University, Stanford, CA, USA
²Department of Radiology, Stanford University, Stanford, CA, USA

Objective: Readmission rates after posterior cervical fusion (PCF) significantly impact patients and healthcare, with complication rates at 15%–25% and up to 12% 90-day readmission rates. In this study, we aim to test whether machine learning (ML) models that capture interfactorial interactions outperform traditional logistic regression (LR) in identifying readmission-associated factors.

Methods: The Optum Clinformatics Data Mart database was used to identify patients who underwent PCF between 2004–2017. To determine factors associated with 30-day readmissions, 5 ML models were generated and evaluated, including a multivariate LR (MLR) model. Then, the best-performing model, Gradient Boosting Machine (GBM), was compared to the LACE (Length patient stay in the hospital, Acuity of admission of patient in the hospital, Comorbidity, and Emergency visit) index regarding potential cost savings from algorithm implementation.

Results: This study included 4,130 patients, 874 of which were readmitted within 30 days. When analyzed and scaled, we found that patient discharge status, comorbidities, and number of procedure codes were factors that influenced MLR, while patient discharge status, billed admission charge, and length of stay influenced the GBM model. The GBM model significantly outperformed MLR in predicting unplanned readmissions (mean area under the receiver operating characteristic curve, 0.846 vs. 0.829; p < 0.001), while also projecting an average cost savings of 50% more than the LACE index.

Conclusion: Five models (GBM, XGBoost [extreme gradient boosting], RF [random forest], LASSO [least absolute shrinkage and selection operator], and MLR) were evaluated, among which, the GBM model exhibited superior predictive performance, robustness, and accuracy. Factors associated with readmissions impact LR and GBM models differently, suggesting that these models can be used complementarily. When analyzing PCF procedures, the GBM model resulted in greater predictive performance and was associated with higher theoretical cost savings for readmissions associated with PCF complications.

Keywords: Machine learning, Predictive modeling, Readmission, Spine, Costs, Posterior cervical fusion
INTRODUCTION

Posterior cervical fusion (PCF) is a common surgical intervention used to treat a variety of cervical spinal pathologies, including spondylosis, spinal tumors, and spinal deformity. However, postoperative complications following PCF are not uncommon; in fact, one literature review reported that patients undergoing PCF have an overall complication rate of 15%–25%. Furthermore, analysis of a spine-specific database for cervical fusion surgeries showed that patients who underwent the posterior approach had unplanned 90-day readmission rates of up to 12%. Such unplanned readmissions contribute to an estimated hospital cost of $10 billion nationally. These costs are increasingly relevant as the rate of cervical fusion surgeries is expected to increase by 13.3% for anterior cervical fusions and 19.3% for PCF.

Given these considerations, there is an ongoing effort to generate predictive models that can successfully identify patients at high risk of readmission following spine surgery. By identifying high-risk patients, model simulations could offer potential interventions that may reduce readmissions and associated healthcare costs. In particular, machine learning (ML) models have shown promise in identifying intervention strategies that can inform how healthcare and hospital systems allocate resources to reduce postoperative readmission rates. These models leverage patient demographic information and perioperative data to determine relevant factors that predict patient's risk of being readmitted following surgery. Here, we build on this literature by using ML correlators, including logistic regression (LR) models, to identify risk factors associated with readmissions following PCF. We hypothesized that, in keeping with previously published work, ML models such as least absolute shrinkage and selection operator (LASSO), random forest (RF), stochastic gradient boosting machine (GBM), or extreme gradient boosting (XGBoost) can out predict and outperform traditional LR models, while also contributing to greater readmission-associated cost savings when implemented.

To test this 2-fold hypothesis, we (1) compared the predictive performance of 4 supervised ML algorithms to a traditionally-implemented ML model, multivariate LR (MLR) and (2) estimated the potential cost savings of reducing readmissions by implementing the best-performing ML model in a clinical setting and comparing it to the LACE (Length patient stay in the hospital, Acuity of admission of patient in the hospital, Comorbidity, and Emergency visit) index. While this study presents an initial effort to use supervised classification and regression ML models to predict unplanned readmissions rates and simulate readmission-associated costs savings, there are inherent limitations or biases of ML models that are unaccounted for and merit recognition. In addition, further clinical evaluation is needed to refine, finetune, and enhance the performance of the models presented in this study.

MATERIALS AND METHODS

1. Cohort

To analyze specific patient utilization, expenditure, and enrollment data between 1/1/2004 and 11/30/2017, the Optum Clinformatics Data Mart database (Optum, Inc., Eden Prairie, MN, USA) was used. Patients were identified using the Current Procedural Terminology code 22600 for posterior cervical decompression with fusion and 22840, 22842, 22843, or 22844 for posterior spinal instrumentation. Patients were subsequently filtered using our eligibility criteria (Fig. 1). To maintain the...
specificity of our study population to PCF and instrumentation, we ensured the exclusion of patients who had undergone anterior or lumbar procedures, as these represent distinct surgical categories with different risk profiles and outcomes. Since this database contains deidentified medical claims, patient consent is not applicable, and the study is exempted from requiring Institutional Review Board approval.

2. Predictors

The predictors used were based on previous studies and included patient demographics, socioeconomic status, procedural service, healthcare utilization, complications and comorbidities\(^{13,14}\) (Table 1). These predictors were selected before analyzing the eligible patient data, and the algorithm implemented to analyze these predictors was in agreement with current medical literature.\(^{15-19}\)

3. Outcomes

The primary outcome assessed was the relative influence of each predictor on the risk for unplanned readmissions, which was normalized using the variable importance score for each model. For secondary outcomes, we measured the performance of each ML model in predicting 30-day readmissions by computing the area under the receiver operating characteristic curve (AUC), which serves as a measure of the model’s ability to estimate the probability of readmission as previously described in Bamber.\(^{16}\) To calculate the potential cost savings associated with implementing these models, we applied the readmission reduction rates projected by the GBM model to the inflation-adjusted, all-payer national estimates of surgical readmissions within 30 days of surgery, as reported in a previously published Healthcare Cost and Utilization Project Statistical Brief.\(^{17}\) The brief provides a comprehensive analysis of 30-day post-surgical readmissions and associated costs across a spectrum of high-volume surgeries, including spinal procedures, in various income demographics. The predictive performance and associated cost savings were compared between the top-performing GBM model and the LACE index, a previously-validated readmission model.\(^{18}\)

Table 1. Univariate logistic regression

<table>
<thead>
<tr>
<th>Variable</th>
<th>Not readmitted</th>
<th>Readmitted</th>
<th>Coefficient (SE)</th>
<th>Pseudo R(^2)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (yr)</td>
<td>60.6 ± 10.9*</td>
<td>69.1 ± 10.6*</td>
<td>0.077 (0.004)</td>
<td>0.098</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Female sex</td>
<td>1,475 (45.3)</td>
<td>397 (45.4)</td>
<td>0.005 (0.077)</td>
<td>&lt; 0.001</td>
<td>0.949</td>
</tr>
<tr>
<td>Medicare</td>
<td>1,411 (43.3)</td>
<td>709 (81.1)</td>
<td>1.726 (0.093)</td>
<td>0.099</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Geographic division</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Atlantic</td>
<td>895 (27.5)</td>
<td>209 (23.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>East North Central</td>
<td>550 (16.9)</td>
<td>174 (19.9)</td>
<td>0.304 (0.116)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>West South Central</td>
<td>398 (12.2)</td>
<td>97 (11.1)</td>
<td>0.043 (0.137)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mountain</td>
<td>325 (10.0)</td>
<td>76 (8.7)</td>
<td>0.001 (0.149)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>West North Central</td>
<td>359 (11.0)</td>
<td>66 (7.6)</td>
<td>-0.239 (0.154)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific</td>
<td>283 (8.7)</td>
<td>93 (10.6)</td>
<td>0.342 (0.142)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle Atlantic</td>
<td>176 (5.4)</td>
<td>75 (8.6)</td>
<td>0.602 (0.158)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New England</td>
<td>132 (4.1)</td>
<td>60 (6.9)</td>
<td>0.666 (0.174)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>East South Central</td>
<td>138 (4.2)</td>
<td>24 (2.7)</td>
<td>-0.295 (0.234)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
<td></td>
<td>0.083</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>POS</td>
<td>1,237 (38.0)</td>
<td>105 (12.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMO</td>
<td>760 (23.3)</td>
<td>271 (31.0)</td>
<td>1.435 (0.124)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>753 (23.1)</td>
<td>401 (45.9)</td>
<td>1.836 (0.119)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPO</td>
<td>227 (7.0)</td>
<td>75 (8.6)</td>
<td>1.359 (0.168)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPO</td>
<td>219 (6.7)</td>
<td>17 (1.9)</td>
<td>-0.089 (0.272)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indemnity</td>
<td>60 (1.8)</td>
<td>5 (0.6)</td>
<td>-0.018 (0.476)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 1. Univariate logistic regression (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Not readmitted</th>
<th>Readmitted</th>
<th>Coefficient (SE)</th>
<th>Pseudo R²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comorbidities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td>41 (1.3)</td>
<td>12 (1.4)</td>
<td>0.088 (0.330)</td>
<td>&lt; 0.001</td>
<td>0.792</td>
</tr>
<tr>
<td>Anemia</td>
<td>111 (3.4)</td>
<td>38 (4.3)</td>
<td>0.253 (0.192)</td>
<td>&lt; 0.001</td>
<td>0.196</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>225 (6.9)</td>
<td>80 (9.2)</td>
<td>0.306 (0.136)</td>
<td>0.001</td>
<td>0.028</td>
</tr>
<tr>
<td>CHF</td>
<td>67 (2.1)</td>
<td>38 (4.3)</td>
<td>0.772 (0.207)</td>
<td>0.003</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Coagulopathy</td>
<td>24 (0.7)</td>
<td>15 (1.7)</td>
<td>0.855 (0.331)</td>
<td>0.001</td>
<td>0.014</td>
</tr>
<tr>
<td>Depression</td>
<td>454 (13.9)</td>
<td>104 (11.9)</td>
<td>-0.182 (0.116)</td>
<td>0.001</td>
<td>0.112</td>
</tr>
<tr>
<td>DM</td>
<td>525 (16.1)</td>
<td>178 (20.4)</td>
<td>0.286 (0.097)</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>DMcx</td>
<td>97 (3.0)</td>
<td>56 (6.4)</td>
<td>0.802 (0.172)</td>
<td>0.005</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Drugs</td>
<td>26 (0.8)</td>
<td>15 (1.7)</td>
<td>0.774 (0.327)</td>
<td>0.001</td>
<td>0.023</td>
</tr>
<tr>
<td>Fluid/electrolyte disorder</td>
<td>163 (5.0)</td>
<td>80 (9.2)</td>
<td>0.648 (0.142)</td>
<td>0.005</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>HTN</td>
<td>1,619 (49.7)</td>
<td>477 (54.6)</td>
<td>0.195 (0.076)</td>
<td>0.002</td>
<td>0.011</td>
</tr>
<tr>
<td>HTNcx</td>
<td>62 (1.9)</td>
<td>39 (4.5)</td>
<td>0.878 (0.208)</td>
<td>0.004</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Hypertrophy</td>
<td>323 (9.9)</td>
<td>73 (8.4)</td>
<td>-0.189 (0.136)</td>
<td>0.001</td>
<td>0.156</td>
</tr>
<tr>
<td>Liver</td>
<td>34 (1.0)</td>
<td>10 (1.1)</td>
<td>0.092 (0.362)</td>
<td>&lt; 0.001</td>
<td>0.800</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>11 (0.3)</td>
<td>1 (0.1)</td>
<td>-1.085 (1.045)</td>
<td>&lt; 0.001</td>
<td>0.228</td>
</tr>
<tr>
<td>Metastatic disease</td>
<td>11 (0.3)</td>
<td>8 (0.9)</td>
<td>1.003 (0.466)</td>
<td>0.001</td>
<td>0.040</td>
</tr>
<tr>
<td>Neurologic (other)</td>
<td>82 (2.5)</td>
<td>50 (5.7)</td>
<td>0.854 (0.184)</td>
<td>0.005</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Obesity</td>
<td>326 (10.0)</td>
<td>104 (11.9)</td>
<td>0.194 (0.120)</td>
<td>0.001</td>
<td>0.110</td>
</tr>
<tr>
<td>Paralysis</td>
<td>60 (1.8)</td>
<td>47 (5.4)</td>
<td>1.108 (0.199)</td>
<td>0.007</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>PHTN</td>
<td>7 (0.2)</td>
<td>10 (1.1)</td>
<td>1.681 (0.494)</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>Psychoses</td>
<td>40 (1.2)</td>
<td>14 (1.6)</td>
<td>0.269 (0.313)</td>
<td>&lt; 0.001</td>
<td>0.400</td>
</tr>
<tr>
<td>Pulmonary</td>
<td>572 (17.6)</td>
<td>151 (17.3)</td>
<td>-0.020 (0.101)</td>
<td>&lt; 0.001</td>
<td>0.841</td>
</tr>
<tr>
<td>PVD</td>
<td>65 (2.0)</td>
<td>34 (3.9)</td>
<td>0.687 (0.215)</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Renal</td>
<td>106 (3.3)</td>
<td>71 (8.1)</td>
<td>0.966 (0.158)</td>
<td>0.008</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Rheumatic</td>
<td>175 (5.4)</td>
<td>39 (4.5)</td>
<td>-0.196 (0.181)</td>
<td>&lt; 0.001</td>
<td>0.272</td>
</tr>
<tr>
<td>Tobacco</td>
<td>775 (23.8)</td>
<td>160 (18.3)</td>
<td>-0.332 (0.097)</td>
<td>0.003</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Tumor</td>
<td>114 (3.5)</td>
<td>26 (3.0)</td>
<td>-0.168 (0.221)</td>
<td>&lt; 0.001</td>
<td>0.438</td>
</tr>
<tr>
<td>Valvular</td>
<td>66 (2.0)</td>
<td>24 (2.7)</td>
<td>0.311 (0.242)</td>
<td>&lt; 0.001</td>
<td>0.209</td>
</tr>
<tr>
<td>Weight loss</td>
<td>8 (0.2)</td>
<td>8 (0.9)</td>
<td>1.322 (0.502)</td>
<td>0.002</td>
<td>0.011</td>
</tr>
<tr>
<td><strong>Medical characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transferred to SNF</td>
<td>54 (1.7)</td>
<td>555 (63.5)</td>
<td>4.636 (0.154)</td>
<td>0.412</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Diagnoses (n)</td>
<td>9.9 ± 4.7</td>
<td>11 ± 5.7</td>
<td>0.045 (0.007)</td>
<td>0.009</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Length of stay (nights)</td>
<td>4.1 ± 4.8</td>
<td>6.5 ± 5.5</td>
<td>0.084 (0.008)</td>
<td>0.031</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prior admissions (n)</td>
<td>0.4 ± 0.7</td>
<td>0.8 ± 1.1</td>
<td>0.520 (0.045)</td>
<td>0.033</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prior ED visits (n)</td>
<td>0.5 ± 1.8</td>
<td>0.9 ± 1.8</td>
<td>0.098 (0.018)</td>
<td>0.007</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Prior major operations (n)</td>
<td>0.3 ± 0.6</td>
<td>0.3 ± 0.6</td>
<td>0.066 (0.064)</td>
<td>&lt; 0.001</td>
<td>0.307</td>
</tr>
<tr>
<td>Prior outpatient visits (n)</td>
<td>17.3 ± 11.4</td>
<td>21.2 ± 14.9</td>
<td>0.022 (0.003)</td>
<td>0.014</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Primary diagnosis</td>
<td></td>
<td></td>
<td></td>
<td>0.012</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Stenosis</td>
<td>1,135 (34.9)</td>
<td>341 (39.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spondylolisthesis</td>
<td>997 (30.6)</td>
<td>305 (34.9)</td>
<td>0.018 (0.090)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Not readmitted</th>
<th>Readmitted</th>
<th>Coefficient (SE)</th>
<th>Pseudo R²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDD</td>
<td>588 (18.1)</td>
<td>118 (13.5)</td>
<td>-0.404 (0.118)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Displacement</td>
<td>206 (6.3)</td>
<td>63 (7.2)</td>
<td>0.018 (0.157)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spondylosis</td>
<td>130 (4.0)</td>
<td>9 (1.0)</td>
<td>-1.468 (0.350)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>88 (2.7)</td>
<td>9 (1.0)</td>
<td>-1.078 (0.355)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiculopathy</td>
<td>50 (1.5)</td>
<td>15 (1.7)</td>
<td>-0.002 (0.301)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>33 (1.0)</td>
<td>8 (0.9)</td>
<td>-0.215 (0.399)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deformity</td>
<td>23 (0.7)</td>
<td>5 (0.6)</td>
<td>-0.324 (0.497)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postlaminectomy syndrome</td>
<td>4 (0.1)</td>
<td>1 (0.1)</td>
<td>-0.184 (1.120)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spondylolysis</td>
<td>2 (0.1)</td>
<td>0 (0)</td>
<td>-11.364 (229.629)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socioeconomic characteristics</td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
<td>0.018</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 12th grade</td>
<td>16 (0.5)</td>
<td>3 (0.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school diploma</td>
<td>1,028 (31.6)</td>
<td>323 (37.0)</td>
<td>0.074 (0.430)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>1,803 (55.4)</td>
<td>456 (52.2)</td>
<td>-0.317 (0.323)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>409 (12.6)</td>
<td>92 (10.5)</td>
<td>0.186 (0.153)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household size</td>
<td></td>
<td></td>
<td></td>
<td>0.063</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>1</td>
<td>1,637 (50.3)</td>
<td>698 (79.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>870 (26.7)</td>
<td>107 (12.2)</td>
<td>-1.128 (0.328)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>290 (8.9)</td>
<td>26 (3.0)</td>
<td>0.621 (0.313)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>269 (8.3)</td>
<td>23 (2.6)</td>
<td>-0.698 (0.259)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>128 (3.9)</td>
<td>16 (1.8)</td>
<td>-0.153 (0.220)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6+</td>
<td>62 (1.9)</td>
<td>4 (0.5)</td>
<td>-0.154 (0.211)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household income (USD)</td>
<td></td>
<td></td>
<td></td>
<td>0.022</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>&lt;40,000</td>
<td>781 (24.0)</td>
<td>341 (39.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40,000–49,999</td>
<td>260 (8.0)</td>
<td>57 (6.5)</td>
<td>-0.591 (0.092)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50,000–59,999</td>
<td>255 (7.8)</td>
<td>84 (9.6)</td>
<td>0.05 (0.095)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60,000–74,999</td>
<td>422 (13.0)</td>
<td>99 (11.3)</td>
<td>-0.229 (0.112)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75,000–99,999</td>
<td>571 (17.5)</td>
<td>123 (14.1)</td>
<td>0.278 (0.120)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;100,000</td>
<td>967 (29.7)</td>
<td>170 (19.5)</td>
<td>-0.266 (0.120)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td>White</td>
<td>2,579 (79.2)</td>
<td>628 (71.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>406 (12.5)</td>
<td>148 (16.9)</td>
<td>0.404 (0.106)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>225 (6.9)</td>
<td>79 (9.0)</td>
<td>0.366 (0.138)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>46 (1.4)</td>
<td>19 (2.2)</td>
<td>0.528 (0.276)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microscope</td>
<td>202 (6.2)</td>
<td>28 (3.2)</td>
<td>-0.692 (0.205)</td>
<td>0.003</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Navigation</td>
<td>210 (6.4)</td>
<td>59 (6.8)</td>
<td>0.049 (0.153)</td>
<td>&lt; 0.001</td>
<td>0.750</td>
</tr>
<tr>
<td>rhBMP</td>
<td>251 (7.7)</td>
<td>52 (5.9)</td>
<td>-0.278 (0.157)</td>
<td>0.001</td>
<td>0.079</td>
</tr>
<tr>
<td>Allograft</td>
<td>1,034 (31.8)</td>
<td>263 (30.1)</td>
<td>-0.078 (0.083)</td>
<td>&lt; 0.001</td>
<td>0.350</td>
</tr>
<tr>
<td>Assistant surgeon</td>
<td>1,247 (38.3)</td>
<td>152 (17.4)</td>
<td>-1.081 (0.096)</td>
<td>0.034</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Increased procedural service</td>
<td>27 (0.8)</td>
<td>9 (1.0)</td>
<td>0.219 (0.387)</td>
<td>&lt; 0.001</td>
<td>0.579</td>
</tr>
</tbody>
</table>

(Continued)
## Table 1. Univariate logistic regression (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Not readmitted</th>
<th>Readmitted</th>
<th>Coefficient (SE)</th>
<th>Pseudo R²</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>452 (13.9)</td>
<td>67 (7.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3–6</td>
<td>2,493 (76.6)</td>
<td>670 (76.7)</td>
<td>1.384 (0.370)</td>
<td></td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>7–12</td>
<td>304 (9.3)</td>
<td>130 (14.9)</td>
<td>0.127 (0.281)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13+</td>
<td>7 (0.2)</td>
<td>7 (0.8)</td>
<td>0.115 (0.145)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple procedures</td>
<td>409 (12.6)</td>
<td>194 (22.2)</td>
<td>0.686 (0.097)</td>
<td>0.011</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Charge (USD)</td>
<td>121,618.6 ± 129,387.3</td>
<td>173,765.3 ± 248,002.8</td>
<td>0 (0)</td>
<td>0.015</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Complications (n)</td>
<td>0.3 ± 0.8</td>
<td>0.4 ± 0.8</td>
<td>0.041 (0.047)</td>
<td>&lt; 0.001</td>
<td>0.396</td>
</tr>
<tr>
<td>CPT codes (n)</td>
<td>11.6 ± 4.0</td>
<td>12.2 ± 4.0</td>
<td>0.040 (0.009)</td>
<td>0.004</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Return to OR</td>
<td>101 (3.1)</td>
<td>39 (4.5)</td>
<td>0.378 (0.193)</td>
<td>0.001</td>
<td>0.056</td>
</tr>
<tr>
<td>Year of surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>50 (1.5)</td>
<td>4 (0.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>64 (2.0)</td>
<td>8 (0.9)</td>
<td>1.958 (0.304)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>119 (3.7)</td>
<td>8 (0.9)</td>
<td>-0.311 (0.291)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>124 (3.8)</td>
<td>13 (1.5)</td>
<td>-0.321 (0.287)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>191 (5.9)</td>
<td>32 (3.7)</td>
<td>0.134 (0.280)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>247 (7.6)</td>
<td>40 (4.6)</td>
<td>0.042 (0.259)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>255 (7.8)</td>
<td>55 (6.3)</td>
<td>-0.054 (0.244)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>272 (8.4)</td>
<td>63 (7.2)</td>
<td>0.395 (0.239)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>264 (8.1)</td>
<td>68 (7.8)</td>
<td>-0.312 (0.231)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>320 (9.8)</td>
<td>104 (11.9)</td>
<td>0.193 (0.226)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>271 (8.3)</td>
<td>102 (11.7)</td>
<td>-0.175 (0.214)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>273 (8.4)</td>
<td>95 (10.9)</td>
<td>-0.017 (0.188)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>395 (12.1)</td>
<td>130 (14.9)</td>
<td>0.115 (0.166)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>411 (12.6)</td>
<td>152 (17.4)</td>
<td>-0.077 (0.151)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute renal failure</td>
<td>20 (0.6)</td>
<td>15 (1.7)</td>
<td>1.039 (0.344)</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>Orthopedic implant</td>
<td>299 (9.2)</td>
<td>36 (4.1)</td>
<td>-0.856 (0.181)</td>
<td>0.006</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>24 (0.7)</td>
<td>4 (0.5)</td>
<td>-0.479 (0.541)</td>
<td>&lt; 0.001</td>
<td>0.349</td>
</tr>
<tr>
<td>Dural tear</td>
<td>23 (0.7)</td>
<td>10 (1.1)</td>
<td>0.487 (0.381)</td>
<td>&lt; 0.001</td>
<td>0.217</td>
</tr>
<tr>
<td>Neurologic</td>
<td>21 (0.6)</td>
<td>6 (0.7)</td>
<td>0.063 (0.465)</td>
<td>&lt; 0.001</td>
<td>0.893</td>
</tr>
<tr>
<td>Pulmonary</td>
<td>92 (2.8)</td>
<td>41 (4.7)</td>
<td>0.526 (0.192)</td>
<td>0.002</td>
<td>0.008</td>
</tr>
<tr>
<td>Wound (other)</td>
<td>6 (0.2)</td>
<td>3 (0.3)</td>
<td>0.624 (0.708)</td>
<td>&lt; 0.001</td>
<td>0.398</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>11 (0.3)</td>
<td>5 (0.6)</td>
<td>0.529 (0.541)</td>
<td>&lt; 0.001</td>
<td>0.346</td>
</tr>
<tr>
<td>Sepsis or septicemia</td>
<td>0 (0)</td>
<td>3 (0.3)</td>
<td>13.885 (187.491)</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Thrombotic</td>
<td>9 (0.3)</td>
<td>3 (0.3)</td>
<td>0.217 (0.668)</td>
<td>&lt; 0.001</td>
<td>0.750</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>32 (1.0)</td>
<td>26 (3.0)</td>
<td>1.128 (0.267)</td>
<td>0.004</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Wound hematoma/hemorrhage</td>
<td>14 (0.4)</td>
<td>12 (1.4)</td>
<td>1.171 (0.395)</td>
<td>0.002</td>
<td>0.004</td>
</tr>
<tr>
<td>Any</td>
<td>501 (15.4)</td>
<td>138 (15.8)</td>
<td>0.031 (0.105)</td>
<td>&lt; 0.001</td>
<td>0.771</td>
</tr>
</tbody>
</table>

Values are presented as number (%) or mean ± standard deviation unless otherwise indicated. SE, standard error; POS, place of service; HMO, health maintenance organization; PPO, preferred provider organization; EPO, exclusive provider network; CHF, congestive heart failure; DM, diabetes mellitus; DMcx, complicated diabetes mellitus; HTN, hypertension; HTNcx, complicated hypertension; PHTN, pulmonary hypertension; PVD, peripheral vascular disease; SNF, skilled nursing facility; ED, Emergency Department; DDD, degenerative disc disease; USD, United States dollar; rhBMP, recombinant human bone morphogenetic protein; CPT, current procedural terminology; OR, operating room.

https://doi.org/10.14245/ns.2347340.670 www.e-neurospine.org
4. Predictive Modeling

Five models were generated and evaluated: a multivariate LR, a penalized LR model chosen based on elastic net variants of the LASSO, RF, stochastic GBM, and XGBoost. The evaluation of these models was based on AUC, sensitivity, and specificity. The parameters relevant to the prediction task were identified from the variable importance scores, which were calculated for each model and used to improve its interpretability. For model generation and tuning details, refer to the Supplementary Text and Table 1.

5. Statistical Analysis

R ver. 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria) was used to perform all statistical analyses. To determine the risk factors associated with readmissions (Table 1), univariate LR models were generated. To determine the strength of each variable on readmission outcomes, the McFadden Pseudo $R^2$ was computed, and the model performance metrics (i.e., AUC, specificity, and sensitivity) were measured using the “caret” statistical package. For all experiments, statistical significance was defined as $p < 0.05$.

### RESULTS

Of the 4,130 patients analyzed in this study, 874 (21.2%) were readmitted within 30 days following surgery. Demographic, medical, socioeconomic, and surgical characteristics were analyzed (Table 1) for all patients who met the selection criteria (Fig. 1). Notably, a univariate LR model identified patient age, insurance plan type, and Medicare use as highly associated with readmissions ($p < 0.001$). Other variables associated with readmissions were household size and income, presence of an assistant surgeon, number of diagnoses, and length of stay ($p < 0.001$). Readmitted patients were, on average, older (69.1 years vs. 60.6 years), more likely to have Medicare (81.1% vs. 43.3%), and experienced greater lengths of stay (5.5% vs. 4.8%, $p < 0.001$). These patients also had greater number of diagnoses (5.7% vs. 4.7%), were significantly more likely to be transferred to a skilled nursing facility (SNF) (63.5% vs. 1.7%), and were more likely to experience postoperative urinary complications (3% vs. 1%, $p < 0.001$).

Based on the risk factors the univariate LR model associated with readmissions, 5 models were generated and evaluated (GBM, XGBoost, RF, LASSO, and MLR). The performance metrics of these 5 models were measured by using a 50% random sample of the data for training and the remaining 50% for testing (Table 2). GBM outperformed all other models, with an AUC of over 15 independent runs (mean ± standard deviation, 0.844 ± 0.015). Based on these findings, new versions of the GBM (top-performing) and MLR (bottom-performing) models were generated. However, this time, instead of using a 50% random sample of data, all the data from 2014–2016 was used to train the models, and all 2017 cohort data was used to test their performance metrics (Table 3). When comparing the AUC mean values (0.846 vs. 0.829, $p < 0.001$), the GBM model significantly outperformed the MLR model. The specificity of the GBM model was also superior to that of the MLR model (0.986 vs. 0.966, $p < 0.001$). The 3 predictors with the greatest relative

---

### Table 2. Performance metrics of models generated using 50% train and 50% test data splits

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBM</td>
<td>0.844 ± 0.015</td>
</tr>
<tr>
<td>XGBoost</td>
<td>0.842 ± 0.016</td>
</tr>
<tr>
<td>RF</td>
<td>0.843 ± 0.015</td>
</tr>
<tr>
<td>LASSO</td>
<td>0.833 ± 0.021</td>
</tr>
<tr>
<td>LR</td>
<td>0.819 ± 0.021</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.

### Table 3. Performance metrics of GBM and LR models generated using data from 2004–2016 for training and data from 2017 for testing

<table>
<thead>
<tr>
<th>Variable</th>
<th>GBM</th>
<th>LR</th>
<th>Difference</th>
<th>95% CI</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.846 ± 0.016</td>
<td>0.829 ± 0.018</td>
<td>0.0170</td>
<td>0.012–0.022</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.610 ± 0.033</td>
<td>0.625 ± 0.037</td>
<td>0.0152</td>
<td>0.021–0.010</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.986 ± 0.002</td>
<td>0.966 ± 0.009</td>
<td>0.0195</td>
<td>0.014–0.025</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation unless otherwise indicated.

GBM, Gradient Boosting Machine; LR, traditional logistic regression; CI, confidence interval; AUC, area under the receiver operating characteristic curve.
influence on the GBM model were the discharge status of the patients, the total billed charge of admission, and the length of stay (Fig. 2). Additional factors associated with increased readmissions included the patient’s age, number of prior outpatient visits, as well as the number of procedural and diagnosis codes for initial admission (Fig. 2).

To determine the outcomes of applying interventions to the top 25% of patients with the highest probability of readmission as determined by the GBM model, we subsequently used the GBM model on 2017 model-naive, inflation-adjusted, all-payer national estimates of 30-day surgical readmissions (Table 4). In the 2017 data, we analyzed 490 patients admitted between January and November, of which 133 (27.2%) had been readmitted within 30 days. Of these, the GBM model flagged 105 patients as constituting the top 25% of patients with the highest readmission probability, 99 of which were accurately identified by the model as having been readmitted, giving a true positive rate of 94% (105 top 25% high-risk patients/99 truly readmitted). Out of the remaining 385 patients who were not in the top 25th percentile of high readmission likelihood, 35 patients were nonetheless eventually readmitted, resulting in a missed patient rate of 9% (35 readmitted patients/385 unlikely readmissions). The GBM model predicted an estimated total costs savings of $803,633 over 11 months as a result of reduced readmissions. These cost savings were calculated assuming that 50% of readmissions were prevented by the interventions.

The LACE index was similarly analyzed: in this case, 62 patients out of the 105 total patients (top 25%) with the highest probability of readmission were correctly identified as having been readmitted (true positive rate of 59%). Out of a total of 385 patients that were not included in the top 25th percentile of high readmission likelihood, 73 patients were eventually readmitted (missed patient rate of 19%). The estimated total cost savings associated with reduced readmissions was $535,755 over an 11-month span. Together, this data shows that the GBM model outperformed the LACE index model when comparing the true positive rates (94% vs. 59%), missed patient rates (9% vs. 19%), and the cost savings ($803,633 vs. $535,755). In fact, the GBM model estimated a 50% decrease in readmission-associated costs when compared to those achieved by the LACE index model (Table 4).

DISCUSSION

In order to develop interventions that reduce readmission rates, it is critical to accurately identify patients who are at high
risk of being readmitted. One strategy is to identify high-risk patients by analyzing the independent risk factors that are associated with readmission probabilities. In this study, we used epidemiological and supervised ML algorithms to analyze 4,130 patients undergoing PCF. We identified the demographic, socioeconomic, clinical, and procedural characteristics associated with patient readmissions within 30 days.

Univariate LR analysis found that patients’ age, their Medicare usage, their insurance plan type, the number of diagnoses, and length of hospital stay were all variables that influenced the readmission rates (Table 1). Interestingly, patients’ discharge to the SNF was strongly associated with readmissions in both multivariate and univariate LR models. Previous studies have identified similarly significant associations between transfer to SNF and readmission rates; in fact, as many as one in 4 SNF patients experience re-hospitalization within 30 days from their initial admission.\(^7,21,22\)

While LR models are commonly used to study and predict unplanned readmission rates, other ML models with the resolution to capture interactions between factors have become popular tools to predict patient outcomes and readmissions.\(^{11,13}\) A growing body of medical literature has probed the potential of these models in supporting clinical decision-making and implementation.\(^6,10\) Here, we used supervised classification and regression ML algorithms to predict readmissions and identify the risk factors that influence these rates. We found that while GBM identified patients’ discharge status, charge of admission, and length of stay as the most influential predictors of readmissions, MLR identified patients with comorbidities and number of procedure codes as the relevant variables. An explanation for this finding is that LR models make linear predictions on readmissions by computing principled estimates of confidence intervals, while other ML algorithms (i.e., GBM, LASSO, etc.), capture interactions between risk factors and nonlinear relationships. The observed differences in predictor weighting between models shows that, depending on which model is emphasized, it is possible to overestimate or underestimate the relevance of readmission predictors. Thus, by leveraging different ML models, it is possible to capture more realistic linear and nonlinear relationships.

Next, we tested the performance of the ML models, specifically, their ability in predicting 30-day readmissions. Compared to previously reported ML models that predicted readmissions following cervical spine surgery with an AUC mean of 0.63–0.81, our GBM model achieved a mean AUC value of 0.865—the highest predictability performance recorded to date.\(^9,10,23-28\) One explanation for this improved performance is that the variables chosen for the analysis and consequentially, the model’s relative weighting of these variables is unique to this study. For instance, the most influential variables for the GBM model

| Table 4. Predictive performance of GBM and LACE index on 2017 test data and associated cost savings |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Month           | Top 25% highest risk patients (n) | True positive rate on top 25% risk patients | Bottom 75% risk patients (n) | Missed patients rate | Estimated cost savings* using GBM (USD) | Average cost savings using GBM over LACE |
|                 | GBM | LACE | GBM | LACE | GBM | LACE | GBM | LACE |
| January         | 12  | 1.00 | 0.60 | 38   | 0.03 | 0.16 | 97,410 | 100% |
| February        | 5   | 1.00 | 0.60 | 18   | 0.17 | 0.26 | 40,588 | 67%  |
| March           | 12  | 1.00 | 0.70 | 36   | 0.22 | 0.35 | 97,410 | 71%  |
| April           | 11  | 1.00 | 0.80 | 36   | 0.17 | 0.26 | 89,293 | 38%  |
| May             | 4   | 1.00 | 0.38 | 27   | 0.04 | 0.11 | 32,470 | 33%  |
| June            | 5   | 1.00 | 0.83 | 21   | 0.14 | 0.24 | 40,588 | 0%   |
| July            | 7   | 1.00 | 0.67 | 33   | 0.09 | 0.13 | 56,823 | 17%  |
| August          | 11  | 1.00 | 0.64 | 39   | 0.10 | 0.24 | 89,293 | 57%  |
| September       | 6   | 1.00 | 0.50 | 39   | 0.05 | 0.09 | 48,705 | 20%  |
| October         | 13  | 0.85 | 0.71 | 48   | 0.15 | 0.16 | 89,293 | 10%  |
| November        | 12  | 0.83 | 0.43 | 52   | 0.06 | 0.14 | 81,175 | 67%  |
| All             | 105 | 0.94 | 0.59 | 385  | 0.09 | 0.19 | 803,633 | 50%  |

GBM, Gradient Boosting Machine; LACE, Length patient stay in the hospital, Acuity of admission of patient in the hospital, Comorbidity, and Emergency visit; USD, United States dollar.

*Cost savings were calculated under the assumption that 50% of readmissions were prevented via targeted interventions for patients correctly identified by the model as likely for readmission.
included the discharge destination (i.e., SNF), total charge billed for admission, length of stay, and patient age. To the best of our knowledge, the patient’s discharge destination has been analyzed as a model predictor for 30-day readmission by only one lumbar spine study, which similarly found that the discharge destination was the most influential variable accounted by the model.\(^7\)

Next, we used the GBM model to determine the top 25% of patients with the greatest probability of having unplanned readmissions. We then simulated the clinical outcomes of implementing an intervention for these flagged patients. We chose the 25% threshold value to account for hospital systems’ varying capacities to apply interventions for high-risk patients. In practice, however, this threshold should be tuned to the individual capacities and resources of different hospital systems. Of all the patients flagged as having high readmission probabilities, 94% were accurately predicted by the GBM model. The bottom 75% of patients with a high risk for readmissions totaled 385 patients. Of these, only 25 patients were eventually readmitted, accounting for a missed patient rate of 9%. When interventions were simulated for the top 25%, the GBM model presented in this study predicted an estimated cost savings of $803,633 over an 11-month period. It is important to note that this estimated cost savings was computed under the assumption that effective intervention(s) led to half of the high-risk patients not being readmitted.

The findings presented here demonstrate that ML models can identify patients with a high risk of readmission and provide targeted interventions that reduce these patient’s probabilities of being readmitted. While certain identified risk factors, such as age are inherently non-modifiable, the proposed interventions are designed to mitigate the risks associated with modifiable factors of patient’s postoperative care. For instance, while we cannot alter a patient’s age, hospital programs can proactively target those discharged to skilled nursing facilities—a predictor identified by the GBM model for readmissions—by increasing follow-up calls, home visits, and telemonitoring practices, all of which have been shown to reduce readmissions.\(^36,39\) This ensures sustained care and strict adherence to postdischarge protocols, effectively mitigating the risk of readmission. Such measures would be particularly important for complex or invasive procedures, which are often associated with higher billed charges of admission, greater number of procedural and diagnostic codes at outset, and higher number of outpatient visits—all of which are factors that our GBM model identified as predictors of greater patient readmissions. Similarly, patient education is also important in the context of complicated diagnoses and procedures. Empirical evidence suggests that communication interventions at the point of discharge, including medication counseling and disease-specific education, can significantly reduce the likelihood of 30-day readmissions.\(^32\) By integrating structured educational programs that specifically target patients with the highest risks of readmissions, as defined by the GBM model, we can equip patients with the knowledge to manage their conditions more effectively, recognize early signs of complications, and understand when to seek medical help. Last, medication management can address the risks associated with polypharmacy and complex medication regimens often seen in patients with multiple codes for procedures and diagnoses. A pharmacist-led approach, encompassing medication reconciliation, a patient-specific medication care plan, discharge counseling, and follow-up contact, can substantially decrease the incidence of medication errors postdischarge, thereby lessening the chances of readmission or emergency department visits.\(^33\) By using ML models to identify patients at high risk of readmission and targeting these patients specifically, resources may be optimally allocated while contributing to a reduction in the costs associated with readmitting these high-risk patients. Nonetheless, future research could benefit from a closer examination of the direct impact of these interventions on readmission rates.

Both our cost savings simulation and previously reported data provide evidence that reducing the number of readmissions, even slightly, can significantly lower the economic burden of unplanned readmissions.\(^34\) With further fine-tuning and customization, these models could (1) aid clinicians in identifying patients with high readmission risks; (2) guide perioperative resource allocation to decrease readmission probabilities; and (3) decrease overall healthcare costs.

In addition to comparing the GBM model to univariate and multivariate LR models, we also analyzed how the performance of our ML algorithm compared to that of clinically-employed predictive models of readmission (i.e., the LACE index model). When we analyzed the top 25% of patients with a high risk for unplanned readmissions, we found that in comparison to the LACE index model, the GBM had a higher true positive rate of readmission, a 50% higher cost savings from readmission prevention, and a lower missed patient rate. Several factors may explain GBM’s higher performance when compared to the LACE index model. For instance, we provided the ML model with an expanded set of predictors to utilize and GBM has an adaptive learning framework that can capture and leverage interactions between variables, including nonlinear relationships.
Currently, there is no consensus on the performance power of the LACE index; while some reports strongly support the model’s ability to discriminate between variables,35–37 others cite the lack of strong predictors38 as reason for the model’s moderate to poor discrimination.38-41

Our study generated models that capture patients at high risk of readmissions and simulates how hypothetical clinical interventions can reduce healthcare utilization costs. However, there are several limitations worth noting when interpreting these results. First, the insurance claims database used was not specific to spine surgery, which can impact accuracy.42,43 Second, we must understand the limitations of using administrative claims data to guide and alter clinical practice.42,43 For example, during the study period, insurers and hospitals attempted to decrease patient’s length of stay by discharging to home. This creates confounds in the associations captured between readmission probabilities and patient’s length of stay. This creates confounds in the associations captured between readmission probabilities and patient’s length of stay. Third, to simulate how intervening of high-risk patients would impact healthcare utilization, we assumed that 50% of the theoretical interventions were effective, which could vary by hospital systems.

These limitations highlight substantial avenues for improvement. Although our study represents an initial effort to utilize ML models to predict readmission rates and mitigate healthcare costs, future prospective studies and clinical trials are needed to refine, validate, and enhance the ML algorithms presented. Moreover, exploring the utilization of alternative data sources beyond administrative claims (i.e., prospective datasets) and predicting 90 days in addition to 30-day readmission rates, could further enhance the robustness and generalizability of our findings.

CONCLUSION

ML models, including MLR, have identified different risk factors associated with patients’ unplanned readmission probabilities. This is potentially a result of each model’s capacity to measure nonlinear relationships and interfactor interactions, and it suggests that models can complement each other when capturing risk predictors for readmissions following PCF. In addition, we found that when comparing 2 ML models, GBM outperformed the MLR model as measured by the mean AUC. These findings support the rationale for the continued generation, improvement, and eventual implementation of ML models in order to reduce readmissions and associated healthcare utilization costs.

NOTES

Supplementary Material: Supplementary Text and Table 1 can be found via https://doi.org/10.14245/ns.2347340.670.

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: ADGS, PGR, DH, JKR, MR, DS, AMD; Data curation: PGR; Formal analysis: PGR; Methodology: PGR, DH, JKR, MR, DS, AMD; Project administration: ADGS, JKR, MR, DS, IJ, AMD; Visualization: ADGS; Writing – original draft: ADGS; Writing – review & editing: ADGS, PGR, SST, IJ, AMD.

ORCID

Aneysis D. Gonzalez-Suarez: 0000-0003-3745-763X
Paymon G. Rezaii: 0000-0002-4803-0853
Daniel Herrick: 0000-0003-1066-4751
Seth Stravers Tigchelaar: 0000-0002-2207-9034
John K. Ratliff: 0000-0003-3452-1907
Mirabela Rusu: 0000-0002-0372-0367
David Scheinker: 0000-0001-5885-8024
Ikchan Jeon: 0000-0002-1293-2724
Atman M. Desai: 0000-0001-8387-3808

REFERENCES


Five different machine learning (ML) models were created and evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity. The models used were logistic regression (LR), penalized LR (chosen through elastic net variants of least absolute shrinkage and selection operator [LASSO]), random forest (RF), stochastic gradient boosting machine (GBM), and extreme gradient boosting (XGBoost). Patients who had missing data for any predictor were excluded from training and testing of all models, which resulted in 3,215 omitted patients. Categorical predictors were transformed into binary format using one-hot encoding.

The data used in the study were divided into 2 sets using a random partitioning technique. The training set consisted of 50% of the data, while the remaining 50% made up the test set. The partitioning was done in a way that ensured both sets had almost equal numbers of patients who were readmitted and those who were not. The training set was used to estimate model parameters and fine-tune the models, while the test set was exclusively used to validate the performance of the models.

To ensure optimal performance, hyperparameters of each model, including RF, GBM, XGBoost, and LASSO, were fine-tuned using a 5-fold cross-validation technique that was repeated 3 times. This involved creating a grid of possible parameter values, with each column representing a specific parameter and each row representing a unique set of parameter values. The training data was then divided into five equal-sized folds, each containing a proportion of readmitted patients similar to the entire training set. The model was trained on four folds and tested on the fifth fold, which was held out to estimate the performance measure (AUC). This process was repeated for each fold, and the average of the five resampled estimates was used as a single 5-fold cross-validation estimate of model performance. The cross-validation procedure was repeated three times to increase precision while maintaining low bias, and the final estimate of model performance was generated by averaging the performance estimates of all three instances of 5-fold cross-validation. The parameter set that produced the best performance estimate was used to define the final tuned model. Further details on the parameter tuning for each model can be found in the Supplementary Table 1.

In each resampling iteration of the 5-fold cross-validation process, a series of preprocessing steps were applied. First, near-
zero variance predictors and highly correlated predictors were eliminated to enhance the subsequent model generation process. Predictors were considered near-zero variance if they met two conditions: (1) they had less than 10% of the total number of samples as the number of distinct values, and (2) the ratio of the frequency of the most common value to the frequency of the second most common value was greater than 19:1. Predictors with pair-wise correlations of 0.9 or higher were deemed highly correlated, and the correlated predictor with the largest mean absolute correlation was removed. Second, predictors were normalized to a mean of zero and a variance of one. Finally, the SMOTE (Synthetic Minority Over-sampling Technique) was used to help optimize the model and specifically, the imbalance between the proportion of readmitted and non-readmitted patients.

The tuned models were compared based on their mean AUCs, and the GBM model was found to have the largest mean AUC. Two additional models were then generated, LR and GBM, using a different training and testing set. This new training set included data from January 1, 2004 up to December 31, 2016, while the test set included data from January 1, 2017 to November 30, 2017. The GBM model was then used to evaluate its applicability in clinical practice. Specifically, the model was used to identify the top 25% of patients at the highest risk of readmission each month, and the prediction accuracy of the model was assessed. The model was evaluated by counting the true positives and calculating the cost savings associated with reducing readmissions, assuming that 50% of interventions on these patients prevented readmission. The same clinical scenario was also applied using the LACE index model, a previously-validated readmission model that uses four variables to predict unplanned 30-day readmission after hospital discharge: length of stay (L), acuity of the admission (A), comorbidity of the patient (C), and emergency department use in the duration of 6 months before admission (E). The cost savings were compared between the LACE index model and the top-performing GBM model.

SUPPLEMENTARY REFERENCES
Original Article

Analyzing Large Language Models’ Responses to Common Lumbar Spine Fusion Surgery Questions: A Comparison Between ChatGPT and Bard

Objective: In the digital age, patients turn to online sources for lumbar spine fusion information, necessitating a careful study of large language models (LLMs) like chat generative pre-trained transformer (ChatGPT) for patient education.

Methods: Our study aims to assess the response quality of Open AI (artificial intelligence)’s ChatGPT 3.5 and Google’s Bard to patient questions on lumbar spine fusion surgery. We identified 10 critical questions from 158 frequently asked ones via Google search, which were then presented to both chatbots. Five blinded spine surgeons rated the responses on a 4-point scale from ‘unsatisfactory’ to ‘excellent.’ The clarity and professionalism of the answers were also evaluated using a 5-point Likert scale.

Results: In our evaluation of 10 questions across ChatGPT 3.5 and Bard, 97% of responses were rated as excellent or satisfactory. Specifically, ChatGPT had 62% excellent and 32% minimally clarifying responses, with only 6% needing moderate or substantial clarification. Bard’s responses were 66% excellent and 24% minimally clarifying, with 10% requiring more clarification. No significant difference was found in the overall rating distribution between the 2 models. Both struggled with 3 specific questions regarding surgical risks, success rates, and selection of surgical approaches (Q3, Q4, and Q5). Interrater reliability was low for both models (ChatGPT: $k = 0.041$, $p = 0.622$; Bard: $k = 0.040$, $p = 0.601$). While both scored well on understanding and empathy, Bard received marginally lower ratings in empathy and professionalism.

Conclusion: ChatGPT3.5 and Bard effectively answered lumbar spine fusion FAQs, but further training and research are needed to solidify LLMs’ role in medical education and healthcare communication.

Keywords: Artificial intelligence, Large language models, Patient education, Lumbar spine fusion, ChatGPT, Bard
INTRODUCTION

Lumbar spine fusion surgery, a pivotal procedure in addressing diverse spinal pathologies, has evolved remarkably over recent years and is one of the most frequently performed neurosurgical procedures worldwide.1 Due to multiple pathologies, presentations, and surgical approaches, patients may often find it daunting to understand the intricacies of lumbar fusion surgery, including its potential risks, benefits, and postoperative trajectories.2

In today's digital era, a substantial number of patients turn to online platforms for surgical information.3 Encouraging patients to access treatment-related online health information can enhance patient compliance and the patient-physician relationship, while also enabling physicians to stay updated on emerging treatments.4 However, the expansive digital domain occasionally presents conflicting, obsolete, or excessively technical data, potentially exacerbating patients’ decision-making conundrum.4 It is critical to understand that the quality of online information and its integration into medical consultations impacts patient care and patient-physician communication.5 This highlights the need for accurate, accessible, and patient-centric online educational tools.

Artificial intelligence (AI) is transforming medicine in many ways, and the advent of large language models (LLMs) such as OpenAI’s chat generative pre-trained transformer (ChatGPT) offers a transformative approach to patient education.6 Leveraging their capability to sift through immense data and produce human-like narratives, LLMs can produce comprehensive, succinct, and individualized information to patients.7 However, ChatGPT’s real-world performance in complex fields like medicine and spine surgery still remains to be seen. Information obtained via LLMs may enable patients to better understand their disease process and treatment options, potentially improving the transparency and trust in the surgical decision-making process.8 But there is also considerable risk to these models, including the potential for inaccurate or biased information that can mislead patients and negatively impact their health. The goal of this study is therefore to evaluate the accuracy, clarity, and comprehensiveness of OpenAI’s ChatGPT 3.5 and Google’s Bard on 10 frequently asked patient questions regarding lumbar fusion surgery.

MATERIALS AND METHODS

A comprehensive Google search was conducted using the search terms “frequently asked questions AND lumbar spine surgery OR lumbar fusion surgery,” yielding approximately 4,610,000 results within 0.51 seconds (September 7th, 2023; region: Germany). For this study, the first 20 Google hits were reviewed, and the following inclusion and exclusion criteria were applied (Table 1).

Concurrently, a research-specific search was executed on PubMed using the term “ChatGPT frequently asked patient questions.” In addition, ChatGPT 4 was directly engaged with the prompt “Suggest a list of the 20 most common frequently asked patient questions about lumbar spine fusion surgery” prompting it to generate a list of questions relevant to our study.

This multiphased approach resulted in a consolidated pool of 158 questions, from which 10 frequently recurring topics emerged (Table 2, Supplementary Material 1). The authors reviewed this topic list and formulated 10 final questions that comprised the most critical and commonly addressed patient concerns on lumbar fusion surgery (Table 3). Fig. 1 presents a flowchart outlining this process to define 10 final questions.

The final questions were then submitted to the AI chatbot ChatGPT 3.5 through its online portal (https://chat.openai.com/chat) on October 21, 2023, using the following prompt (Answer Set #1): “Act as an expert spine surgeon who is up to date with the latest scientific research and has years of experience counseling patients with empathy and clarity. Provide a comprehensive and easily understandable answer to the following question about lumbar spine fusion surgery. Limit your answer to 150 words and focus on the most important aspects.” The same questions and prompt were also presented to Google’s Bard (https://bard.google.com/chat) on the same date (Answer Set #2). For each question, a

Table 1. Inclusion and exclusion criteria for questions

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Published after January 1st 2017</td>
<td>Nongeneralizable information, e.g., provider or implant specific details</td>
</tr>
<tr>
<td>Published in English language</td>
<td>Emphasis on nonneurosurgical aspects, e.g., anesthesiological information</td>
</tr>
<tr>
<td>Information presented in FAQ or Q&amp;A sections</td>
<td></td>
</tr>
</tbody>
</table>

FAQ, frequently asked questions; Q&A, questions and answers.
new window was created in the respective Chatbot to avoid bias from the prior questions. A list of all answers provided by the 2 Chatbots can be found in Supplementary Material 2. ChatGPT, by OpenAI, utilizes the generative pre-trained transformer (GPT) architecture, focusing on generating broad, versatile responses across various topics through deep learning techniques. The model is pretrained on a diverse range of internet texts, allowing it to generate responses across a wide array of topics. Its iterative training and updates, such as the transition from GPT 3 to GPT 3.5 and beyond, focus on improving its understanding, accuracy, and ability to generate human-like text based on the input it receives. Google's Bard, on the other hand, leverages language model for dialogue applications (LaMDA), a system designed specifically to handle conversational applications. LaMDA's training regime includes a blend of reinforcement learning from human feedback and other methods to fine-tune its performance in dialogue-based tasks. This focus aims to produce more relevant and contextually appropriate responses.
especially in conversational settings. This technical divergence—ChatGPT’s wide-ranging generative capabilities versus Bard’s conversational precision—highlights their potential differences in applicability to medical education and patient communication.

Five blinded spine surgeons (either in spine-fellowship or spine-fellowship trained attending spine surgeons who did not know that these were LLM-generated responses) rated each response using a previously published rating system: ‘excellent response not requiring clarification’, ‘satisfactory requiring minimal clarification’, ‘satisfactory requiring moderate clarification’, or ‘unsatisfactory requiring substantial clarification’. Satisfactory responses conveyed primarily factual data, largely devoid of inaccuracies, albeit necessitating some elucidation. Responses warranting ‘minimal clarification’ were factually accurate but either lacked comprehensive information or failed to capture nuances from the literature. Those necessitating ‘moderate clarification’ relayed obsolete or irrelevant data. A response was deemed unsatisfactory if it encompassed data that was either outdated or overly generic, rendering it susceptible to misinterpretation.

11

The blinded raters also responded to two additional questions (Table 4) on a 5-point Likert scale (from “I strongly disagree” to “I strongly agree”) to assess whether the responses were easy to understand and communicated clearly, as well as whether they address patient concerns empathetically and professionally.

Data are presented using absolute values, percentages, mean and standard deviations for descriptive purposes. The word counts from ChatGPT 3.5, and Bard were compared using an independent t-test. To assess the relationship between word count and median ratings for each model, Pearson correlation coefficient (r) was calculated. The distribution of answer ratings across predefined categories was analyzed using the Wilcoxon signed-rank test, where ‘W’ is the test statistic. This test was chosen to compare the differences in ratings between ChatGPT and Bard.
3.5 and Bard. The Mann-Whitney U-test was used to compare the ratings between the 2 models for each of the 10 questions. The interrater reliability for the ratings was assessed using Cohen kappa. All statistical procedures were performed using IBM SPSS Statistics ver. 28.0 (IBM Co., Armonk, NY, USA) and Excel version 2302 (Microsoft 365, Microsoft, Redmond, WA, USA). The level of statistical significance was set at p < 0.05.

This study was exempt from Institutional Review Board review.

RESULTS

Despite being prompted to limit the answer to 150 words each, Bard’s answers had a significantly higher word count of 202.6 ± 42.9 (range, 138–287) compared to ChatGPT 3.5 (mean, 158.9 ± 18.1; range, 127–189; p < 0.05). ChatGPT’s word count was positively correlated with the median rating (r = 0.735, p < 0.05), while Bard’s word count was negatively correlated with the median rating (r = -0.68, p < 0.05).

Across both models and all 10 questions, 97% of answers were rated as satisfactory, and only 3% were unsatisfactory. Specifically, 64% (n = 64) were excellent without any clarification needed, 28% (n = 28) were satisfactory requiring minimal clarification, 5% (n = 5) were satisfactory requiring moderate clarification, and 3% (n = 3) were unsatisfactory requiring substantial clarification (Fig. 2).

For ChatGPT across all 10 questions, 62% of responses were excellent without any clarification needed, 32% were satisfactory requiring minimal clarification, 2% were satisfactory requiring moderate clarification, and 4% were unsatisfactory requiring substantial clarification. Bard had a slightly, but not statistically significantly, better performance, with only 2% of responses being rated as unsatisfactory requiring substantial clarification. For Bard, 66% of responses were excellent, 24% were satisfactory requiring minimal clarification, and 8% were satisfactory requiring moderate clarification. There was no statistically significant difference in the overall distribution of ratings between the 2 answer sets (ChatGPT 3.5 vs. Bard, W = 12; p = 1).

For the ChatGPT model, 7 out of 10 questions received median ratings of “excellent”. Questions Q3, Q4, and Q5 had the lowest median ratings of “satisfactory requiring minimal clarification” (Fig. 3). The lowest median ratings were seen in Q3 and Q5 for the Bard responses, both of which had median respons-

https://doi.org/10.14245/ns.2448098.049

Fig. 2. Pie chart with the distribution of overall ratings, expressed in percentages, for the combined question set across the 2 large language models.

Fig. 3. Distribution of rater evaluations for ChatGPT and Bard across 10 lumbar surgery questions. The bars represent the percentage of raters assigning each category. ChatGPT, chat generative pre-trained transformer.
Fig. 4. Median ratings comparison between ChatGPT 3.5 and Bard. Bars show median scores; error bars show range (minimum–maximum scoring). No significant differences in ratings across all questions. ChatGPT, chat generative pre-trained transformer.

There was no statistically significant difference in the ratings between the 10 answers from ChatGPT compared to Bard (Fig. 4; Q1: p = 0.42, Q2: p = 0.18, Q3: p = 1.0, Q4: p = 1.0, Q5: p = 0.08, Q6: p = 1.0, Q7: p = 0.52, Q8: p = 0.60, Q9: p = 0.60, and Q10: p = 0.52). Both model responses had poor interrater reliability (ChatGPT: k = 0.041, p = 0.622; Bard: k = -0.040, p = 0.601).

For the questions assessing clarity/easiness to understand and empathy/professionalism, the median ratings for both ChatGPT and Bard were 5/5. This indicates that on average, raters found the answers from both models to be easy to understand, clearly communicated, and addressing patient concerns empathetically and professionally. There was no statistically significant difference in the median scores between ChatGPT and Bard for these 2 questions. However, a select few individual raters rated Bard slightly lower on empathy and professionalism in addressing patient concerns.

**DISCUSSION**

AI, particularly language models like ChatGPT, is increasingly being explored for its potential in patient education. Recent studies have begun to elucidate the capabilities and limitations of AI in different domains of medicine with a special focus on patient education.\(^1\) In our current study, we systematically identified common patient questions on lumbar spine fusion surgery and evaluated the answers given by ChatGPT 3.5 and Google Bard based on ratings by “blinded” spine-fellowship trained surgeons.

ChatGPT adhered more closely to the specified 150-word limit than Bard. Notably, a positive correlation existed between ChatGPT’s word count and higher median ratings, whereas Bard’s word count showed a negative correlation with high ratings. This suggests that ChatGPT’s longer responses were received more favorably, indicating that its additional content was perceived as valuable. Conversely, Bard’s longer answers were seen as less effective or relevant, negatively impacting their reception.

The majority of responses for both models were excellent and did not require any clarification. Bard slightly outperformed ChatGPT in the proportion of excellent ratings, although this difference was not statistically significant. In addition, and very importantly, both models also achieved very high scores in empathy and professionalism. The slightly lower individual ratings for Bard on empathy and professionalism could imply a perceived difference in tone or response style.

Our results are consistent with findings from Ayers et al.,\(^12\) who reported very high quality and empathy of AI-generated responses. In fact, in their cross-sectional study of 195 randomly drawn patient questions from a social media forum, they found that AI-generated responses had higher quality and empathy than physician-generated ones. This suggests that in the future, AI may be used to draft initial responses to patient queries, which can then be reviewed and personalized by clinicians. Such an approach could revolutionize patient-physician interactions and improve efficiency. In our current study, we did not compare the answers given by the LLMs with answers by physicians but focused on assessing the accuracy and reliability of the LLM-provided answers. The continual alignment of AI-generated information with expert knowledge, especially in the sensitive field of medicine, will pose an ongoing challenge for research in the coming years.

Questions Q3 (risks and complications of surgery), Q4 (surgical success rate), and Q5 (approach/method) for both models received lower median ratings, falling into the ‘satisfactory requiring minimal clarification’ category. This suggests specific areas where both models need to be improved and fine-tuned. These questions represent areas where the application of AI in medical contexts faces limitations due to the need for highly individualized, context-sensitive, and up-to-date information, which can be challenging for an AI model to synthesize and communicate effectively.

1. Surgical Risks and Complications

Question 3 focuses on the risks and complications of lumbar fusion surgery. The complexity in answering this question lies
in the multifaceted nature of surgical risks and complications. Lumbar spine fusion surgery involves numerous variables, including patient-specific factors (age, health status, underlying conditions), surgical techniques, and postoperative care. AI models may struggle to integrate and personalize this vast, variable data into a concise, patient-specific response. This was summarized in ChatGPT's answer: "...While these risks exist, they are relatively rare, and we strive to minimize them. Your individual health, surgical technique, and postoperative care all play a role in the outcome. We prioritize your well-being and aim to provide the best possible results." Similarly, Bard summarized: "...The risks of lumbar spine fusion surgery vary depending on several factors, such as the patient's age, overall health, and the extent of the surgery. Your surgeon will discuss the risks and benefits of surgery with you in detail before you make a decision about whether or not to proceed." Indeed, the challenges faced by AI models like ChatGPT and Bard in addressing questions about risks and complications stem from the inherent multidimensionality of the topic.

2. Surgical Success Rate

Question 4 regarding surgical success rate was another question with overall lower ratings. The success rate of lumbar spine fusion surgery can be challenging to quantify due to varying definitions of "success," such as pain relief, functional improvement, spinal stability, appropriate hardware placement, and/or spinal bony fusion. These outcomes can differ significantly based on individual patient characteristics, the specific pathology being treated, and the surgical techniques used. It is also important to consider from whose perspective we are determining the success of the surgery, as this could be different for the surgeon versus the patient themselves versus the patient's caregiver. AI models may find it difficult to provide a generalized answer that accurately reflects the nuanced outcomes of different patient scenarios. This complexity is acknowledged in ChatGPT's answer for Q4: "Lumbar spine fusion surgery can be highly successful, but success rates vary depending on individual factors and the specific condition being treated. On average, the success rate falls between 60% to 90%. Factors influencing success include patient age, overall health, the surgeon's skill, and the reason for surgery..." Bard provided a similar answer: "The success rate of lumbar spine fusion surgery is generally considered to be high, with studies reporting success rates ranging from 70% to 90%. Success is typically defined as a significant reduction in pain and improvement in function. The success rate of lumbar spine fusion surgery can vary depending on a number of factors, such as the underlying condition being treated, the surgical approach used, the skill of the surgeon, and the patient's overall health..."

Both ChatGPT's and Bard's responses accurately capture the spectrum of factors that can influence the success of a patient's outcomes from spine fusion surgery.

3. Surgical Approach

Question 5 on the best surgical approach is inherently complex due to the evolving nature of surgical techniques and the lack of a universally accepted 'best' approach. Factors influencing the choice of surgical approach include the specific spinal pathology, surgeon's expertise, and technological advancements. AI models may struggle to make definitive recommendations without considering the individualized context of each patient's case. ChatGPT's answer for this question was: "The 'best' approach depends on your diagnosis, overall health, and surgeon's expertise. It's vital to discuss these options with your surgeon, who will consider your specific case and recommend the most appropriate approach." Bard listed posterior lumbar interbody fusion, anterior lumbar interbody fusion, and transforminal lumbar interbody fusion and pointed out the growing trend of minimally invasive spine surgery. It went on to say: "Again, it has been underlined, that the best surgical approach for lumbar spine fusion depends on a number of factors, including the specific condition being treated, the number of levels being fused, and the patient's anatomy..."

Neither ChatGPT nor Bard included lateral lumbar interbody fusion (LLIF) in their responses. LLIF has emerged as a viable surgical approach, particularly noted for its less invasive nature and reduced disruption of posterior spinal elements. Studies have shown LLIF to be effective in achieving spinal fusion with potentially lower complication rates and less postoperative pain, especially beneficial in multilevel spine disorders and deformities.

Both AI models' responses, while broad and general, aptly capture the multifactorial and individualized nature of determining the 'best' surgical approach. Studies have shown that the complication rates of a surgeon's surgical approach depend on several factors, including the patient's unique pathology, the number of spine levels involved, or the specific surgical devices and techniques employed—all of which shape the risk and outcome profile of the surgical intervention. This underscores the need for a personalized approach in surgical planning to optimize patient outcomes.

Our study is limited in that it is a small study with only 5 rat-
ers. We had poor interrater reliability, suggesting that there may be differences in subjective interpretation of the models’ answers, and/or different expectations for what constitutes a clear and comprehensive answer. In addition, our raters were all practicing spine surgeons, and we did not include any patients themselves. As a result, our findings do not necessarily reflect the patient perspective on the clarity and utility of the Chatbot responses to patient questions on lumbar fusion surgery. In addition, newer versions of LLMs have already been developed since this analysis was performed, and will continue to evolve at a rapid rate, potentially further improving the AI-generated responses. It’s important to note our study intentionally focused on evaluating LLMs based on real-world, patient-posed questions from frequently asked question (FAQ) sections, rather than optimizing inquiries for maximal LLM performance. This approach aims to provide insights into the actual advice patients might encounter online, acknowledging a potential trade-off in analytical precision. Further, our study did not assess the consistency of LLM responses to the same question asked multiple times. This decision was based on our aim to evaluate the LLMs’ performance in typical, real-life single-query interactions rather than exploring the variability of responses.

Recognizing that not all individuals with spinal conditions may be proficient in using AI, our study specifically evaluates LLM efficacy for users who engage with these platforms. This focus aims to shed light on AI’s capabilities and constraints in enriching patient education among digitally inclined segments of the population.

In the future, incorporating individualized patient data, as well as data from specific surgeons and spine centers, into AI-based LLMs could significantly enhance the precision and relevance of the information provided by AI, making it a more effective tool in patient education and decision-making. Personalized data would allow for more tailored responses regarding risks and outcomes, while center-specific data can inform patients about the practices and success rates of particular surgeons or facilities. This approach could not only aid in informed decision-making but also facilitates quality improvement and benchmarking in medical practices. However, ensuring data privacy and ethical use of this data will be essential in this process.

A recent study by Rajjoub et al. assessed ChatGPT’s responses against the 2011 North American Spine Society Clinical Guideline for lumbar spinal stenosis (LSS). This comparative analysis revealed that ChatGPT’s responses were congruent with the current literature on LSS. Specifically, the study found alignment in ChatGPT’s answers regarding the definition, diagnostic tests, and both nonsurgical and surgical interventions for LSS.

The authors suggested that ChatGPT can effectively support the decision-making process for LSS diagnosis and treatment, potentially making it a valuable tool in the context of lumbar spine fusion surgery education.

Beyond specific surgical contexts, AI’s role in patient education spans various areas. An article from the American Medical Association’s Journal of Ethics highlights the potential for AI to enhance patient-clinician relationships. The authors suggested, that by automating routine inquiries and administrative tasks, AI could allow clinicians to focus more on patient interaction and relationship-building. This aspect is particularly relevant for patient education, where AI could provide detailed and personalized information about treatment options, thus facilitating shared decision-making.

CONCLUSION

LLMs like ChatGPT and Bard hold significant promise for patient education in lumbar spine fusion surgery and broader medical contexts. In this study, we find that both LLMs produce accurate, clear, and empathetic responses to the most commonly asked questions about spinal fusion surgery. Nevertheless, human oversight remains crucial to ensure the effective and appropriate use of AI in healthcare. Training the models with patient-, surgeon-, and center-specific data may potentially increase their value. Future research should continue to explore and refine AI’s role, aiming for a harmonious integration of technology and human expertise in patient care and education.

NOTES

Supplementary Material: Supplementary materials 1-2 can be found via https://doi.org/10.14245/ns.2448098.049.

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author contribution: Conceptualization: SPL, CCZ; Formal analysis: SPL, ETY, ADGS, RK, PF, KW, NM, MNS; Methodology: SPL, ETY, ADGS, RK, PF, KW, NM, MNS, CCZ; Project administration: SPL, CCZ; Writing – original draft: SPL, ETY, ADGS, CCZ; Writing – review & editing: SPL, ETY, ADGS, RK, PF, KW, NM, MNS, CCZ.
REFERENCES


Inhibition of Ferroptosis by Mesenchymal Stem Cell-Derived Exosomes in Acute Spinal Cord Injury: Role of Nrf2/GCH1/BH4 Axis

Yixin Chen*, Bingfa Li*, Jing Quan, Zhe Li, Yan Li, Yinbo Tang
Department of Rehabilitation, Xiangya Hospital of Central South University, Changsha, China

Objective: The therapeutic benefits of exosomes obtained from mesenchymal stem cells (MSCs) in acute spinal cord injury (SCI) have been demonstrated in recent years, but the precise mechanisms remain unknown. In this study, the efficacy and mechanisms of MSC-derived exosomes (MSC-Exo) in acute SCI were investigated.

Methods: By utilizing a BV2 ferroptosis cellular model and an SCI rat model, we investigated the effects of MSC-Exo on iron death related indicators and NF-E2 related factor 2 (Nrf2)/GTP cyclolase I (GCH1)/5,6,7,8-tetrahydrobiopterin (BH4) signaling axis, as well as their therapeutic effects on SCI rats.

Results: The results revealed that MSC-Exo effectively inhibited the production of ferrous iron, lipid peroxidation products malonaldehyde and reactive oxygen species, and ferroptosis-promoting factor prostaglandin-endoperoxide synthase 2. Concurrently, they upregulated ferroptosis suppressors FTH-1 (ferritin heavy chain 1), SLC7A11 (solute carrier family 7 member 11), FSP1 (ferroptosis suppressor protein 1), and GPX4 (glutathione peroxidase 4), contributing to enhanced neurological recovery in SCI rats. Further analysis showed the Nrf2/GTP/BH4 signaling pathway's critical role in suppressing ferroptosis. Additionally, MSC-Exo was found to inhibit lipopolysaccharide-induced ferroptosis in BV2 cells and SCI rats by activating the Nrf2/GCH1/BH4 axis.

Conclusion: In summary, the study demonstrates that MSC-Exo mitigates microglial cell ferroptosis via the Nrf2/GCH1/BH4 axis, showing potential for preserving and restoring neurological function post-SCI.

Keywords: Spinal cord injury, Ferroptosis, Mesenchymal stem cell, Exosome

INTRODUCTION

Spinal cord injury (SCI), characterized by high mortality and disability rates, commonly results in limb dysfunction, restricted mobility, and can be fatal. Traumatic events are the predominant cause of SCI, while non-traumatic instances arise from ailments like infections or tumors.1 Acute SCI leads to the demise of neurons and glial cells, accompanied by oxidative stress, inflammation, and ischemia.2-4 These conditions contribute to sensory loss, motor function impairment, and severe outcomes like paraplegia, tetraplegia, or even early mortality;5 with the extent of functional impairment varying based on the lesion's severity and location. Despite notable progress in SCI management,6,7 it remains an incurable neurological disorder, necessitating lifelong patient care and treatment. Understanding the intricate regulatory mechanisms of SCI is crucial for developing innovative therapies to combat this debilitating condition.

Ferroptosis, a recently identified type of nonapoptotic cell demise, is distinguished by iron-dependent and accumulation of intracellular reactive oxygen species, and it had been proven to be implicated in various diseases such as ischemia-reperfusion injury, cancers, neurodegenerative disorders and SCI.8,9
Ferroptosis has been found to be a major cause of secondary injury of SCI, and deferoxamine facilitates traumatic SCI recovery through suppression of ferroptosis. Intraperitoneal administration of SRS16-86, a ferroptosis inhibitor, significantly ameliorates astrogliosis and promotes recovery of SCI in rats. Building upon our prior investigations, we have established the involvement of ferroptosis in SCI. Notably, administration of sodium selenite has been observed to foster neurological recuperation in SCI-affected rat models, principally through the suppression of ferroptosis. Despite these advancements, the comprehensive understanding of ferroptosis's functionalities and regulatory pathways in SCI context remains an area of ongoing exploration.

In the realm of SCI therapeutics, mesenchymal stem cells (MSCs) emerge as a potent candidate. Their efficacy is attributed to their capacity to alleviate secondary injuries. This is achieved through a multi-pronged approach involving the suppression of inflammatory responses, secretion of paracrine factors, and their potential to differentiate into neural cell lineages, presenting a multifaceted therapeutic mechanism in SCI treatment.

Recent investigations have shed light on the therapeutic potential of MSC-derived exosomes (MSC-Exo) in addressing acute central nervous system injuries. Focusing on SCI, studies have indicated that MSC-derived extracellular vesicles contribute to spinal cord functional recuperation. This is achieved by enhancing angiogenesis and axonal regeneration, diminishing cellular apoptosis, and attenuating inflammation and immune reactions. Despite these advancements, the specific role of ferroptosis in these processes remains an under-explored domain.

Divalent metal transporter 1 (DMT1) mediates the translocation of free iron from the endosome to the cytoplasm, thereby creating an “unstable iron pool” that predisposes cells to ferroptosis. Recent research indicated that exosomes derived from human umbilical cord blood MSCs inhibited ferroptosis by downregulating DMT1, thereby attenuating myocardial injury. Motivated by these findings, we hypothesize that the modulation of ferroptosis by MSC-Exo could be a critical factor in SCI recovery. This suggests a potential avenue for therapeutic intervention in SCI treatment.

Ferroptosis, fundamentally driven by iron-dependent lipid peroxidation, implicates oxidative stress regulatory pathways as pivotal players. Within this context, the GTP cyclolase I (GCH1)/tetrahydrobiopterin (BH4) signaling axis, linked to oxidative stress, emerges as notable. BH4, essential for nitric oxide synthases, when deficient, leads to superoxide generation. GCH1, as the rate-limiting enzyme in BH4 synthesis, governs the GTP to BH4 conversion, with its suppression amplifying oxidative stress in vivo. Intriguingly, recent discoveries position GCH1/BH4 as an independent ferroptosis regulatory system, separate from the GSH/glutathione peroxidase 4 (GPX4) axis. Under the oxidative stress induced by the accumulation of reactive oxygen species (ROS), overexpression of NF-E2 related factor 2 (Nrf2) transcriptionally activated the GCH1/BH4 pathway. This upregulation of GCH1 expression facilitated the restoration of intracellular BH4 levels, effectively mitigating ROS generation triggered by radiation exposure and thereby ameliorating oxidative stress. Consequently, the activation of Nrf2 leads to an antioxidant defense mechanism via the GCH1/BH4 axis. Enhanced Nrf2 expression correlates with inflammatory response moderation and ferroptosis resistance post-RSL3 stimulation in microglia and macrophages. Additionally, emerging evidence indicates that the overexpression of heme oxygenase 1 (HO-1) increases intracellular iron, disrupting redox balance and leading to ferroptosis, while Nrf2 mitigates oxidative stress and ferroptosis in SCI through the downregulation of HO-1. However, the exact role of the Nrf2/GCH1/BH4 pathway in ferroptosis, and its potential interplay with MSC-derived exosomes, remains an area ripe for investigation.

This study is designed to delve into the role of MSC-derived exosomes in ferroptosis within the context of SCI, specifically probing the mechanisms of the Nrf2/GCH1/BH4 signaling pathway.

MATERIALS AND METHODS

1. Cell Culture, Transfection, and Treatment

The isolation and cultivation of bone marrow MSCs were conducted following the methodologies outlined in earlier research. In brief, C57BL/6 mice (male, 4-week-old) were euthanized, and the hind limbs were dissected and placed on ice in DMEM (ThermoFisher Scientific, Waltham, MA, USA). Connective and muscle tissues were meticulously excised from the femur and tibia. Subsequently, bone marrow cells were extracted from these bones using a 27-gauge needle affixed to a 10-mL syringe, ensuring a thorough and efficient retrieval process. Bone marrow cells, at a concentration of 2.5 × 10⁶ cells/mL, were subjected to a cultivation process in 10-cm dishes at a temperature of 37°C for a duration of 3 hours. During this period, non-adherent cells were systematically removed through medium replacement. The culture was maintained for an additional 72 hours, with the medium being refreshed at 8-hour intervals.
The adherent cells underwent washing and received fresh culture medium every 3 days. Following a 2-week period, these cells were washed and treated with 0.25% trypsin/1 mM ethylenediaminetetraacetic acid for 2 minutes. The dissociated cells were then harvested and transferred to a 25 cm² flask for further culture. After 2–3 passages, the MSCs harvested exhibited typical spindle shape morphology with a diameter of 25–30 µm (Supplementary Fig. 1A). The features of MSCs surface markers include positive CD44 and negative CD11b (Supplementary Fig. 1B). After induction, the cells differentiated into adipocytes, osteoblasts, and cartilage (Supplementary Fig. 1C–E).

This study was conducted with the approval of the Experimental Animal Ethics Committee of Xiangya Hospital Central South University and approved (approval number: 202109058), adhering to all relevant ethical guidelines.

As the primary supportive cells for neurons, microglia play a crucial role in the nutrition and regeneration of spinal cord neuronal cells. Microglia rapidly adapt to changes in the microenvironment following SCI, thereby also playing a vital role in secondary damage and subsequent recovery of SCI. Therefore, we chose the most common microglial cell, the BV2 cells, for the cellular experiments. The advantages of BV2 cells include characteristics of primary microglia, immortality, and the ability for sustained growth, while limitations include their derivation from mouse microglia, which may not fully replicate the features and complexity of spinal cord microglia in vivo. Previous studies have shown that lipopolysaccharide (LPS) induced ferroptosis in various tissue and cell models, as well as BV2 cells. In addition to ferroptosis, LPS also significantly induce cellular inflammatory responses and other forms of cell death, which is similar to the pathological mechanisms of SCI. Therefore, we chose LPS-induced ferroptosis in BV2 cells for the mechanism study related to SCI.

In this study, BV2 cells, sourced from the Cell Bank of the Chinese Academy of Science in Shanghai, China, were subjected to a 24-hour treatment with LPS at a concentration of 100 ng/mL. For Nrf2 knockdown, Nrf2 siRNA (si-Nrf2) was obtained from RiboBio (Guangzhou, Guangdong, China), and BV2 cells were transfected with si-Nrf2 using Lipofectamine RNAiMAX (ThermoFisher Scientific). After 72 hours, cells were collected for subsequent assays. In some assays, BV2 cells were cocultured with MSC-derived exosomes at 1 µg/mL, tert-butylhydroquinone (TBHQ, Selleck Chemicals, Houston, TX, USA) at 25 µM and FIN56 (Selleck Chemicals, Houston, TX, USA) at 5 µM, respectively.

2. Exosome Isolation and Characterization
MSCs were cultured in a serum-depleted environment, following which exosomes were isolated utilizing the Total Exosome Isolation reagent from ThermoFisher Scientific. The characterization of these exosomes was meticulously performed using nanoparticle tracking analysis (NTA) conducted by Malvern Instruments (Westborough, MA, USA) and transmission electron microscopy (TEM) provided by ThermoFisher Scientific. To further validate the exosome identity, the presence of markers such as CD63, CD81, and TSG101 was confirmed through Western blot analysis.

3. Transmission Electron Microscopy
The examination of exosomes via TEM was carried out in accordance with previous study. In brief, exosomes were concentrated by centrifuging at 100,000 g for a period of 90 minutes. After the removal of the supernatant, the exosome pellets were fixed using a 2.5% glutaraldehyde solution, maintained at 4°C for an hour. Subsequently, the glutaraldehyde solution was discarded, and the pellets were subjected to a triplicate wash using a 0.1-M sodium cacodylate solution. Pellets were post-fixed in 2% Osmium tetroxide at 4°C for 1 hour and washed 3 times in 0.1-M sodium cacodylate solution. A series of gradient acetone solution (50%, 60%, 70%, 80%, 90%, 95%, and 100%) were prepared, and pellets were incubated in each acetone solution for 10 minutes. The mixture of acetone and low viscosity (1:1, 1:1, and 1:3) was prepared, and pellets were consecutively incubated in the mixture of acetone and low viscosity (3:1, 1:1, and 1:3) for 30 minutes. Finally, pellets were incubated in pure low viscosity embedding mixture overnight, embed and baked at 65°C for 24 hours. Sections (60 nm) were sliced and stained with 2% uranyl acetate for 20 minutes and lead citrate for 10 minutes followed by observation under TEM.

4. Immunofluorescence Staining
BV2 cells underwent fixation in a 4% paraformaldehyde solution and were permeabilized using 0.2% Triton X-100. Following washing and blocking, the cells were incubated overnight with rabbit anti-GPX4 antibodies (2 µg/mL) and anti-FSP1 antibodies (1 µg/mL), both sourced from Abcam (Cambridge, UK). The cells were then thoroughly rinsed and treated with an Alexa Fluor 488-conjugated secondary antibody for one hour. Post-staining with 4',6-diamidino-2-phenylindole (DAPI) provided by Beyotime (Shanghai, China), the cells were prepared for imaging. Observations were made using a confocal microscope from Nikon (Tokyo, Japan). The results of immunofluorescence
staining were analyzed using Image J software (National Institutes of Health, Bethesda, MD, USA) and semiquantitative analysis was performed based on the average optical density.

5. Methylthiazolyl diphenyl-Tetrazolium Bromide Assay
In the prescribed treatment protocol for BV2 cells, the culture medium was initially discarded. Thereafter, a mixture of 100 µL of fresh medium and 10 µL of methylthiazolyl diphenyl-tetrazolium bromide (MTT) reagent was added to each culture. The cells were then incubated for a duration of 4 hours. Following this incubation period, dimethyl sulfoxide was introduced into the cultures. The absorbance at 540 nm was subsequently measured to assess cell viability. The MTT reagent used in this procedure was procured from Sigma (St. Louis, MO, USA).

6. Real-Time Quantitative Reverse Transcription-Polymerase Chain Reaction

Total RNA was isolated from BV2 cells employing the Trizol reagent, sourced from Beyotime. This RNA was then subjected to reverse transcription into cDNA using the QuantiTect Reverse Transcription Kit (QIAGEN, Germantown, MD, USA). Ferritin heavy chain 1 (FTH1), prostaglandin-endoperoxide synthase 2 (PTGS2), solute carrier family 7 member 11 (SLC7A11), Nrf2, and GCH1 were examined by quantitative polymerase chain reaction with SYBR Green (ThermoFisher Scientific) and normalized to GAPDH. Primers were shown in Table 1. The 2−ΔΔCt method was used for calculation.

Table 1. Quantitative real-time polymerase chain reaction primers

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTH1</td>
<td>Forward: 5′-CAGACCGTGATGACTGGGAG-3′&lt;br&gt;Reverse: 5′-CTCAATGAAGTCACATAAGTGGGG-3′</td>
</tr>
<tr>
<td>PTGS2</td>
<td>Forward: 5′-GCTTTCAACAGTTTTCTCTACACAA-3′&lt;br&gt;Reverse: 5′-CATTCTTCTCCCCCAGCAAC-3′</td>
</tr>
<tr>
<td>SLC7A11</td>
<td>Forward: 5′-ATCTCCCCCAAGGGCATTAC-3′&lt;br&gt;Reverse: 5′-GCATAGGACAGGGCTCCAAA-3′</td>
</tr>
<tr>
<td>Nrf2</td>
<td>Forward: 5′-GAGCAGCATGATTTAAGC-3′&lt;br&gt;Reverse: 5′-CAGCGAGCTGCTTTTC-3′</td>
</tr>
<tr>
<td>GCH1</td>
<td>Forward: 5′-TGCTTTACCTCGTCCATCCTGC-3′&lt;br&gt;Reverse: 5′-CCTTTGACATCACCATCTGGC-3′</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Forward: 5′-GTCTTCTGGGCAAGACAGTA-3′&lt;br&gt;Reverse: 5′-CTGGACAGAAACCCACTTC-3′</td>
</tr>
</tbody>
</table>

FTH1, ferritin heavy chain 1; PTGS2, prostaglandin-endoperoxide synthase 2; SLC7A11, solute carrier family 7 member 11; Nrf2, NF-E2 related factor 2; GCH1, GTP cyclohydrolase I.

7. Experimental Rat SCI Model
As the most commonly used animal model for SCI currently, Sprague-Dawley (SD) rats offer the benefits of low cost, easy availability, and similarities in electrophysiology and morphology to the human spinal cord. Therefore, this study selected SD rats for the animal experiments. However, considering the complexity of human SCI, the SD rat SCI model cannot fully simulate the pathological process of human SCI. SD rats (male, 8-week-old) were purchased from SJA Laboratory Animal (Changsha, Hunan, China) and blindly divided into control, SCI, SCI+ normal saline and SCI+MSC-Exo groups. SCI rats were anesthetized, and the spinal cord was processed for laminectomy at the T9/10 position. Then, a contusion injury was performed via dropping a mass of 10 g at a vertical height of 40 mm on the spinal cord. Control rats received same procedures without the contusion injury. Rats in SCI+MSC-Exo were intrathecally injected with MSC-Exo (100 µg). After 14 days, spinal cord tissues (T10) were collected for pathological examination, Western blotting and oxidative stress analysis. In addition, we conducted Basso, Beattie, and Bresnahan (BBB) score for all the rats at 1, 3, 7, 14 days after SCI. All experimental procedures involving rats received approval from the Animal Care and Use Committee at Xiangya Hospital, Central South University, ensuring compliance with ethical standards.

8. Western Blotting
Protein (20 µg) extracted from BV2 cells and spinal cord tissues was electrophoresed and transferred to PVDF membranes from Bio-Rad (Hercules, CA, USA). Subsequently, membranes were incubated with rabbit antibodies against CD63 (0.5 µg/mL, Abcam), CD81 (0.1 µg/mL, Abcam), GPX4 (2 µg/mL, Abcam), and FSP1 (1 µg/mL, Abcam) overnight. Membranes were rinsed and incubated with an horseradish peroxidase-conjugated secondary antibody (ThermoFisher Scientific). Enhanced chemiluminescence substrate (Abcam) was added to visualize bands. The grayscale values of each band were analyzed using Image J software, then the relative quantitative analysis was performed by calculating the ratio of each group to the internal reference.

9. Enzyme-Linked Immunosorbent Assay
BH4 were determined with enzyme-linked immunosorbent assay (ELISA) kits following manuals. Tetrahydrobipterin (BH4) ELISA kit was bought from ElAab (Wuhan, Hubei, China).
10. Iron Assay
To determine Fe\(^{2+}\) level, BV2 cells were lysed on ice, and supernatants were harvested after centrifugation. The level of Fe\(^{2+}\) in BV2 cells was examined with the Iron Assay Kit (Sigma) following the manual.

11. Oxidative Stress Examination
Superoxide dismutase (SOD) and malonaldehyde (MDA) were examined with SOD (Sigma) and MDA (BioVision, Milpitas, CA, USA) assay kits, respectively. For ROS staining, cells and spinal cord tissues were stained with CellRox green (ThermoFisher Scientific) at 1 µM for 30 minutes, washed and stained with DAPI (Beyotime).

12. Hematoxylin and Eosin Staining
The spinal cord samples were embedded in paraffin and cut into 5-µm thick sections. The sections were stained with hematoxylin-eosin reagents (Servicebio, Wuhan, China), and the hematoxylin and eosin (H&E) staining process included 5 steps: dewaxing, staining, dehydration, transparency, and sealing. When finished, the number of neural cells, arrangement regularity, and cavity sizes were observed using an optical microscope (Eclipse E100; Nikon).

### RESULTS

1. LPS-Induced Ferroptosis in BV2 Cells
Upon treatment with LPS, BV2 cells demonstrated notable alterations in cell morphology, indicative of damage, along with a discernible reduction in cell count (Fig. 1A). Moreover, LPS obviously impaired BV2 cell proliferation (Fig. 1B). To demonstrate that the form of cell death induced by LPS was ferroptosis,\(^{29,30}\) we examined ferroptosis-related factors GPX4, PTGS2, FTH1, and SLC7A11. LPS-treated BV2 cells showed significantly increased expression of PTGS2 and decreased expression of GPX4, FTH1, and SLC7A11 (Fig. 1C). In addition, LPS-treated BV2 cells showed markedly elevated Fe\(^{2+}\), ROS and MDA (Fig. 1D–F). These observations implied that LPS could induce ferroptosis in BV2 cells.

---

**Fig. 1.** Lipopolysaccharide (LPS) induced ferroptosis in BV2 cells. BV2 cells were treated with LPS at 100 ng/mL for 24 hours. (A) Morphological observation of BV2 cells (scale bar, 100 µm). (B) Cell proliferation was evaluated with methylthiazolyldiphenyl-tetrazolium bromide (n = 3). (C) Quantitative real-time polymerase chain reaction analysis of GPX4, PTGS2, FTH1, and SLC7A11 (n = 3). (D) The relative Fe\(^{2+}\) level (n = 3). (E) Level of reactive oxygen species (ROS) (n = 3). (F) Level of malonaldehyde (MDA) (n = 3). GPX4, glutathione peroxidase 4; PTGS2, prostaglandin-endoperoxide synthase 2; FTH1, ferritin heavy chain 1; SLC7A11, solute carrier family 7 member 11. *p < 0.05. **p < 0.01.
2. Characterization of MSC-Derived Exosomes

Bone marrow MSCs were isolated from the femurs and tibias and cultured for exosome isolation. The morphology of passage 1, 2, and 3 MSCs was examined, and the classical spindle-shaped morphology was observed (Fig. 2A). Subsequently, exosomes were isolated from bone marrow MSCs followed by exosome characterization. NTA and TEM validated that MSC-derived exosomes (MSC-Exo) showed typical round or round-like vesicle structure with a diameter of 40 to 200 nm (Fig. 2B and C). Moreover, high expression of exosome markers CD63, CD81, and TSG101 was observed in MSC-derived exosomes, but not in MSCs (Fig. 2D). These results confirmed typical exosome characterization of MSC-derived exosomes.

3. MSC-Exo Suppressed Ferroptosis in LPS-Treated BV2 Cells

To investigate the activity of MSC-Exo in ferroptosis, BV2 cells were treated with LPS, LPS+MSC-Exo, or LPS+exosome-depleted supernatant from MSCs. Control cells were treated with saline.

BV2 cells were treated with LPS at 0, 25, 50, or 100 ng/mL for 48 hours, and cell proliferation was analyzed with MTT (n = 3). LPS-induced BV2 cells were cocultured with MSC-Exo at 0.5, 1, or 1.5 μg/mL for 48 hours, and cell proliferation was analyzed with MTT. As shown in Fig. 3A and B, BV2 cell proliferation was inhibited by LPS but promoted by MSC-Exo in a dose-dependent manner. Therefore, we chose the concentration of 100 ng/mL and 1.5 μg/mL for subsequent experiments.

LPS downregulated FTH1, SLC7A11, and GPX4, but upregulated PTGS2 (Fig. 3C). MSC-Exo enhanced the expression of FTH1, SLC7A11, and GPX4, and inhibited the expression of PTGS2, while exosome-depleted supernatant showed no effect (Fig. 3C). Furthermore, LPS-induced ROS was attenuated by MSC-Exo in BV2 cells (Fig. 3D), and the elevated level of MDA were reversed by MSC-Exo (Fig. 3E). No significant change was observed in BV2 cells treated with exosome-depleted supernatant (Fig. 3D and E). Furthermore, we found that LPS raised Fe²⁺ level in BV2 cells, but the effect was reversed by MSC-Exo. LPS-induced Fe²⁺ upregulation was not affected by exosome-depleted supernatant (Fig. 3F). As shown in Fig. 3G, BV2 cell proliferation was obviously inhibited by LPS but MSC-Exo partially alleviated this inhibition, while no similar change was observed in cells treated with exosome-depleted supernatant.

As GPX4 and FSP1 are key regulators in ferroptosis, we addi-
Fig. 3. Mesenchymal stem cell-derived exosomes (MSC-Exo) suppressed ferroptosis in BV2 cells. (A) BV2 cells were treated with lipopolysaccharide (LPS) at 0, 25, 50, or 100 ng/mL for 48 hours. Cell proliferation was analyzed with methylthiazolyldiphenyl-tetrazolium bromide (MTT) (n = 3). (B) LPS-induced BV2 cells were cocultured with MSC-Exo at 0.5, 1, or 1.5 μg/mL for 24 hours. Cell proliferation was analyzed with MTT (n = 3). BV2 cells were treated with LPS at 100 ng/mL, LPS+MSC-Exo at 1.5 μg/mL or LPS+exosome-depleted supernatant. (C) Quantitative real-time polymerase chain reaction analysis of GPX4, PTGS2, FTH1, and SLC7A11 (n = 3). (D) The relative ratio of reactive oxygen species (ROS)-positive BV2 cells (n = 3). (E) Levels of malonaldehyde (MDA) (n = 3). (F) The relative Fe$^{2+}$ level (n = 3). (G) Cell proliferation was analyzed with MTT (n = 3). GPX4, glutathione peroxidase 4; PTGS2, prostaglandin-endoperoxide synthase 2; FTH1, ferritin heavy chain 1; SLC7A11, solute carrier family 7 member 11. *p < 0.05. **p < 0.01. ***p < 0.001.
expression of GPX4 and FSP1, but simultaneous MSC-Exo treatment increased their expression (Fig. 4).

Collectively, these findings demonstrated that MSC-Exo attenuated LPS-induced ferroptosis in BV2 cells.

4. MSC-Exo Restrained Ferroptosis Through Activation of the Nrf2/GCH1/BH4 Signaling in BV2 Cells

As the Nrf2/GCH1/BH4 signaling plays key roles in the regulation of ferroptosis, we examined the expression of Nrf2 and GCH1 and BH4 generation in BV2 cells. We found that Nrf2 and GCH1 were downregulated and BH4 generation was reduced in LPS-treated BV2 cells when compared with the control group (Fig. 5A and B). TBHQ, an activator of Nrf2, significantly elevated the expression of Nrf2 and GCH1 and BH4 generation (Fig. 5A and B). In contrast, knockdown of Nrf2 further enhanced LPS-mediated downregulation of Nrf2, GCH1, and BH4 (Fig. 5A and B). Moreover, elevated ROS-positive cells were reduced by simultaneous TBHQ treatment but further upregulated by Nrf2 knockdown (Fig. 5C). LPS-mediated upregulation of PTGS2 and downregulation of FTH1, SLC7A11, and GPX4 in BV2 cells were partly abrogated by TBHQ treatment but reinforced by Nrf2 knockdown (Fig. 5D). Our data suggested that the Nrf2/GCH1/BH4 signaling suppressed ferroptosis in BV2 cells.

Furthermore, we examined whether MSC-Exo regulated the Nrf2/GCH1/BH4 signaling to modulate ferroptosis. Reduced levels of Nrf2, GCH1, and BH4 in LPS-treated BV2 cells were partially enhanced by MSC-Exo treatment (Fig. 6A and B). In addition, MSC-Exo inhibited LPS-induced ROS, and the inducer of ferroptosis FIN56 markedly promoted ROS in LPS-treated BV2 cells (Fig. 6C). LPS-mediated suppression of GPX4, FSP1, FTH1, and SLC7A11 was reversed by MSC-Exo, and FIN56-treated cells showed lowest expression (Fig. 6D and E). LPS-induced expression of PTGS2 and Fe²⁺ was partially suppressed by MSC-Exo, and FIN56 treatment dramatically enhanced the expression of PTGS2 and Fe²⁺ (Fig. 6E). Taken together, our findings indicated that MSC-Exo may suppressed ferroptosis via activating the Nrf2/GCH1/BH4 signaling.

5. Administration of MSC-Exo Significantly Improved Neurological Rehabilitation After SCI in Rats

Following the establishment of a rat model of SCI via laminectomy and contusion injury, MSC-Exo were administered intrathecally. A 2-week observation period revealed a significant enhancement in the BBB locomotor rating scale in the hind limbs of the SCI rats, as depicted in Fig. 7A. Immediately following SCI, rats in the SCI, SCI+normal saline, and SCI+
MSCs-Exo inhibited ferroptosis by Nrf2/GCH1/BH4 axis in SCI

Chen Y, et al.

MSC-Exo groups exhibited severe hind limb paralysis. Hind limb motor function began to recover at 3 days after SCI and showed a gradual increase during the 14-day experimental period. The BBB scores of the SCI+MSC-Exo group were higher relative to those of the SCI group and SCI+normal saline group at 14 days after SCI. TEM analyses of the injured spinal cord tissues from both SCI and MSC-Exo groups revealed varying degrees of ferroptosis, characterized by phenomena such as hemolysis, mitochondrial shrinkage, and increased membrane density, as illustrated in Fig. 7B. Comparative analysis with control rats indicated pronounced neural cell damage and heightened inflammatory cell infiltration in the SCI group, as observed through H&E staining. However, the administration of MSC-Exo markedly mitigated these damages and reduced inflammatory cell infiltration (Fig. 7C). Similarly, SCI obviously increased the cavity area of spinal cord tissue, while MSC-Exo administration effectively reduced the cavity area (Fig. 7D). Additional-ly, the expression levels of GPX4, Nrf2, and GCH1 were found to be downregulated in SCI rats, with an upregulation of PTGS2, which was reversed upon MSC-Exo treatment, as shown in Fig. 7E. Furthermore, elevated levels of ROS, MDA, and Fe²⁺ in the SCI group were normalized post MSC-Exo treatment (Fig. 7F–H). These findings suggest that MSC-Exo significantly enhance neurological rehabilitation following SCI in rats, po-

Fig. 5. The Nrf2/GCH1/BH4 signaling suppressed ferroptosis in BV2 cells. BV2 cells were treated with lipopolysaccharide (LPS), LPS in combination with tert-butylhydroquinone (TBHQ) at 25 µM for 24 hours. Nrf2-knockdown cells were treated with LPS. (A) Quantitative real-time polymerase chain reaction (qRT-PCR) analysis of Nrf2 and GCH1 (n = 3). (B) BH4 generation was examined using enzyme-linked immunosorbent assay (n = 3). (C) Reactive oxygen species (ROS) (green) staining in BV2 cells. Scale bar, 200 µm. (D) qRT-PCR analysis of GPX4, PTGS2, FTH1, and SLC7A11 (n = 3). Nrf2, NF-E2 related factor 2; GCH1, GTP cyclase I; BH4, tetrahydrobiopterin; GPX4, glutathione peroxidase 4; PTGS2, prostaglandin-endoperoxide synthase 2; FTH1, ferritin heavy chain 1; SLC7A11, solute carrier family 7 member 11. *p < 0.05. **p < 0.01. ***p < 0.001.
MSCs-Exo inhibit ferroptosis by Nrf2/GCH1/BH4 axis in SCI

Chen Y, et al.

MSCs-Exo inhibit ferroptosis by Nrf2/GCH1/BH4 axis in SCI

Chen Y, et al.

https://doi.org/10.14245/ns.2448038.019

www.e-neurospine.org

651

Fig. 6. Mesenchymal stem cell-derived exosomes (MSC-Exo) restrained ferroptosis through activation of the Nrf2/GCH1/BH4 signaling. BV2 cells were treated with LPS, LPS+MSC-Exo, or LPS+FIN56 (5 µM) for 24 hours. (A) Quantitative real-time polymerase chain reaction analysis of Nrf2 and GCH1 (n = 3). (B) BH4 generation was examined using enzyme-linked immunosorbent assay (n = 3). (C) Reactive oxygen species (ROS) (green) staining in BV2 cells. Scale bar, 200 µm. (D) Protein levels of GPX4 and FSP1 (n = 3). (E) Expression analysis of Fe2+, PTGS2, FTH1, and SLC7A11 (n = 3). Nrf2, NF-E2 related factor 2; GCH1, GTP cyclohydrolase I; BH4, tetrahydrobiopterin; GPX4, glutathione peroxidase 4; FSP1, ferroptosis suppressor protein 1; PTGS2, prostaglandin-endoperoxide synthase 2; FTH1, ferritin heavy chain 1; SLC7A11, solute carrier family 7 member 11. *p < 0.05. **p < 0.01. ***p < 0.001.

Potentially through the inhibition of ferroptosis via the Nrf2/GCH1 signaling pathway.

DISCUSSION

SCI detrimentally impacts spinal cord functionality and contributes to a heightened mortality rate among affected individuals. In the realm of SCI management, therapies based on MSCs have emerged as a promising avenue. Recent experimental investigations in SCI have underscored the neuroprotective potential of MSC-Exo, positioning them as an innovative therapeutic strategy. Notably, exosomes derived from bone marrow MSCs have been documented to repair traumatic SCI by curbing the activation of A1 neurotoxic reactive astrocytes. In our current investigation, we discerned that MSC-Exo safeguard LPS-treated BV2 cells by impeding ferroptosis. Furthermore, the in vivo administration of these MSC-Exo was observed to effectively alleviate SCI symptoms.

Owing to their capacity for multilineage differentiation, MSCs can evolve into nerve cells, aiding in the repair of damaged neural tissues. However, recent research increasingly attributes the therapeutic efficacy of MSCs to their paracrine functions. MSCs discharge paracrine factors such as growth and inflammatory mediators in the form of exosomes, which play a pivotal role in modulating cell function and ultimately contribute to the overall therapeutic outcome.
Fig. 7. Mesenchymal stem cell-derived exosomes (MSC-Exo) administration improved neurological rehabilitation after spinal cord injury (SCI) in rats. Rats were divided into control, SCI, SCI+normal saline, and SCI+MSC-Exo groups (n = 3). (A) Basso, Beattie, and Bresnahan (BBB) scores of the hind limbs of SCI rats. (B) Varying degrees of ferroptosis such as hemolysis, mitochondrial shrinkage, and increased membrane density were observed in the injured spinal cord tissues. Scale bar, 2 µm. (C) Hematoxylin and eosin staining of spinal cord tissues. Scale bars: upper, 100 µm; lower, 25 µm. (D) Cavity area of spinal cord tissues (n = 3). (E) Protein levels of GPX4, PTGS2, Nrf2, and GCH1 in spinal cord tissues (n = 3). (F) The reactive oxygen species (ROS)-positive rate in spinal cord tissues (n = 3). (G) Levels of malonaldehyde (MDA) in spinal cord tissues (n = 3). (H) The relative Fe²⁺ level in spinal cord tissues (n = 3). GPX4, glutathione peroxidase 4; PTGS2, prostaglandin-endoperoxide synthase 2; Nrf2, NF-E2 related factor 2; GCH1, GTP cyclohydrolase I. *p < 0.05. **p < 0.01. ***p < 0.001.
role in modulating oxidative stress, inflammation, and apoptosis, thereby exhibiting neuroprotective activities. The remarkable biocompatibility and low immunogenicity of exosomes enhance their potential for clinical applications across various diseases. In the context of cancers and neurodegenerative disorders, MSC-Exo have been identified as key players. The use of exosomes from human Wharton's jelly MSCs has emerged as a promising approach in the prevention and treatment of perinatal brain injury. A recent study revealed that MSC-Exo combat ferroptosis in acute liver injury by preserving SLC7A11 function, thus promoting liver repair. Shao et al. have documented that MSC-Exo impedes ferroptosis in neuronal cells via the IncGm36569/miR-5627-5p/FSP1 axis in acute SCI. In our study, we observed that MSC-Exo hindered LPS-induced ferroptosis in BV2 cells and enhanced neurorecovery post-SCI in rats, reinforcing the view that MSC-Exo hold neuroprotective capabilities.

Accumulating evidence suggests that targeting ferroptosis suppression could evolve into an efficacious strategy for enhancing SCI treatment. Consequently, a deep understanding of the regulatory mechanisms governing ferroptosis is vital for its strategic manipulation. Several studies have identified key regulators within ferroptosis metabolic pathways, such as GPX4, FSP1, PTGS2, FTH1, and SLC7A11. Our findings indicate that LPS diminishes the expression of GPX4, FTH1, and SLC7A11, while augmenting PTGS2 expression in BV2 cells. Notably, these alterations were partially reversed by the application of MSC-Exo.

Antioxidant synthesis stands as a primary defense against ferroptosis. BH4, a critical antioxidant, mitigates lipid peroxidation and confers cellular protection against ferroptosis. The synthesis of BH4 relies on the activity of GCH1, which catalyzes the transformation of guanosine triphosphate into dihydronicotinamide triphosphate, eventually leading to BH4 production. The GCH1/BH4 signaling pathway has been recognized for its role in ferroptosis suppression. Nrf2, a pivotal transcription factor in antioxidant defense, stimulates GCH1 transcription and boosts BH4 synthesis, suggesting the implication of the Nrf2/GCH1/BH4 axis in ferroptosis regulation. In our study, we observed that the Nrf2/GCH1/BH4 signaling pathway could counteract ferroptosis in LPS-treated BV2 cells. Moreover, we report that MSC-Exo could reduce LPS-induced ferroptosis potentially through activating the Nrf2/GCH1/BH4 pathway. This observation was corroborated by similar findings obtained in our SCI rat model experiments, further validating the neuroprotective role of MSC-Exo in SCI contexts through modulation of ferroptosis.

There is also a close association between Nrf2/GCH1/BH4 signaling axis and key ferroptosis regulators, which are mainly mediated by Nrf2. That is to say, Nrf2 regulate ferroptosis through multiple signal axes. For instance, in liver injury, Nrf2 inhibit ROS production by activating the HO-1/GPX4 axis, thereby inhibiting ferroptosis. Nrf2/GPX4 axis effectively inhibit ferroptosis in DOX induced cardiopathy. In the PTGS2/COX-2/Nrf2 signaling, there is an interaction between PTGS2/COX-2 and Nrf2 through electrophilic oxo-derivative (EFOX). PTGS2/COX-2 regulates EFOX synthesis, while EFOX induce Nrf2 dissociation, nuclear translocation, and other functions, thereby regulating oxidative stress.

However, current methods for the isolation and identification of exosomes vary widely, with no standardized approach. There is also considerable debate regarding how to define the dosage for clinical applications of exosomes. The prevalent storage methods commonly utilize isotonic buffers to prevent pH changes during storage, in order to maintain the integrity of exosomes. Due to the lack of data on the impact of processing, storage duration, and preservatives on the structural stability and functional efficacy of exosomes, further research is necessary to establish a gold standard for exosome storage. It should also be recognized that the reproducibility of MSC-Exo-based therapies for SCI in humans may be poor. This is attributed to the differences between animal and human spinal cords, which may lead to less than optimal outcomes when exosome therapies effective in animals are applied to human SCI.

To validate the clinical translatability of MSC-Exo in SCI treatment, further research should progress in the following directions: continue studies on the mechanisms and potential side effects of MSC-Exo in the treatment of SCI, providing more theoretical basis for clinical use; further standardize the acquisition, storage, and utilization methods of MSC-Exo, laying the foundation for future clinical treatments of SCI; prior to clinical application for SCI treatment, rigorous, large-scale animal experiments and preclinical studies are required to ensure the safety and clinical efficacy of MSC-Exo therapy.

CONCLUSION

The findings of this research elucidate that exosomes derived from MSCs play a pivotal role in inhibiting ferroptosis, thereby facilitating functional recovery following SCI. It appears that the regulatory influence of MSC-Exo on ferroptosis might be contingent upon the Nrf2/GCH1/BH4 signaling axis. This study not only provides an initial insight into the modulation of ferroptosis in BV2 cells and SCI rat models but also paves the way...
for potential exosome-based therapeutic approaches for the treatment of SCI.

NOTES

Supplementary Material: Supplementary Fig. 1 can be found via https://doi.org/10.14245/ns.2448038.019.

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study was supported by Hunan Provincial Natural Science Foundation Project (No. 2022JJ40772) and Changsha Municipal Natural Science Foundation (No. kq2014285).

Author Contribution: Conceptualization: YC, BL, JQ, YL, YT; Data curation: YC, JQ, YL; Formal analysis: YC, BL, JQ, ZL; Funding acquisition: YC; Methodology: YC, BL, YT; Project administration: BL; Visualization: YC, ZL, YL, YT; Writing – original draft: YC, JQ, ZL, YL, YT; Writing – review & editing: YC, JQ, YT.

ORCID
Yixin Chen: 0000-0002-5511-2275
Bingfa Li: 0009-0003-9973-3772
Jing Quan: 0009-0003-8357-8544
Zhe Li: 0009-0009-6983-5114
Yan Li: 0009-0005-5121-4652
Yinbo Tang: 0000-0001-8764-640X

REFERENCES
MSCs-Exo Inhibite Ferroptosis by Nrf2/GCH1/BH4 Axis in SCI

Chen Y., et al.

MSCs-Exo Inhibite Ferroptosis by Nrf2/GCH1/BH4 Axis in SCI
Chen Y., et al.

https://doi.org/10.14245/ns.2448038.019

www.e-neurospine.org

655


Supplementary Fig. 1. Characterization of the mesenchymal stem cells (MSCs). (A) Photos of MSCs cells on day 7 of the first-generation culture, day 3 of the second-generation culture, and day 3 of the third-generation culture, respectively. Scale bar, 100 μm. (B) Flow cytometry was used to detect the number of positive cells for CD44 and CD11b in MSCs, and control serves as a negative control for flow cytometry testing. (C) Oil red O staining was used to detect the adipogenic differentiation ability of MSCs. Scale bar, 25 μm. (D) Alkaline phosphatase staining was used to detect the osteogenic differentiation ability of MSCs. Scale bar, 100 μm. (E) Alcian blue staining to detect the chondrogenic differentiation ability of MSCs. Scale bar, 100 μm.
Comparison of Single or Double Titanium Mesh Cage for Anterior Reconstruction After Total En Bloc Spondylectomy for Thoracic and Lumbar Spinal Tumors

Ao Leng*, Qi Wang*, Jiacheng Li, Yu Long, Song Shi, Lingzhi Meng, Mingming Guo, Hailong Yu, Liangbi Xiang

Department of Orthopaedics, General Hospital of Northern Theater Command, Shenyang, China

Objective: To compare the clinical efficacy of anterior column reconstruction using single or double titanium mesh cage (TMC) after total en bloc spondylectomy (TES) of thoracic and lumbar spinal tumors.

Methods: A retrospective cohort study was performed involving 39 patients with thoracic or lumbar spinal tumors. All patients underwent TES, followed by anterior reconstruction and screw-rod instrumentation via a posterior-only procedure. Twenty-two patients in group A were treated with a single TMC to reconstruct the anterior column, whereas 17 patients in group B were reconstructed with double TMCs.

Results: The overall follow-up is 20.5 ± 4.6 months. There is no significant difference between the 2 groups regarding age, sex, body mass index, tumor location, operative time, and intraoperative blood loss. The time for TMC placement was significantly shortened in the double TMCs group (5.2 ± 1.3 minutes vs. 15.6 ± 3.3 minutes, p = 0.004). Additionally, postoperative neural complications were significantly reduced with double TMCs (5/22 vs. 0/17, p = 0.046). The kyphotic Cobb angle and mean intervertebral height were significantly corrected in both groups (p ≤ 0.001), without obvious loss of correction at the last follow-up in either group. The bone fusion rates for single TMC and double TMCs were 77.3% and 76.5%, respectively.

Conclusion: Using 2 smaller TMCs instead of a single large one eases the placement of TMC by shortening the time and avoiding nerve impingement. Anterior column reconstruction with double TMC is a clinically feasible, and safe alternative following TES for thoracic and lumbar tumors.

Keywords: Spinal neoplasms, Surgical decompression, Instrumentation, Titanium alloy

INTRODUCTION

Vertebral tumors, though relatively rare, present a complex and challenging spectrum of conditions that significantly impact spinal health and overall well-being. These tumors, arising either within the spinal column (primary) or spreading from other parts of the body (metastatic), pose intricate clinical and therapeutic considerations owing to their potential for causing spinal instability, neurological compromise, and associated morbidity. In general, spinal tumor treatment imposes a substantial economic burden on both patients and society. The average cost of the hospital admission was estimated to be $55,801, with a 90-day readmission rate reaching up to 11.6%.

Total en bloc spondylectomy (TES) is intended to completely...
remove the tumor while preserving the integrity of the tumor border. It has been reported to decrease the local recurrence and prolong survival in appropriately selected patients. The operation involves the complete removal of the vertebral body and surrounding ligament structure, resulting in severe instability of the spine. Rigid reconstruction is of vital importance after tumor resection. Titanium mesh cage (TMC) is most commonly used for anterior column reconstruction. By filling the space left by the tumor, TMC restores vertebral height and helps maintain proper alignment, reducing the risk of spinal deformity or collapse. In the meanwhile, TMC serves as a container for bone graft material, promoting bone fusion and integration, aiding in the reconstruction of the affected vertebral segment.

In order to achieve solid stability, a large TMC filled with allografts or autografts is typically used to reconstruct the anterior column. Studies suggest that the most suitable diameter of a TMC equals the diameter of the lower endplate of the adjacent cephalad vertebra. However, the placement of a large TMC carries the risk of iatrogenic injury to the surrounding neural elements. Especially in the thoracic spine, this often requires significant traction or even ligation of unilateral nerve root to accommodate the TMC adequately. In contrast, while a cage of smaller diameter is safer to insert, it provides insufficient bone graft contact area, which will increase the risk of nonfusion, subsidence, or endplate fracture. Therefore, we adopted a different strategy, opting to use 2 smaller TMCs instead of a single large one for anterior reconstruction, aiming to minimize the risk of neural complications while ensuring adequate stability. The aim of this study is to evaluate the clinical efficacy and long-term safety of anterior reconstruction using double TMCs by comparing its outcomes with those associated with a single TMC.

MATERIALS AND METHODS

1. Patient Data

Between 2012 and 2022, a retrospective review was conducted on 51 patients diagnosed with thoracic or lumbar spinal tumors who underwent TES and anterior column reconstruction using TMC. 45 (88.2%) were occupied with single-level TES. Prior to 2018, our center routinely utilized a single large TMC for anterior reconstruction following TES. Since 2019, we have adopted a different strategy, opting to use 2 smaller TMCs instead of a single large one for anterior reconstruction, aiming to minimize the risk of neural complications while ensuring adequate stability. The inclusion criteria of this study include: (1) pathologically confirmed single-segment thoracic or lumbar spinal tumors (primary or metastatic), (2) patients receiving TES and anterior reconstruction using single or double TMC, and (3) patients who were followed for at least a year. The exclusion criteria were: (1) tumors located in the cervical or sacral spine, (2) recurrent tumors, (3) patients undergoing subtotal corpectomy or total piecemeal spondylectomy, and (4) patients who were lost to follow-up. Patients were divided into 2 groups based on the implant used. After screening, 39 patients were included in this retrospective cohort study, with written informed consent obtained from all patients or their legal guardians.

Data regarding the patient's age, sex, body mass index (BMI), and preoperative symptoms were obtained from a review of clinical notes. The severity of back pain was evaluated using the visual analogue scale (VAS), while the American Spinal Injury Association (ASIA) impairment scale was assessed for motor function. Operative duration, estimated blood loss, instrumentation method, and any complications associated with the surgery were obtained from operative notes. Plain radiographs were examined to evaluate anterior body height compression, kyphotic Cobb angle, mean intervertebral height (MIH), TMC subsidence or dislocation, bony fusion, and the presence of instrumentation failure. The exposure distance was universally set at 100 cm, with a current of 630 mA and a voltage of 80 kv. The exposure time varies among different patients, depending on factors such as the thickness of the chest wall and the inflation of the lungs, ranging from 32 msec to 46 msec.

This study was approved by the Institutional Review Board of General Hospital of Northern Theater Command (Y(2022)180) where the experiment was performed. A written informed consent was obtained from all patients or their legal guardians.

2. Measurement of Radiographic Parameters

- **MIH** was defined as the average of anterior intervertebral height (AIH) and posterior intervertebral height (PIH) on the sagittal plain radiograph. Whereas AIH was the distance between the most anterior point of the inferior edge of the upper vertebra and the superior edge of the lower vertebra, while PIH was the distance between the most posterior point of the inferior edge of the upper vertebra and the superior edge of the lower vertebra. **MIH** was calculated before surgery, immediately after operation, and at 1-year follow-up (Fig. 1).
- **Kyphotic Cobb angle** was measured as the angle between the superior endplate of the vertebral body above the affected level and the inferior endplate of the vertebral body below the affected level (Fig. 1).

https://doi.org/10.14245/ns.2448052.026
Anterior body height compression was calculated in patients with pathologic fracture according to the following formula: \( b_\alpha = (1-2^\alpha A_0)/(A_\beta + A_\delta) \times 100\% \), where \( b_\alpha \) was the percentage of anterior body height compression; \( A_0 \) was the anterior body height of the compromised vertebra; and \( A_\beta \) and \( A_\delta \) were the anterior body heights of the proximal and distal levels.\(^{10} \)

Bone fusion was evaluated according to the radiologic criteria of Bridwell et al.,\(^{12} \) whereas Bridwell I and II were indicated as satisfied fusion whereas mature bony trabeculae that bridged across the cage between the adjacent upper and lower endplates without radiolucent line were observed.

TMC subsidence was defined as the loss of the MIH more than 3 mm at each time point compared to the measurement taken preoperatively.\(^{11} \)

TMC oblique was defined as a mismatch of > 10° between the adjacent endplates (or osteotomy planes) on serial postoperative radiographs.\(^{13} \)

Other internal fixation failures include loosening, pulling out, or breakage of the screws and rods, as well as TMC dislodgement.

All measurements were independently performed by 2 authors (AL and JL). If there was no significant difference, the values were averaged. In cases of significant difference, a third author was consulted to perform additional measurements and make a final determination (Supplementary Table 1).

3. Surgical Procedure

All surgeries were performed by the same group of surgeons at our center. The patients were placed in the prone position following administration of general anesthesia. A posterior midline incision was made over the affected level, and the lamina, facet joints, and transverse processes were meticulously exposed and visualized. Poly-axial pedicle screws were placed in 2 segments above and below the index vertebra.\(^{14} \) The diameter and length of pedicle screws were initially measured using preoperative radiographs and adjusted during surgery with the assistance of intraoperative fluoroscopy C-arm. A temporary rod was placed on the less affected side to stabilize the spine, reducing the risk of spinal cord injury due to instability during decompression. The surgical procedure involved blunt dissection to separate the vertebral body and anterior structures. Either a thread-wire saw or an ultrasonic osteotome was used to cut off bilateral vertebral pedicles, without sacrifice of the nerve root, followed by the extraction of posterior structures. Adjacent intervertebral discs and endplates were dissected before the vertebral body was rotated and removed in one piece. After tumor resection, a vernier caliper was used to measure the distance between the lower endplate of the superior vertebrae body and the upper endplate of the inferior vertebrae body. TMCs were then appropriately sized and packed with morselized artificial bone graft made of β-tricalcium phosphate, without the application of growth factors. Two TMCs were inserted obliquely into the corpectomy defect via bilateral corridors without traction of the nerve roots or the dural sac. Two rods of Ti6Al4V (diameter, 6.0 mm) were connected to each pedicle screw and the posterior instrumentation was adjusted to slightly compress the inserted cage. Distilled water was applied to the surgical site for 5 minutes before closure and the placement of drainage catheters. In all cases, a rigid spinal brace was used for a postoperative period of at least 3 months. Follow-up evaluation was performed at 3 months, 6 months, and 1 year postoperatively with plain radiographs, computed tomography (CT), ASIA classification, and complication assessment (Fig. 2).

4. Statistical Analysis

Statistical analyses were conducted using IBM SPSS Statistics.
ver. 22.0 (IBM Co., Armonk, NY, USA). Continuous variables were presented as mean ± standard deviation. The normality of distribution was assessed using the Shapiro-Wilk test, and the homogeneity of variances was evaluated using Levene test. Pillai’s Trace multivariate test was conducted following Mauchly test of sphericity. Categorical variables were presented as percentages. One-sided Fisher exact test was used to compare binary data between independent groups. Statistical significance is assumed at a p-value of < 0.05.

RESULTS

1. Preoperative Information
Among the 39 patients enrolled in our study, there were 20 males and 19 females, with an average age of 52.1 ± 13.4 years and a range of 23 to 76 years. Of these, 14 received treatment for primary spinal tumors, consisting of 5 chondrosarcomas, 4 invasive hemangiomas, 3 synovial sarcomas, and 2 fibromas. The remaining 25 patients were admitted for metastatic spinal tumors, and the tumor histology was as follows: 9 lymphomas, 3 prostate cancers, 4 breast cancers, 3 renal cancers, 5 lung cancers, and 1 multiple myeloma. The average age for patients with primary or metastatic diseases was 50.7 ± 12.9 years and 52.8 ± 13.7 years, respectively. In 12 patients (31%), tumors were located in the lumbar spine, while 27 patients (69%) had tumors in the thoracic spine. Sixteen patients developed pathologic fractures, with an average anterior body height compression of 36.3%. Preoperative evaluation using VAS indicates an average score of 6.5 ± 1.0 and 6.1 ± 0.8, respectively, for each group. At the time of admission, 11 patients (28.2%) displayed normal

Fig. 2. A 67-year-old female patient (case #33) was admitted with motor dysfunction and mild back pain. (A) Preoperative plain radiographs reveal compression of the eighth thoracic vertebra. (B–D) To further clarify the cause of vertebral compression and observe any neurological involvement, an enhanced magnetic resonance imaging (MRI) was ordered for the patient. Through MRI, we noticed abnormal long T1 and long T2 signals in the eighth and twelfth thoracic vertebrae, with significant enhancement on the contrasted sequence. Meanwhile, we noticed high signal intensity in corresponding thoracic spinal cord, and nodular high signal at the level of thoracic vertebrae 4–7, with significant enhancement on the contrasted sequence. (E) Through preoperative computed tomography, we noticed an eccentric lytic lesion without calcification inside the tumor. (F) A total en bloc spondylectomy was performed at the T8 level, while vertebroplasty was performed in T4 and T12. (G, H) Two appropriately sized titanium mech filled with morselized artificial bone graft were inserted via bilateral corridors without traction of the nerve roots or the dural sac. (I) Postoperative plain radiographs demonstrated satisfactory alignment and restoration of intervertebral height. (J–M) Follow-up evaluation at 1 year showed no instrumentation failure or local recurrence.
motor function, 24 patients retained partial motor function, and 5 patients were paralytic. There was no significant difference regarding the baseline information between the 2 groups (Tables 1 and 2).

2. Perioperative Findings

All patients underwent TES, anterior reconstruction with TMC, and posterior instrumentation. Pedicle screws were placed 2 levels above and below the excised vertebra. Only in 1 patient involving a tumor in the thoracolumbar junction area (L1) with severe osteoporosis, the instrumentation was extended to 6 segments. The average operation time and estimated blood loss showed no significant difference between the 2 groups. However, the time for TMC placement was significantly shortened in the double TMC group (5.2 ± 1.3 minutes vs. 15.6 ± 3.3 minutes, p = 0.004). Ten patients developed surgery-related complications, including wound infection in 2 patients, nerve root disturbance in 5 patients, pleural effusion in 2 patients, and dural tear in 1 patient. While none of the patients reconstructed with double TMC experienced nerve root disturbance, 5 out of 22 patients treated with single TMC experienced either transient numbness or a decline in muscle strength after surgery (p = 0.046). Symptomatic treatment including glucocorticoids, methylcobalamin, as well as physical rehabilitation were prescribed, and all patients achieved full recovery at last follow-up. For patients with localized primary disease, postoperative radiation therapy is not typically prescribed. In the meanwhile, all patients diagnosed with metastatic tumors underwent chemotherapy or targeted therapy as evaluated by an oncologist or a specialist.

3. Postoperative Evaluation

After a mean follow-up of 20.5 ± 4.6 months, all patients with motor impairment achieved complete or partial improvement, except for 1 patient with preoperative grade B who remained the same after surgery (Table 2). In the meanwhile, the average VAS score dropped significantly from 6.5 ± 1.0 to 2.1 ± 1.0, and 6.1 ± 0.8 to 2.4 ± 1.0 for patients reconstructed with single or double TMC, respectively. Although ASIA and VAS scores improved in the last follow-up in both groups compared with preoperative assessment, there was no significant difference between the 2 groups. The radiographic outcomes are summarized in Table 3. In patients with kyphosis secondary to pathological fractures, the average preoperative Cobb angle significantly improved from 19.1° ± 4.9° before surgery to 7.1° ± 2.0° immediately postoperatively (p < 0.001). This correction was sustained at 8.4° ± 2.0° during the final follow-up, with no evident loss of correction (p = 0.291). Notably, no significant difference was observed between the 2 groups (p = 0.291) (Supplementary Fig. 1A). Meanwhile, the MIH was 34.5 ± 11.5 mm preoperatively, showing a significant increase to 39.1 ± 8.3 mm immediately after surgery (p = 0.001), but decreased to 37.0 ± 8.5 mm at the

### Table 1. General information of patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>Single TMC (n = 22)</th>
<th>Double TMC (n = 17)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td></td>
<td></td>
<td>0.485</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>50.7 ± 14.2</td>
<td>53.8 ± 12.1</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>31−76</td>
<td>23−67</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td>0.517</td>
</tr>
<tr>
<td>Male</td>
<td>11 (50)</td>
<td>11 (64.7)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>11 (50)</td>
<td>6 (35.3)</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>24.4 ± 2.4</td>
<td>23.7 ± 2.2</td>
<td>0.358</td>
</tr>
<tr>
<td>Tumor histology</td>
<td></td>
<td></td>
<td>0.318</td>
</tr>
<tr>
<td>Primary</td>
<td>9 (40.9)</td>
<td>4 (23.5)</td>
<td></td>
</tr>
<tr>
<td>Metastatic</td>
<td>13 (59.1)</td>
<td>13 (76.5)</td>
<td></td>
</tr>
<tr>
<td>Tumor location</td>
<td></td>
<td></td>
<td>0.730</td>
</tr>
<tr>
<td>Thoracic</td>
<td>16 (72.7)</td>
<td>11 (64.7)</td>
<td></td>
</tr>
<tr>
<td>Lumbar</td>
<td>6 (27.3)</td>
<td>6 (35.3)</td>
<td></td>
</tr>
<tr>
<td>Presence of pathologic fracture</td>
<td>10 (45.5)</td>
<td>5 (29.4)</td>
<td>0.343</td>
</tr>
<tr>
<td>Preoperative VAS score</td>
<td>6.5 ± 1.0</td>
<td>6.1 ± 0.8</td>
<td>0.198</td>
</tr>
<tr>
<td>Postoperative VAS score</td>
<td>2.1 ± 1.0</td>
<td>2.4 ± 1.0</td>
<td>0.401</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>14 (63.6)</td>
<td>14 (82.4)</td>
<td>0.177</td>
</tr>
<tr>
<td>Total operative time (min)</td>
<td>299.6 ± 41.3</td>
<td>301.2 ± 48.7</td>
<td>0.913</td>
</tr>
<tr>
<td>Time for TMC placement (min)</td>
<td>15.6 ± 3.3</td>
<td>5.2 ± 1.3</td>
<td>0.004</td>
</tr>
<tr>
<td>Estimated blood loss (mL)</td>
<td>1,431.8 ± 552.4</td>
<td>1,647.1 ± 1,010.6</td>
<td>0.401</td>
</tr>
<tr>
<td>Nerve disturbance</td>
<td>5 (22.7)</td>
<td>0 (0)</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%) unless otherwise indicated.

TMC, titanium mesh cage; SD, standard deviation; BMI, body mass index; VAS, visual analogue scale.

### Table 2. Neurologic status evaluated by the ASIA impairment scale (AIS)

<table>
<thead>
<tr>
<th>AIS</th>
<th>Single TMC (n = 22)</th>
<th>Double TMC (n = 17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative</td>
<td>Final follow-up</td>
<td>Preoperative</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>E</td>
<td>6</td>
<td>14</td>
</tr>
</tbody>
</table>

ASIA, American Spinal Injury Association; TMC, titanium mesh cage.
last follow-up. Yet, there was no significant difference noted between the 2 groups (p = 0.351) (Supplementary Fig. 1B). Regarding subsidence, the average measured was 2.0 ± 1.5 mm, with a subsidence rate of 17.9% (n = 7) at the final follow-up (Table 4). No significant difference was found regarding subsidence between single TMC and double TMCs (p = 0.650). Meanwhile, the bone fusion rates for single TMC and double TMCs were 77.3% and 76.5%, respectively, showing no significant differences between the groups (p = 0.623). Notably, no hardware-related complication including loosening, pulling out, or breakage of the screws and rods, TMC dislodgement, or TMC oblique occurred in both groups.

### DISCUSSION

In spine surgery, vertebrectomy is a common procedure employed for addressing traumatic, infectious, or neoplastic conditions. Subsequent to vertebrectomy, anterior column reconstruction becomes imperative to maintain stability and reinstate weight-bearing function. Various options have emerged for anterior reconstruction, including bone grafts, bone cement, TMC, carbon fiber stackable cage, artificial vertebral body, etc. Among these, TMC is the most commonly used, since it offers superior stability and demonstrates better cost-effectiveness. The advantages of the TMC are that it provides robust structural support, varies in diameter and height, allows for more space for mercerized bone grafts, and offers exceptional strength while being lightweight. Additionally, the radiolucent property of titanium aids in postoperative imaging, enabling better visualization during follow-up assessments.

Despite these advantages, inserting a large, fixed-height TMC between vertebrae poses technical challenges and risks of impinging on the neural structures. In some instances, nerve roots are even sacrificed to accommodate larger cages for better anterior reconstruction. On the contrary, cages with smaller diameters may mitigate nerve injury but elevate the risk of postoperative subsidence. Based on a finite element analysis, the diameter of the TMC is correlated with its future stability, and the most suitable diameter for reconstruction equals 1/1 the diameter of the lower endplate of the adjacent cephalad vertebra. As a result, expandable cages have been developed and introduced to ease insertion. However, there remain some limitation. These include instrumentation failure related to overexpansion, especially in patients with osteopenia or osteoporosis. Moreover, they might induce stress-shielding and lack adequate space for

---

**Table 3. Postoperative evaluations on MIH and Cobb angle**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preoperative</th>
<th>Postoperative</th>
<th>Follow-up</th>
<th>Repeated measures ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIH</td>
<td></td>
<td></td>
<td>F-value</td>
</tr>
<tr>
<td>Single TMC</td>
<td>33.4 ± 10.9</td>
<td>37.8 ± 7.3</td>
<td>35.8 ± 7.6</td>
<td>11.8</td>
</tr>
<tr>
<td>Double TMC</td>
<td>35.8 ± 12.5</td>
<td>40.7 ± 9.4</td>
<td>38.7 ± 9.5</td>
<td></td>
</tr>
</tbody>
</table>

|              |               |               |           | Group | 0.9        | 0.351     | 0.024 |
|              |               |               |           | Time*group | 0.1      | 0.830     | 0.001 |

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preoperative</th>
<th>Postoperative</th>
<th>Follow-up</th>
<th>Repeated measures ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIH</td>
<td></td>
<td></td>
<td>F-value</td>
</tr>
<tr>
<td>Single TMC</td>
<td>17.8 ± 5.3</td>
<td>6.9 ± 2.1</td>
<td>8.2 ± 2.3</td>
<td></td>
</tr>
<tr>
<td>Double TMC</td>
<td>21.3 ± 3.5</td>
<td>7.6 ± 2.0</td>
<td>8.6 ± 1.8</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4. Postoperative evaluation on subsidence rate and bone fusion rate**

<table>
<thead>
<tr>
<th>Rate</th>
<th>Single TMC (n = 22)</th>
<th>Double TMC (n = 17)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subsidence rate</td>
<td>4 (18.2)</td>
<td>3 (17.6)</td>
<td>0.650</td>
</tr>
<tr>
<td>Bone fusion rate</td>
<td>17 (77.3)</td>
<td>13 (76.5)</td>
<td>0.623</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. MIH, mean intervertebral height; TMC, titanium mesh cage; ANOVA, analysis of variance. *p < 0.05, follow-up compared with preoperative.

Values are presented as number (%). TMC, titanium mesh cage.
bone grafting, potentially leading to pseudarthrosis. Furthermore, a significant drawback of expandable cages is its high cost. Therefore, in this study, we adopted a different strategy, opting to use 2 smaller TMCs instead of a single large one for anterior reconstruction, aiming to minimize the risk of neural complications while ensuring adequate stability.

Overall, all patients in this study experienced partial or complete improvement after decompression, with a significant decrease in the mean VAS score. Meanwhile, the time for TMC placement was significantly shortened with double TMC (5.2 ± 1.3 minutes vs. 15.6 ± 3.3 minutes, p = 0.004). In addition, 5 of 22 patients who received a single TMC experienced iatrogenic nerve root disturbances, displaying numbness and motor dysfunction after surgery. Conversely, among patients who received double TMCs, no nerve complications were observed. In addition to symptom alleviation, a primary concern is whether the use of double TMC can offer rigid stability, since the main drawback of TMC is its susceptibility to subsidence. According to Van Jonbergen et al., a reduction in postoperative intervertebral height exceeding 3 mm constitutes TMC subsidence. Based on their criteria, an incidence of 20%–40% has been reported in literature. In our study, 17.9% of patients developed subsidence, with an average of 2.0 ± 1.5 mm. However, there was no significant difference in the subsidence rate between the 2 groups (p = 0.65). This demonstrates the safety and feasibility of applying double TMC of smaller size. According to the study by Hou and Luo, the double TMC with off-center positions were closer to the cortical rim of the vertebral body, which was the strongest part of the endplate. Furthermore, even if subsidence does occur, its impact on bone fusion and clinical outcome remains controversial. The systematic review of Karikari et al. concluded that subsidence did not significantly affect successful fusion or clinical outcomes. Similarly, Yan et al. found no correlation between cage subsidence, clinical outcomes, sagittal alignment, or fusion rate. However, Matsumoto et al. identified cage subsidence > 5 mm as a risk factor for instrumentation failure. In our study, 2 out of 7 patients with subsidence > 3 mm developed chronic back pain, without neural function deterioration. Symptom relief was achieved through non-steroidal anti-inflammatory drugs, physical therapy, and bracelet protection.

Based on previous studies, some risk factors associated with TMC subsidence have been identified. Patient-related factors include a BMI > 28 kg/m², perioperative radiotherapy, and poor bone density. Instrumentation-related risk factors involve selection of an excessively long TMC, overdistraction during insertion, aggressive correction of spinal curvature, and positioning the TMC obliquely. Additionally, a finite analysis points out that a mismatch between TMC and adjacent endplates also affects the biomechanical properties and increases the risk of internal fixation failure. Therefore, to potentially prevent postoperative subsidence, measures such as avoiding overexpanding the intervertebral height, optimizing TMC placement, and initiating antosteoporosis treatments 6 months before surgery might be beneficial. However, due to the limited number of patients with subsidence in our study, we were unable to conduct a comprehensive Logistic analysis for risk factors. Nonetheless, we observed that patients who experienced subsidence had a significantly higher average age compared to those without subsidence (63.4 ± 7.7 vs. 49.6 ± 13.0, p = 0.01).

Certain limitations should be acknowledged. In this study, we exclusively performed TMC reconstruction for a single-level spondylectomy, the assessment of multilevel reconstruction was not performed. Additionally, although we compared baseline information between the 2 groups and found no significant difference, the retrospective nature of our study inherently introduces some selection bias. Meanwhile, the small number of patients and relatively short follow-up period of minimum 1-year limit our ability to conduct extensive statistical analysis. Furthermore, despite the superior accuracy of CT in assessing bone fusion, we opted for plain radiographs considering radiation exposure and financial burden. Additionally, since dual-energy x-ray absorptiometry tests are not routinely conducted in our center, we were unable to provide bone mineral density data for each patient. To enhance our understanding of TMC reconstruction, finite element analysis for biomechanical insights and validating these findings with larger clinical sample sizes would be beneficial.

CONCLUSION

Using 2 smaller TMCs instead of a single large one eases the placement of TMC by shortening the time and avoiding nerve impingement. Anterior column reconstruction with double TMC is a clinically feasible, and safe alternative treatment following TES for thoracic and lumbar tumors.

NOTES

Supplementary Materials: Supplementary Table 1 and Fig. 1 can be found via https://doi.org/10.14245/ns.2448052.026.
Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This work was supported by the Natural Science Foundation of Liaoning Province (Grant number: 2022-YGJC-09).

Author Contribution: Conceptualization: QW, AL; Formal analysis: AL, JL; Data curation: QW, JL; Methodology: YL, SS; Visualization: LM, MG; Project administration: HY, LX; Writing – original draft: AL; Writing – review & editing: HY, LX.

ORCID
Ao Leng: 0000-0001-9499-8286
Qi Wang: 0000-0002-0744-0880
Jiacheng Li: 0009-0000-7505-8870
Yu Long: 0000-0002-5817-8296
Song Shi: 0000-0002-0757-8142
Lingzhi Meng: 0000-0003-0027-7889
Mingming Guo: 0000-0002-2806-2196
Hailong Yu: 0000-0001-5442-3285
Liangbi Xiang: 0000-0002-9484-1710

REFERENCES
## Supplementary Table 1. Radiologic measurement of mean intervertebral height, Kyphotic Cobb angle, and bone fusion

<table>
<thead>
<tr>
<th></th>
<th>Preoperation</th>
<th>Postoperation</th>
<th>Follow-up</th>
<th>Preoperation</th>
<th>Postoperation</th>
<th>Follow-up</th>
<th>Bone fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 Final</td>
<td>1 2 Final</td>
<td>1 2 Final</td>
<td>1 2 Final</td>
<td>1 2 Final</td>
<td>1 2 Final</td>
<td></td>
</tr>
<tr>
<td>Mean intervertebral height</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.1</td>
<td>30.7</td>
<td>30.4</td>
<td>29.1</td>
<td>29.5</td>
<td>29.3</td>
<td>28.8</td>
<td>28.6</td>
</tr>
<tr>
<td>18.2</td>
<td>18.2</td>
<td>18.2</td>
<td>30.4</td>
<td>30.4</td>
<td>30.4</td>
<td>29.4</td>
<td>28.8</td>
</tr>
<tr>
<td>38.0</td>
<td>37.8</td>
<td>37.9</td>
<td>37.0</td>
<td>37.4</td>
<td>37.2</td>
<td>32.2</td>
<td>32.6</td>
</tr>
<tr>
<td>50.2</td>
<td>49.4</td>
<td>49.8</td>
<td>48.6</td>
<td>48.8</td>
<td>48.7</td>
<td>46.0</td>
<td>46.0</td>
</tr>
<tr>
<td>15.5</td>
<td>15.5</td>
<td>15.5</td>
<td>30.2</td>
<td>30.2</td>
<td>30.2</td>
<td>27.7</td>
<td>25.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kyphotic Cobb angle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.1</td>
<td>24.5</td>
<td>24.3</td>
<td>8.0</td>
<td>8.6</td>
<td>8.3</td>
<td>8.9</td>
<td>9.3</td>
</tr>
<tr>
<td>48.2</td>
<td>12.4</td>
<td>47.0</td>
<td>9.5</td>
<td>9.5</td>
<td>9.5</td>
<td>11.0</td>
<td>10.4</td>
</tr>
<tr>
<td>35.6</td>
<td>35.6</td>
<td>35.6</td>
<td>34.9</td>
<td>34.9</td>
<td>34.9</td>
<td>30.8</td>
<td>30.8</td>
</tr>
<tr>
<td>24.3</td>
<td>24.3</td>
<td>24.3</td>
<td>30.2</td>
<td>30.4</td>
<td>30.3</td>
<td>28.5</td>
<td>28.5</td>
</tr>
<tr>
<td>48.6</td>
<td>49.0</td>
<td>48.8</td>
<td>49.5</td>
<td>49.9</td>
<td>49.7</td>
<td>48.3</td>
<td>48.9</td>
</tr>
<tr>
<td>38.3</td>
<td>38.1</td>
<td>38.2</td>
<td>38.6</td>
<td>38.0</td>
<td>38.3</td>
<td>37.5</td>
<td>37.7</td>
</tr>
<tr>
<td>34.1</td>
<td>34.1</td>
<td>34.3</td>
<td>48.7</td>
<td>48.3</td>
<td>48.5</td>
<td>46.3</td>
<td>45.9</td>
</tr>
<tr>
<td>30.4</td>
<td>30.4</td>
<td>30.4</td>
<td>29.6</td>
<td>29.6</td>
<td>29.6</td>
<td>29.5</td>
<td>29.5</td>
</tr>
<tr>
<td>23.4</td>
<td>23.4</td>
<td>23.4</td>
<td>29.6</td>
<td>30.0</td>
<td>29.8</td>
<td>28.8</td>
<td>28.2</td>
</tr>
<tr>
<td>22.3</td>
<td>22.3</td>
<td>22.8</td>
<td>28.8</td>
<td>29.0</td>
<td>28.9</td>
<td>28.9</td>
<td>28.9</td>
</tr>
<tr>
<td>29.4</td>
<td>30.1</td>
<td>29.8</td>
<td>37.8</td>
<td>37.6</td>
<td>37.7</td>
<td>32.6</td>
<td>32.4</td>
</tr>
<tr>
<td>16.2</td>
<td>16.5</td>
<td>16.4</td>
<td>38.1</td>
<td>38.1</td>
<td>38.1</td>
<td>37.7</td>
<td>37.3</td>
</tr>
<tr>
<td>21.6</td>
<td>22.1</td>
<td>21.9</td>
<td>45.5</td>
<td>44.9</td>
<td>45.2</td>
<td>40.6</td>
<td>40.6</td>
</tr>
<tr>
<td>45.0</td>
<td>45.0</td>
<td>45.0</td>
<td>46.8</td>
<td>46.8</td>
<td>46.8</td>
<td>46.3</td>
<td>46.3</td>
</tr>
<tr>
<td>19.2</td>
<td>19.8</td>
<td>19.5</td>
<td>44.4</td>
<td>44.4</td>
<td>44.2</td>
<td>41.1</td>
<td>41.7</td>
</tr>
<tr>
<td>16.8</td>
<td>17.1</td>
<td>17.0</td>
<td>29.3</td>
<td>29.5</td>
<td>29.4</td>
<td>27.0</td>
<td>26.2</td>
</tr>
<tr>
<td>24.2</td>
<td>24.2</td>
<td>24.2</td>
<td>30.5</td>
<td>30.5</td>
<td>30.5</td>
<td>25.0</td>
<td>25.4</td>
</tr>
<tr>
<td>35.4</td>
<td>35.6</td>
<td>35.4</td>
<td>35.4</td>
<td>35.4</td>
<td>35.4</td>
<td>34.0</td>
<td>34.0</td>
</tr>
<tr>
<td>30.3</td>
<td>30.5</td>
<td>30.4</td>
<td>37.3</td>
<td>38.3</td>
<td>38.0</td>
<td>36.5</td>
<td>35.9</td>
</tr>
<tr>
<td>54.4</td>
<td>53.8</td>
<td>54.1</td>
<td>59.1</td>
<td>59.1</td>
<td>59.1</td>
<td>57.7</td>
<td>57.1</td>
</tr>
<tr>
<td>43.1</td>
<td>43.1</td>
<td>43.1</td>
<td>44.0</td>
<td>44.0</td>
<td>44.0</td>
<td>42.5</td>
<td>42.5</td>
</tr>
<tr>
<td>46.7</td>
<td>46.3</td>
<td>46.5</td>
<td>45.6</td>
<td>46.0</td>
<td>45.8</td>
<td>44.4</td>
<td>44.6</td>
</tr>
<tr>
<td>44.5</td>
<td>44.5</td>
<td>44.5</td>
<td>47.0</td>
<td>47.0</td>
<td>47.0</td>
<td>46.2</td>
<td>46.4</td>
</tr>
<tr>
<td>31.0</td>
<td>30.5</td>
<td>30.8</td>
<td>29.3</td>
<td>28.9</td>
<td>29.1</td>
<td>28.9</td>
<td>28.9</td>
</tr>
<tr>
<td>39.1</td>
<td>39.3</td>
<td>39.2</td>
<td>40.4</td>
<td>40.6</td>
<td>40.5</td>
<td>37.7</td>
<td>37.9</td>
</tr>
<tr>
<td>42.0</td>
<td>42.6</td>
<td>42.3</td>
<td>43.0</td>
<td>43.0</td>
<td>43.0</td>
<td>41.6</td>
<td>42.0</td>
</tr>
<tr>
<td>35.3</td>
<td>35.3</td>
<td>35.3</td>
<td>35.8</td>
<td>35.6</td>
<td>35.7</td>
<td>32.4</td>
<td>33.2</td>
</tr>
<tr>
<td>50.2</td>
<td>50.2</td>
<td>50.2</td>
<td>48.5</td>
<td>48.1</td>
<td>48.3</td>
<td>47.3</td>
<td>47.3</td>
</tr>
<tr>
<td>38.5</td>
<td>38.9</td>
<td>38.7</td>
<td>39.6</td>
<td>40.0</td>
<td>39.8</td>
<td>38.2</td>
<td>38.2</td>
</tr>
<tr>
<td>16.3</td>
<td>16.3</td>
<td>16.3</td>
<td>28.2</td>
<td>27.8</td>
<td>28.0</td>
<td>23.6</td>
<td>23.4</td>
</tr>
<tr>
<td>54.5</td>
<td>54.3</td>
<td>54.4</td>
<td>53.3</td>
<td>53.3</td>
<td>53.3</td>
<td>52.3</td>
<td>51.7</td>
</tr>
<tr>
<td>32.0</td>
<td>32.6</td>
<td>32.3</td>
<td>37.3</td>
<td>37.9</td>
<td>37.6</td>
<td>35.2</td>
<td>35.2</td>
</tr>
<tr>
<td>45.1</td>
<td>45.1</td>
<td>45.1</td>
<td>44.6</td>
<td>44.0</td>
<td>44.3</td>
<td>44.0</td>
<td>43.6</td>
</tr>
</tbody>
</table>
Supplementary Fig. 1. Results of repeated measures analysis of variance. (A) The average preoperative Cobb angle significantly improved from 19.1° ± 4.9° preoperatively to 7.1° ± 2.0° immediately postoperatively and was maintained at 8.4° ± 2.0° at the last follow-up, without obvious loss of correction (p < 0.001). No significant difference was observed between the 2 groups (p = 0.291). (B) The mean intervertebral height was 34.5 ± 11.5 mm preoperatively and significantly increased to 39.1 ± 8.3 mm immediately after surgery, but decreased to 37.0 ± 8.5 mm at the last follow-up (p = 0.001). However, no significant difference was noticed between the 2 groups (p = 0.351). TMC, titanium mesh cage.
The Quantitative Evaluation of Automatic Segmentation in Lumbar Magnetic Resonance Images

Objective: This study aims to overcome challenges in lumbar spine imaging, particularly lumbar spinal stenosis, by developing an automated segmentation model using advanced techniques. Traditional manual measurement and lesion detection methods are limited by subjectivity and inefficiency. The objective is to create an accurate and automated segmentation model that identifies anatomical structures in lumbar spine magnetic resonance imaging scans.

Methods: Leveraging a dataset of 539 lumbar spinal stenosis patients, the study utilizes the residual U-Net for semantic segmentation in sagittal and axial lumbar spine magnetic resonance images. The model, trained to recognize specific tissue categories, employs a geometry algorithm for anatomical structure quantification. Validation metrics, like Intersection over Union (IOU) and Dice coefficients, validate the residual U-Net's segmentation accuracy. A novel rotation matrix approach is introduced for detecting bulging discs, assessing dural sac compression, and measuring yellow ligament thickness.

Results: The residual U-Net achieves high precision in segmenting lumbar spine structures, with mean IOU values ranging from 0.82 to 0.93 across various tissue categories and views. The automated quantification system provides measurements for intervertebral disc dimensions, dural sac diameter, yellow ligament thickness, and disc hydration. Consistency between training and testing datasets assures the robustness of automated measurements.

Conclusion: Automated lumbar spine segmentation with residual U-Net and deep learning exhibits high precision in identifying anatomical structures, facilitating efficient quantification in lumbar spinal stenosis cases. The introduction of a rotation matrix enhances lesion detection, promising improved diagnostic accuracy, and supporting treatment decisions for lumbar spinal stenosis patients.

Keywords: Automatic segmentation, Lumbar spine, Magnetic resonance imaging, Residual U-Net, Spinal stenosis
INTRODUCTION

Degenerative lumbar spondylolisthesis or degenerative disc disease are potential causes of lumbar spinal stenosis. This condition may manifest with clinical symptoms like low back pain, sciatica, and neurological claudication, restricting movement and disrupting daily posture. While medication or physical therapy proves effective for most patients, surgery becomes necessary for those unresponsive to conservative treatments. Surgical intervention involves decompression and stabilization to address instability. Preoperative assessment for surgical indications and outcomes relies on radiological examinations, including magnetic resonance imaging (MRI) scans.

Degenerative changes in lumbar regions, as revealed by MRI scans, encompass disc height reduction, osteophyte formation, and degenerative disc herniation, constituting over 90% of lumbar spinal central and lateral stenosis cases. The integrity of intervertebral discs, comprised of the nucleus pulposus, ring-shaped cartilaginous endplates, and collagenous annulus-fibrosis layers, plays a pivotal mechanical role. The extracellular matrix in these discs manages tensile strength and osmotic pressure, facilitating load transmission across the spine column in response to body weight and daily activities, but this diminishes with age. Intervertebral disc degeneration, exacerbated by accumulated compressive loads, is implicated in lumbar spinal stenosis observed in MRI scans. T2-weighted MRI emerges as a potent tool for detecting morphologic changes in intervertebral disc degeneration, including height loss and water-intensity loss. Degeneration at this level may correlate with adjacent endplate degeneration. Signal intensity changes in MRI scans appear to signify a spectrum of vertebral body marrow changes associated with lumbar degenerative disease. In the previous studies, intervertebral disc could be measured in quantitative parameters that showed a greater ability to reflect the aging effect of degeneration. While prior studies employed quantitative parameters to measure intervertebral discs, limitations like manual segmentation, feature point marking, subjectivity, and human eye capacity hindered consistency and efficiency. Consequently, treatment decisions for lumbar spinal stenosis patients still rely largely on self-reports and physical examinations. Advances in clinical imaging techniques promise more robust measurements, overcoming the drawbacks of time-consuming processes, labor intensity, and the need for specialized domain knowledge.

To address the challenges posed by clinically evaluated lumbar spinal stenosis, advanced spine indices measurement methods employing segmentation-based approaches and deep learning algorithms were employed. Traditionally, semisegmentation software was conventionally used to separate different spinal tissues and measure anatomical structures. However, these manual processes proved time-consuming, labor-intensive, and required specialized expertise. The integration of deep learning algorithms, particularly convolutional neural networks (CNNs), introduced automated segmentation for identifying lumbar spine anatomical structures. Utilizing a 3D CNN on 23 patients MRI, an automatic supervised segmentation approach for vertebral body formation was developed. The accuracy and superiority of MRI-based 3-dimensional (3D)-CNN images were compromised without the detection of soft tissue such as intervertebral discs, dural sac, and yellow ligaments. In a separate study, a fully automated cervical segmentation architecture utilized a deep, fully connected CNN to reduce misdiagnosis. The centroid of the vertebrae was located using probabilistic spatial regression, achieving a Dice similarity coefficient of 0.84 and a shape error of 1.69 mm. To extend segmentation to all vertebrae, including cervical, thoracic, and lumbar spines, a 3D U-Net and deconvolution network were employed in an iterative segmentation model. Cross-entropy was applied as the loss function for multi-label classification. Midsagittal diameter and cross-sectional area of the dural sac and intervertebral discs were automatically determined using a CNN with cross-space distance-preserving regularization, yielding reliable performance (mean absolute error range: 1.04 ± 0.09 mm to 3.54 ± 0.28 mm).

Furthermore, a pretrained U-Net-based CNN network, incorporating a region proposal network and postprocessing with a thresholding method for distinguishing fat and muscular tissue, demonstrated the applicability of pretrained networks for limited image datasets. It produced favorable results for segmenting paraspinal muscles and exhibited excellent Intersection over Union (IOU) and differential scanning calorimetry for segmenting intervertebral discs. While CNNs have become widely utilized in medical image recognition and detection, the complexity of optimization increases with depth. To mitigate these issues, residual convolution was introduced, leading to the development of a residual U-Net.

Despite the time saved by automated segmentation, manual landmarks annotation remained a time-consuming aspect. Many methods focused on landmarks detection in spine indices measurement, using forecasted heatmaps to construct landmarks from which spine indices were derived. These landmarks, identified as locations with the highest peak values in the forecasted heatmaps, facilitated the definition and calculation of the con-
tour and anatomy of intervertebral discs, vertebrae, dura sac, and lamina. Previous studies demonstrated various landmarks-based analyses.\textsuperscript{11,22} Although CNN models performed well in spine segmentation, there is a notable gap in evidence supporting the application of this technique in MRI with lesion detection, particularly for conditions such as lumbar spinal stenosis, to enable robust detection and support imaging diagnosis.

This study utilizes MRI images from patients with lumbar spinal stenosis to implement an automatic segmentation technique for model training. The primary objective of the model is to identify various anatomical structures such as the intervertebral disc, vertebral body, and yellow ligament. Furthermore, the model aims to provide global quantitative measurements for both normal and abnormal levels at each lumbar spine segment.

**MATERIALS AND METHODS**

1. **Data Preparation**

This study conducted a retrospective analysis of axial and sagittal T2-weighted/T1-weighted (T2W/T1W) MRI studies at a single medical center. The inclusion criteria encompassed patients undergoing lumbar spine MRI due to lumbar spinal stenosis, while individuals under the age of 20, those with severe artifacts in lumbar MRI, vertebral fractures, postsurgery conditions, and those with trauma or malignancy were excluded from the dataset. MRI images were acquired in 3-Tesla scanner (MR750, GE Medical Systems, Milwaukee, WI, USA).

To formulate the lumbar spine semantic segmentation model, the dataset underwent division into training, validation, and testing sets. On average, labeling each slice takes between 10 to 15 minutes, posing a significant challenge to this process. To streamline this, 3 professional clinicians were enlisted to annotate the ground truth images using the open-source software ITK-SNAP.\textsuperscript{23} We utilized the ‘polygon’ and ‘paintbrush’ features within ITK-SNAP for this purpose. The polygon tool allowed for shape editing by adjusting vertices on the image, while the paintbrush tool facilitated quick drawing and refinement using mouse input, accommodating masks of various shapes and sizes. To reduce variability in MRI scans, we adjusted the contrast and intensity of images to ensure accurate delineation of each anatomical structure. The annotation process involved identifying key anatomical structures such as the intervertebral disc, vertebral body, and dural sac in sagittal views (Fig. 1A), as well as the intervertebral disc, dural sac, lamina, and yellow ligament in axial views (Fig. 1B).

To identify the appropriate slices in axial view, a dual-view setup within ITK-SNAP was utilized to ensure precise alignment of corresponding locations across axial and midsagittal views. The selected axial slice of the disc was defined as the one cutting closest to the half-height of the disc in sagittal view. Five appropriate axial slices were extracted, each from the lumbar level discs: L1/L2, L2/L3, L3/L4, L4/L5, and L5/S1 in the axial MRI scan.

The study was performed according to the Helsinki Declaration (http://www.wma.net/en/30publications/10policies/b3/) and approved by the Institutional Review Board (IRB) of Taipei Veterans General Hospital (IRB No. 2023-01-019AC) where the experiment was performed.

2. **The Architecture of the Residual U-Net**

The lumbar spine tissue semantic segmentation task employed the residual U-Net, whose architecture is depicted in Fig. 2A. The encoding phase was initiated by inputting 320 × 320 pixel lumbar spine T1W MR images into the residual U-Net featuring a sequential series of a convolution layer, followed by a batch normalization layer with the Rectified Linear Unit (ReLU) activating function, and a second convolution layer with an identity map addition. Two residual units followed this, each comprising 2 stacks of batch normalization layers with ReLU activating function, followed by a convolution layer and an identity map addition.

Subsequently, the encoding features traversed the bridge compartment before entering the decoding phase. The decoding
phase comprised 3 stacks of up-sampling layers, concatenate layers, and residual units, with additional corresponding skip connections to mitigate gradient vanishing during deep network training. The output segmentation mask results were classified into 4 classes (intervertebral disc, dural sac, ligamentum flavum, and lamina) in axial view lumbar spine MRIs, and 3 classes (intervertebral disc, vertebral body, and dural sac) in sagittal view lumbar magnetic resonance (MR) images.

The residual U-Net model underwent training with a batch size of 10 for 2,000 epochs, utilizing the Adadelta optimizer with a learning rate of 0.001. Training data were augmented with adjustments to brightness, contrast, rotation, and translation to enhance model generalization. TensorFlow (https://www.tensorflow.org) was used to implement all networks, and the codes executed on a server equipped with an RTX 2080 Ti GPU.

After automatic segmentation, we identified anatomical structures in both sagittal and axial views. Fig. 2B summarizes the detected lesions within each of these anatomical structures.

### 3. The Measurement of Performance Evaluations

Within the domain of supervised learning prediction, particularly in the context of deep learning methodologies, the proficiency of the model under scrutiny was meticulously evaluated across both training and testing datasets through performance assessments. This evaluation entailed a detailed comparative analysis between the predicted output image and the ground truth image, aiming to precisely identify and measure the discrepancies between the predicted and true labeled regions. To facilitate a nuanced and intuitive gauge of the segmentation quality, the Dice coefficient score was employed as a pivotal metric. Dice coefficient score, serving as a quantifiable measure, computes the percentage of overlap between 2 images on a scale from 0 to 1, where a Dice coefficient score of 1 denotes a perfect and complete overlap between the segmented output and the ground

---

**Fig. 2.** The schema of automatic spinal measurement. (A) The flow chart of automatic segmentation according to residual U-Net was illustrated. (B) After automatic segmentation, anatomical structures in both sagittal and axial views were identified to detect lesions within each of these anatomical structures. MRI, magnetic resonance imaging.
Traditionally, the Dice coefficient score has found extensive utility in the evaluation of models within the realm of medical imaging, establishing itself as a standard metric for assessing the efficacy of medical image segmentation endeavors. The Dice coefficient effectively quantifies the extent of spatial overlap between the predicted segmentation \( A \) and the ground truth segmentation \( B \), providing a critical insight into the performance of segmentation. Equation 1 defines the formula to get the Dice coefficient score,

\[
\text{Dice coefficient score } (A, B) = \frac{2|A \cap B|}{|A| + |B|}\tag{1}
\]

where \(||\) represents the number of voxels. The overlap values were between 0 and 1, where 1 denotes an identical pair of masks and 0 indicates no match between the two. Such quantification not only underscores the precision of the segmentation but also encapsulates the ability of model to accurately demarcate and identify relevant anatomical features within MR images.

The IOU also plays a pivotal role in quantifying the accuracy of segmentation models performance. IOU, also referred to as the Jaccard index, is a measure used to evaluate the extent of overlap between the predicted segmentation and the ground truth, offering a clear and concise quantification of predicted performance in delineating targeted anatomical structures within MR images. This metric calculates the ratio of the intersection of the predicted and true segments to their union, providing values that range from 0, indicating no overlap, to 1, signifying perfect congruence between the predicted and actual images. Equation 2 demonstrated the formula to get the IOU index.

\[
\text{IOU } (A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}\tag{2}
\]

### 4. The Measurement of Intervertebral Disc Geometry Index

The computer vision method identified landmarks of the intervertebral disc to measure the geometry of each lumbar spine intervertebral disc (Fig. 3A). In the sagittal view of lumbar MR images, the leftmost and rightmost points of the intervertebral disc were identified to calculate the width \( W_{\alpha} \), Fig. 3B). The height \( H_{\alpha} \), Fig. 3B) was defined as the points intersected with the border of the intervertebral disc, using the normal vector of the line connecting the leftmost point, rightmost point, and intervertebral disc centroid. The intervertebral disc area was estimated based on the segmented pixel results, while the volume was integrated into the area in each slice.

Lumbar alignment was measured by detecting each vertebral structure. Spinal listhesis was detected by measuring the distance \( L \), Fig. 3C) between 2 reference lines derived from the edges of the upper and lower vertebrae in the sagittal view. Reference lines were derived from the axial images on a scale from 0 to 1, where a value of 0 indicates no overlap, to 1, signifying perfect congruence between the predicted and actual images. The overlap values were between 0 and 1, where 1 denotes an identical pair of masks and 0 indicates no match between the two. Such quantification not only underscores the precision of the segmentation but also encapsulates the ability of model to accurately demarcate and identify relevant anatomical features within MR images.

The computer vision method identified landmarks of the intervertebral disc to measure the geometry of each lumbar spine intervertebral disc (Fig. 3A). In the sagittal view of lumbar MR images, the leftmost and rightmost points of the intervertebral disc were identified to calculate the width \( W_{\alpha} \), Fig. 3B). The height \( H_{\alpha} \), Fig. 3B) was defined as the points intersected with the border of the intervertebral disc, using the normal vector of the line connecting the leftmost point, rightmost point, and intervertebral disc centroid. The intervertebral disc area was estimated based on the segmented pixel results, while the volume was integrated into the area in each slice.

Lumbar alignment was measured by detecting each vertebral structure. Spinal listhesis was detected by measuring the distance \( L \), Fig. 3C) between 2 reference lines derived from the edges of the upper and lower vertebrae in the sagittal view. Reference lines were derived from the axial images on a scale from 0 to 1, where a value of 0 indicates no overlap, to 1, signifying perfect congruence between the predicted and actual images. The overlap values were between 0 and 1, where 1 denotes an identical pair of masks and 0 indicates no match between the two. Such quantification not only underscores the precision of the segmentation but also encapsulates the ability of model to accurately demarcate and identify relevant anatomical features within MR images.
Intervertebral disc herniation was quantified by measuring the diameter of the dural sac ($D_a$) in sagittal and axial views. In the sagittal view (Fig. 3D), the diameters of the corresponding dural sacs of each lumbar intervertebral disc were calculated. $D_a$ was determined based on the distance between the intersection of the normal vector passing through the intervertebral disc centroid and the dural sac. The compression of the dural sac caused by the bulging disc was defined by the distance between the line of posterior endplates and the edge of the dural sac ($\text{Dist}$, Fig. 3C). In the axial view (Fig. 3E), $D_a$ with the corresponding lumbar intervertebral disc was evaluated with the intersected length with the dural sac and the line connecting the intervertebral disc centroid and the bottom of the lamina. To assess noncentral line herniated discs, a sweep of ±30° measuring the dural sac diameter was also applied (Fig. 3F). Whether the herniated disc occurred other than the central line herniated could be verified by the decreased dural sac diameter and the corresponding degree of the site. Furthermore, the thickness of yellow ligament was also measured (Fig. 3G). The yellow ligament was split into middle ($\text{YL}_m$) and lateral ($\text{YL}_l$) part. To locate the yellow ligament precisely, the $\text{YL}_m$ was defined as the length of intersection between dural sac centroid and the lowest point of yellow ligament. As the $\text{YL}_m$ was specify, the line connecting dural sac centroid and the lowest point was considered as the reference and swept ±30° to measure the thickness of yellow ligament. The $\text{YL}_l$ was determined as the longest distance between dural sac centroid and the intersection points of yellow ligament.

5. The Measurement of Disc Hydration

The degree of intervertebral disc degeneration is closely tied to the loss of water content in the nucleus pulposus. To assess this degeneration, we calculated the signal intensity difference (ΔSI) within intervertebral disc areas. Both T1W/T2W MR images of the lumbar spine in sagittal view were utilized, obtaining intervertebral disc signal intensity through a pre-generated intervertebral disc mask created by the residual U-Net model. The T1W/T2W pixel intensities of the intervertebral disc served as features for the K-means model, facilitating clustering into hydrated and dehydrated groups. The centroids of these groups were identified, and their distance calculated. Additionally, the mean intensity of cerebrospinal fluid (CSF) was computed as a normalized factor to minimize signal intensity variation under different MR imaging conditions. Equation 3 expresses the formula for ΔSI between the centroids.

$$\Delta SI = \frac{SI_{\text{hydration}} - SI_{\text{dehydration}}}{SI_{\text{CSF}}}$$

Where $SI_{\text{CSF}}$ stands for the signal intensity corresponding to the distance of origin and cluster center for the CSF area, $SI_{\text{hydration}}$ and $SI_{\text{dehydration}}$ were the distance of origin between cluster center of hydration and dehydration, respectively.

6. Statistical Analysis

In the course of statistical analysis, we employed confidence intervals (CIs) to interpret the observed results. Setting a significance level of 0.05, these intervals offer an estimated range. Within this range, it can be asserted with 95% confidence that the true population parameter exists. CIs act as a robust metric for elucidating the statistical uncertainty linked to the estimates, offering a nuanced comprehension of the precision and reliability of experimental findings.

RESULTS

1. Demography of Data

A total of 539 patients were included from the Taipei Veterans General Hospital. The dataset consisted of 268 males and 271 females. The age distribution was 59.4 ± 18.9 years old (range, 43–78 years) in males, and 65.3 ± 14.3 years old (range, 46–81 years) in females. The training data set consisted of 207 males and 171 females, while 32 males and 21 females were included for validation dataset at training procedure. To evaluate the performance of the trained model, a testing dataset consisting of 56 males and 52 females was used.

2. The Performance of Segmentation by Using Residual U-Net

To address challenges related to gradient explosion, vanishing gradients, and optimization complexity, the residual U-Net was developed. Additionally, skip connections were employed to mitigate overfitting, recognized as a more effective strategy than randomly deactivating units. In this study, the residual U-Net model was chosen for the semantic segmentation task on lumbar spine MR images. The sagittal lumbar spine MR images involved the segmentation of 3 tissue categories: intervertebral disc, vertebra, and dural sac. To validate the segmentation results, Table 1 reports the mean IOU, mean Dice coefficient, and their 95% CIs. The mean IOU values were 0.91 for...
intervertebral disc, 0.93 for vertebra, and 0.93 for dural sac. The corresponding mean Dice coefficients were 0.91, 0.93, and 0.87, respectively.

In axial view lumbar MRIs, segmentation involved 4 tissue categories: intervertebral disc, lamina, dural sac, and yellow ligament. The overall mean IOU values were 0.82 for intervertebral disc, 0.85 for lamina, 0.82 for dural sac, and 0.83 for yellow ligaments. The mean Dice coefficients were 0.82 for intervertebral disc, 0.84 for lamina, 0.82 for dural sac, and 0.83 for the yellow ligaments.

3. The Quantification of Anatomical Structure in Lumbar Spine

The automated quantification system for lumbar spine anatomical structures was utilized for measurements following tissue segmentation by the residual U-Net model. In the sagittal view, the dimensions of the intervertebral disc height ($H_{dc}$) and width ($W_{dc}$) were individually measured at each level, with mean values of 10.8 mm and 33.1 mm, respectively. The diameter of the dural sac ($D_{ds}$) was also measured, yielding an average of 10.6 mm. Sagittal view calculations were also performed to assess disc hydration, resulting in an average hydration level of 55.4%. In the axial view, measurements included the central line diameter ($D_{cs}$) and middle thickness ($Y_{Lm}$) and lateral thickness of the yellow ligament ($Y_{Ll}$), with average values of 10.7 mm, 2.1 mm, and 2.5 mm, respectively (Table 2). A comparison between the training and testing datasets revealed no significant differences in anatomical measurements, as indicated by p-values between each pair of groups ($p > 0.05$).

4. Case Illustration

A 65-year-old man experienced persistent back pain radiating to both lower limbs for several months. Despite attempts at conservative treatment, a lumbar MRI scan was conducted, revealing a bulging disc at the L4–5 levels resulting in lumbar spinal stenosis. The axial and sagittal images from the scan were utilized as testing data in an automatic segmented model. In the sagittal view, the compression of the dural sac (Dist) at L4–5 was measured at 5.3 mm. The disc hydration at the index level was 18%, indicating a value below the normal range (Fig. 4A). Additionally, the diameter of the dural sac ($D_{ds}$) at L4–5 was significantly smaller than at other levels (Fig. 4B). In the axial view, the rotation matrix displayed a smooth inverted U-shaped curve at the normal level (Fig. 4C), which deteriorated at the L4–5 level, suggesting lumbar spinal stenosis with dural sac compression (Fig. 4C). Measurements indicated that the yellow ligament was 2.5 mm thick.

Table 1. Performance of segmentation in anatomical structures

<table>
<thead>
<tr>
<th>Segmented structure</th>
<th>Mean IOU</th>
<th>95% CI</th>
<th>Mean Dice coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagittal view</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc</td>
<td>0.91</td>
<td>0.87–0.95</td>
<td>0.91</td>
</tr>
<tr>
<td>Vertebral body</td>
<td>0.93</td>
<td>0.89–0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>Dural sac</td>
<td>0.93</td>
<td>0.89–0.97</td>
<td>0.87</td>
</tr>
<tr>
<td>Axial view</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disc</td>
<td>0.82</td>
<td>0.80–0.84</td>
<td>0.82</td>
</tr>
<tr>
<td>Lamina</td>
<td>0.85</td>
<td>0.83–0.87</td>
<td>0.84</td>
</tr>
<tr>
<td>Dural sac</td>
<td>0.82</td>
<td>0.80–0.84</td>
<td>0.82</td>
</tr>
<tr>
<td>Yellow ligament</td>
<td>0.83</td>
<td>0.82–0.85</td>
<td>0.83</td>
</tr>
</tbody>
</table>

IOU, Intersection over Union; CI, confidence interval.

Table 2. Normal range of segmented structure in sagittal and axial views

<table>
<thead>
<tr>
<th>Variable</th>
<th>Level</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagittal view</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H_{dc}$ (mm)</td>
<td>L1–2</td>
<td>9.2–9.8</td>
</tr>
<tr>
<td>$W_{dc}$ (mm)</td>
<td>L2–3</td>
<td>10.0–10.8</td>
</tr>
<tr>
<td>Disc hydration (%)</td>
<td>L3–4</td>
<td>10.1–10.7</td>
</tr>
<tr>
<td>$D_{ds}$ (mm)</td>
<td>L4–5</td>
<td>11.7–12.3</td>
</tr>
<tr>
<td>$D_{cs}$ (mm)</td>
<td>L5–S1</td>
<td>11.4–11.8</td>
</tr>
<tr>
<td>Axial view</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_{cs}$ (mm)</td>
<td></td>
<td>10.8</td>
</tr>
<tr>
<td>$Y_{Lm}$ (mm)</td>
<td></td>
<td>33.1</td>
</tr>
<tr>
<td>$Y_{Ll}$ (mm)</td>
<td></td>
<td>55.4</td>
</tr>
</tbody>
</table>

The normal range was defined by the 95% confidential interval.

$H_{dc}$, disc height; $W_{dc}$, disc width; $D_{ds}$, diameter of dural sac; $Y_{Lm}$, medial thickness of yellow ligament; $Y_{Ll}$, lateral thickness of yellow ligament.
ligament thickness was 3.1 mm laterally and 3.9 mm medially. After quantitative analysis, it was concluded that the bulging disc at L4–5, coupled with a hypertrophic yellow ligament, contributed to lumbar spinal stenosis at the same level, aligning with clinical observations.

**DISCUSSION**

The lumbar degenerative disease encompasses disc degeneration, facet hypertrophy, and hypertrophic yellow ligaments, leading to lumbar spinal stenosis and associated clinical symptoms. Instabilities like retrolisthesis or spondylolisthesis can exacerbate these clinical conditions. Regarding disc degeneration, the Modic classification categorizes disc conditions into 4 types based on MRI scans: type 0 signifies a normal disc and vertebral body appearance; type I involves bone marrow edema within the vertebral body and hypervascularization; type II indicates fatty replacements of the red bone marrow within the vertebral body, and type III involves subchondral bone sclerosis. Additionally, disc-only degeneration can be assessed using the Pfirrmann grading scale, which categorizes discs into grade I (normal disc), grade II (inhomogeneous disc with normal height and clear nucleus/annulus distinction), grade III (inhomogeneous gray disc with a blurred border between the nucleus and annulus and normal to slightly decreased height), grade IV (inhomogeneous hypointense dark gray disc with significant height loss), and grade V (inhomogeneous black disc with disc space collapse). While these classification systems offer varied disc degeneration grading, they lack specific quantitative scales. The introduction of deep learning models allows for the quantitative evaluation of central and lateral lumbar stenosis. However, there remains a need for a comprehensive evaluation of each anatomical structure in the lumbar spine and the identification of the specific components causing lumbar spinal stenosis.

This study employed a combination of the residual U-Net and a geometry algorithm to establish an automatic system for the quantitative analysis of spinal anatomical structures. The residual U-Net’s segmentation capability effectively distinguished various labeled tissues, facilitating subsequent measurements. While numerous methods, such as the K-means clustering algorithm, deep CNN, and U-Net, have been utilized for medical segmentation in recent years, they often faced challenges in handling extensive hidden layers and larger images. These limitations were primarily attributed to computational resource constraints and the “vanishing gradient” problem. To address the vanishing gradient issue, the residual unit was introduced. Previous research has demonstrated the effectiveness of the residual U-Net in enhancing the performance of deep convolutional networks and addressing class imbalance issues when compared to traditional U-Net and other improved variants.

There were 2 different views as the input in this model. In the
Automatic Segmentation in Lumbar MR Images

Liang YW, et al.

In clinical practice, lumbar degeneration is assessed using various grading systems tailored to clinical presentations. The Meyerding classification grades spondylolisthesis in patients with lumbar instability via lateral radiographs. MRI examinations aid in evaluating disc degeneration by assessing morphological changes in intervertebral discs and endplates. Another retrospective radiologic study aimed to establish a qualitative grading system for lumbar spinal stenosis and determine its reliability and clinical relevance. Dural sac cross-sectional area differs significantly between symptomatic and asymptomatic individuals. The study introduced a 7-grade classification based on dural sac morphology observed on T2 axial MRI, involving 95 subjects. Results showed substantial intra- and moderate interobserver agreement, with surgical patients exhibiting smaller dural sac cross-sectional areas and a higher proportion of higher grades. Various factors contribute to lumbar degeneration and stenosis, including disc generation and facet joint and yellow ligament hypertrophy. While previous grading systems focused on single anatomical degeneration, this study compared a 3D-CNN model, suggesting that a global assessment across different anatomical structures could enhance clinical diagnosis. While experienced clinical experts may find neural network introduction limiting for diagnosis, it could aid residency training. Prior studies indicate neural networks assist trainees in identifying urgent findings and gaining diagnostic confidence. Especially in the field of neurosurgery, a deep learning artificial intelligence model significantly enhanced the diagnostic abilities of novice physicians when it came to identifying pediatric skull fractures on plain radiographs. Moreover, it effectively discerned the traits of hydrocephalus from computed tomography images of the brain, automating the analysis process for junior doctors. Overall, the proposed model offers a comprehensive assessment of lumbar spinal stenosis, benefiting both experienced experts and residents in training.

This study involved a retrospective review of MR images focusing on patients diagnosed with lumbar degenerative disease or lumbar spinal stenosis. In the data training phase, levels that had undergone surgery were excluded to establish the normal range for each anatomical structure. However, all images were retrospectively reviewed from patients who had undergone lumbar surgery, potentially introducing bias in case selection for data training. During data testing and validation, all levels were utilized to distinguish between normal and abnormal findings. To enhance clinical applicability, a grading system will be developed, employing a global evaluation of each anatomical structure to establish correlations with the severity of lumbar spinal stenosis. This system aims to support future surgical decision-making processes.

CONCLUSION

This study introduces an advanced approach to lumbar degenerative diseases, particularly lumbar spinal stenosis, utilizing a residual U-Net and deep learning algorithm for automated segmentation of lumbar spine MRI scans. The residual U-Net demonstrates high accuracy in semantic segmentation, yielding precise measurements of intervertebral discs, vertebral bodies, dural sac, lamina, and yellow ligaments. The proposed method, applied to a dataset of 539 lumbar spinal stenosis patients, showcases its potential for clinical use. Additionally, a novel algorithm...
employing a rotation matrix detects bulging discs and dural sac compression, offering valuable insights into pathology. The findings highlight the efficiency and reliability of this automated segmentation technique, providing a promising avenue for improving diagnostic accuracy and guiding treatment decisions in lumbar spinal stenosis.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This work is financially supported by Taiwan Food and Drug Administration under contract numbers of MOHW109-FDA-M-114-000747, MOHW110-FDA-D-114-000741, MOHW111-FDA-H-114-000741, MOHW112-FDA-H-114-000741, and MOHW113-FDA-D-113-000741.

Author Contribution: Conceptualization: YWL, TCL, YYC; Formal analysis: YTF, TCL, CRY, HWH; Data curation: YWL, CCC, HKC, CCK, THT, LYF, JCW, WCH; Methodology: YTF, TCL; Project administration: JCW, WCH, YYC, CHK; Writing – original draft: YWL, TCL, CHK; Writing – review & editing: JCW, CHK.

ORCID
Yao-Wen Liang: 0009-0008-0112-5338
Yu-Ting Fang: 0000-0001-7917-6874
Ting-Chun Lin: 0009-0004-4767-1749
Cheng-Ru Yang: 0009-0007-9545-3852
Chih-Chang Chang: 0000-0003-2785-0921
Hsuan-Kan Chang: 0009-0004-6893-187X
Chin-Chu Ko: 0000-0003-3061-1059
Tsung-Hsi Tu: 0000-0002-0550-5855
Li-Yu Fay: 0000-0004-0914-8776
Jau-Ching Wu: 0000-0002-6996-3409
Wen-Cheng Huang: 0000-0002-6996-3409
Hsiang-Wei Hu: 0009-0005-7297-2217
You-Yin Chen: 0000-0003-4869-3857
Chao-Hung Kuo: 0000-0001-5528-204X

REFERENCES


Original Article

Corresponding Author

Rujun Li

https://orcid.org/0000-0002-7831-5868

Department of Neurosurgery, The Second Affiliated Hospital of Soochow University, Soochow University, Sanxiang Road, Gusu District, Suzhou 215004, China
Email: newjun_li@163.com

Received: January 15, 2024
Revised: April 9, 2024
Accepted: April 17, 2024

A Nomogram for Predicting Overall Survival of Patients With Primary Spinal Cord Glioblastoma

Yao Wang¹, Qingchun Mu¹, Minfeng Sheng¹, Yanming Chen¹, Fengzeng Jian², Rujun Li¹

¹Department of Neurosurgery, The Second Affiliated Hospital of Soochow University, Soochow University, Suzhou, China
²Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Beijing, China

Objective: Primary spinal cord glioblastoma (PSCGBM) is a rare malignancy with a poor prognosis. To date, no prognostic nomogram for this rare disease was established. Hence, we aimed to develop a nomogram to predict overall survival (OS) of PSCGBM.

Methods: Clinical data of patients with PSCGBM was retrospectively collected from the neurosurgery department of Soochow University Affiliated Second Hospital and the Surveillance Epidemiology and End Results database. Information including age, sex, race, tumor extension, extent of resection, adjuvant treatment, marital status, income, year of diagnosis and months from diagnosis to treatment were recorded. Univariate and multivariate Cox regression analyses were used to identify independent prognostic factors for PSCGBM. A nomogram was constructed to predict 1-year, 1.5-year, and 2-year OS of PSCGBM.

Results: A total of 132 patients were included. The 1-year, 1.5-year, and 2-year OS were 45.5%, 29.5%, and 18.9%, respectively. Four variables: age groups, tumor extension, extent of resection, and adjuvant therapy, were identified as independent prognostic factors. The nomogram showed robust discrimination with a C-index value for the prediction of 1-year OS, 1.5-year OS, and 2-year of 0.71 (95% confidence interval [CI], 0.61–0.70), 0.72 (95% CI, 0.62–0.70), and 0.70 (95% CI, 0.61–0.70), respectively. The calibration curves exhibited high consistencies between the predicted and observed survival probability in this cohort.

Conclusion: We have developed and internally validated a nomogram for predicting the survival outcome of PSCGBM for the first time. The nomogram has the potential to assist clinicians in making individualized predictions of survival outcome of PSCGBM.

Keywords: Spinal cord, Glioblastoma, Nomogram, Rare diseases, Prognostic factors

INTRODUCTION

Primary spinal cord glioblastoma (PSCGBM) is a rare disease, accounting for approximately 1.5% of intraspinal tumors.¹ To date, no standard treatment algorithm of PSCGBM was established. In contrast to its intracranial counterpart, PSCGBM was reported to have poorer survival outcome. The median survival time for PSCGBM is just about 9 months.² ⁴

Several factors might contribute to its worse prognosis. Firstly, gross total resection (GTR) of PSCGBM is a great challenge due to dense nerve fiber in spinal cord and no definite margin between normal spinal cord and tumor.³ As a result, patients with PSCGBM frequently had high postoperative residual tumor burden. Secondly, infertile blood supply of spinal cord might lead to insufficient chemotherapeutic drug permeability. Thirdly, MGMT promoter methylation, which is a prognostic marker for benefit from temozolomide (TMZ), infrequently occurred in spinal cord astrocytoma.⁶ ⁷ In conclusion, the cur-
rent investigation is still ongoing to determine the precise efficacy of TMZ in the treatment of PSCGBM.1,8-10

In previous studies, sex, ages, adjuvant treatment and surgical treatment were found to be prognostic factors of PSCGBM in Cox proportional hazards model.11-14 However, Cox model could not be used to predict individual survival outcome and quantify survival probability. In recent years, nomograms are widely used for cancer prognosis.15-17 As compared with traditional Cox regression model, a nomogram is a simple, visual and personalized scoring system for the prognostic prediction and can be used to predict individual survival probability. To date, no nomogram for predicting survival outcome of PGCGBM is established. Here, we have developed a nomogram for PSCGBM in our study to predict OS based on a large cohort.

MATERIALS AND METHODS

1. Study Population

Data were extracted from Surveillance Epidemiology and End Results (SEER) database (the Incidence-SEER 8Regs Custom Data, Nov 2021 Sub [1975–2019 varying] and Incidence-SEER 17Regs Custom Data, Nov 2021 Sub [2000–2019 varying]). Overlapped data between the 2 subdatabases were identified based on unique patient ID. Only patients who were diagnosed with glioblastoma (ICD-O-3 code: 9940, 9941) and lesions located at the spinal cord or cauda equina (ICD-O-3 code: C72.0 for “spinal cord”; C72.1 for “cauda equina”) were included. Additionally, patients who were diagnosed with PSCGBM and underwent surgery at the Department of Neurosurgery, the Second Affiliated Hospital of Soochow University, were also included. Patients meet the following criteria would be excluded: (1) metastasis, instead of primary lesion, which could be identified by sequence number and primary tumors were marked with “one primary only” or “1st of 2 or more primaries”; (2) no surgery was performed or surgical strategy was unknown; (3) death from other causes or cause was unknown; (4) survival time was not available; (5) diagnostic confirmation was not based on pathological examination. The following data were collected: age groups, sex, race, tumor extension, extent of resection, adjuvant therapy, year of diagnosis, marital status at diagnosis, median household income (MHI) adjusted for inflation to 2019 and months from diagnosis to treatment. Detailed screening flow chart is shown in Fig. 1.

The study was approved by the Institutional Review Board (IRB) of the Second Affiliated Hospital of Soochow University (IRB No. JD-HG-2024-047).

2. Definition of Variables

Age distribution was categorized into 3 groups: 5–17, 18–64, and ≥ 65 years. Race was divided into white, black, and other/unknown. Tumor invasion was stratified to localized, distant and unknown. Extent of resection was classified as biopsy, partial resection (PR), GTR, and unknown. Adjuvant treatment was divided into none, radiotherapy (RT) only, chemotherapy (CT) only, radiochemotherapy, and unknown. Years of diagnosis were categorized into 3 groups at 20-year intervals. Marital status was stratified to single (never married), married (including common law), and divorced/widowed. MHI inflation adjusted to 2019 was categorized as ≤ $50,000, $50,000–$59,999, $60,000–$69,999, ≥ $70,000 and unknown. Months from diagnosis to treatment were categorized into 2 groups based on whether patients received immediate treatment within 1 month. Survival outcome was dichotomized into alive and cancer-specific death.

3. Data Analysis and Diagnostic Prediction Model Building

Continuous variables were reported as mean ± standard deviation or median ± interquartile range (IQR), as appropriate. Categorical data were presented as the frequency (percentage). Two-tailed t-test was used for normally distributed continuous variables and Mann-Whitney U-test for nonnormally distributed continuous variables; chi-square test or Fisher test was used for categorical variables. The Kaplan-Meier method was applied to calculate survival time and rates. The nomograms were built based on the results of multivariable Cox analyses of OS. The final model selection for the nomograms was performed by a backward step-down selection process using the Akaike information criterion. The performance of the nomogram was measured by the C-index. Calibration of the nomogram for 1-, 1.5-, 2-year survival was done by comparing the predicted with the observed survival. In the present study, the nomogram was subjected to 1,000 bootstrap resamples for internal validation. Furthermore, decision curve analysis (DCA) was performed to finalize the ranges of threshold probabilities within which the nomograms were clinically valuable by rmda (risk model decision analysis package). A significance level of p < 0.05 was used to denote statistical significance. All statistical analysis was performed using R software (ver. 4.3.2, R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

A total of 132 patients with PSCGBM were included (SEER
Fig. 1. The final diagnosis flowchart for patient enrollment. SEER, Surveillance Epidemiology and End Results; GBM, glioblastoma.

The median survival time was 11 months (Fig. 2). Kaplan-Meier survival curves by variable categories showed that age groups, extent of resection and adjuvant therapy were associated with survival outcome (Fig. 3A–J).

Multivariate Cox regression analysis demonstrated that age groups (18–64 years: hazard ratio [HR], 0.59; 95% CI, 0.38–0.92; p = 0.021), tumor extension (distant: HR, 2.71; 95% CI, 1.33–5.52; p = 0.006), extent of resection (GTR: HR, 0.36; 95% CI, 0.14–0.93; p = 0.034), adjuvant therapy (radiochemothera-
Table 1. Demographic and treatment characteristics of patients with primary spinal cord glioblastoma

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total (n = 132)</th>
<th>Alive (n = 24)</th>
<th>Cancer-specific death (n = 108)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr), median (IQR)</td>
<td>30 (15–46)</td>
<td>37.5 (24.0–53.5)</td>
<td>28 (14.8–46.0)</td>
<td>0.102</td>
</tr>
<tr>
<td>Age groups (yr)</td>
<td></td>
<td></td>
<td></td>
<td>0.106</td>
</tr>
<tr>
<td>0–17</td>
<td>39 (29.5)</td>
<td>3 (12.5)</td>
<td>36 (33.3)</td>
<td></td>
</tr>
<tr>
<td>18–64</td>
<td>83 (62.9)</td>
<td>18 (75.0)</td>
<td>65 (60.2)</td>
<td></td>
</tr>
<tr>
<td>≥65</td>
<td>10 (7.6)</td>
<td>3 (12.5)</td>
<td>7 (6.5)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td>0.967</td>
</tr>
<tr>
<td>Female</td>
<td>61 (46.2)</td>
<td>11 (45.8)</td>
<td>50 (46.3)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>71 (53.8)</td>
<td>13 (54.2)</td>
<td>58 (53.7)</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td>0.194</td>
</tr>
<tr>
<td>White</td>
<td>75 (56.8)</td>
<td>10 (41.7)</td>
<td>65 (60.2)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>7 (5.3)</td>
<td>1 (4.2)</td>
<td>6 (5.6)</td>
<td></td>
</tr>
<tr>
<td>Other/unknown</td>
<td>50 (37.9)</td>
<td>13 (54.2)</td>
<td>37 (34.3)</td>
<td></td>
</tr>
<tr>
<td>Tumor extension</td>
<td></td>
<td></td>
<td></td>
<td>0.074</td>
</tr>
<tr>
<td>Localized</td>
<td>104 (78.8)</td>
<td>23 (95.8)</td>
<td>81 (75.0)</td>
<td></td>
</tr>
<tr>
<td>Distant</td>
<td>10 (7.6)</td>
<td>0 (0)</td>
<td>10 (9.3)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>18 (13.6)</td>
<td>1 (4.2)</td>
<td>17 (15.7)</td>
<td></td>
</tr>
<tr>
<td>Extent of resection</td>
<td></td>
<td></td>
<td></td>
<td>0.087</td>
</tr>
<tr>
<td>Biopsy</td>
<td>14 (10.6)</td>
<td>0 (0)</td>
<td>14 (13.0)</td>
<td></td>
</tr>
<tr>
<td>Partial resection</td>
<td>71 (53.8)</td>
<td>12 (50.0)</td>
<td>59 (54.6)</td>
<td></td>
</tr>
<tr>
<td>Gross total resection</td>
<td>47 (35.6)</td>
<td>12 (50.0)</td>
<td>35 (32.4)</td>
<td></td>
</tr>
<tr>
<td>Adjuvant therapy</td>
<td></td>
<td></td>
<td></td>
<td>0.936</td>
</tr>
<tr>
<td>None</td>
<td>10 (7.6)</td>
<td>2 (8.3)</td>
<td>8 (7.4)</td>
<td></td>
</tr>
<tr>
<td>Radiotherapy only</td>
<td>18 (13.6)</td>
<td>4 (16.7)</td>
<td>14 (13.0)</td>
<td></td>
</tr>
<tr>
<td>Chemotherapy only</td>
<td>2 (1.5)</td>
<td>0 (0)</td>
<td>2 (1.9)</td>
<td></td>
</tr>
<tr>
<td>Radiochemotherapy</td>
<td>88 (66.7)</td>
<td>15 (62.5)</td>
<td>73 (67.6)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>14 (10.6)</td>
<td>3 (12.5)</td>
<td>11 (10.2)</td>
<td></td>
</tr>
<tr>
<td>Year of diagnosis</td>
<td></td>
<td></td>
<td></td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>1975–1994</td>
<td>10 (7.5)</td>
<td>0 (0)</td>
<td>10 (9.3)</td>
<td></td>
</tr>
<tr>
<td>1995–2014</td>
<td>63 (47.7)</td>
<td>4 (16.7)</td>
<td>59 (54.6)</td>
<td></td>
</tr>
<tr>
<td>2014–2023</td>
<td>59 (44.7)</td>
<td>20 (83.3)</td>
<td>39 (36.1)</td>
<td></td>
</tr>
<tr>
<td>Marital status at diagnosis</td>
<td></td>
<td></td>
<td></td>
<td>0.191</td>
</tr>
<tr>
<td>Single (never married)</td>
<td>73 (55.3)</td>
<td>10 (41.7)</td>
<td>63 (58.3)</td>
<td></td>
</tr>
<tr>
<td>Married (including common law)</td>
<td>50 (37.9)</td>
<td>13 (54.2)</td>
<td>37 (34.3)</td>
<td></td>
</tr>
<tr>
<td>Divorced/widowed</td>
<td>9 (6.8)</td>
<td>1 (4.2)</td>
<td>8 (7.4)</td>
<td></td>
</tr>
<tr>
<td>MHI inflation adjusted to 2019 (USD)</td>
<td></td>
<td></td>
<td></td>
<td>0.397</td>
</tr>
<tr>
<td>&lt;50,000</td>
<td>40 (30.3)</td>
<td>10 (41.7)</td>
<td>30 (27.8)</td>
<td></td>
</tr>
<tr>
<td>50,000–59,999</td>
<td>21 (15.9)</td>
<td>5 (20.8)</td>
<td>16 (14.8)</td>
<td></td>
</tr>
<tr>
<td>60,000–69,999</td>
<td>27 (20.5)</td>
<td>4 (16.7)</td>
<td>23 (21.3)</td>
<td></td>
</tr>
<tr>
<td>&gt;70,000</td>
<td>36 (27.3)</td>
<td>5 (20.8)</td>
<td>31 (28.7)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>8 (6.1)</td>
<td>0 (0)</td>
<td>8 (9.3)</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Although the median survival time of PSCGBM is relatively short, of note, we observed that 45 patients had a survival time of at least 18 months. Therefore, it will be helpful for clinician to discriminate patients with relative better survival outcome from those with relative poorer survival outcome by using a nomogram. Based on independent predictors of OS identified by multivariable Cox analysis, a nomogram for predicting 1-year OS, 1.5-year OS, and 2-year OS was constructed (Fig. 4).

In terms of discrimination of the nomogram model, the C-index value was 0.71 (95% CI, 0.61–0.70), 0.70 (95% CI, 0.61–0.70), and 0.72 (95% CI, 0.62–0.72) for the prediction of 1-year OS, 1.5-year OS, and 2-year OS, respectively and comparable C-index values were confirmed through bootstrapping validation (C-index for the prediction of 1-year OS, 1.5-year OS, and 2-year OS: 0.74, 0.74, 0.75) (Fig. 5A–C). As respect to calibration, the calibration curves of the score system showed high consistencies between the predicted and observed survival probability in this cohort (Fig. 5D–F). Finally, DCA was performed to evaluate the clinical usefulness of the nomogram. When the predicted threshold probability was 80%–100% for 1-year OS, 1.5-year OS, and 2-year OS, application of this mod-

### Table 1. Demographic and treatment characteristics of patients with primary spinal cord glioblastoma (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total (n = 132)</th>
<th>Alive (n = 24)</th>
<th>Cancer-specific death (n = 108)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Months from diagnosis to treatment</td>
<td></td>
<td></td>
<td></td>
<td>0.003*</td>
</tr>
<tr>
<td>Immediate treatment within 1 mo</td>
<td>102 (77.3)</td>
<td>13 (54.2)</td>
<td>89 (82.4)</td>
<td></td>
</tr>
<tr>
<td>No immediate treatment within 1 mo</td>
<td>30 (22.7)</td>
<td>11 (45.8)</td>
<td>19 (17.6)</td>
<td></td>
</tr>
</tbody>
</table>

Values are presented as number (%) unless otherwise indicated.
IQR, interquartile range; MHI, median household income; USD, United States dollar.
*p < 0.05, the groups exhibited statistically significant differences.

Fig. 2. In primary spinal cord glioblastoma, the median survival time is 11 months. The overall survival rates at 1, 1.5, and 2 years were 45.5%, 29.5%, and 18.9%, respectively.
Fig. 3. In the Kaplan-Meier survival curve analysis, we considered variables such as age groups (A), sex (B), race (C), tumor extension (D), extent of resection (E), adjuvant therapy (F), year of diagnosis (G), marital status at diagnosis (H), median household income (MHI) inflation adjusted to 2019 (I), and months from diagnosis to treatment (J). The analysis revealed significant differences only in age groups, adjuvant therapy and extent of resection.
Fig. 3. In the Kaplan-Meier survival curve analysis, we considered variables such as age groups (A), sex (B), race (C), tumor extension (D), extent of resection (E), adjuvant therapy (F), year of diagnosis (G), marital status at diagnosis (H), median household income (MHI) inflation adjusted to 2019 (I), and months from diagnosis to treatment (J). The analysis revealed significant differences only in age groups, adjuvant therapy and extent of resection. (Continued)
el to predict survival outcome could add more benefit than the treat-all or treat-none strategy (Fig. 5G–I).

**DISCUSSION**

PSCGBM accounts for only 1%–5% of central nervous system glioblastomas and 1.5% of all spinal cord tumors.18 Consistent with the findings of our study, previous studies reported that the disease has a poor prognosis with a median survival time of 12–14 months.1,19 Due to its rarity, prognostic factors associated with OS of PSCGBM are not well understood. In our study, univariate and multivariate Cox regression analysis indi-
### Table 2. Univariate and multivariate Cox analysis of cancer-specific survival

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>95% CI</td>
</tr>
<tr>
<td><strong>Age groups (yr)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–17</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>18–64</td>
<td>0.64</td>
<td>0.42–0.97</td>
</tr>
<tr>
<td>≥ 65</td>
<td>1.26</td>
<td>0.55–2.85</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.70</td>
<td>0.30–1.63</td>
</tr>
<tr>
<td>Other/unknown</td>
<td>0.96</td>
<td>0.64–1.45</td>
</tr>
<tr>
<td><strong>Tumor extension</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Localized</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Distant</td>
<td>2.34</td>
<td>1.20–4.56</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.48</td>
<td>0.88–2.51</td>
</tr>
<tr>
<td><strong>Extent of resection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biopsy</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Partial resection</td>
<td>0.62</td>
<td>0.35–1.12</td>
</tr>
<tr>
<td>Gross total resection</td>
<td>0.38</td>
<td>0.20–0.72</td>
</tr>
<tr>
<td><strong>Adjuvant therapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Radiotheraphy only</td>
<td>0.48</td>
<td>0.20–1.17</td>
</tr>
<tr>
<td>Chemotherapy only</td>
<td>1.85</td>
<td>0.39–8.80</td>
</tr>
<tr>
<td>Radiochemotherapy</td>
<td>0.43</td>
<td>0.20–0.91</td>
</tr>
<tr>
<td>Unknown</td>
<td>0.39</td>
<td>0.15–1.02</td>
</tr>
<tr>
<td><strong>Year of diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1975–1994</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>1995–2014</td>
<td>0.57</td>
<td>0.29–1.12</td>
</tr>
<tr>
<td>2014–2023</td>
<td>0.46</td>
<td>0.23–0.93</td>
</tr>
<tr>
<td><strong>Marital status at diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single (never married)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Married (including common law)</td>
<td>0.76</td>
<td>0.51–1.15</td>
</tr>
<tr>
<td>Divorced/widowed</td>
<td>0.89</td>
<td>0.42–1.86</td>
</tr>
<tr>
<td><strong>MHI inflation adjusted to 2019 (USD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50,000</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>50,000–59,999</td>
<td>0.92</td>
<td>0.05–1.68</td>
</tr>
<tr>
<td>60,000–69,999</td>
<td>1.53</td>
<td>0.89–2.63</td>
</tr>
<tr>
<td>&gt; 70,000</td>
<td>0.83</td>
<td>0.49–1.39</td>
</tr>
<tr>
<td>Unknown</td>
<td>1.74</td>
<td>0.79–3.81</td>
</tr>
<tr>
<td><strong>Months from diagnosis to treatment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate treatment within 1 mo</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>No immediate treatment within 1 mo</td>
<td>0.63</td>
<td>0.38–1.03</td>
</tr>
</tbody>
</table>

HR, hazard ratio; CI, confidence interval; MHI, median household income; USD, United States dollar.

*p < 0.05, a statistical difference with the first subgroup within the group.
cated that age groups, tumor extension, extent of resection and adjuvant therapy were predictors of PSCGBM. Likewise, a multicenter study by Inoue et al.\(^\text{11}\) suggested that adolescent and young adult (HR, 3.53; 95% CI, 1.17–10.64), intracranial dissemination (HR, 4.30; 95% CI, 1.29–14.36), and no radiation therapy (HR, 57.34; 95% CI, 6.73–488.39) were risk factors for mortality of patients with PSCGBM.

In contrast to intracranial GBM, PSCGBM frequently occurred in younger aged population and showed no sex predilection.\(^\text{3,19,20}\) The median age in our cohort was 30 years (IQR, 16–45). Consistently, the study by Konar et al.\(^\text{21}\) revealed that mean age of PSCGBM was 27 years and 51% of patients were below 18 years of age. However, in their study, age was not identified as an independent predictor of mortality. A retrospective study by Moinuddin et al.\(^\text{14}\) included 190 patients with PSCGBM, the mean age was 40.8 ± 22.3 and age was found to be significantly associated with OS (p = 0.046).

The majority of PSCGBM primarily occurs in the cervical, thoracic and conus medullaris regions.\(^\text{22,23}\) In line with our findings, Moinuddin et al.\(^\text{14}\) found that extended lesion was associated with unfavorable survival outcome of PSCGBM. The study conducted by Ardeshiri et al.\(^\text{24}\) also revealed that patients with more than 3 segments involved exhibited a higher likelihood of experiencing neurological deterioration in comparison to those with only one segment involved. It was reported that 40%–50% patients with PSCGBM could develop cerebrospinal fluid (CSF) dissemination and CSF dissemination was significantly related to poor OS.\(^\text{10,11,21}\) Closer anatomic proximity of spinal tumors to the subarachnoid space compared to their intracranial counterparts might contribute to the high rate of CSF dissemination.

As to extent of resection, the role of surgical resection in PSCGBM is not well understood. The study of Lam et al.\(^\text{13}\) indicated no statistically significant effect of the extent of resection on the length of survival among patients with PSCGBM. Even worse, McGirt et al.\(^\text{3}\) and Wolff et al.\(^\text{25}\) found that radical resection could worsen survival outcome of PSCGBM. Conversely, a large cohort study involving 208 PSCGBMs by Chalif et al.\(^\text{26}\) demonstrated that GTR independently conferred a survival benefit to patients with PSCGBM (HR, 0.194; p < 0.001). Consistently, Kahn et al.\(^\text{27}\) and Corradini et al.\(^\text{28}\) also found that GTR could improve survival outcome of PSCGBM. Additionally, cordectomy, which is a more radical type of GTR and viewed as a salvage treatment, is expected to improve long-term survival by restricting or delaying intracranial dissemination of PSCGBM.\(^\text{3,28-36}\) In our study, GTR was identified as a protective factor of favorable survival outcome (HR, 0.36; 95% CI, 0.14–0.93; p = 0.034). Due to no definite margin between normal spinal cord and PSCGBM, GTR without neurological compromise is a
Fig. 5. The area under the curve (AUC) represents the area under the blue curve, providing a method for predicting annual overall survival (OS) accuracy. Specifically, the AUC values for 1 year (A), 1.5 years (B), and 2 years (C) were determined as 0.71, 0.70, and 0.72, respectively. Importantly, an AUC exceeding 0.70 indicates a superior predictive effect. In plots for 1 year (D), 1.5 years (E), and 2 years (F), greater alignment between the red line segment and the blue dashed line signifies higher accuracy in survival prediction by the model. Finally, the decision curve analysis was employed to validate the clinical efficacy of 1- (G), 1.5- (H), and 2-year OS rates (I). CI, confidence interval.
great challenge. Although GTR can improve patient outcomes, we need to make a balance between the patient's neurological function and survival prognosis. In our view, for patients with intact neurological function (Mc Cormick grade ≤ 3), we do not recommend pursuing complete tumor resection at the risk of functional compromise, while for paraplegic patients (Mc Cormick grade = 5), we recommend GTR or even cordectomy in the case of preoperative informed consent is available and intraoperative frozen pathology indicates high-grade glioma.

The role of radiochemotherapy in PSCGBM is controversial. An aggressive approach involving whole-brain irradiation along with focal spinal irradiation has been suggested, even in the absence of evidence indicating intracranial dissemination. However, Chalif et al. found that radiation was not independently associated with improved survival of PSCGBM, while chemotherapy was significantly related to improved survival in patients with PSCGBM. In the retrospective study of Kaley et al., it was found that both TMZ and bevacizumab could improve the survival rate. However, the impact of CT on PSCGBM is less pronounced as compared to cerebral GBM. The study by Hernandez-Duran et al. did not show a significant relationship between TMZ and prolonged survival. Although the effect of RT on PSCGBM is not well established, RT was frequently prescribed to PSCGBM following surgical treatment, which repurposed the treatment strategy of intracranial GBM. It is worth noting that the study by Inoue et al. demonstrated that RT was associated with prolonged survival time of PSCGBM, but chemotherapy did not. In our study, only radiochemotherapy showed protective effect in patients with PSCGBM, while radiation only or chemotherapy only did not confer survival benefit. Likewise, the study by Cheng et al. revealed that radiation plus TMZ could prolong survival time as compared to TMZ only or none (p = 0.002). Additionally, immunotherapy showed favorable efficacy in hematological malignant tumor.

In recent years, immunotherapy was tested in brain glioblastoma and shown promise in intracranial gliomas with some research suggesting benefit for spinal cord gliomas. The application of immunotherapy, such as immune checkpoint inhibitors or chimeric antigen receptor T-cell therapy, to PSCGBM stands for future direction of research.

In our study, a nomogram was constructed to predict individual survival outcome of PSCGBM. In recent years, nomograms were widely used to predict survival outcome of other tumors, such as gastric cancer, non–small-cell lung cancer, hepatocellular Carcinoma and proved to be an excellent tool of predicting individual survival outcome in recent years. As to generalizability, the development of our nomogram was based on a large public dataset and patients treated in our institute, which contained different populations. In addition, the nomogram showed good discrimination and calibration. Therefore, our nomogram has good generalizability. Although several Cox proportional hazards models of PSCGBM were developed and some survival predictors were found, the advantage of our nomogram over traditional Cox hazard-proportional model is that nomogram could be used to predict individual survival outcome of PSCGBM and quantify the survival probability of an individual with PSCGBM. Therefore, this tool could be used to guide clinical decision-making based on quantized survival outcome and to discriminate patients with relatively better survival outcome and patients with relatively poorer survival outcome. Moreover, this tool could be applied to facilitate preoperative clinician-patient communication.

However, several limitations should be noted. Firstly, preoperative neurological findings and spinal lesion levels were not recorded in SEER database, which might influence the interpretation of our findings. Secondly, the dosage of RT and chemotherapy drug was not indicated in SEER database. Thirdly, H3 K27M mutation, a diagnostic and prognostic marker of diffuse midline glioma, H3 K27-altered, World Health Organization (WHO) grade 4, frequently occurred in primary spinal cord astrocytoma, but was not recorded in SEER database. PSCGBM with H3 K27M mutation will be assigned an integrated diagnosis of diffuse midline glioma, H3 K27-altered, WHO grade 4. Therefore, integrated analysis without selecting out these cases with H3 K27M might confound our final result. Finally, due to the rarity of PSCGBM, no extra samples are available for an external validation. A multicenter study with large sample size is warranted to validate our findings.

**CONCLUSION**

A robust population-based survival-predicting model for PSCGBM is established and internationally validated. This nomogram offers clinicians a simple-to-use method for assessing mortality risk in patients with PSCGBM.

**NOTES**

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Acknowledgments: The authors express our sincere gratitude to Dr. Lei Cheng from Xuanwu Hospital, Capital Medical University for his invaluable contributions in project conception, language editing, and image enhancement.

Author Contribution: Conceptualization: YW, MS, FJ, RL; Formal analysis: YW, QM; Investigation: YW, MS, YC; Methodology: YW, QM, FJ, RL; Project administration: YW, QM, RL; Writing – original draft: YW, QM, MS; Writing – review & editing: YW, QM, YC, FJ, RL.

ORCID
Yao Wang: 0009-0006-9762-3045
Qingchun Mu: 0000-0001-7578-5381
Minfeng Sheng: 0000-0003-0241-6248
Yanning Chen: 0000-0001-9194-0580
Fengzeng Jian: 0000-0001-7860-278X
Rujun Li: 0000-0002-7831-5868

REFERENCES
A Comparative Factor Analysis and New Magnetic Resonance Imaging Scoring System for Differentiating Pyogenic Versus Tuberculous Spondylodiscitis

Terdpong Tanaviriyachai, Patchara Pornsopanakorn, Kongtush Choovongkomol, Tada Virathepsuporn, Urawit Piyapromdee, Sarut Jongkittanakul, Weera Sudprasert, Sirichai Wiwatrojanagul

Department of Orthopedic Surgery, Maharat Nakhon Ratchasima Hospital, Nakhon Ratchasima, Thailand

Department of Radiology, Rayong Hospital, Rayong, Thailand

Department of Medicine, Maharat Nakhon Ratchasima Hospital, Nakhon Ratchasima, Thailand

Objective: This study aimed to compare and analyze differences in clinical and magnetic resonance imaging (MRI) findings between tuberculous spondylodiscitis (TbS) and pyogenic spondylodiscitis (PyS), and to develop and validate a simplified multiparameter MRI-based scoring system for differentiating TbS from PyS.

Methods: We compared predisposing factors in 190 patients: 123 with TbS and 67 with PyS, confirmed by laboratory tests, culture, or pathology. Data encompassing patient demographics, clinical characteristics, laboratory results, and MRI findings were collected between 2015 and 2020. Data were analyzed using logistic regression methods, and selected coefficients were transformed into an MRI-based scoring system. Internal validation was performed using bootstrapping method.

Results: Univariate analysis revealed that the significant risk factors associated with TbS included thoracic lesions, vertebral destruction > 50%, intraosseous abscess, thin-walled abscess, well-defined paravertebral abscess, subligamentous spreading, and epidural abscess. Multivariate analysis revealed that only thoracic lesions, absence of epidural phlegmon, subligamentous spreading, intraosseous abscesses, well-defined paravertebral abscesses, epidural abscesses, and absence of facet joint arthritis were independent predictive factors for TbS (all p < 0.05). These potential predictors were used to derive an MRI-based scoring system. Total scores ≥ 14/29 points significantly predicted the probability of TbS, with a sensitivity of 97.58%, specificity of 92.54%, and an area under the curve of 0.96 (95% confidence interval, 0.925–0.995).

Conclusion: This simplified MRI-based scoring system for differentiating TbS from PyS helps guide appropriate treatment when the causative organism is not identified.

Keywords: Spondylodiscitis, Tuberculous spondylodiscitis, Pyogenic spondylodiscitis, MRI scoring system, Predictive scoring system

INTRODUCTION

Infectious spondylodiscitis (IS) is a septic inflammation of the spine involving vertebral bodies and paraspinal structures. During the progression of the disease, the formation of abscesses or edema can destroy vertebrae or cause neurologic disorders. The overall incidence of spinal infection in adults is approximately 2.2 per 100,000 per year, with a slowly increasing...
New MRI Scoring System for Differentiating Pyogenic vs. Tuberculous Spondylodiscitis

Tanaviriyachai T, et al.

**MATERIALS AND METHODS**

We retrospectively collected medical records of patients diagnosed with IS admitted to the Maharat Nakhon Ratchasima Hospital between January 2015 and December 2020. Cases with microbiologically and pathologically documented evidence were included in this study. PyS was diagnosed when the etiological organism was identified through percutaneous vertebral biopsy, surgical drainage, or blood culture (a minimum of 2 separate sets). TbS was diagnosed based on pathological samples, tissue cultures, and polymerase chain reaction (PCR) tests. Patients with spondylodiscitis caused by other pathogens (e.g., fungal, or parasitic), unconfirmed spondylodiscitis (if no pathogens were isolated), lack of pretreatment MRI, absence of gadolinium administration during MRI, or lack of T1-weighted or fluid-sensitive sequences were excluded.

Clinical data included age, sex, predisposing factors and/or associated illnesses, onset of the symptoms, fever, Frankel grading, and causative organisms. Laboratory data comprised white blood cell (WBC) count, proportion of neutrophils, C-reactive protein (CRP) levels, erythrocyte sedimentation rate (ESR), and serum alkaline phosphatase (ALP) levels. All MRI examinations followed a standard protocol, including axial and sagittal T1-weighted sequences, axial and sagittal fluid-sensitive sequences, including T2-weighted with fat-saturation (T2w fat-sat) or short tau inversion recovery sequences, and axial and sagittal T1-weighted sequences after gadolinium administration. MRI findings were evaluated by consensus between a 5-year-experienced spine surgeon and a musculoskeletal radiologist. Details of each finding were evaluated and are described in Table 1. The infection in the thoracolumbar region is classified as a thoracic or lumbar lesion based on the extent of vertebral body destruction, with thoracic areas being more severe.

This study was performed in accordance with the Helsinki Declaration and approved by the Maharat Nakhon Ratchasima Hospital Institutional Review Board (MNRH IRB No. 089/2020). The patients were informed that the data concerning their cases would be submitted for publication and provided their consent.

Statistical analyses were performed using Stata Statistical Software (ver. 14; StataCorp LP, College Station, TX, USA). After a descriptive study of the variables, t-tests were used to compare continuous variables. Chi-square tests were used to compare predisposing factors and associated illnesses. All tests were 2-sided, and a p-value of 0.05 was considered significant. All variables that were significant in the chi-square test were included in a multivariate logistic regression analysis using stepwise backward elimination for the derived independent variables.

The diagnostic accuracy of the reduced multivariate model was evaluated in terms of calibration and discrimination. Calibration was performed using Hosmer-Lemeshow goodness of fit statistics. A calibration plot comparing the agreement between the disease probabilities estimated using the model and the observed disease data is also presented. Discriminative power was evaluated using the area under the receiver operating characteristic (ROC) curve. Internal validation was performed using a bootstrapping procedure with 1,000 replicates. Bootstrap resampling is a statistical technique used for estimating the sampling distribution of a statistic by resampling with replacement from the observed data. This resampling is applicable in various situations, offering versatility in statistical problems like parameter estimation and hypothesis testing. It requires minimal assumptions and is easy to implement, making it a practical way to assess statistic variability without complex

---

Trend worldwide in recent years, IS is potentially life-threatening, with a mortality rate of 3%–20%. Common causes of IS include pyogenic spondylodiscitis (PyS) and tuberculous spondylodiscitis (TbS), which account for 40%–80% and 17%–40% of all IS cases, respectively. Insufficient specific signs and symptoms might cause delayed diagnosis and treatment, leading to disastrous consequences.

It is critical to distinguish between TbS and PyS to provide appropriate treatment. However, the identification of these 2 entities is challenging because of their nonspecific signs and symptoms. Microbiological diagnosis is the gold standard for differentiating between TbS and PyS. However, identifying the microbes is difficult. Previous reports on patients with PyS showed a negative culture rate ranging from 10% to 30%. In contrast, obtaining a positive culture for TbS typically requires 3 weeks, with a success rate ranging between 50% and 70%.

When microbiological identification is impossible, clinical, laboratory, and magnetic resonance imaging (MRI) findings may aid in identifying a potential causative microorganism. Previous studies have distinguished radiological findings between TbS and PyS. However, few studies have developed a scoring system that uses predictive factors to stratify the probability of TbS from PyS.

In the present study, we aimed to compare and analyze the differences in the clinical, laboratory, and MRI findings between TbS and PyS, and to develop and validate a simplified multiparameter MRI-based scoring system for differentiating between TbS and PyS.
Table 1. Description of individual MRI features

<table>
<thead>
<tr>
<th>No.</th>
<th>Radiological parameter</th>
<th>MRI features in identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pathologic vertebral body signal</td>
<td>Classified as hypointense, isointense, hyperintense or heterogeneous, compared to the unaffected vertebrae.</td>
</tr>
<tr>
<td>2</td>
<td>Pathologic intervertebral disc signal</td>
<td>Classified as hypointense, hyperintense or isointense, compared to unaffected discs.</td>
</tr>
<tr>
<td>3</td>
<td>Vertebral endplate involvement</td>
<td>Classified as: - eroded: vertebral body destruction with reduction in body height of less than half. - destroyed: vertebral body destruction with reduction in body height of more than half.</td>
</tr>
<tr>
<td>4</td>
<td>Extent of vertebral destruction</td>
<td>Classified as: - minimal: confined to vertebral endplate. - severe: extended to vertebral body including vertebral body abscess or vertebral collapse.</td>
</tr>
<tr>
<td>5</td>
<td>Extent of intervertebral disc destruction</td>
<td>Classified as: - none/mild, moderate, or severe including complete disc destruction or disc abscess.</td>
</tr>
<tr>
<td>6</td>
<td>Pathologic vertebral body contrast enhancement</td>
<td>Classified as marginal, homogeneous, or heterogeneous*.  *Heterogenous enhancement is defined as abnormal bone marrow signals combined with hypo- and hypersignal intensity in the coronal, sagittal, and axial planes.</td>
</tr>
<tr>
<td>7</td>
<td>Pathologic intervertebral disc contrast enhancement</td>
<td>Classified as diffuse, marginal, focal, or absent.</td>
</tr>
<tr>
<td>8</td>
<td>Presence of intraosseous or intervertebral disc abscess</td>
<td>Presence of collections within the vertebra or intervertebral disc space.</td>
</tr>
<tr>
<td>9</td>
<td>Paravertebral abscess</td>
<td>Presence of collections within the adjacent paravertebral soft tissues classified as absent, well-defined* or ill-defined. *A well-defined paraspinal abscess is defined as regular and smooth abscess wall.</td>
</tr>
<tr>
<td>10</td>
<td>Paravertebral abscess wall characteristics</td>
<td>Classified as thin wall (a wall thickness &lt; 2 mm) or thick wall (a wall thickness ≥ 2 mm).</td>
</tr>
<tr>
<td>11</td>
<td>Epidural abscess</td>
<td>Epidural encroachment or indentation by pus.</td>
</tr>
<tr>
<td>12</td>
<td>Epidural phlegmon</td>
<td>Epidural encroachment or indentation by granulation tissue.</td>
</tr>
<tr>
<td>13</td>
<td>Septic facet joint arthritis</td>
<td>Presence of collections within facet joint or soft tissue surround the facet joint.</td>
</tr>
<tr>
<td>14</td>
<td>Subligamentous spreading</td>
<td>Presence of anterior subligamentous bone signal alterations and abscess which extends more than 2 vertebral levels.</td>
</tr>
<tr>
<td>15</td>
<td>Spinal cord compression</td>
<td>Epidural encroachment or indentation to the spinal cord.</td>
</tr>
</tbody>
</table>

MRI, magnetic resonance imaging.

mathematical derivations. However, this resampling procedure has several drawbacks, including its reliance on the original sample, its inability to accurately represent population variability in small samples, and its assumption of stationary data distribution, which may not be suitable in dynamic environments.

Subsequently, a simplified risk score transformation was generated. Each item was assigned a specific score based on the logistic regression coefficients of the multivariate model. To achieve this, the regression coefficient of each item was divided by its lowest coefficient, the result was rounded to the closest integer. The total scores were then categorized into 2 groups (TbS and PyS) for clinical applicability. Sensitivity and specificity were calculated separately for each group using a population-analog approach. Calibration and discrimination were assessed using a score-based multivariate logistic model.

**RESULTS**

Among the 420 patients diagnosed with IS, 190 had a confirmed diagnosis, matched all inclusion criteria, and were retro-
**Table 2. General demographic data**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tuberculosis (n = 123)</th>
<th>Pyogenic (n = 67)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>56.89 ± 15.38</td>
<td>56.86 ± 11.52</td>
<td>0.989</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>58</td>
<td>48</td>
<td>0.001*</td>
</tr>
<tr>
<td>Female</td>
<td>65</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Duration of symptoms (wk)</td>
<td></td>
<td></td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>&lt;2</td>
<td>9 (7.31)</td>
<td>26 (39.39)</td>
<td></td>
</tr>
<tr>
<td>2–4</td>
<td>12 (9.76)</td>
<td>15 (22.73)</td>
<td></td>
</tr>
<tr>
<td>&gt;4</td>
<td>102 (82.93)</td>
<td>26 (39.39)</td>
<td></td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>18 (14.63)</td>
<td>20 (29.85)</td>
<td>0.013*</td>
</tr>
<tr>
<td>Frankel grade</td>
<td></td>
<td></td>
<td>0.294</td>
</tr>
<tr>
<td>A</td>
<td>15 (12.20)</td>
<td>5 (7.46)</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>15 (12.20)</td>
<td>8 (11.94)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>30 (24.39)</td>
<td>14 (20.90)</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>31 (25.20)</td>
<td>12 (17.91)</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>32 (26.01)</td>
<td>28 (41.79)</td>
<td></td>
</tr>
<tr>
<td>Underlying disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td></td>
<td></td>
<td>0.008*</td>
</tr>
<tr>
<td>Yes</td>
<td>12 (9.76)</td>
<td>16 (23.88)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>111 (90.24)</td>
<td>51 (76.12)</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td></td>
<td></td>
<td>0.890</td>
</tr>
<tr>
<td>Yes</td>
<td>34 (27.65)</td>
<td>19 (28.36)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>89 (72.35)</td>
<td>48 (71.64)</td>
<td></td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td></td>
<td></td>
<td>0.665</td>
</tr>
<tr>
<td>Yes</td>
<td>5 (4.07)</td>
<td>1 (1.64)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>118 (95.93)</td>
<td>66 (98.51)</td>
<td></td>
</tr>
<tr>
<td>Cirrhosis</td>
<td></td>
<td></td>
<td>0.282</td>
</tr>
<tr>
<td>Yes</td>
<td>1 (0.81)</td>
<td>2 (2.99)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>122 (99.19)</td>
<td>65 (97.01)</td>
<td></td>
</tr>
<tr>
<td>White blood cells (/mm$^3$)</td>
<td></td>
<td></td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>≤10,000</td>
<td>85 (69.11)</td>
<td>21 (31.34)</td>
<td></td>
</tr>
<tr>
<td>&gt;10,000</td>
<td>38 (30.89)</td>
<td>46 (68.66)</td>
<td></td>
</tr>
<tr>
<td>Neutrophil (&gt;75%)</td>
<td>32 (26.01)</td>
<td>36 (53.73)</td>
<td>0.008*</td>
</tr>
<tr>
<td>Peak ESR (&gt;40 mm/hr)</td>
<td>100 (81.30)</td>
<td>53 (82.81)</td>
<td>0.886</td>
</tr>
<tr>
<td>Peak CRP (&gt;5 mg/dL)</td>
<td>112 (91.05)</td>
<td>62 (96.88)</td>
<td>0.224</td>
</tr>
<tr>
<td>alkaline phosphatase (&gt;120 IU/L)</td>
<td>40 (32.52)</td>
<td>38 (59.38)</td>
<td>&lt;0.001*</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%). ESR, erythrocyte sedimentation rate; CRP, C-reactive protein. *p < 0.05, statistically significant differences.

spectively enrolled. The characteristics of the 190 patients are summarized in Table 2. The mean age at diagnosis was 56.8 years (range, 18–84 years), with 106 males and 84 female patients. Data were collected from 67 patients with PyS and 123 patients with TbS. Among the 67 patients with PyS, the causative organism was confirmed by culture of percutaneous spinal biopsy and surgical drainage in 64.2% (n = 43), and blood culture in 35.8% (n = 24). *Staphylococcus aureus* was the most common microorganism identified in 56.7% (n = 38), followed by *Streptococcus* spp. (19.4%, n = 13), *Escherichia coli* (13.4%, n = 9), *Bacillus* spp. (2.9%, n = 2), *Klebsiella pneumoniae* (2.9%, n = 2), *Brucellosis* (1.4%, n = 1), *Burkholderia pseudomallei* (1.4%, n = 1), and *Pseudomonas aeruginosa* (1.4%, n = 1). Among the 123 patients with TbS, the diagnosis was confirmed by percutaneous spinal biopsy and surgical drainage in 25.2% (n = 31) of patients. The remaining patients with TbS were confirmed by positive results for PCR of *Mycobacterium tuberculosis* and pathology demonstrating caseous granulomatous inflammation.

Clinically, back pain was the most common symptom observed in both groups with 94.3% (n = 116) among TbS and 94% (n = 63) among PyS patients. The duration of symptoms lasted >4 weeks in 102 TbS patients (82.93%) and 26 PyS patients (39.39%) (p < 0.01). The number of patients with diabetic mellitus was 12 (9.76%) in the TbS group and 16 (23.88%) in the PyS group (p < 0.01). Laboratory findings of the 2 groups are shown in Table 2. PyS was more frequently associated with the following parameters: WBC > 10,000/mm$^3$, a higher proportion of neutrophils >75%, and ALP > 120 IU/L (p < 0.01).

As shown in Table 3, thoracic involvement was significantly more frequent in TbS than in PyS (61.78% vs. 22.39%, p < 0.001), while lumbar involvement was more common in PyS than in TbS (85.07% vs. 56.91%, p < 0.001). No significant differences were observed in cervical or sacral involvement. Moreover, no differences were found in the number of involved vertebrae, involvement of the posterior elements, or posterior wall retropulsion. On T1-weighted MRI, the vertebral body signal was typically hypointense in both groups. However, the TbS group exhibited a proportionately more heterogeneous intensity than the PyS group (8.13% vs. 0%, p < 0.03). Destruction of vertebral endplates and vertebral destruction >50% were more common in the TbS group than in the PyS group (26.83% vs. 8.96%, p < 0.001 and 60.16% vs. 19.4%, p < 0.001, respectively). No differences were found in the disc signal or extent of disc destruction between the groups. On T1-weighted gadolinium MRI, the vertebral body was more frequently heterogeneously enhanced in the TbS group (89.43% vs. 38.81%, p < 0.001). Moreover, vertebral intraosseous abscesses were more frequent in TbS compared to PyS (69.1% vs. 7.46%, p < 0.001). No significant differences were reported between the intervertebral disc involvement.
Table 3. General MRI parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tuberculous (n = 123)</th>
<th>Pyogenic (n = 67)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location of lesion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cervical</td>
<td></td>
<td></td>
<td>0.060</td>
</tr>
<tr>
<td>Yes</td>
<td>4 (3.25)</td>
<td>10 (14.93)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>119 (96.75)</td>
<td>57 (85.07)</td>
<td></td>
</tr>
<tr>
<td><strong>Thoracic</strong></td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>76 (61.78)</td>
<td>15 (22.39)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>47 (37.90)</td>
<td>52 (77.61)</td>
<td></td>
</tr>
<tr>
<td><strong>Lumbar</strong></td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>70 (56.91)</td>
<td>57 (85.07)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>53 (43.09)</td>
<td>10 (14.93)</td>
<td></td>
</tr>
<tr>
<td><strong>Sacral</strong></td>
<td>0.051</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>14 (11.38)</td>
<td>15 (22.39)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>109 (88.62)</td>
<td>52 (77.61)</td>
<td></td>
</tr>
<tr>
<td><strong>No. of vertebrae involved</strong></td>
<td>0.517</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤2</td>
<td>68 (55.28)</td>
<td>34 (50.75)</td>
<td></td>
</tr>
<tr>
<td>&gt;2</td>
<td>55 (44.72)</td>
<td>33 (49.25)</td>
<td></td>
</tr>
<tr>
<td><strong>Vertebral involvement</strong></td>
<td>0.106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Continuous</td>
<td>115 (93.50)</td>
<td>58 (86.57)</td>
<td></td>
</tr>
<tr>
<td>Noncontinuous</td>
<td>8 (6.50)</td>
<td>9 (13.43)</td>
<td></td>
</tr>
<tr>
<td><strong>Involvement posterior element</strong></td>
<td>0.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>54 (43.90)</td>
<td>37 (55.22)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>69 (56.10)</td>
<td>30 (44.78)</td>
<td></td>
</tr>
<tr>
<td><strong>Posterior somatic wall retropulsion</strong></td>
<td>0.168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>20 (16.26)</td>
<td>6 (8.96)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>103 (83.74)</td>
<td>61 (91.04)</td>
<td></td>
</tr>
<tr>
<td><strong>Parameters on T1-weighted MRI</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vertebral body signal</strong></td>
<td>0.033*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypointense</td>
<td>110 (89.43)</td>
<td>65 (97.01)</td>
<td></td>
</tr>
<tr>
<td>Isointense</td>
<td>3 (2.44)</td>
<td>2 (2.99)</td>
<td></td>
</tr>
<tr>
<td>Hyperintense</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>10 (8.13)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Disc signal</strong></td>
<td>0.531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypointense</td>
<td>102 (82.92)</td>
<td>60 (89.55)</td>
<td></td>
</tr>
<tr>
<td>Isointense</td>
<td>15 (12.20)</td>
<td>5 (7.46)</td>
<td></td>
</tr>
<tr>
<td>Hyperintense</td>
<td>6 (4.88)</td>
<td>2 (2.99)</td>
<td></td>
</tr>
<tr>
<td><strong>Vertebral endplate involvement</strong></td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>8 (6.50)</td>
<td>16 (23.88)</td>
<td></td>
</tr>
<tr>
<td>Eroded</td>
<td>82 (66.67)</td>
<td>45 (67.16)</td>
<td></td>
</tr>
<tr>
<td>Completely destroyed</td>
<td>33 (26.83)</td>
<td>6 (8.96)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. General MRI parameters (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tuberculous (n = 123)</th>
<th>Pyogenic (n = 67)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of vertebral destruction</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50%</td>
<td>49 (39.84)</td>
<td>54 (80.60)</td>
<td></td>
</tr>
<tr>
<td>&gt; 50%</td>
<td>74 (60.16)</td>
<td>13 (19.40)</td>
<td></td>
</tr>
<tr>
<td>Extent of disc destruction</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>24 (19.51)</td>
<td>7 (10.45)</td>
<td></td>
</tr>
<tr>
<td>&lt; 50%</td>
<td>53 (43.09)</td>
<td>22 (32.84)</td>
<td></td>
</tr>
<tr>
<td>&gt; 50%</td>
<td>46 (37.40)</td>
<td>38 (56.72)</td>
<td></td>
</tr>
<tr>
<td><strong>Parameters evaluated on T1-weighted+gadolinium MRI (intravertebral involvement)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertebral body contrast enhancement</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>2 (1.63)</td>
<td>9 (13.43)</td>
<td></td>
</tr>
<tr>
<td>Homogeneous</td>
<td>11 (8.94)</td>
<td>32 (47.76)</td>
<td></td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>110 (89.43)</td>
<td>26 (38.81)</td>
<td></td>
</tr>
<tr>
<td>% Vertebral body enhancement</td>
<td>0.774</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50%</td>
<td>10 (8.13)</td>
<td>4 (5.97)</td>
<td></td>
</tr>
<tr>
<td>&gt; 50%</td>
<td>113 (91.87)</td>
<td>63 (94.03)</td>
<td></td>
</tr>
<tr>
<td>Intravertebral disc contrast enhancement</td>
<td>0.093</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diffuse</td>
<td>34 (27.64)</td>
<td>10 (14.93)</td>
<td></td>
</tr>
<tr>
<td>Marginal</td>
<td>33 (26.83)</td>
<td>28 (41.79)</td>
<td></td>
</tr>
<tr>
<td>Focal</td>
<td>45 (36.59)</td>
<td>22 (32.84)</td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>11 (8.94)</td>
<td>7 (10.45)</td>
<td></td>
</tr>
<tr>
<td>Vertebral Intraosseous abscess</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>85 (69.10)</td>
<td>5 (7.46)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>38 (30.90)</td>
<td>62 (92.54)</td>
<td></td>
</tr>
<tr>
<td>Disc abscess</td>
<td>0.619</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>16 (13.00)</td>
<td>7 (10.45)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>107 (87.00)</td>
<td>60 (89.55)</td>
<td></td>
</tr>
<tr>
<td><strong>Parameters evaluated on T1-weighted+gadolinium MRI (extravertebral involvement)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paravertebral tissue</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absent</td>
<td>2 (1.63)</td>
<td>3 (4.48)</td>
<td></td>
</tr>
<tr>
<td>Well-defined</td>
<td>112 (91.06)</td>
<td>2 (2.99)</td>
<td></td>
</tr>
<tr>
<td>Ill-defined</td>
<td>9 (7.31)</td>
<td>62 (92.54)</td>
<td></td>
</tr>
<tr>
<td>Paravertebral abscess</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>102 (82.93)</td>
<td>30 (44.78)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>21 (17.07)</td>
<td>37 (55.22)</td>
<td></td>
</tr>
<tr>
<td>Paravertebral abscess wall characteristics</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thin and regular</td>
<td>100 (81.30)</td>
<td>2 (2.99)</td>
<td></td>
</tr>
<tr>
<td>Thick and irregular</td>
<td>2 (1.63)</td>
<td>28 (43.28)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>21 (17.07)</td>
<td>37 (55.22)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. General MRI parameters (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tuberculous (n = 123)</th>
<th>Pyogenic (n = 67)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidural abscess</td>
<td></td>
<td></td>
<td>0.001*</td>
</tr>
<tr>
<td>Yes</td>
<td>83 (67.48)</td>
<td>29 (43.28)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>40 (32.52)</td>
<td>38 (56.72)</td>
<td></td>
</tr>
<tr>
<td>Epidural phlegmon</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>6 (4.88)</td>
<td>54 (80.60)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>117 (95.12)</td>
<td>13 (19.40)</td>
<td></td>
</tr>
<tr>
<td>Facet joint arthritis</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>57 (46.34)</td>
<td>54 (80.60)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>66 (53.66)</td>
<td>13 (19.40)</td>
<td></td>
</tr>
<tr>
<td>Subligamentous spreading</td>
<td>&lt; 0.001*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>116 (94.30)</td>
<td>28 (41.79)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>7 (5.70)</td>
<td>39 (58.21)</td>
<td></td>
</tr>
<tr>
<td>Spinal cord compression</td>
<td>0.101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>101 (82.11)</td>
<td>61 (91.04)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>22 (17.89)</td>
<td>6 (8.96)</td>
<td></td>
</tr>
</tbody>
</table>

MRI, magnetic resonance imaging.
*p < 0.05, statistically significant differences.

contrast enhancement and disc abscesses. In terms of paravertebral involvement, the TbS group exhibited a higher prevalence of well-defined paravertebral abscesses (82.93% vs. 44.78%, p < 0.001), abscesses with thin and regular walls (81.3% vs. 2.99%, p < 0.001), epidural abscesses (67.48% vs. 43.28%, p = 0.001), and anterior longitudinal subligamentous spreading (94.3% vs. 41.79%, p < 0.001). However, the presence of epidural phlegmon (4.88% vs. 80.6%, p < 0.001) and facet joint arthritis (46.34% vs. 80.6%, p < 0.001) was strongly associated with PyS. No significant differences were observed between the presence or absence of spinal cord compression.

The duration of symptoms > 4 weeks (odds ratio [OR], 8.19; 95% confidence interval [CI], 4.10–16.33; p < 0.001), Diabetes mellitus (OR, 0.49; 95% CI, 0.15–0.77; p = 0.01), WBC > 10,000/mm³ (OR, 0.21; 95% CI, 0.11–0.40; p < 0.001), neutrophil proportion > 75% (OR, 0.35; 95% CI, 0.19–0.65; p = 0.001), ALP > 120 IU/L (OR, 0.31; 95% CI, 0.17–0.58; p < 0.001), presence of thoracic lesions (OR, 5.67; 95% CI, 2.87–11.20; p < 0.001), severe vertebral destruction (OR, 6.35; 95% CI, 3.14–12.86; p < 0.001), heterogenous contrast-enhanced vertebral body (OR, 19.21; 95% CI, 3.91–94.26; p < 0.001), presence of vertebral intraosseous abscess (OR, 28.06; 95% CI, 10.44–75.36; p < 0.001), well-defined paravertebral enhancement (OR, 84.75; 95% CI, 8.74–820.87; p < 0.001), presence of epidural abscess (OR, 2.75; 95% CI, 1.49–5.07; p = 0.001), absence of facet joint arthritis (OR, 4.88; 95% CI, 2.42–9.84; p < 0.001), and anterior longitudinal subligamentous spreading (OR, 23.28; 95% CI, 9.42–57.49; p < 0.001) were identified as possible risk factors for TbS in the univariate analysis (p < 0.2) in our study (Table 4). No significant differences were found in temperature > 38°C (OR, 0.49; 95% CI, 0.24–1.02; p = 0.01), peak ESR > 40 mm/hr (OR, 0.96; 95% CI, 0.43–2.14; p = 0.93), peak CRP > 5 mg/dL (OR, 0.35; 95% CI, 0.07–1.65; p = 0.19), or ill-defined paravertebral enhancement (OR, 0.21; 95% CI, 0.03–1.49; p = 0.12). Multiple logistic

Table 4. Results of a univariate analysis of possible risk factors for tuberculous spondylodiscitis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tuberculous</th>
<th>Pyogenic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of symptoms (wk)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2</td>
<td>Ref</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2–4</td>
<td>2.29</td>
<td>0.75–6.96</td>
<td>0.144</td>
</tr>
<tr>
<td>&gt; 4</td>
<td>12.15</td>
<td>4.93–29.94</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Temperature (&gt; 38°C)</td>
<td>0.49</td>
<td>0.24–1.02</td>
<td>0.057</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>0.34</td>
<td>0.15–0.77</td>
<td>0.01*</td>
</tr>
<tr>
<td>White blood cells (/mm³)</td>
<td>0.21</td>
<td>0.11–0.40</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Neutrophil (&gt; 75%)</td>
<td>0.35</td>
<td>0.19–0.65</td>
<td>0.001*</td>
</tr>
<tr>
<td>Peak ESR (&gt; 40 mm/hr)</td>
<td>0.96</td>
<td>0.43–2.14</td>
<td>0.927</td>
</tr>
<tr>
<td>Peak CRP (&gt; 5 mg/dL)</td>
<td>0.35</td>
<td>0.07–1.65</td>
<td>0.185</td>
</tr>
<tr>
<td>Alkaline phosphatase (&gt; 120 IU/L)</td>
<td>0.31</td>
<td>0.17–0.58</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Location of lesion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cervical</td>
<td>0.19</td>
<td>0.05–0.63</td>
<td>0.007*</td>
</tr>
<tr>
<td>Thoracic</td>
<td>5.67</td>
<td>2.87–11.20</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Lumbar</td>
<td>0.23</td>
<td>0.10–0.50</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Severe vertebral destruction</td>
<td>6.35</td>
<td>3.14–12.86</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Heterogenous contrast-enhanced vertebral body</td>
<td>19.21</td>
<td>3.91–94.26</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Vertebral intraosseous abscess</td>
<td>28.06</td>
<td>10.44–75.36</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Well-defined paravertebral enhancement</td>
<td>84.75</td>
<td>8.74–820.87</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Ill-defined paravertebral enhancement</td>
<td>0.21</td>
<td>0.03–1.49</td>
<td>0.120</td>
</tr>
<tr>
<td>Epidural abscess</td>
<td>2.75</td>
<td>1.49–5.07</td>
<td>0.001*</td>
</tr>
<tr>
<td>Epidural phlegmon</td>
<td>0.012</td>
<td>0.004–0.033</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>No facet joint arthritis</td>
<td>4.88</td>
<td>2.42–9.84</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>Subligamentous spreading</td>
<td>23.28</td>
<td>9.42–57.49</td>
<td>&lt; 0.001*</td>
</tr>
</tbody>
</table>

CI, confidence interval; ESR, erythrocyte sedimentation rate; CRP, C-reactive protein.
*p < 0.05, statistically significant differences.
regression analysis showed that thoracic lesion (OR, 819.81; 95% CI, 6.84–98,313.95; p = 0.006), absence of epidural phlegmon (OR, 900.86; 95% CI, 31.39–25,857.73; p < 0.001), anterior longitudinal subligamentous spreading (OR, 185.78; 95% CI, 7.92–4,360.64; p = 0.001), presence of vertebral intraosseous abscess (OR, 19.59; 95% CI, 1.75–219.70; p = 0.016), well-defined paravertebral enhancement (OR, 10.79; 95% CI, 1.28–90.80; p = 0.029), presence of epidural abscess (OR, 9.69; 95% CI, 0.78–121.06; p = 0.038), and absence of facet joint arthritis (OR, 7.25; 95% CI, 0.91–57.94; p = 0.042) were independent predictive factors for TbS (Table 5).

1. MRI Scoring Transformation

Each potential predictor of TbS in the multivariate model was assigned a specific score derived from the logistic regression coefficient: thoracic lesion, 7 points, no epidural phlegmon 7 points, subligamentous spreading 5 points, intraosseous abscess 3 points, well-defined paravertebral abscess 2.5 points, epidural abscess 2.5 points, and no facet joint arthritis 2 points (Table 5). The scoring scheme, with a total score ranging from 0 to 29, included categories for differentiation. This cutoff point was based on a calibration plot of sensitivity and specificity. For discriminative ability, the area under the parametric ROC curve for the score-based logistic regression model was 0.96 (95% CI, 125.40–3,257.95) (Fig. 1). The calibration was illustrated using a calibration plot, with a p-value of < 0.001. The predicted probability of TbS increased as the score increased, with a high level of agreement between actual and predicted diseases (Fig. 2). The total score was significantly different between the groups (> 14 points, p < 0.001), with a sensitivity of 97.58% and specificity of 92.54%. The application of this MRI scoring transformation is illustrated in Figs. 3 and 4.

Table 5. Results of multiple logistic regression analysis and scoring system for tuberculous spondylodiscitis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds ratio</th>
<th>Significance</th>
<th>95% CI</th>
<th>Coefficients</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thoracic lesion</td>
<td>819.81</td>
<td>0.006</td>
<td>6.84–98,313.95</td>
<td>6.71</td>
<td>7</td>
</tr>
<tr>
<td>No epidural phlegmon</td>
<td>900.86</td>
<td>&lt; 0.001</td>
<td>31.39–25,857.73</td>
<td>6.80</td>
<td>7</td>
</tr>
<tr>
<td>Subligamentous spreading</td>
<td>185.78</td>
<td>0.001</td>
<td>7.92–4,360.64</td>
<td>5.22</td>
<td>5</td>
</tr>
<tr>
<td>Vertebral Intraosseous abscess</td>
<td>19.59</td>
<td>0.016</td>
<td>1.75–219.70</td>
<td>2.97</td>
<td>3</td>
</tr>
<tr>
<td>Well-defined paravertebral enhancement</td>
<td>10.79</td>
<td>0.029</td>
<td>1.28–90.80</td>
<td>2.38</td>
<td>2.5</td>
</tr>
<tr>
<td>Epidural abscess</td>
<td>9.69</td>
<td>0.038</td>
<td>0.78–121.06</td>
<td>2.27</td>
<td>2.5</td>
</tr>
<tr>
<td>No facet joint arthritis</td>
<td>7.25</td>
<td>0.042</td>
<td>0.91–57.94</td>
<td>1.98</td>
<td>2</td>
</tr>
</tbody>
</table>

CI, confidence interval.

Fig. 1. Using the prediction probability of a multivariate logistic regression model, the receiver operator characteristic (ROC) analysis of magnetic resonance imaging scores for separated tuberculous spondylodiscitis from pyogenic spondylodiscitis is presented.

Fig. 2. Model calibration plots illustrating the predicted probability of tuberculous spondylodiscitis increased as the total score > 14 points with a high level of agreement between actual and predicted risks (sensitivity, 97.58%; specificity, 92.54%).
DISCUSSION

Despite the number of earlier studies\textsuperscript{10-15} that have described the clinical, laboratory, and MRI features of PyS and TbS, we obtained significant distinguishing characteristics from our comparison of the 2 groups. However, no single result can distinguish between these circumstances. Compared to other studies,\textsuperscript{10-16,19} our study represents the largest series comparing microbiologically confirmed cases of PyS and TbS. A longer symptom duration (> 4 weeks) and the absence of fever (Table 2) were more frequently associated with TbS than with PyS. According to Yoon et al.,\textsuperscript{20} TbS risk factors included a median latency to spondylodiscitis diagnosis of > 7 days, and patients with TbS experienced fever less frequently than those with...
PyS. The diagnosis of spinal infection is highly sensitive to inflammatory indicators, such as WBC count, neutrophil count, ESR, and CRP level. Our results were in agreement with the findings of Kim et al., who reported that high levels of ALP (> 120 IU/L) and neutrophil predominance (> 75%) in leukocytosis (> 10,000/mm³) were more commonly predictive of PyS. Similarly, Lertudomphonwanit et al. observed that a neutrophil fraction <78% and WBC count < 9,700/mm³ were highly suggestive diagnostic clues for differentiating patients with TbS from those with PyS. Compared with previous studies, Kim et al. found that ESR > 40 mm³ and CRP > 5 mg/dL were more frequently associated with PyS. In contrast to our study, these biomarker cutoff values were incapable of differentiating PyS from TbS. According to Lertudomphonwanit et al., ESR levels of < 92 mm/hr were highly suggestive indicators of TbS. Nevertheless, CRP level was not shown to be a predictive factor in their study. The demographics of our patient group may partially explain this outcome. As demonstrated in this study, both groups had delayed time to diagnosis; therefore, ESR and CRP levels may have been more significant at the time of diagnosis.

MRI has substantially improved the diagnosis of spinal infections. Even in the early stages of spinal infections, the increased sensitivity of MRI allows the identification of pathogenic alterations in the spine. According to a previous study, contrast-enhanced MRI is a reliable method for differentiating TbS from PyS. Our study demonstrated the presence of thoracic lesions, intrasosseous abscesses, anterior longitudinal subligamentous spreading, and well-defined paravertebral enhancement as predictive factors for TbS, which corresponded well with the review by Lee. Tuberculous spondylitis typically begins in the anterior cancellous bone of the vertebral body. This is followed by the destruction of the vertebral body, extending beneath the anterior longitudinal ligament, leading to the formation of an abscess near the vertebral body. The thoracic spine is the region most frequently affected by this process. MRI findings are supported by a recent study that indicated a well-defined paraspinous abscess as one of the hallmarks of TbS, whereas PyS typically exhibits more widespread, ill-defined areas of enhancement. According to a recent study by Kanna et al., large abscesses with a thin wall are one of the MRI findings that are strongly predictive of TbS.

Epidural soft-tissue thickening, also known as epidural phlegmon, manifests as a diffuse and homogeneous contrast-enhancing process. This presentation may indicate an inflammatory process before turning into an epidural abscess and is less amenable to surgical drainage. Our study revealed a higher prevalence of epidural phlegmon among PyS patients. According to Zhang et al., patients with PyS presented with phlegmon characterized by ill-defined boundaries and occasional small abscesses with thick and irregular walls (97% in PyS vs. 37% in TbS). Conversely, the epidural abscesses that were more common in the TbS group were larger and had well-defined borders. They are more likely to develop into a ring-shaped, thin, smooth-walled, polysoluble abscess. These results were also observed in our study. Patients with TbS more frequently have a slow-growing infection that finally results in an epidural abscess at a later stage.

Septic arthritis of the facet joint was diagnosed based on ero-

![Image](https://doi.org/10.14245/ns.2448120.060)
sion, edema, and enlargement of the facet joint space (Fig. 5). Associated inflammatory changes in the epidural space or adjacent paraspinal muscles can be seen with gadolinium-enhanced T1-weighted MRI. In our study, this finding was a reliable predictor of PyS. Due to the aggressiveness of the organism and its propensity to spread outside the facet capsule, which can cause synovitis, perisynovial inflammation, and erosive changes to the articular surface, the synthesis of a proteolytic enzyme is implicated in the inflammatory process of PyS. Similar to Harada’s results, the enhancement of soft tissues around the facet joints was more frequent in PyS than in TbS.

To the best of our knowledge, there are few diagnostic prediction tools available to effectively distinguish TbS from PyS. Zhang et al. analyzed the MRI findings of spinal infections (32 cases of PyS, 38 cases of TbS), to identify key distinguishing features between PyS and TbS, and establish a systematic scoring method. Using the scoring system, the correct coincidence rate was 95.23%, with a sensitivity of 91.67%, and specificity of 100%. However, the predictive parameters for detecting PyS and TbS were separated using a prediction tool. In our study, we developed a simplified MRI scoring system for the diagnostic prediction of TbS, based primarily on predictive factors (Table 5). Total scores ≥ 14 points may significantly predict the risk of TbS, with a sensitivity of 97.58% and specificity of 92.54%. The discriminative ability of the score-based logistic regression model was 0.96, as indicated by the ROC curve. Figs. 3 and 4 provide demonstrations of the scheme used. As the average duration of symptoms was 3 months in patients with TbS and 4 weeks in patients with PyS, this scoring system is a valuable diagnostic tool that can help distinguish between TbS and PyS, particularly in the subacute to chronic stages of the disease.

The clinical predictive model of this study provided significant advantages. We included all diagnostically relevant variables in the model and transformed the regression equation into a scoring system for use in clinical settings. Nevertheless, this study has some limitations. First, the study, a retrospective cohort study, included patients from a single hospital. However, a data imbalance occurred between groups, which was corrected using multivariable logistic regression analysis. Adjusting the data ratio could potentially reduce the study’s power. Second, the clinical or laboratory data regarding the onset of symptoms may have been biased because our center was a referral center. Larger population studies are required to assess the clinical relevance of these findings. Third, although internal validation was performed in our study, the reproducibility of the scoring remains unknown until a prospective external validation study is conducted in another setting or at a different time. Finally, because we did not include individuals with other low-virulence causative organisms, such as fungi, in our investigation, the generalizability of our findings for these patients may be limited.

CONCLUSION

This study validated the predictive parameters for differentiating TbS from PyS and developed a simplified MRI-based scoring system to predict the likelihood of TbS, assisting clinicians in making judgments when the causative pathogen remains elusive.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: TT, SW; Formal analysis: KC; Investigation: PP, TV, SJ; Methodology: UP; Project administration: WS; Writing – original draft: TT; Writing – review & editing: TT.

ORCID
Terdpong Tanaviriyachai: 0000-0002-6587-0344
Patchara Pornsopanakorn: 0009-0009-5678-6018
Kongtush Choovongkomol: 0000-0002-6999-4547
Urawit Piyapromdee: 0000-0002-8451-5989
Sarut Jongkittanakul: 0000-0002-5459-0953
Weera Sudprasert: 0000-0003-0697-8053

REFERENCES

The Role of Spinal Cord Compression in Predicting Intraoperative Neurophysiological Monitoring Events in Patients With Kyphotic Deformity: A Magnetic Resonance Imaging-Based Study

Zhen Jin1±, Jie Li±±, Hui Xu1, Zongshan Hu1, Yanjie Xu1, Ziyang Tang1, Yong Qiu1,2, Zhen Liu1,2, Zezhang Zhu1,2

1Division of Spine Surgery, Department of Orthopedic Surgery, Nanjing Drum Tower Hospital, The Clinical College of Nanjing Medical University, Nanjing, China
2Division of Spine Surgery, Department of Orthopedic Surgery, Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China

Objective: To establish a novel classification system for predicting the risk of intraoperative neurophysiological monitoring (IONM) events in surgically-treated patients with kyphotic deformity.

Methods: Patients with kyphotic deformity who underwent surgical correction of cervico-thoracic, thoracic, or thoracolumbar kyphosis in our center from July 2005 to December 2020 were recruited. We proposed a classification system to describe the morphology of the spinal cord on T2-weighted sagittal magnetic resonance imaging: type A, circular/symmetric cord with visible cerebrospinal fluid (CSF) between the cord and vertebral body; type B, circular/oval/symmetric cord with no visible CSF between the cord and vertebral body; type C, spinal cord that is fattened/deformed by the vertebral body, with no visible CSF between the cord and vertebral body. Furthermore, based on type C, the spinal cord compression ratio (CR) < 50% was defined as the subtype C-, while the spinal cord CR ≥ 50% was defined as the subtype C+. IONM event was documented, and a comparative analysis was made to evaluate the prevalence of IONM events among patients with diverse spinal cord types.

Results: A total of 294 patients were reviewed, including 73 in type A; 153 in type B; 53 in subtype C- and 15 in subtype C+. Lower extremity transcranial motor-evoked potentials and/or somatosensory evoked potentials were lost intraoperatively in 41 cases (13.9%), among which 4 patients with type C showed no return of spinal cord monitoring data. The 14 subtype C+ patients (93.3%) had IONM events. Univariate logistic regression analysis showed that patients with a type C spinal cord (subtype C-: odds ratio [OR], 10.390; 95% confidence interval [CI], 2.215–48.735; p = 0.003; subtype C+, OR, 497.000; 95% CI, 42.126–5,863.611; p < 0.001) are at significantly higher risk of a positive IONM event during deformity correction compared to those with a type A. In further multiple logistic regression analysis, the spinal cord classification (OR, 5.371; 95% CI, 2.966–9.727; p < 0.001) was confirmed as an independent risk factor for IONM events.

Conclusion: We presented a new spinal cord classification system based on the relative position of the spinal cord and vertebrae to predict the risk of IONM events in patients with kyphotic deformity. In patients with type C spinal cord, especially those in C+ cases, it is essential to be aware of potential IONM events, and adopt standard operating procedures to facilitate neurological recovery.

Keywords: Intraoperative neurophysiological monitoring, Kyphosis, Spinal cord classification
INTRODUCTION

Kyphosis is a progressive sagittal spinal deformity caused by congenital vertebrate defect, degeneration, infection, and trauma of the spine.1,2 Severe kyphosis often causes significant spinal instability, pain, and neurological compromise, which entails surgical correction to address the deformity and mitigate these symptoms. In corrective surgery, one particular concern is the risk of neurological complications. Despite huge efforts to improve surgical techniques and intraoperative neurophysiological monitoring (IONM), the reported risk of neurological complications is still as high as 2.8%.3

Global kyphosis > 90°, vertebral column resection, cervicothoracic/thoracic osteotomy, blood loss > 3,000 mL, and preoperative neurologic deficits are reportedly associated with an elevated risk for neurological complications.4 Given the high risk, IONM has now become indispensable for monitoring neurological function during correction surgery.5 Loss of IONM data is an undesirable event during surgery, with a reported incidence ranging from 3%–27%. Moreover, despite established procedures for tackling IONM events, 3.8% of these events ultimately result in permanent neurological deficits.6,7 Hence, the investigation into the risk factors for IONM events is still warranted clinically.

The relative spacing of the spinal cord within the canal is a critical determinant for spinal cord compression, and kyphotic spine carries a substantial risk of neurological compromise. In thoracic spinal stenosis caused by ossification of the thoracic ligamentum flavum (OLF), the ossified nodules are close to the dura mater, which increases the risk of intraoperative spinal cord injury.8 In lumbar spine disease, the presence of a dural sac exhibiting a homogeneous gray signal devoid of cerebrospinal fluid (CSF) signal is related to an increased likelihood of conservative treatment failure.9 In addition to the encroaching of the spine canal, the distraction of the spinal cord originating from kyphotic deformity potentially precipitate neurological functional impairment.10 Despite the established link between spinal cord compression and neurological function in degenerative spinal disease, its implications for spinal deformities are not elucidated. Based on the presence or absence of visible CSF signals in patients with scoliosis, Sielatycki et al.11 presented a risk classification utilizing axial-T2 magnetic resonance imaging (MRI) of the apex of the curve. This classification indicated the 11.7% of patients type 3 spinal cords, characterized by flattening or deformation caused by the apical concave pedicle or vertebral body, and the absence of visible CSF is at risk for IONM.11

These findings outline a small percentage of patients with coronal deformity with high risk of IONM. But for patient with kyphotic deformity the percentage of severe spinal cord encroachment and distraction can be higher. Originates from protruded vertebral bodies, discs, and posterior elements in each etiology, the spinal cord compression presents with various forms of compression. In extreme cases, the severe compression can cause deformation of cord and complete loss of CSF flow. Despite established knowledge of the relative location of the spinal cord in coronal deformity, its implications for kyphotic deformities are poorly understood. Therefore, in this study, we present a novel classification of spinal cord morphology shape based on sagittal MRI to predict the risk of IONM events in patients with kyphosis.

MATERIALS AND METHODS

1. Patients

Patients who underwent surgical correction of cervicothoracic, thoracic, or thoracolumbar kyphosis in our center from July 2005 to December 2020 were recruited (Table 1). Informed consent was obtained from all patients before participation in

Table 1. Patient characteristics (n = 294)

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>30.2 ± 16.3</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>186 (63.3)</td>
</tr>
<tr>
<td>Female</td>
<td>108 (36.7)</td>
</tr>
<tr>
<td>Etiology</td>
<td></td>
</tr>
<tr>
<td>Scheuermann disease</td>
<td>49 (16.7)</td>
</tr>
<tr>
<td>Congenital kyphosis</td>
<td>79 (26.9)</td>
</tr>
<tr>
<td>Degenerative kyphosis</td>
<td>17 (5.8)</td>
</tr>
<tr>
<td>Old fracture</td>
<td>20 (6.8)</td>
</tr>
<tr>
<td>Spinal tuberculosis</td>
<td>24 (8.2)</td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>105 (35.7)</td>
</tr>
<tr>
<td>Spinal cord classification</td>
<td></td>
</tr>
<tr>
<td>Type A</td>
<td>73 (24.8)</td>
</tr>
<tr>
<td>Type B</td>
<td>153 (52.0)</td>
</tr>
<tr>
<td>Type C</td>
<td>68 (23.1)</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%). Type A, circular/symmetric cord with visible cerebrospinal fluid (CSF) between the cord and vertebral body; type B, circular/oval/symmetric cord with no visible CSF between the cord and vertebral body; type C, spinal cord that is fattened/deformed by the vertebral body, with no visible CSF between the cord and vertebral body.
the study. The inclusion criteria included: (1) patients without obvious scoliosis (Cobb angle < 10°); (2) with preoperative full-spine radiographs and MRIs; and (3) with IONM data including somatosensory evoked potentials (SEPs), motor evoked potentials (MEPs), free running electromyography. Patients with a history of spinal surgery, or with intraspinal lesions such as Chiari malformation, syringomyelia, tethered cord, etc. were excluded. Given the presence of preoperative neurological deficits was identified as an independent risk factor for IONM events,12,13 patients with American Spinal Injury Association (ASIA) A, B, or C were also excluded. Comorbidities were quantified for each patient based on the Charlson Comorbidity Index (CCI).14 The study was approved by the Institutional Review Board of Nanjing Drum Tower Hospital (2021-398-01).

2. Radiographic Parameters

The coronal Cobb angle of the main curve and sagittal global kyphosis were measured on preoperative standing whole spinal x-rays at the PACS (picture archiving and communications systems) workstation. The spinal cord compression ratio (CR) was calculated in terms of anteroposterior/lateral spinal cord diameter (Fig. 1).15 The ossification of the ligamentum flavum is also diagnosed through MRI. Preoperative MRIs were obtained within 6 months of the date of surgery. All radiographic parameters were conducted using Surgimap (v2.3.2.1). Two spinal surgeons (ZJ & JL) who were independent of the operations measured the radiographic assessments, and the mean values were calculated for analysis.

3. IONM Parameters

In our practice, IONM is routinely applied to each case with spine deformity and IONM events are documented in database. The criteria for identifying IONM events were met if one or more of the following conditions occurred: (1) SEP latency increased by more than 10%, (2) SEP amplitude decreased by over 50%, or (3) MEP amplitude decreased by more than 80%.4

4. Classification System

We devised a simple classification system to describe the morphology of the spinal cord as seen on T2-weighted sagittal MRI (Fig. 2).11 Type A cord was defined as a circular/symmetric cord with visible CSF between the cord and the vertebral body. Type B cord was defined as a circular/oval/symmetric cord with no visible CSF between the vertebral body and the cord. Type C cord was defined as a spinal cord that is fattened/deformed by the vertebral body, with no visible CSF between the vertebral body and the cord. Based on type C, given the significantly elevated risk of severe spinal cord deformation or compression leading to potential neurological deficits, the spinal cord CR < 50% was further defined as the subtype C-, while the spinal cord CR ≥ 50% was defined as the subtype C+.15-17

5. Statistical Analysis

The Student t-test and analysis of variance were employed to evaluate discrepancies in continuous variable between the types A, B, subtypes C- and C+ spinal cord groups. Pearson test was used to compare the distribution of demographic, radiological, and surgical parameters among those with IONM alerts and those without IONM alerts. Univariate logistic regression analysis was performed to screen for potential variables related to IONM alerts. In order to avoid multicollinearity in the multivariate regression, one variable was eliminated if the Spearman correlation coefficient was > 0.5.18 A backward elimination method was used to exclude the nonsignificant variables. A post hoc analysis of the power of logistic regression was performed according to Faul et al.19 All statistical analyses were performed using IBM SPSS Statistics ver. 21.0 (IBM Co., Armonk, NY, USA). A p-value of < 0.05 was considered to be statistically significant.
RESULTS

A cohort of 294 patients (186 males and 108 females) was enrolled in this study. Table 1 presents the demographic data of the included patients. The age of the participants was 30.2 years (range, 2–79 years). The etiology of kyphosis includes Scheuermann disease (n = 49), congenital kyphosis (n = 79), degenerative kyphosis (n = 17), old fracture (n = 20), spinal tuberculosis (n = 24), and ankylosing spondylitis (n = 105).

The classification of spinal cords revealed that among all cases, 73 (24.8%) were classified as type A, 153 (52.0%) as type B, 53 (18.0%) as subtype C- and 15 (5.1%) as subtype C+ spinal cords. During the corrective surgical procedures, IONM events were observed in a subset of patients. Specifically, 2 patients (2.7%) with type A spinal cords, 13 patients (8.5%) with type B spinal cords, 12 patients (22.6%) with subtype C- spinal cords, and 14 patients (93.3) with subtype C+ spinal cords experienced IONM events (p < 0.001). Notably, only 2 type B patients with OLF exhibited IONM events (Fig. 3). Additionally, 26 patients (35.6%) with type A spinal cords, 31 patients (20.3%) with type B spinal cords, 22 patients (41.5%) with subtype C- spinal cords, and 10 patients (66.7%) preoperative ASIA is D grade (p < 0.001). Moreover, the patients with type A spinal cords were found to be older compared to those with type B, subtypes C- and C+ spinal cords, with mean ages of 39.4 years versus 26.3 years, 31.3 years and 21.3 years (p < 0.001) (Table 2).

Furthermore, the spinal osteotomy grade of patients with types B and C spinal cords was significantly higher than that of patients with type A spinal cords (p < 0.001). Additionally, patients with type C spinal cords displayed a significantly larger sagittal Cobb angle (subtype C-, 83.9°; subtype C+, 93.6°), in comparison to patients with types A and B spinal cords, who

---

**Fig. 2.** Spinal cord risk classification system. CSF, cerebrospinal fluid; CR, compression ratio.
had Cobb angles of 69.1° and 75.9° (p < 0.001). Notably, patients with type C spinal cords displayed a significantly larger postoperative changes in Cobb angle (subtype C-, 45.4°; subtype C+, 46.8°), in comparison to patients with types A and B spinal cords, who had 37.7° and 41.9° (p = 0.035) (Table 2).

A standardized operating procedure was implemented in response to IONM events, which included instructions to elevate mean arterial pressure to > 90 mmHg to optimize spinal cord perfusion. Any surgical corrective action corresponding to a specific warning signal was promptly reversed. A thorough assessment was conducted to identify spinal cord compression and manage additional mechanical injuries. Following the resolution of IONM alerts, all patients underwent successful completion of the surgery. Subsequently, 36 of those 41 patients exhibited a return of spinal cord monitoring data to baseline after surgery. Notably, all 4 patients whose spinal cord monitoring data did not revert to baseline had a type C spinal cord. The dorsal aspect ratio (DAR) was significantly larger in type C (subtype C-, 12.8°; subtype C+, 12.7°) vs. types A (8.8°) and B (10.5°) spinal cord patients (p < 0.001) (Table 2).

We further tracked the neurological status of the 14 patients with the subtype C+, none of whom experienced worsened neurological damage postoperatively, with neurological function improving in 4 patients (Table 3). Corrective surgery was indicated in these cases to halt kyphosis correction and prevent further deterioration of neurological function.

The risk factors associated with IONM events are shown in Table 4. Pearson tests showed that sagittal DAR (p < 0.001), neurological impairment (ASIA grade D) (p < 0.001), spinal cord classification (p < 0.001), etiology (p < 0.001), and spinal osteotomy classification (p < 0.001) differed significantly between with IONM events and without IONM events groups. There was no significant association between patient’s sex, age, CCI, body mass index, main lesion site, postoperative changes in Cobb angle (Table 4). Univariate logistic regression analysis showed 4 potential risk factors, including larger sagittal DAR (75%–100%: OR, 4.583; 95% CI, 1.731–12.134, p = 0.002), preoperative neurological impairment (ASIA grade D, p < 0.001), type C spinal cord (subtype C-: OR, 10.390; 95% CI, 2.215–48.735; p = 0.003; subtype C+: OR, 497.000; 95% CI, 42.126–5,863.611; p < 0.001), higher spinal osteotomy grade (grade 4: OR, 6.375; 95% CI, 1.322–30.753; p = 0.021; grade 5: OR, 8.870; 95% CI, 2.668–29.482; p < 0.001; grade 6: OR, 5.667; 95% CI, 1.195–26.874; p = 0.029). Due to the lack of significant differences observed among various etiologies, we have opted not to include them in the multiple regression analysis (Table 4). The 4 variables with were selected for multiple logistic regression with backwards stepwise selection (Table 5). In step 1, the sagit-

Fig. 3. Two patients with ossification of the thoracic ligamentum flavum. Neuromonitoring events occurred with passive deformity correction after posterior column osteotomies. Full return of data was seen after apical pediclectomies were done for circumferential spinal cord decompression. (A) A 55-year-old male patient with old fracture. (B) A 57-year-old male patient with congenital kyphosis.
Table 2. Spinal cord classification and results (n = 294)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type A (n = 73)</th>
<th>Type B (n = 153)</th>
<th>Subtype C- (n = 53)</th>
<th>Subtype C+ (n = 15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (yr)</td>
<td>39.4</td>
<td>26.3</td>
<td>31.3</td>
<td>21.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.4 ± 2.9</td>
<td>22.9 ± 4.0</td>
<td>22.8 ± 2.5</td>
<td>22.6 ± 3.4</td>
<td>0.604</td>
</tr>
<tr>
<td>CCI</td>
<td>0.5 ± 0.7</td>
<td>0.7 ± 0.8</td>
<td>0.4 ± 0.6</td>
<td>0.5 ± 0.7</td>
<td>0.070</td>
</tr>
<tr>
<td>Sagittal Cobb</td>
<td>69.1 ± 16.9</td>
<td>75.9 ± 20.1</td>
<td>83.9 ± 21.1</td>
<td>93.6 ± 24.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sagittal DAR</td>
<td>8.8 ± 3.1</td>
<td>10.5 ± 4.4</td>
<td>12.8 ± 4.6</td>
<td>12.7 ± 7.0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Postoperative changes in Cobb</td>
<td>37.7 ± 12.3</td>
<td>41.9 ± 16.2</td>
<td>45.4 ± 18</td>
<td>46.8 ± 26.2</td>
<td>0.035</td>
</tr>
<tr>
<td>Etiology (n)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheuermann disease</td>
<td>13</td>
<td>28</td>
<td>7</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Congenital kyphosis</td>
<td>11</td>
<td>44</td>
<td>15</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Degenerative kyphosis</td>
<td>11</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Old fracture</td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Spinal tuberculosis</td>
<td>3</td>
<td>15</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>25</td>
<td>63</td>
<td>14</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Spinal osteotomy classification (n)</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Grade 1</td>
<td>16</td>
<td>26</td>
<td>12</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Grade 2</td>
<td>25</td>
<td>21</td>
<td>18</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Grade 3</td>
<td>25</td>
<td>67</td>
<td>14</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>3</td>
<td>23</td>
<td>8</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Grade 6</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Preoperative ASIA score</td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Grade D</td>
<td>26 (35.6)</td>
<td>31 (20.3)</td>
<td>22 (41.5)</td>
<td>10 (66.7)</td>
<td></td>
</tr>
<tr>
<td>Grade E</td>
<td>47 (64.4)</td>
<td>122 (79.7)</td>
<td>31 (58.5)</td>
<td>5 (33.3)</td>
<td></td>
</tr>
<tr>
<td>Main lesion site (n)</td>
<td></td>
<td></td>
<td></td>
<td>0.344</td>
<td></td>
</tr>
<tr>
<td>Thoracic spine</td>
<td>65</td>
<td>127</td>
<td>48</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Lumbar spine</td>
<td>8</td>
<td>26</td>
<td>5</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Intraoperative neuromonitoring events</td>
<td>2 (2.7)</td>
<td>13 (8.5)</td>
<td>12 (22.6)</td>
<td>14 (93.3)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation or number (%) unless otherwise indicated.

Type A, circular/symmetric cord with visible cerebrospinal fluid (CSF) between the cord and vertebral body; type B, circular/oval/symmetric cord with no visible CSF between the cord and vertebral body; subtype C-, spinal cord compression ratio (CR) < 50%; subtype C+, spinal cord CR ≥ 50%; BMI, body mass index; CCI, Charlson Comorbidity Index; DAR, dorsal aspect ratio; ASIA, American Spinal Injury Association.

Table 3. The neurologic function changes in 14 patients with subtype C+ undergoing IONM events

<table>
<thead>
<tr>
<th>Preoperative ASIA score</th>
<th>Postoperative ASIA score</th>
<th>Final follow-up ASIA score</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>D</td>
<td>D</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
<td>E</td>
<td>4</td>
</tr>
<tr>
<td>E</td>
<td>E</td>
<td>E</td>
<td>4</td>
</tr>
</tbody>
</table>

Subtype C+, spinal cord compression ratio ≥50%; IONM, intraoperative neurophysiological monitoring; ASIA, American Spinal Injury Association.
Table 4. Analysis of the parameters related to IONM events

<table>
<thead>
<tr>
<th>Variable</th>
<th>Without IONM events</th>
<th>With IONM events</th>
<th>p-value</th>
<th>OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td>0.983</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>160</td>
<td>26</td>
<td></td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>93</td>
<td>15</td>
<td>0.993 (0.500–1.969)</td>
<td>0.983</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.209</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 25%</td>
<td>52</td>
<td>14</td>
<td>Reference</td>
<td>0.222</td>
<td></td>
</tr>
<tr>
<td>25%–49.9%</td>
<td>65</td>
<td>11</td>
<td>0.629 (0.263–1.500)</td>
<td>0.295</td>
<td></td>
</tr>
<tr>
<td>50%–74.9%</td>
<td>69</td>
<td>9</td>
<td>0.484 (0.195–1.205)</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td>75%–100%</td>
<td>67</td>
<td>7</td>
<td>0.388 (0.146–1.031)</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>CCI</td>
<td>0.668</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>129</td>
<td>24</td>
<td>Reference</td>
<td>0.670</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>102</td>
<td>14</td>
<td>0.738 (0.363–1.498)</td>
<td>0.400</td>
<td></td>
</tr>
<tr>
<td>≥ 2</td>
<td>22</td>
<td>3</td>
<td>0.733 (0.203–2.643)</td>
<td>0.635</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>0.224</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>137</td>
<td>21</td>
<td>Reference</td>
<td>0.440</td>
<td></td>
</tr>
<tr>
<td>Below normal</td>
<td>8</td>
<td>0</td>
<td>0 (0)</td>
<td>0.999</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>87</td>
<td>19</td>
<td>1.425 (0.725–2.802)</td>
<td>0.305</td>
<td></td>
</tr>
<tr>
<td>Obese</td>
<td>21</td>
<td>1</td>
<td>0.311 (0.040–2.433)</td>
<td>0.266</td>
<td></td>
</tr>
<tr>
<td>Etiology (1)</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old fracture and degenerative kyphosis</td>
<td>37</td>
<td>0</td>
<td>Reference</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>94</td>
<td>11</td>
<td>0 (0)</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>Scheuermann’s disease</td>
<td>46</td>
<td>3</td>
<td>0 (0)</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>Spinal tuberculosis</td>
<td>20</td>
<td>4</td>
<td>0 (0)</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>Congenital kyphosis</td>
<td>56</td>
<td>23</td>
<td>0 (0)</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>Etiology (2)</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congenital kyphosis</td>
<td>56</td>
<td>23</td>
<td>Reference</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Old fracture and degenerative kyphosis</td>
<td>37</td>
<td>0</td>
<td>0.000 (0.000–0.000)</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>94</td>
<td>11</td>
<td>0.285 (0.129–0.628)</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Scheuermann’s disease</td>
<td>46</td>
<td>3</td>
<td>0.159 (0.045–0.562)</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>Spinal tuberculosis</td>
<td>20</td>
<td>4</td>
<td>0.487 (0.150–1.582)</td>
<td>0.231</td>
<td></td>
</tr>
<tr>
<td>Main lesion site</td>
<td>0.438</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thoracic spine</td>
<td>217</td>
<td>37</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumbar spine</td>
<td>36</td>
<td>4</td>
<td>0.652 (0.219–1.939)</td>
<td>0.441</td>
<td></td>
</tr>
<tr>
<td>Sagittal DAR</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 25%</td>
<td>65</td>
<td>6</td>
<td>Reference</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>25%–49.9%</td>
<td>71</td>
<td>5</td>
<td>0.763 (0.222–2.619)</td>
<td>0.667</td>
<td></td>
</tr>
<tr>
<td>50%–74.9%</td>
<td>65</td>
<td>8</td>
<td>1.333 (0.438–4.058)</td>
<td>0.612</td>
<td></td>
</tr>
<tr>
<td>75%–100%</td>
<td>52</td>
<td>22</td>
<td>4.583 (1.731–12.134)</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Preoperative ASIA score</td>
<td>&lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade E</td>
<td>190</td>
<td>15</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade D</td>
<td>63</td>
<td>26</td>
<td>5.228 (2.605–10.490)</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 4. Analysis of the parameters related to IONM events (Continued)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Without IONM events</th>
<th>With IONM events</th>
<th>p-value</th>
<th>OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postoperative changes in Cobb</td>
<td></td>
<td></td>
<td>0.102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 25%</td>
<td>56</td>
<td>10</td>
<td>Reference</td>
<td>0.117</td>
<td></td>
</tr>
<tr>
<td>25%–49.9%</td>
<td>67</td>
<td>9</td>
<td>0.752 (0.286–1.980)</td>
<td>0.564</td>
<td></td>
</tr>
<tr>
<td>50%–74.9%</td>
<td>71</td>
<td>6</td>
<td>0.473 (0.162–1.381)</td>
<td>0.171</td>
<td></td>
</tr>
<tr>
<td>75%–100%</td>
<td>59</td>
<td>16</td>
<td>1.519 (0.636–3.627)</td>
<td>0.347</td>
<td></td>
</tr>
<tr>
<td>Spinal cord classification</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type A</td>
<td>71</td>
<td>2</td>
<td>Reference</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Type B</td>
<td>140</td>
<td>13</td>
<td>3.296 (0.724–15.009)</td>
<td>0.123</td>
<td></td>
</tr>
<tr>
<td>Subtype C-</td>
<td>41</td>
<td>12</td>
<td>10.390 (2.215–48.735)</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>Subtype C+</td>
<td>1</td>
<td>14</td>
<td>497.000 (42.126–5,863.611)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Spinal osteotomy classification</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 1</td>
<td>51</td>
<td>4</td>
<td>Reference</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Grade 2</td>
<td>64</td>
<td>1</td>
<td>0.199 (0.022–1.838)</td>
<td>0.155</td>
<td></td>
</tr>
<tr>
<td>Grade 3</td>
<td>98</td>
<td>12</td>
<td>1.561 (0.479–5.086)</td>
<td>0.460</td>
<td></td>
</tr>
<tr>
<td>Grade 4</td>
<td>8</td>
<td>4</td>
<td>6.375 (1.322–30.753)</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>Grade 5</td>
<td>23</td>
<td>16</td>
<td>8.870 (2.668–29.482)</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Grade 6</td>
<td>9</td>
<td>4</td>
<td>5.667 (1.195–26.874)</td>
<td>0.029</td>
<td></td>
</tr>
</tbody>
</table>

IONM, intraoperative neurophysiological monitoring; OR, odds ratio; CI, confidence interval; CCI, Charlson Comorbidity Index; BMI, body mass index; Type A, circular/symmetric cord with visible cerebrospinal fluid (CSF) between the cord and vertebral body; Type B, circular/oval/symmetric cord with no visible CSF between the cord and vertebral body; subtype C-, spinal cord compression ratio (CR) < 50%; subtype C+, spinal cord CR ≥ 50%.

1Pearson test. 2Univariate logistic regression analysis.

Table 5. Multivariate logistic regression analysis related to IONM events

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)*</th>
<th>p-value</th>
<th>OR (95% CI)*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal cord classification</td>
<td>5.181 (2.851–9.414)</td>
<td>&lt;0.001</td>
<td>5.371 (2.966–9.727)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Spinal osteotomy classification</td>
<td>1.701 (1.278–2.264)</td>
<td>&lt;0.001</td>
<td>1.739 (1.304–2.319)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Preoperative ASIA score D (vs. E)</td>
<td>3.037 (1.291–7.140)</td>
<td>0.011</td>
<td>3.221 (1.376–7.538)</td>
<td>0.007</td>
</tr>
<tr>
<td>Sagittal DAR (per 25 percentage interval)</td>
<td>1.204 (0.790–1.835)</td>
<td>0.389</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

IONM, intraoperative neurophysiological monitoring; OR, odds ratio; CI, confidence interval; ASIA, American Spinal Injury Association; DAR, dorsal aspect ratio; N/A, not available.

*All variables entered in the first step; each 2 of the 4 variables showed Spearman correlation coefficient < 0.50. †Sagittal DAR variable was removed.

DISCUSSION

In this study, we introduce a novel spinal cord classification system based on sagittal plane MRI for predicting the risk of IONM events in patients with kyphotic deformity. Our findings demonstrate a markedly increased risk of IONM events among patients with subtype C- spinal cord (OR, 10.390), and an even higher risk in subtype C+ (OR, 497.000) during deformity correction procedures, compared to those with a normal spinal cord morphology and CSF presence between the spinal cord and vertebral body (type A). In our multifactorial regression analysis, we have further identified the spinal cord classification as an independent risk factor. This system is a straightforward and effective approach to evaluating the hazard of neuromonitoring events before kyphosis correction. In cases with type C (including subtypes C- and C+) spinal cords, it is of paramount importance to avoid too aggressive correction of deformity. Instead, decompression of the cord and the preservation of neurological function should be prioritized.

On top of all, we identified cases with type C spinal cord had...
the higher risk of neurological complications. In such cases, the spinal cord is subjected to both the compression of the protruded spinal column and the distraction force that alters the regular shape of the cord. Radiographically, the preoperative radiographic parameter known as the DAR was introduced as a predictor of which patient groups were at higher risk for IONM events. The DAR was determined as the coronal Cobb angle divided by the number of vertebrae encompassed within the curve. An elevated DAR value indicated a more severe, acutely angled curve with an increased likelihood of IONM changes.

Our findings indicate that patients with a type C spinal cord demonstrate a markedly larger sagittal DAR value, resulting in deformation of the spinal cord and subsequent impairment of neurological function. Furthermore, in these cases, the distraction and torsion of the spinal cord may cause paraplegia and hypoxia of the cord. Notably, all patients included in this study had a DAR larger than 7, indicating a severe kyphotic deformity. In our multifactorial regression analysis, we have further determined that DAR is not an independent risk factor. In this particular, the relative position and morphology of the spinal cord may serve as a more relevant measure that accounts for the risk of IONM in corrective surgery than the severity of deformity.

Due to the relatively severe spinal cord compression, the incidence of IONM events in the subtype C+ case is as high as 93.3% (Table 2). Notably, we also observed some unique cases with concurrent IONM events. In 3 patients with congenital chondrodysplasia, the spinal cord was compressed by the vertebral body, with no CSF between the spinal cord and vertebral body or lamina (Fig. 4A). Two patients with ankylosing spondylitis developed idiopathic spinal cord herniation (Fig. 4B). Given that the subtype C+ spinal cord indicates a highly stretched condition, and further elevated risk of IONM based on our analysis, we believe it to be a crucial risk factor for IONM events that require our heightened vigilance.

Our analysis revealed that type B spinal cords (OR, 0.752; p = 0.564) did not exhibit a significantly elevated risk of IONM compared to type A. Our hypothesis suggests that type B spinal cords maintain regular spinal morphology and sufficient blood supply, thus minimizing the risk of injury during correction. However, our investigation revealed that IONM events occurred in only 2 type B spinal cord patients with OLF. Previous study has identified OLF as a significant risk factor for IONM events and can lead to a range of postoperative complications. Thus, it is crucial that we remain vigilant in our preparation and proactively take measures to prevent IONM events in patients with OLF.

The primary strength of this classification system is its simplicity, allowing for quick and easy assessment of spinal cord morphology on sagittal MRI preoperatively, without the need for invasive examinations. Furthermore, the identification of type C spinal cords serves as a strong indicator of the increased risk of experiencing an IONM event during deformity correction. Establishing a standardized operating procedure in response to type C spinal cord presence can facilitate prompt management of the issue through direct spinal cord decompression or by accepting a lesser curve correction. Corrective surgery was indicated in these cases to prevent kyphosis worsening and further deterioration of neurological function, suggesting that patients with type C spinal cord conditions are not contraindicated for surgery. Comparatively, type A patients have a lower overall risk of IONM events. For these cases, care should be taken to restore the physiological spinal morphology, with caution paid to cases exhibiting large sagittal DAR.

Fig. 4. (A) A 24-year-old male patient with congenital chondrodysplasia. The spinal cord is fattened by the vertebral body, with no visible cerebrospinal fluid between the spinal cord and vertebral body or lamina. (B) A 48-year-old male patient suffered from ankylosing spondylitis with idiopathic spinal cord herniation. The 2 conditions are defined as C+ type spinal cord.
We believe that this classification system is simple yet informative, providing valuable guidance for alerting the risk of neurological deficit. Furthermore, the classification system presented by Sielatycki et al.\textsuperscript{11} could be integrated with our approach to evaluating more complex spinal deformities, thereby enabling a more comprehensive prediction of the risk of neurological complications. However, this study had some limitations, including a relatively multiple etiology and a heterogeneous population. Additionally, the number of patients with OLF was relatively small. Furthermore, the study was conducted in a single surgeon’s practice, working with experienced assistants and an intraoperative neuromonitoring team. Therefore, the results may not be generalizable to all spinal surgery practices.

CONCLUSION

Here we presented a novel sagittal MRI-based spinal cord classification system to predict the risk of IONM event in patients with kyphotic deformity. Specifically, our observations revealed that patients with type C spinal cord, particularly those with the subtype C+, are subject to a significantly elevated risk of experiencing IONM events. These findings have important implications for patients undergoing kyphosis correction, as it is imperative to prepare for potential IONM events by implementing vertebral column resection and/or circumferential spinal cord decompression. Our novel classification system explicitly outlines the different severity and pattern of spinal cord compression in the context of kyphotic deformity, thereby offering a valuable tool for clinicians to identify high-risk patients and implement appropriate surgical correction.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study was supported by National Natural Science Foundation of China (NSFC) (No. 82272545), 333 High Level Talents Cultivation Project of Jiangsu Province ((2022)3-1-238), Clinical Trials from the Affiliated Drum Tower Hospital, Medical School of Nanjing University (2022-LCYJ-MS-22).

Author Contribution: Conceptualization: YQ, ZL, ZZ; Data curation: ZJ, ZT; Formal analysis: ZJ, JL, HX, ZH, YX, ZT, ZZ; Funding acquisition: HX, ZL, ZZ; Methodology: ZJ, JL, HX, ZH, YX, YQ, ZL, ZZ; Project administration: ZL, ZZ; Visualization: JL, ZH, ZL; Writing – original draft: ZJ, JL, ZH, YX, ZT; Writing – review & editing: YQ, ZL, ZZ.

ORCID

Zhen Jin: 0000-0003-3647-9845
Jie Li: 0000-0003-1096-9328
Hui Xu: 0000-0002-6401-8483
Zongshan Hu: 0000-0002-9907-4389
Yanjie Xu: 0000-0002-3128-5710
Ziyang Tang: 0000-0003-1243-4634
Yong Qiu: 0000-0002-2444-4744
Zhen Liu: 0000-0003-2027-8088
Zezhang Zhu: 0000-0001-8569-3595

REFERENCES

classification of spinal cord shape and CSF presence at the curve apex to assess risk of intraoperative neuromonitoring data loss with thoracic spinal deformity correction. Spine Deform 2020;8:655-61.


Prediction of Screw Loosening After Dynamic Pedicle Screw Fixation With Lumbar Polyetheretherketone Rods Using Magnetic Resonance Imaging-Based Vertebral Bone Quality Score

Guozheng Jiang1,*, Luchun Xu1,*, Yukun Ma1, Jianbin Guan2, Yongdong Yang1, Wenqing Zhong1, Wenhai Li1, Shibo Zhou1, JiaWei Song1, Ningning Feng1, Ziyu Qiu1, Zeyu Li1, YiShu Zhou1, Letian Meng1, Yi Qu1, Xing Yu1

1Department of Neurosurgery, Dongzhimen Hospital, Beijing University of Traditional Chinese Medicine, Beijing, China
2Department of Neurosurgery, Honghui-Hospital, Xi’an Jiaotong University, Xi’an, China

Objective: To investigate the correlation between magnetic resonance imaging-based vertebral bone quality (VBQ) score and screw loosening after dynamic pedicle screw fixation with polyetheretherketone (PEEK) rods, and evaluate its predictive value.

Methods: A retrospective analysis was conducted on the patients who underwent dynamic pedicle screw fixation with PEEK rods from March 2017 to June 2022. Data on age, sex, body mass index, hypertension, diabetes, hyperlipidemia history, long-term smoking, alcohol consumption, VBQ score, L1–4 average Hounsfield unit (HU) value, surgical fixation length, and the lowest instrumented vertebra were collected. Logistic regression analysis was employed to assess the relationship between VBQ score and pedicle screw loosening (PSL).

Results: A total of 24 patients experienced PSL after surgery (20.5%). PSL group and non-PSL group showed statistical differences in age, number of fixed segments, fixation to the sacrum, L1–4 average HU value, and VBQ score (p < 0.05). The VBQ score in the PSL group was higher than that in the non-PSL group (3.56 ± 0.45 vs. 2.77 ± 0.31, p < 0.001). In logistic regression analysis, VBQ score (odds ratio, 3.425; 95% confidence interval, 1.552–8.279) were identified as independent risk factors for screw loosening. The area under the receiver operating characteristic curve for VBQ score predicting PSL was 0.819 (p < 0.05), with the optimal threshold of 3.15 (sensitivity, 83.1%; specificity, 80.5%).

Conclusion: The VBQ score can independently predict postoperative screw loosening in patients undergoing lumbar dynamic pedicle screw fixation with PEEK rods, and its predictive value is comparable to HU value.

Keywords: Vertebral bone quality, Pedicle screw loosening, Lumbar degenerative, Dynamic pedicle fixation, Bone quality

INTRODUCTION

The titanium rod-assisted lumbar fusion surgery, combined with pedicle screws, is currently considered the “gold standard” for treating degenerative diseases of the lumbar spine.1,2 However, after lumbar interbody fusion, increased mobility and abnormal stress in adjacent segments ultimately lead to adjacent segment degeneration (ASD).3,4 With the advancement of bio-
mechanical research and material development, various instruments and surgical techniques that preserve mobility have emerged clinically, such as the Dynesys system,\textsuperscript{14,15} Isobar TTL system,\textsuperscript{16} and polyetheretherketone (PEEK) rods system.\textsuperscript{17} These dynamic fixation technologies aim to prevent or slow down the occurrence of ASD.

Postoperative screws in dynamic fixation are subjected to long-term stress at the bone interface, and pedicle screw loosening (PSL) is one of the most common complications following dynamic fixation, especially in patients with osteoporosis.\textsuperscript{10-12} A precise preoperative assessment of the patient’s bone quality is clinically significant for identifying the critical threshold range of vertebral bone quality associated with dynamic fixation-related PSL.

Although dual-energy x-ray absorptiometry (DEXA) is currently recognized as the gold standard for diagnosing osteoporosis, it still has certain limitations.\textsuperscript{13} Recently, the VBQ score based on magnetic resonance imaging (MRI) has been proposed as a more efficient and accurate method for detecting osteoporosis.\textsuperscript{14} The latest research indicates that the VBQ score is an independent risk factor for PSL after interbody fusion and lumbar rigid fixation surgery.\textsuperscript{13,15,16}

Currently, there is no research assessing the correlation between VBQ scores and postoperative PSL in dynamic fixation with the PEEK rods system. This study retrospectively analyzes the correlation between dynamic fixation-related PSL and VBQ scores, aiming to explore the critical range of VBQ scores for predicting PSL.

\section*{MATERIALS AND METHODS}

\subsection*{1. Patients}

This study conducted a retrospective analysis of patients who underwent lumbar spine dynamic fixation surgery using the PEEK rods system due to degenerative diseases from March 2017 to June 2022. The study obtained approval from Institutional Review Board of Dongzhimen Hospital Affiliated to Beijing university of Chinese Medicine (2022DZMEC-085-04). As a retrospective study, informed consent was deemed unnecessary.

Inclusion criteria were as follows: (1) clinically symptomatic and radiologically diagnosed with degenerative diseases of the lumbar spine (such as lumbar disc herniation, lumbar spinal stenosis, degenerative spondylolisthesis or instability of the lumbar spine), with ineffective conservative treatment for over 6 months, or effective conservative treatment for over 3 months but experiencing recurrent symptoms severely impacting daily activities and normal work; (2) patients retaining a certain degree of mobility in the lumbar spine segments involved; (3) undergoing posterior lumbar surgery using the pedicle screw system with PEEK rods for dynamic fixation, with lumbar MRI performed in our hospital within one month before surgery, and lumbar computed tomography (CT) scan within 3 months before surgery.

Exclusion criteria included: (1) a history of previous lumbar spine surgery; (2) lumbar spondylolysis, structural retroversion, severe intervertebral space narrowing; (3) severe osteoporosis, coronal imbalance of the spine, lateral displacement of the vertebral body > 1 cm, and coronal vertebral body wedge compression > 1/3; (4) patients with severe menopausal symptoms, mental disorders.

\subsection*{2. Evaluation of Clinical and Radiological Parameters}

All Individual data recorded patient age, sex, body mass index (kg/m$^2$), history of hypertension, diabetes, hyperlipidemia, and long-term smoking and drinking. Additionally, VBQ score and L1–4 average HU values were documented. Surgical data included the length of surgical fixation (number of segments) and the lowest instrumented vertebra (LIV).

\subsection*{3. Calculation of the VBQ Score}

Following the method proposed by Ehresman et al.,\textsuperscript{17} VBQ scoring was conducted using T1-weighted images from lumbar spine MRI. On the midsagittal plane of the lumbar spine, the region of interest (ROI) was placed between the trabecular part of the L1–4 vertebral bodies and the cerebrospinal fluid (CSF) gap at the level of L3. The average signal intensity (SI) within each ROI of the vertebral bodies and the average SI of CSF at the L3 level were recorded (Fig. 1). The VBQ score was calculated as the median SI of L1–4 vertebral bodies divided by the average SI of CSF [VBQ score = median (SIL1-4)/SICSF]. In cases where the midsagittal plane could not accurately measure the ROI (vascular tumors, changes in spinal lateral curvature), a parasagittal plane was used as a substitute for the midsagittal plane ROI. If the entire vertebral body could not be measured, that vertebra was excluded. If the CSF at the L3 level was completely obstructed, the ROI for CSF was placed at the L2 or L4 level.

\subsection*{4. Measurement of HU Value}

Three transverse sections parallel to the upper and lower endplates and the midhorizontal plane of the vertebral body were
selected from CT images. ROI were delineated within the trabecular bone, avoiding cortical bone, areas of local bone hyperplasia, and the vertebral venous plexus. The HU value within the ROI was then obtained (Fig. 2). The mean values from the 3 planes were calculated to represent the HU value of the vertebral body. The L1–4 average HU value was calculated as the average HU value of the L1–4 vertebrae.

5. Pedicle Screws Loosening

We conducted lumbar spine x-ray examinations (anterior-posterior, lateral, flexion-extension views) for follow-up patients to assess screw loosening. Screw loosening was defined as the presence of a radiolucent zone of 1 mm or more around the pedicle screw.18 For suspected loosening cases, lumbar spine CT scans were performed, and in the coronal reconstruction images with bone window settings, the presence of a ring-shaped low-density shadow around the screw was considered indicative of screw loosening.19 Based on the follow-up results, patients were categorized into the PSL group and the non-PSL group.

Two experienced surgeons independently measured bone quality parameters (HU score, VBQ score) and extracted other clinical data. To assess reliability, 1 month later, each observer remeasured the same 30 cases.

6. Statistical Analysis

All data were statistically analyzed using IBM SPSS Statistics ver. 25.0 (IBM Co., Armonk, NY, USA). The intraclass correlation coefficient (ICC) was employed to assess intraobserver and interobserver consistency. All continuous variables are presented as mean ± standard deviation and analyzed using Student t-test. Categorical variables are reported as frequencies and percentages, and evaluated using the chi-square test. Logistic regression analysis was employed to identify independent risk factors for screw loosening. Receiver operating characteristic (ROC) curve analysis was utilized to assess the predictive value of VBQ scores for PSL, and the Youden index was used to establish the optimal threshold for VBQ scores. A p-value < 0.05 was considered significant.

RESULTS

This study included 117 patients undergoing dynamic fixation with PEEK rods for lumbar spine diseases, including 69 males and
48 females, with an average age of 59.4 ± 9.1 years (43–71 years). The follow-up period was 37.5 ± 11.4 months (14–55 months). All patients were divided into the PSL group (24 cases, 20.5%) and the non-PSL group (93 cases, 79.5%). There were no revision surgeries required for either group during the 12 months postoperatively.

1. Demographics and Operative Profiles

The age, range of fixation, LIV at S1, VBQ score, and HU value exhibited statistically significant differences between the PSL group and the non-PSL group (p < 0.05). Additionally, patients in the PSL group had lower HU values (p < 0.001) and higher VBQ scores (p < 0.001) than those in the non-PSL group (Table 1).

2. Prediction of PSL Using VBQ Scores

Variables with p < 0.05 in Table 1 were considered potential risk factors influencing screw loosening and were included in the multivariate logistic regression analysis (Table 2).

The results indicated that fixation length (odds ratio [OR], 3.749; 95% confidence interval [CI], 1.312–9.136), LIV at S1 (OR, 3.338; 95% CI, 1.655–7.091), average vertebral CT value of L1–4 (OR, 1.031; 95% CI, 1.014–1.263), and VBQ score (OR, 3.425; 95% CI, 1.552–8.279) were independent risk factors for screw loosening; age was not a predictive factor for screw loosening. Furthermore, there was a significant negative correlation between VBQ score and HU values (Fig. 3). The VBQ score and HU value was shown to have excellent interrater reliability with an ICC > 0.8.

3. Evaluation of Predictive Scoring Measures

The area under the ROC for VBQ scores and HU values in predicting screw loosening were 0.819 and 0.793, respectively (Fig. 4). The optimal threshold for predicting screw loosening was determined by the maximum Youden index, and it was found to be 3.15 for VBQ scores (sensitivity, 0.831; specificity, 0.805).

DISCUSSION

PSL is a common complication after lumbar fixation surgery.
using the pedicle screw system, particularly with a significantly higher incidence in patients with osteoporosis. Studies have shown that the proportion of screw loosening in patients with normal bone density ranges from 1% to 27%, while in osteoporotic patients, the proportion can be as high as 60%. Therefore, accurate bone quality assessment in patients undergoing lumbar fixation with the pedicle screw system is crucial for guiding preoperative planning and selecting postoperative measures to mitigate bone loss.

The most commonly used tool in clinical practice for assessing bone mineral density (BMD) is DEXA. However, in patients with degenerative lumbar diseases, factors such as osteophyte formation, facet joint degeneration, and aortic calcification can make DEXA less accurate in evaluating the trabecular bone BMD in the vertebral bodies. Additionally, not all patients undergo routine DEXA examinations before surgery. Therefore, there is a need for a convenient and efficient alternative to DEXA for bone quality assessment.

Previous research has found that CT-HU value can be used for bone quality assessment and have good predictive value for PSL after lumbar rigid fixation surgery. A recently developed MRI-based method predicts PSL, offering a new tool for evaluating PSL after lumbar rigid fixation fusion surgery, alongside HU values.

Histological analysis reveals that osteoporotic bone is typically characterized by the replacement of local fat cells and the predominant feature of trabecular bone atrophy. Additionally, an increase in fat infiltration and excessive trabecular bone atrophy can result in higher SI on T1-weighted MRI images. Based on this theory, Ehresman et al. first proposed the VBQ score, which introduces CSF signal to eliminate differences in individual patient and scanner baseline signals compared to previous MRI-based scores. Due to its simplicity, feasibility, and resistance to confounding factors, VBQ score has rapidly become a practical tool for assessing vertebral bone quality in clinical settings. Previous studies have indicated that VBQ score serves as an independent predictor for PSL after lumbar rigid fixation surgery.

Currently, instrumentation for posterior lumbar fixation via the pedicle screw system mainly comprises 2 categories: dynamic rods, which allow for some degree of motion, and titanium rods, which offer minimal to no motion. PEEK rods belong to the category of instruments for dynamic lumbar fixation and are applied in nonfusion lumbar surgeries in this study, whereas titanium rods are commonly utilized in fusion lumbar surgeries. These 2 types of instrumentation are guided by distinct core principles, resulting in variations in screw loosening rates. Titanium rods prioritize robust fixation of the operative segment to promote fusion, while dynamic rods aim to preserve lumbar mobility while ensuring segmental stability to mitigate complications arising from diminished lumbar mobility. Presently, consensus is lacking regarding the comparative postoperative screw loosening rates between dynamic and rigid fixation. We underscore the importance of addressing postoperative screw loosening in dynamic fixation procedures. Following complete fusion of the operative segment in rigid fixation surgeries, the function of screws and titanium rods is essentially fulfilled. In contrast, dynamic rods continue to provide lumbar support and withstand sustained mechanical stress. Should screw loosening occur, it could significantly compromise long-term lumbar stability and potentially expedite degeneration of adjacent segments. Consequently, compared to fusion fixation, greater attention should be directed toward bone quality assessment in patients undergoing dynamic lumbar fixation surgeries.

To the best of our knowledge, this is the first study to assess the correlation between VBQ score and PSL after dynamic fixation surgery. In this study, PSL was identified using a combination of x-ray and CT methods, with the highest incidence occurring approximately 1 year postoperatively during follow-up. Interestingly, as the follow-up period lengthened, the rate of screw loosening gradually decreased, with some loosening patients showing gradual improvement and eventual nonloosening status, a phenomenon consistent with a recent report by Shu et al.
Patients in the PSL group were older, had more fixed segments, had a higher proportion of fixation to S1, lower HU values of L1–4 vertebral bodies, and significantly higher VBQ scores. Multivariable logistic regression analysis indicated that the number of fixed segments, LIV at S1, HU values of L1–4 vertebral bodies, and VBQ scores were independent risk factors for screw loosening, while age was not an independent risk factor for screw loosening in this study, which may be related to the relatively small number of cases included. The predictive value of HU values and VBQ scores was relatively ideal. These results are generally consistent with previous studies on screw loosening after lumbar rigid fixation surgery.\textsuperscript{13,27,29,34}

The longer the fixed segment stabilized within the lumbar spine through pedicle screw fixation, the more likely the occurrence of PSL. In this study, the proportion of 3-segment and 4-segment PSL groups was 45.8%, significantly higher than the proportion in the non-PSL group, which was 19.4% (Fig. 5). This may be attributed to the increased stress on the screws at the proximal and distal ends with a greater number of fixed segments, leading to a higher likelihood of screw loosening.\textsuperscript{35}

The unique location and anatomical structure of the sacrum may contribute to a higher postoperative loosening rate of sacral screws. Literature reports the occurrence rate of sacral PSL to be in the range of 15.6%–46.5%, significantly higher than the loosening rate observed in lumbar PSL.\textsuperscript{36,37} L5–S1, serving as the transitional zone between the lumbar spine and sacrum, not only bears a substantial portion of the body weight but also experiences increased shear forces at the lumbosacral angle, further intensifying local loads.\textsuperscript{36,39} The uneven distribution of trabecular bone density within the sacrum poses challenges, with higher trabecular bone density near the anterior cortex region and lower density in the posterior cortex regions adjacent to the sacral wings.\textsuperscript{30} Consequently, sacral screws often require sufficient depth of insertion to obtain an adequate grip. The first sacral segment exhibits a wider pedicle inner diameter and weaker cortical support for screw fixation. Therefore, identifying predictive indicators for sacral screw loosening after sacral fixation becomes particularly crucial.

Current research indicates that osteoporosis is one of the risk factors influencing screw loosening.\textsuperscript{41} Bone quality is a primary determinant of stability for pedicle screws since it governs the strength of the screw-bone interface.\textsuperscript{42} The HU values of lumbar vertebral bodies can accurately reflect the strength of trabecular bone density, demonstrating good practicality in assessing bone density. Some literature has confirmed that HU values can serve as one of the predictive factors for postoperative screw loosening in the lumbar spine.\textsuperscript{33,44}

While bone density is generally considered a component of skeletal strength, it cannot be the sole indicator of skeletal mechanical performance. During the process of osteoporosis, local fat cells gradually replace trabecular bone, weakening the overall bone structure.\textsuperscript{45} VBQ score reflects the extent of fat infiltration in vertebral trabecular bone, offering another perspective on bone quality. This study reveals that higher VBQ scores in most patients correlate with a higher occurrence rate of screw loosening. Additionally, there is a moderate negative correlation with HU values (p < 0.001), indicating that patients with higher VBQ scores tend to have correspondingly lower HU values. The predictive value of VBQ scores for screw loosening is generally comparable to HU values. Furthermore, this study suggests that VBQ score is consistently reliable, easy to operate, less influenced by confounding factors, and serves as a practical tool for auxiliary assessment of bone quality.

This study also has some limitations. Different field strengths (1.5/3.0 T) may have a significant impact on VBQ scores. The study of Lin et al.\textsuperscript{46} showed that VBQ1.5T scores were more sensitive for identifying osteoporosis compared to VBQ3.0T scores. Considering these factors, all patient images included in this study were obtained under a 1.5 T magnetic field, which may introduce bias into the final results. All cases in this study were sourced from a single center, and the relatively limited number of cases may render some risk factors influencing PSL, as reported in previous studies, statistically insignificant in this study. Additionally, some patients in this study exhibited lower HU values (< 50 HU) or higher VBQ scores (> 3.5), but did not experience screw loosening during the postoperative follow-up. PSL is a result influenced by multiple factors, and the included influencing factors in this study are still not comprehensive enough.

**CONCLUSION**

The VBQ score can independently predict postoperative PSL...
in patients undergoing dynamic fixation with lumbar pedicle screws and PEEK rods, and its predictive value is comparable to HU values.

NOTES

Conflict of Interest: The authors have nothing to disclose.
Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: GJ, XY; Data curation: GJ, WZ, JS, NF; Formal analysis: YM, WL, JS, NF, YZ; Methodology: JG, YY, ZQ; Project administration: LX, SZ, YQ, XY; Visualization: NF, ZL, LM; Writing – original draft: GJ, LX; Writing – review & editing: JG, XY.

ORCID
Guozheng Jiang: 0000-0002-6336-3795
Luchun Xu: 0000-0001-7736-007X
Yukun Ma: 0009-0004-7405-2011
Jianbin Guan: 0009-0008-9046-1965
Yongdong Yang: 0000-0002-1732-7742
Wenqing Zhong: 0009-0008-8657-940X
Wenhai Li: 0000-0001-6091-0346
Shibo Zhou: 0000-0003-1260-0807
JiaWei Song: 0009-0008-1413-4567
Ningning Feng: 0000-0001-6464-3599
Ziye Qiu: 0009-0007-8461-6230
Zeyu Li: 0009-0001-4872-3959
YiShu Zhou: 0009-0002-9550-3606
Xing Yu: 0000-0003-3598-3492

REFERENCES

44. Li J, Zhang Z, Xie T, et al. The preoperative Hounsfield unit value at the position of the future screw insertion is a better predictor of screw loosening than other methods. Eur Radiol 2023;33:1526-36.
Comparison of Surgical Burden, Radiographic and Clinical Outcomes According to the Severity of Baseline Sagittal Imbalance in Adult Spinal Deformity Patients

Se-Jun Park¹, Jin-Sung Park¹, Dong-Ho Kang¹, Hyun-Jun Kim¹, Yun-Mi Lim¹, Chong-Suh Lee²

¹Department of Orthopedic Surgery, Spine Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea
²Department of Orthopedic Surgery, Haeundae Bumin Hospital, Busan, Korea

Objective: To determine the clinical impact of the baseline sagittal imbalance severity in patients with adult spinal deformity (ASD).

Methods: We retrospectively reviewed patients who underwent ≥ 5-level fusion including the pelvis, for ASD with a ≥ 2-year follow-up. Using the Scoliosis Research Society-Schwab classification system, patients were classified into 3 groups according to the severity of the preoperative sagittal imbalance: mild, moderate, and severe. Postoperative clinical and radiographic results were compared among the 3 groups.

Results: A total of 259 patients were finally included. There were 42, 62, and 155 patients in the mild, moderate, and severe groups, respectively. The perioperative surgical burden was greatest in the severe group. Postoperatively, this group also showed the largest pelvic incidence minus lumbar lordosis mismatch, suggesting a tendency towards undercorrection. No statistically significant differences were observed in proximal junctional kyphosis, proximal junctional failure, or rod fractures among the groups. Visual analogue scale for back pain and Scoliosis Research Society-22 scores were similar across groups. However, severe group’s last follow-up Oswestry Disability Index (ODI) scores significantly lower than those of the severe group.

Conclusion: Patients with severe sagittal imbalance were treated with more invasive surgical methods along with increased the perioperative surgical burden. All patients exhibited significant radiological and clinical improvements after surgery. However, regarding ODI, the severe group demonstrated slightly worse clinical outcomes than the other groups, probably due to relatively higher proportion of undercorrection. Therefore, more rigorous correction is necessary to achieve optimal sagittal alignment specifically in patients with severe baseline sagittal imbalance.

Keywords: Adult spinal deformity, Baseline severity, Sagittal imbalance, Clinical outcome, Radiographic outcome, Deformity correction

INTRODUCTION

Adult spinal deformity (ASD) is a debilitating condition associated with sagittal malalignment causing substantial pain and functional disability.¹⁻⁴ It is well known that increased sagittal deformity leads to worse health-related quality of life.⁵⁻⁷ Therefore, the optimal restoration of spino-pelvic malalignment has been a cornerstone of surgical management for ASD for
achieving good clinical outcomes.\(^*\)\(^9\)\(^\) Several authors have suggested the optimal surgical targets, including Scoliosis Research Society (SRS)-Schwab classification, age-adjusted sagittal alignment goals, and Global Alignment and Proportion (GAP) score.\(^\)\(^{10-12}\) Although these systems have their own correction targets, the common determinant factors are patient’s age and pelvic incidence (PI). Therefore, the current guidelines will propose the same surgical target without considering of the severity of baseline sagittal imbalance in patients of the same age and PI.

ASD is a disease entity with a wide spectrum of severity. For patients with mild sagittal deformity, only a small gap exists between the current sagittal imbalance status and the surgical target. Therefore, less-morbid surgery may be sufficient to achieve optimal sagittal correction. In contrast, patients with severe sagittal imbalance will have a larger gap to the desired correction target from the current deformity status, frequently necessitating more complicated surgery, thereby increasing the perioperative burden such as operation time, perioperative morbidity, and length of hospital stay.\(^\)\(^13\) However, it is undetermined how the effect of severity of baseline sagittal imbalance on the clinical outcomes after corrective surgery for ASD remained undetermined. We hypothesized the clinical outcomes would not be inferior, even in patients with severe baseline sagittal imbalance, if the correction was performed successfully. In the current study, we aimed to determine the clinical impact of the baseline sagittal imbalance severity by comparing various perioperative and postoperative outcomes among the patients with mild, moderate, and severe baseline sagittal imbalances.

**MATERIALS AND METHODS**

This study was approved by the Institutional Review Board of Samsung Medical Center (2024-03-027). The requirement for informed consent was waived due to the retrospective nature of this study.

1. **Study Cohort**

This was a retrospective case series study based on records retrieved from a prospective ASD database at Samsung Medical Center. The study cohort included consecutive patients who underwent surgery for degenerative-type ASD between 2012 and 2021. Patient inclusion criteria were as follows: ≥60 years of age; ASD radiographically defined by C7 sagittal vertical axis (SVA) ≥50 mm, PI–lumbar lordosis (LL) mismatch ≥10°, or pelvic tilt (PT) ≥25° or coronal Cobb angle ≥30°; and ≥5 fused vertebral levels from the sacrum, all including the pelvis with iliac fixation. The severity of baseline sagittal imbalance was determined based on the SRS-Schwab classification. The SRS-Schwab classification consists of 3 sagittal modifiers of PI–LL mismatch, SVA, and PT.\(^\)\(^12\) Each sagittal modifier was graded as 0 (<10°), + (10°–20°), ++ (>20°) for PI–LL mismatch, 0 (<40 mm), + (40–95 mm), ++ (>95 mm) for SVA, and 0 (<20°), + (20°–30°), ++ (>30°) for PT. Scores were assigned to each item of the sagittal modifiers, for example, 0 points for grade 0; 1 point for grade +; and 2 points for grade ++. By modifying the previously reported categorization of baseline sagittal imbalance using the SRS-Schwab classification,\(^\)\(^14\)\(^\)\(^15\) patients were classified into 3 groups: mild (score: 1 or 2 points), moderate (score: 3 or 4 points), and severe (score: 5 or 6 points). No patients had a total score of 0 point.

More than 2-years of follow-up with complete radiographic, and patient-reported outcome measure (PROM) data were required for inclusion. Patients were excluded if they lacked appropriate radiographs; had not completed the PROM questionnaire at the final follow-up; had undergone previous thoracic or lumbar fusion surgery; or had syndromic, neuromuscular, inflammatory, or other pathological, rather than degenerative, conditions.

2. **Collected Data**

The demographic data included age, sex, body mass index (BMI), T score, and American Society of Anesthesiologists (ASA) physical status classification grade. Variables related to the surgical technique included total fusion level, oblique lumbar interbody fusion (OLIF), anterior column realignment (ACR), usage of additional rods, cement augmentation in uppermost instrumented vertebra (UIV), and 3-column osteotomy. The perioperative variables included operation time, estimated blood loss, number of red blood cells (RBCs) transfused, intensive care unit (ICU) admission, length of hospital stay, incidence and causes of return to the operating room during the hospital stay, and postoperative medical complications.

Standing posteroanterior and lateral whole-spine radiographs were analyzed at baseline and immediately after surgery (approximately 1 week postoperatively) to measure the following radiographic parameters: PI, LL, PI–LL mismatch, sacral slope (SS), PT, thoracic kyphosis (TK), T1 pelvic angle (TPA), and SVA. For the posteroanterior and lateral whole-spine radiographs, all patients positioned their hands on their shoulders. In addition to postoperative comparison of absolute values of sagittal parameters, the appropriateness of surgical correction
was evaluated with regard to how much the postoperative sagittal alignment met the correction target of the legacy systems such as SRS-Schwab classification, age-adjusted sagittal alignment goals, and GAP score. SRS-Schwab classification was previously described in study cohort section. The ideal age-adjusted PI–LL was calculated using a previously reported formula: PI–LL = (age––55 years)/2+3. Then, based on the offset value between actual PI–LL and ideal PI–LL values, the patients were divided into the following 3 groups: undercorrection (offset > 10°), matched correction (offset within ± 10°), and overcorrection (offset < -10°). Finally, the GAP score is expressed as the total score of relative pelvic version, relative lumbar, lordosis distribution index, relative spinopelvic alignment, and age, ranging from 0 to 13 points. Three groups were created according to the total score as follows: proportioned (score, 0–2), moderately disproportioned (score, 3–6), and severely disproportioned (score, ≥ 7).

Mechanical failures such as proximal junctional complications and rod fractures were recorded. Proximal junctional kyphosis (PJK) was defined as a proximal junctional angle (PJA) of ≥ 10° and increase of PJA ≥ 10° compared to preoperative PJA. Proximal junctional failure (PJF) indicated fracture at the UIV or UIV+1, failure of UIV fixation, myelopathy, or any reasons of revision surgery.

Clinical outcomes were compared using 3 PROM questionnaires, namely, the visual analogue scale (VAS) for the back pain, Oswestry Disability Index (ODI), and the SRS-22 questionnaire (SRS-22) scores. Preoperative and final PROM questionnaires were used for analysis. In addition, we compared the proportion of patients achieving minimal clinically important difference (MCID) in VAS, ODI, and SRS-22 at the last follow-up. The MCID values used in the current study were 1.2 for VAS, 12.8 for ODI. For SRS-22, the MCID values were 1.05 for function, 0.85 for pain, 1.05 for appearance, 0.70 for mental, and 1.05 for subtotal, respectively.

3. Statistical Analysis

Data are presented as frequencies with percentages for categorical variables and as means with standard deviations for continuous variables. Comparisons of variables among the 3 groups were performed using chi-square or Fisher exact tests for categorical variables and analysis of variance with a post hoc test (Tukey test) for continuous variables. Statistical analyses were conducted by professional statisticians using IBM SPSS Statistics ver. 27.0 (IBM Co., Armonk, NY, USA). A p-value < 0.05 was considered statistically significant.

RESULTS

A total of 259 patients met the inclusion criteria and were included in the study cohort. The mean age was 69.0 years and 225 patients (86.9%) were female. There were 42, 62, and 155 patients in the mild, moderate, and severe groups, respectively (Table 1). There were more female patients, and the T score was significantly less, in the severe group. There were no differences in age, BMI, or ASA physical status classification grade among the 3 groups. With regard to operative variables, the number of fusion levels differed significantly among the 3 groups (6.1, 7.2, and 7.5, respectively; p = 0.001). Significantly more patients underwent OLIF surgery at L5–S1, ACR, cement augmentation in UIV, and 3-column osteotomy as the severity of baseline sagittal imbalance increased (p = 0.048, p < 0.001, p = 0.041, and p = 0.001, respectively). The operation time, total number of RBC transfusion, and number of patients requiring ICU care were significantly greater in the severe group (p = 0.004, p = 0.031, and p = 0.027, respectively). The length of hospital stay was significantly longer in the severe group than in the mild group (p = 0.049). None of the patients in the mild group required revision surgery during their hospital stay; however, the inpatient revision rate was not statistically significant. There were no cases of revision surgery due delayed complication other than PJF or rod fracture after discharge in both groups. Moreover, there were no significant differences in postoperative medical complications among the 3 groups.

With regard to radiographic parameters, the PI was significantly smaller in the mild group than in the severe group (50.9° vs. 55.2°, p = 0.023) (Table 2). Other preoperative sagittal parameters showed significant differences among the 3 groups in terms of LL, PI–LL, SS, PT, TK, TPA, and SVA. There were no significant differences in the postoperative LL, SS, PT, TPA, or SVA. However, the postoperative PI–LL mismatch was significantly greater in the severe group (2.6°, 4.9°, and 8.4°, respectively; p = 0.003), and the postoperative TK was the smallest in the severe group. Postoperative changes in all sagittal parameters were significantly greater in the severe group. With regard to SRS-Schwab classification, significantly more patients achieved a sagittal modifier grade 0 of PI–LL mismatch in the mild group than the other groups (p = 0.033) (Fig. 1). However, there were no differences in number of patients with regard to sagittal modifier grades of PT or SVA. There were more patients with undertecorrection relative to age-adjusted PI–LL targets in the severe group (Fig. 2). No significant differences were found in patient distribution relative to the GAP score (Fig. 3).
There was a trend of increasing PJK and PJF as the baseline severity increased (Table 3). However, no statistically significant differences were found among the 3 groups in terms of PJK and PJF development or revision surgery for PJF (p = 0.270, p = 0.162, and p = 0.799, respectively). The incidence of rod fractures, as well as the revision rate for rod fractures, did not differ among the 3 groups (p = 0.569 and p = 0.265, respectively).

There were no significant differences in the preoperative VAS scores for back pain, ODI, or SRS-22 scores among the 3 groups (Table 4). There were also no differences in the scores at the last follow-up or their postoperative changes in the VAS scores for back pain and SRS-22 scores. However, the ODI score at the last follow-up was significantly lower in the mild group than in the severe group (29.6 vs. 37.0, p = 0.019). ODI improvement was
also higher in the mild group than in the other groups (28.5, 17.8, and 20.4, respectively; p = 0.029). Regarding the MCID, there were no significant differences in the number of patients to achieve MCID in VAS, ODI, and all components of SRS-22 such as activity, pain, appearance, mental, and subtotal domains (Table 5). In subgroup analyses, a significantly higher proportion of patients in the mild group achieved the MCID in the ODI compared to the moderate group (78.6% vs. 56.5%, p = 0.02). Similarly, a greater percentage of patients in the severe group reached MCID in the appearance score of the SRS-22 questionnaire than in the moderate group (78.1% vs. 56.5%, p = 0.025). In the severe group without rod fracture, patients with undercorrection exhibited a higher ODI score at the last follow-up than those with matched or overcorrection; however, this difference was not statistically significant, likely due to the small sample size (41.8 vs. 34.0, p = 0.063), and the SRS-22 total score

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mild group</th>
<th>Moderate group</th>
<th>Severe group</th>
<th>p-value</th>
<th>p-value (subanalyses between groups)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>50.9 ± 10.4</td>
<td>51.8 ± 10.2</td>
<td>55.2 ± 11.2</td>
<td>0.023*</td>
<td>B: 0.024*, C: 0.038*</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>50.8 ± 10.9</td>
<td>51.9 ± 10.8</td>
<td>55.2 ± 10.5</td>
<td>0.018*</td>
<td>B: 0.017*, C: 0.038*</td>
</tr>
<tr>
<td>LL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>39.2 ± 11.1</td>
<td>27.2 ± 10.1</td>
<td>7.0 ± 15.9</td>
<td>&lt; 0.001**</td>
<td>A: &lt; 0.001**, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>48.1 ± 10.2</td>
<td>46.9 ± 11.1</td>
<td>46.9 ± 12.3</td>
<td>0.834</td>
<td>NA</td>
</tr>
<tr>
<td>Change (°)</td>
<td>8.9 ± 8.8</td>
<td>19.7 ± 9.7</td>
<td>39.9 ± 18.3</td>
<td>&lt; 0.001**</td>
<td>A: 0.001**, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>PI–LL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>11.7 ± 7.1</td>
<td>24.6 ± 2.8</td>
<td>48.2 ± 13.5</td>
<td>&lt; 0.001**</td>
<td>A: &lt; 0.001**, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>2.6 ± 7.6</td>
<td>4.9 ± 9.7</td>
<td>8.4 ± 11.5</td>
<td>0.003*</td>
<td>B: 0.002*, C: 0.029*</td>
</tr>
<tr>
<td>SS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>28.9 ± 9.2</td>
<td>25.3 ± 9.3</td>
<td>19.0 ± 11.0</td>
<td>&lt; 0.001**</td>
<td>A: 0.048*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>33.1 ± 7.9</td>
<td>33.7 ± 8.1</td>
<td>35.4 ± 10.0</td>
<td>0.231</td>
<td>NA</td>
</tr>
<tr>
<td>Change (°)</td>
<td>4.1 ± 6.9</td>
<td>8.4 ± 7.5</td>
<td>16.4 ± 10.9</td>
<td>&lt; 0.001**</td>
<td>A: 0.027*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>PT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>22.0 ± 7.1</td>
<td>26.5 ± 6.5</td>
<td>36.5 ± 9.8</td>
<td>&lt; 0.001**</td>
<td>A: 0.011*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>17.7 ± 7.2</td>
<td>18.1 ± 7.4</td>
<td>19.7 ± 9.3</td>
<td>0.277</td>
<td>NA</td>
</tr>
<tr>
<td>Change (°)</td>
<td>-4.3 ± 6.5</td>
<td>-8.4 ± 8.0</td>
<td>-16.8 ± 10.8</td>
<td>&lt; 0.001**</td>
<td>A: 0.032*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>TK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>29.2 ± 10.2</td>
<td>21.4 ± 10.6</td>
<td>8.3 ± 12.2</td>
<td>&lt; 0.001**</td>
<td>A: 0.001**, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>31.1 ± 8.3</td>
<td>28.8 ± 10.2</td>
<td>23.6 ± 10.6</td>
<td>&lt; 0.001**</td>
<td>B: &lt; 0.001**, C: 0.001**</td>
</tr>
<tr>
<td>Change (°)</td>
<td>1.9 ± 6.7</td>
<td>7.4 ± 9.4</td>
<td>15.3 ± 12.5</td>
<td>&lt; 0.001**</td>
<td>A: 0.012*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>TPA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (°)</td>
<td>19.7 ± 6.1</td>
<td>24.8 ± 5.5</td>
<td>36.0 ± 10.5</td>
<td>&lt; 0.001**</td>
<td>A: 0.017*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>14.3 ± 6.2</td>
<td>14.9 ± 8.3</td>
<td>15.8 ± 9.0</td>
<td>0.517</td>
<td>NA</td>
</tr>
<tr>
<td>Change (°)</td>
<td>-4.1 ± 5.7</td>
<td>-9.9 ± 7.6</td>
<td>-20.1 ± 11.8</td>
<td>&lt; 0.001**</td>
<td>A: 0.016*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>SVA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative (mm)</td>
<td>35.7 ± 29.9</td>
<td>53.7 ± 42.9</td>
<td>88.5 ± 52.3</td>
<td>&lt; 0.001</td>
<td>A: 0.045*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
<tr>
<td>Immediate PO (°)</td>
<td>20.9 ± 34.0</td>
<td>17.0 ± 30.7</td>
<td>17.7 ± 29.4</td>
<td>0.751</td>
<td>NA</td>
</tr>
<tr>
<td>Change (mm)</td>
<td>-14.8 ± 40.2</td>
<td>-36.7 ± 50.3</td>
<td>-71.5 ± 52.7</td>
<td>&lt; 0.001**</td>
<td>A: 0.030*, B: &lt; 0.001**, C: &lt; 0.001**</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation.
P1, pelvic incidence; PO, postoperative; LL, lumbar lordosis; SS, sacral slope; PT, pelvic tilt; TK, thoracic kyphosis; TPA, T1 pelvic angle; SVA, sagittal vertical axis; NA, not available; A, mild vs. moderate; B, mild vs. severe; C, moderate vs. severe.
*p < 0.05. **p < 0.01.
at the last follow-up was significantly lower in patients with undercorrection compared to those with matched or overcorrection (2.8 vs. 3.5, \( p = 0.013 \)).

Representative cases for patients in the mild and severe groups are illustrated in Figs. 4 and 5, respectively.

Table 3. Comparison of the mechanical failure among the 3 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mild group</th>
<th>Moderate group</th>
<th>Severe group</th>
<th>p-value</th>
<th>p-value (subanalyses between groups)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PJK</td>
<td>7 (16.7)</td>
<td>19 (30.6)</td>
<td>41 (26.5)</td>
<td>0.270</td>
<td>NA</td>
</tr>
<tr>
<td>PJF</td>
<td>5 (11.9)</td>
<td>15 (24.2)</td>
<td>40 (25.8)</td>
<td>0.162</td>
<td>NA</td>
</tr>
<tr>
<td>Revision surgery for PJF</td>
<td>4 (9.5)</td>
<td>7 (11.3)</td>
<td>13 (8.4)</td>
<td>0.799</td>
<td>NA</td>
</tr>
<tr>
<td>Time to revision for PJF (mo)</td>
<td>43.6 ± 30.7</td>
<td>31.8 ± 29.9</td>
<td>39.7 ± 44.8</td>
<td>0.869</td>
<td>NA</td>
</tr>
<tr>
<td>Rod fracture</td>
<td>9 (21.4)</td>
<td>19 (30.6)</td>
<td>44 (28.4)</td>
<td>0.569</td>
<td>NA</td>
</tr>
<tr>
<td>Revision surgery for rod fractures</td>
<td>1 (2.4)</td>
<td>4 (6.5)</td>
<td>15 (10.3)</td>
<td>0.265</td>
<td>NA</td>
</tr>
<tr>
<td>Time to revision for rod fracture (mo)</td>
<td>29.5 ± 21.9</td>
<td>37.4 ± 33.9</td>
<td>34.5 ± 31.3</td>
<td>0.664</td>
<td>NA</td>
</tr>
</tbody>
</table>

Values are presented as number (%) or mean ± standard deviation.
PJK, proximal junctional kyphosis; PJF, proximal junctional failure; NA, not available.

Fig. 1. Comparison of patient distribution relative to the postoperative SRS-Schwab classification among the 3 groups. SRS, Scoliosis Research Society; PI, pelvic incidence; LL, lumbar lordosis; PT, pelvic tilt; SVA, sagittal vertical axis. *\( p < 0.05 \).

Fig. 2. Comparison of patient distribution relative to the postoperative age-adjusted PI–LL target among the 3 groups. PI, pelvic incidence; LL, lumbar lordosis. *\( p < 0.05 \).

Fig. 3. Comparison of patient distribution relative to the postoperative Global Alignment and Proportion (GAP) score among the 3 groups.
DISCUSSION

Given that a positive sagittal imbalance leads to poor clinical outcome,6,7 optimal restoration of spinopelvic malalignment is a key factor in achieving good clinical outcomes. Considering that the alignment target is largely determined by the patient’s PI and age, patients with a more severe sagittal imbalance may require a more aggressive surgical strategy to reach the desired surgical target. In the current study, we observed that patients in the severe group had a higher probability of undergoing more invasive surgeries, such as OLIF at L5–S1, ACR, and 3-column osteotomy. Neuman et al.13 reported a surgical invasiveness threshold to predict the odds of major complication. They found that surgical variables, such as 3-column osteotomy, anterior interbody fusion (vs. posterior interbody fusion), iliac fixation, and revision surgery, significantly increased the risk of surgical and medical complications. Samuel et al.20 conducted a similar study to investigate perioperative morbidity. They observed that a longer operative time was a better predictor of inpatient complications than surgical invasiveness itself. Song et al.21 also reported that operation time was associated with a higher rate of 30-day morbidity and blood transfusion.

Table 4. Comparison of the clinical outcomes among the 3 groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mild group</th>
<th>Moderate group</th>
<th>Severe group</th>
<th>p-value</th>
<th>p-value (subanalyses between groups)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS for the back pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>64.8 ± 24.4</td>
<td>68.2 ± 20.9</td>
<td>71.6 ± 22.4</td>
<td>0.185</td>
<td>NA</td>
</tr>
<tr>
<td>At the last follow-up</td>
<td>34.3 ± 26.1</td>
<td>36.5 ± 25.7</td>
<td>35.9 ± 26.3</td>
<td>0.914</td>
<td>NA</td>
</tr>
<tr>
<td>Change</td>
<td>-30.5 ± 35.3</td>
<td>-31.8 ± 30.7</td>
<td>-35.7 ± 30.7</td>
<td>0.525</td>
<td>NA</td>
</tr>
<tr>
<td>ODI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>58.1 ± 14.9</td>
<td>54.4 ± 16.4</td>
<td>57.4 ± 15.2</td>
<td>0.372</td>
<td>NA</td>
</tr>
<tr>
<td>At the last follow-up</td>
<td>29.6 ± 16.6</td>
<td>36.6 ± 18.2</td>
<td>37.0 ± 17.9</td>
<td>0.057</td>
<td>B: 0.019*</td>
</tr>
<tr>
<td>Change</td>
<td>-28.5 ± 22.2</td>
<td>-17.8 ± 20.0</td>
<td>-20.4 ± 20.5</td>
<td>0.029*</td>
<td>A: 0.010*, B: 0.025*</td>
</tr>
<tr>
<td>SRS-22 total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>2.5 ± 0.5</td>
<td>2.3 ± 0.5</td>
<td>2.3 ± 0.5</td>
<td>0.592</td>
<td>NA</td>
</tr>
<tr>
<td>At the last follow-up</td>
<td>3.5 ± 0.8</td>
<td>3.3 ± 0.8</td>
<td>3.4 ± 0.7</td>
<td>0.515</td>
<td>NA</td>
</tr>
<tr>
<td>Change</td>
<td>0.9 ± 0.9</td>
<td>0.8 ± 0.8</td>
<td>1.1 ± 0.7</td>
<td>0.163</td>
<td>NA</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. VAS, visual analogue scale; ODI, Oswestry Disability Index; SRS-22, Scoliosis Research Society-22 questionnaire; NA, not available; A, mild vs. moderate; B, mild vs. severe.
*p < 0.05.

Table 5. Comparison of the number of patients achieving MCID for VAS, ODI, and SRS-22 at the last follow-up

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mild group</th>
<th>Moderate group</th>
<th>Severe group</th>
<th>p-value</th>
<th>p-value (subanalyses between groups)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAS for the back pain (threshold = 1.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>27 (64.3)</td>
<td>48 (77.4)</td>
<td>121 (78.1)</td>
<td>0.170</td>
<td>NA</td>
</tr>
<tr>
<td>ODI (threshold = 12.8)</td>
<td>33 (78.6)</td>
<td>35 (56.5)</td>
<td>106 (68.4)</td>
<td>0.055</td>
<td>A: 0.020*</td>
</tr>
<tr>
<td>SRS-22 activity (threshold = 0.85)</td>
<td>18 (42.9)</td>
<td>23 (37.1)</td>
<td>58 (37.4)</td>
<td>0.921</td>
<td>NA</td>
</tr>
<tr>
<td>SRS-22 pain (threshold = 0.90)</td>
<td>18 (42.9)</td>
<td>30 (48.6)</td>
<td>84 (54.2)</td>
<td>0.676</td>
<td>NA</td>
</tr>
<tr>
<td>SRS-22 appearance (threshold = 1.05)</td>
<td>30 (71.4)</td>
<td>35 (56.5)</td>
<td>121 (78.1)</td>
<td>0.065</td>
<td>C: 0.025*</td>
</tr>
<tr>
<td>SRS-22 mental (threshold = 0.70)</td>
<td>24 (57.1)</td>
<td>32 (51.6)</td>
<td>82 (52.9)</td>
<td>0.937</td>
<td>NA</td>
</tr>
<tr>
<td>SRS-22 subtotal (threshold = 1.05)</td>
<td>21 (50.0)</td>
<td>28 (45.2)</td>
<td>92 (59.4)</td>
<td>0.388</td>
<td>NA</td>
</tr>
</tbody>
</table>

Values are presented as number (%). MCID, minimal clinically importance difference; VAS, visual analogue scale; ODI, Oswestry Disability Index; SRS-22, Scoliosis Research Society-22 questionnaire; NA, not available; A, mild vs. moderate; B, mild vs. severe; C, moderate vs. severe.
*p < 0.05.
Surgical Outcome for ASD According to Deformity Severity

Park SJ, et al.

Fig. 4. Representative case of mild sagittal imbalance. A 69-year-old female presented with persistent back pain due to lumbar kyphoscoliosis. Her preoperative sagittal parameters are as follows: PI = 40, LL = 22, PI-LL = 18, PT = 24, SVA = 20 mm (sum of sagittal modifier score = 2). She underwent the corrective surgery using oblique lumbar interbody fusion at L3–5 and posterior lumbar interbody fusion at L5–S1 with T10-pelvis fixation. Her postoperative sagittal parameters are as follows: LL = 45, PI-LL = -5, PT = 6, SVA = 34 mm (sum of sagittal modifier score = 0). PI, pelvic incidence; LL, lumbar lordosis; PT, pelvic tilt; SVA, sagittal vertical axis.

Fig. 5. Representative case of mild sagittal imbalance. A 73-year-old female presented with persistent back pain due to lumbar kyphoscoliosis. Her preoperative sagittal parameters are as follows: PI = 60, LL = -7, PI-LL = 67, PT = 36, SVA = 194 mm (sum of sagittal modifier score = 6). She underwent the corrective surgery using oblique lumbar interbody fusion at L2–3, L4–S1 and corner osteotomy at L3–4 with T10-pelvis fixation. Her postoperative sagittal parameters are as follows: LL = 49, PI-LL = 11, PT = 21, SVA = 24 mm (sum of sagittal modifier score = 2). PI, pelvic incidence; LL, lumbar lordosis; PT, pelvic tilt; SVA, sagittal vertical axis.

In the current study, the operation time was significantly longer in the severe group than in the mild group, but the gap between groups was not large with just 1.6 hours. In the current study, there were no cases of return to the operating room due to inpatient surgical complications in the mild group. Although approximately 5% of patients in the moderate and severe groups required revision surgery during their hospital stay, all complications were treated successfully, leaving no permanent deficit. There were no significant differences in medical complications between the groups. Therefore, although surgery in the severe group increased the surgical burden with regard to surgical invasiveness and inpatient morbidity, the complication rate and its treatability were within acceptable ranges.

In the current study, the severity of baseline sagittal imbalance and postoperative changes of all sagittal parameters were clearly distinguished among the 3 groups. We observed that PJK and PJF developed less frequently in the mild group than in the other groups. However, no statistical significance was found for the development of PJK and PJF or revision surgery. The severity of baseline sagittal imbalance and subsequent
postoperative greater change in sagittal deformity are known risk factors for PJK and PJF development. However, the appropriateness of postoperative sagittal correction is equally crucial. Our findings indicate no significant differences in achieving matched correction postoperatively among mild, moderate, and severe groups. Furthermore, the severe group had a lower incidence of overcorrection, compared to mild and moderate groups. Considering that overcorrection increases PJF risk, the lower rate of postoperative overcorrection in severe group could have decreased the incidence of PJF. However, several studies have published contradictory results showing that the amount of correction or the final sagittal alignment did not affect PJK or PJF development. Further follow-up studies are required to clarify this discrepancy. The incidence of rod fractures showed a trend similar to that of PJK and PJF. The incidence of rod fracture and the revision rate were lowest in the mild group, but these results did not reach statistical significance in the Fisher exact test or in the intergroup subanalyses. It is currently understood that mechanical failure after ASD surgery is closely associated with the shape of sagittal alignment such as the GAP score, rather than the absolute value of radiographic parameters. In the current study, we found that GAP score categories did not differ among the groups; therefore, our findings can explain the negative findings of mechanical complication occurrence among the 3 groups.

We observed the greatest postoperative PI–LL value, the fewest patients achieving grade 0 in the SRS-Schwab PI–LL modifier, and the more patients with undercorrection relative to the age-adjusted PI–LL in the severe group. It is well known that undercorrection has been associated with poor clinical outcomes in ASD surgery. Our findings indicate that a higher proportion of patients in the severe group experienced undercorrection after surgery and had significantly elevated ODI scores and SRS-22 total scores at the final follow-up, consistent with previous studies. The optimal restoration of sagittal malalignment is crucial for success after ASD surgery. Lee et al. reported that the strict correction relative to all sagittal modifiers of SRS-Schwab classification ensures better SRS-22 scores even in a long-term follow-up of 90.3 months. Park et al. also demonstrated that matched correction relative to the age-adjusted PI–LL target is necessary to achieve good clinical outcomes and to reduce PJK development. Patients with severe sagittal imbalance are likely to be undercorrected compared to those with mild and moderate sagittal imbalance. Therefore, greater efforts are required to achieve adequate correction as the severity of baseline sagittal imbalance increases.

This study has a few limitations. First, an inherent limitation of this study is the retrospective nature, which allows for the possibility of selection bias. Second, the results of this study may lack the generalizability considering the heterogeneous nature of patients with ASD because we only included patients with degenerative-type ASD. However, we applied strict inclusion criteria, such as narrow age group (≥ 60 years), main preoperative diagnosis with sagittal imbalance, and pelvic fixation in all cases, during patient selection to reduce such heterogeneity. Third, there was no investigation regarding the history of lower limb joint replacement surgery. Severe knee arthritis and similar conditions necessitating lower limb joint replacement can impact compensatory mechanisms in patients with ASD. However, we routinely check the range of motion of hip or knee, and patients with severe flexion contracture in these joints, irrespective of undergoing total joint surgery, were not included. Finally, we adopted the SRS-Schwab classification to group patients according to the baseline severity of sagittal malalignment. Different results may be obtained if other criteria, such as an age-adjusted alignment target or GAP score, are applied. However, the SRS-Schwab classification is currently the most popular tool in the current literatures for defining the severity of sagittal deformity.

CONCLUSION

Patients with more severe sagittal imbalance were treated with more invasive surgical methods, with an increased perioperative surgical burden. Regardless to severity of baseline sagittal imbalance, all patients exhibited significant radiological and clinical improvements after surgery. However, in term of ODI, the severe group demonstrated slightly worse clinical outcomes compared to the other groups, probably due to relatively higher proportion of undercorrection. Therefore, more rigorous correction is necessary to achieve optimal sagittal alignment specifically in patients with severe baseline sagittal imbalance.

NOTES

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: SJP, JSP, DHK; Data curation: SJP, JSP, DHK, HJK, YML; Formal analysis: HJK, YML; Methodology: SJP, DHK; Project administration: JSP,
ORCID
Se-Jun Park: 0000-0002-2412-9437
Jin-Sung Park: 0000-0001-6517-8609
Dong-Ho Kang: 0000-0002-9114-0150
Hyun-Jun Kim: 0009-0005-8151-1084
Yun-Mi Lim: 0000-0001-5007-8699
Chong-Suh Lee: 0000-0003-1373-4815

REFERENCES

https://doi.org/10.14245/ns.2448250.125 www.e-neurospine.org 731
Preliminary Clinical and Radiological Outcomes of the “No-Punch” Decompression Techniques for Unilateral Biportal Endoscopic Spine Surgery

Jwo-Luen Pao
Department of Orthopedic Surgery, Far Eastern Memorial Hospital, New Taipei City, Taiwan

Objective: To avoid the most offending surgical instrument for dural tears, we develop a “no-punch” decompression technique for unilateral biportal endoscopic (UBE) spine surgery.

Methods: This retrospective study enrolled 68 consecutive patients with degenerative lumbar spinal stenosis segments. The treatment results were evaluated using the visual analogue scale (VAS) for low back and leg pain, the Japanese Orthopaedic Association (JOA) scores, and the Oswestry Disability Index (ODI). Radiological outcomes were evaluated using the preoperative and postoperative magnetic resonance imaging.

Results: This study included 36 male and 32 female patients who received 109 segments of decompression, with an average age of 68.7 (37–90 years). The average operation time was 52.2 minutes. The average hospital stay was 3.1 days. There were no dural tears but 3 minor surgical complications, all treated conservatively. The VAS for low back and leg pain improved from 4.6 and 7.0 to 0.8 and 1.2. The JOA score improved from 16.2 to 26.8, with an improvement rate of 82.0%. The ODI improved from 50.1 to 18.7. All these improvements were statistically significant. The cross-sectional dural area improved from 61.1 to 151.3 mm², with an average increase of 90.2 mm² and 205.3%. 87.1% of the ipsilateral facet joints and 84.7% of the contralateral facet joints were preserved. In 61% of the decompressed segments, the ipsilateral facet joints were preserved better than the contralateral facet joints.

Conclusion: The UBE “no-punch” decompression technique effectively avoids the dural tears. It provides effective neural decompression, excellent facet joint preservation, and good treatment outcomes.

Keywords: Minimally invasive surgery, Biportal endoscopy, Dural tears, Complications, Treatment outcomes

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified

INTRODUCTION

Minimally invasiveness is a trend in every surgical field. Minimally invasive spine surgery has rapidly evolved, and endoscopic spine surgery has gained worldwide popularity because of the rapid advancements in endoscopic techniques and surgical instruments. In recent years, the unilateral biportal endoscopic (UBE) technique has drawn much attention from spine surgeons. This minimally invasive technique is a revolutionary endoscopic technique performed through 2 independent portals with continuous irrigation of normal saline. The normal saline provides hydrostatic pressure to suppress bleeding, carrying away bone debris and oozing. Combined with a high-resolution endoscopic system, UBE delivers a clear, bright, and magnified
surgical field, enabling surgeons to perform delicate surgical procedures without excessive soft tissue damage. UBE has been applied to address a variety of spinal pathologies, such as discectomy for lumbar disc herniation, laminotomy for degenerative lumbar spinal stenosis (DLSS), lumbar interbody fusion for disc degeneration or spondylolisthesis, and cervical foraminal stenosis, all of which demonstrated good clinical efficacy and satisfactory outcomes.¹⁻⁷

Endoscopic spine surgery is technically demanding, with a steep learning curve. A high incidence of complications usually occurs in the early learning curve.⁶⁻¹⁰ The UBE technique cannot be exempted from this. The most common complication is incidental dural tear, followed by epidural hematoma and incomplete decompression.¹¹ Most studies claim that dural tears in endoscopic spine surgeries are usually small and can be managed conservatively with few sequelae.¹²⁻¹³ However, severe neurological sequelae that impair the treatment outcomes do occur. In contrast to traditional open surgeries, dural tears are difficult to repair under the endoscope. As practice does not guarantee perfect results, we need a thorough understanding of the injury mechanisms of dural tears and new surgical techniques to prevent their occurrence.

The Kerrison punch was designed by an English physician, Robert Maters Kerrison, in the eighteenth century. It has been used as the most crucial surgical instrument in modern spine surgery for the decompression of neural elements. However, in our experience of more than 4,000 cases of minimally invasive spine surgeries, we realized that the Kerrison punch was the most offending surgical instrument to cause dural tears or nerve root injuries. Therefore, we developed the “no-punch” technique for UBE surgeries, trying not to use the Kerrison punch while performing the decompressive procedures.

This study aims to introduce the novel UBE “no-punch” decompression technique and review its clinical and radiological outcomes for the treatment of DLSS.

MATERIALS AND METHODS

1. Patient Selection

The study was conducted after obtaining proof from the Research Ethics Review Committee of Far Eastern Memorial Hospital (No. 113052-E). Sixty-eight patients who received UBE “no-punch” decompression surgery for DLSS from March 2022 to June 2023 were included in this retrospective study.

The indications for UBE spine surgeries are axial back pain, radicular leg pain, single or multiple lumbar radiculopathies, and neurogenic intermittent claudication due to DLSS. All patients have at least 3 months of conservative treatment before surgical intervention. All patients had plain x-rays of anteroposterior, lateral, and dynamic lateral views before and 6 months after the surgery. Magnetic resonance imaging (MRI) was arranged before and 6 months after the surgery. All the diagnoses must have radiological evidence corresponding to the patient’s clinical presentation. The author performed all the surgeries in a single institute. We excluded patients who had lumbar disc herniation, segmental instability with more than 4-mm translation on the dynamic lateral x-rays, more than grade II spondylolisthesis, scoliosis of more than 20°, infection, and a history of prior lumbar spine surgeries.

2. Evaluation of Clinical Data and Outcomes

We retrieved demographic data, clinical data, surgical complications, and treatment outcomes from chart reviews. We also reviewed all the operation notes and video records to examine the occurrence, mechanism, offending surgical instrument, and management of surgical complications. All patients had at least 6 months of follow-up after the surgery. The treatment outcomes were evaluated using the Japanese Orthopaedic Association (JOA) score and the Oswestry Disability Index (ODI) before surgery and 1, 3, and 6 months after the surgery.¹⁴⁻¹⁵

We used the Schizas grading system to evaluate the stenosis severity. The Schizas system classifies stenosis severity into 4 grades using the axial T2-weighted MRI images. Grade A is no or minor stenosis with an oval spinal canal and clear cerebral spinal fluid space. Grade D is the most severe, with a collapsed spinal canal and no cerebral spinal fluid space.¹⁶ When the post-operative MRI was available, we measured the crosse-sectional dural area and the width of ipsilateral and contralateral facet joints using the ImageJ software (https://imagej.net/) to evaluate the decompression efficacy and facet joint preservation.¹

The results of patient-reported outcome measures were analyzed using Wilcoxon signed-rank test. The results of radiological measurements were analyzed using the paired t-test. A p-value of < 0.05 was considered statistically significant.

3. Surgical Techniques

We will use the UBE unilateral laminotomy for bilateral decompression (ULBD) for DLSS at L4–5 from the left-side approach as an example to describe the “no-punch” decompression technique in detail.

The surgery is performed under endotracheal general anesthesia with the patient placed in the prone position on a radio-
lucent surgical table. The surgical field is disinfected in the usual manner. Because the UBE surgery is performed with continuous saline irrigation, a water-tight draping with a sound drainage system is essential to prevent soaking and resultant hypothermia of the patient.

The initial target area is the spinolaminar junction, which is the junction of the lower margins of the spinous process and the lamina of L4 on the left side. The skin incisions are localized at the intersection of the medial pedicle line and the lower pedicle lines of the L4 and L5 pedicles on the left side. For a right-handed surgeon, the cranial skin incision (about 6 mm long) will be the entry for the endoscope, and the caudal skin incision (about 10 mm long) will be the entry for the surgical instruments (Fig. 1A).

We prefer transverse skin incisions for better cosmesis. After marking the skin landmarks under the fluoroscope, we incise the skin and deep fascia using a No. 11 scalpel. We insert the endoscopic sheath and its trocar into the endoscopic portal and a blunt dilator into the working portal. These 2 instruments meet at the spinolaminar junction to establish the triangulation, confirmed using the fluoroscope (Fig. 1B). The soft tissues at the spinolaminar junction are gently dissected. Then, the trocar and blunt dilator are replaced by the endoscope (4 mm × 30°, ConMed, Largo, FL, USA) and the radiofrequency wand (ArthroCare, Austin, TX, USA) with the inflow of normal saline. We use the radiofrequency wand to ablate the soft tissue and coagulate the bleeders in the muscles to create a clear endoscopic surgical field. The saline bags are hung about 30 cm higher than the level of the surgical site. The circulating nurse monitors and adjusts the saline bags’ height to maintain adequate hydrostatic pressure. A semitubular tube at the working portal is very helpful in maintaining a good saline outflow (Fig. 1C). A good control of saline inflow/outflow is mandatory for hemostasis and a clear surgical field, especially when using the high-speed drill.

We prefer using the high-speed drill with a 4-mm coarse diamond ball tip (Primado II, NSK, Tokyo, Japan) as the primary instrument for removing bone. We design a set of osteotomes with 3 different curved angles: 0°, 10°, and 20° for the “no-punch” decompression technique. The osteotomes are 4 mm wide and 2 mm thick with a symmetric tapered shape to the tip (Fig. 1D). The surgical procedure will be explained step-by-step as follows:

1. Use the high-speed drill to start the laminotomy from the spinolaminar junction at L4 (Fig. 2). The laminotomy is widened medially and cranially until the underlying cranial end of the ligamentum flavum is free from its attachment and the epidural fat is exposed (Fig. 3).

2. Widen the laminotomy laterally and move the drill caudally. The joint capsule covering the facet joint should be pre-
served as much as possible. Then, use the blunt elevator to dissect and remove the superficial part of ligamentum flavum away from the cranial margin of the L5 lamina (Fig. 4). Use the drill to trim out the cranial border of the L5 lamina and the base of the L5 spinous process until the caudal end of the deep part of the ligamentum lamina is free from its attachment (Fig. 5).

(3) Advance the drill contralaterally along the upper margin of the L5 lamina. If possible, remove the superficial part of the contralateral ligamentum flavum. This will provide more working space for contralateral decompression.

(4) Identify the inferior margin of the spinous process. Use the blunt elevator to separate the ligamentum flavum from the spinous process and contralateral lamina of L4 (Fig. 6). Advance the drill into the space between the contralateral lamina and ligamentum flavum to perform sublaminar decompression. The drill can be advanced very deep to the contralateral lateral recess if the ligamentum flavum remains, which protects the underlying neural tissues (Fig. 7). Keep the ligamentum flavum

![Image 1](https://doi.org/10.14245/ns.2448376.188)
intact until the end of bony decompression. The fluted cutting drill is not recommended in this step because it may destroy the ligamentum flavum and lead to catastrophic neural injury. Then, the drilling work is done.

(5) Use the curved osteotome to decompress the ipsilateral lateral recess. Find the medial border of the L5 superior articular process (SAP) and the L4 inferior articular process (IAP). The curved osteotome is used to undercut the fact joint from L5 SAP to L4 IAP (Fig. 8). Then, the bony fragments are separated using twisting maneuvers. Use the angled curette to separate the bony fragments further and detach the ipsilateral ligamentum flavum from underneath the L4 IAP and lamina (Fig. 9A). Then, the micropituitary rongeur is used to grasp the bony fragments and take them out along with the ipsilateral ligamentum flavum as a whole piece.

(6) Use the curved osteotome to chop the cranial margin of the contralateral L5 lamina. Follow the lamina to identify the contralateral L5 SAP. Use the osteotomes to undercut the contra-
lateral L5 SAP (Fig. 9B). Use the elevator or curette to separate the bony fragments and the contralateral ligamentum flavum from its attachment. Elevate the ligamentum flavum and release the underlying epidural adhesion. Then, use the micropituitary rongeur to grasp the ligamentum flavum firmly and take it out along with the bony fragments as a whole piece (Fig. 10).

(7) Check for residual stenosis. Use the osteotomes and curettes to remove residual osteophytes and ligamentum flavum. Use the radiofrequency wand to coagulate the bleeders. Use bone wax to seal all the cutting surfaces of the bone. Stop the saline irrigation temporarily to check for any occult bleeding. A negative suction drain is indicated when the hemostasis is in doubt.

RESULTS

There were 68 patients receiving 109 segments of UBE “no-punch” decompression for DLSS, including 36 male and 32 female patients, with an average age of 68.7 ± 9.9 (37–90) years. All the patients had at least 6 months of follow-up after the surgery.

Of the 68 patients, 54 patients had pure DLSS (79.4%), 9 patients also had low-grade spondylolisthesis (13.2%), and 5 patients also had degenerative scoliosis of less than 20° (7.4%). One-segment decompression was performed in 38 patients (55.9%), 2-segment in 19 patients (27.9%), and 3-segment in 11 patients (16.2%). The decompression was most frequently performed at L4–5 (56 patients, 51.4%), followed by L3–4 (31 patients, 28.4%) and L2–3 (14 patients, 12.8%). Schizas grade D stenosis was observed in 62 segments (56.9%), grade C in 23 segments (21.1%),

Table 1. Clinical data (n = 68)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>36 (52.9)</td>
</tr>
<tr>
<td>Female</td>
<td>32 (47.1)</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>68.7 ± 9.9 (37–90)</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
</tr>
<tr>
<td>DLSS</td>
<td>54 (79.4)</td>
</tr>
<tr>
<td>Spondylolisthesis</td>
<td>9 (13.2)</td>
</tr>
<tr>
<td>Scoliosis</td>
<td>5 (7.4)</td>
</tr>
<tr>
<td>Segments of decompression (n = 109)</td>
<td></td>
</tr>
<tr>
<td>1-Segment</td>
<td>38 (55.9)</td>
</tr>
<tr>
<td>2-Segment</td>
<td>19 (27.9)</td>
</tr>
<tr>
<td>3-Segment</td>
<td>11 (16.2)</td>
</tr>
<tr>
<td>Level of decompression (n = 109)</td>
<td></td>
</tr>
<tr>
<td>L1–2</td>
<td>1 (0.9)</td>
</tr>
<tr>
<td>L2–3</td>
<td>14 (12.8)</td>
</tr>
<tr>
<td>L3–4</td>
<td>31 (28.4)</td>
</tr>
<tr>
<td>L4–5</td>
<td>56 (51.4)</td>
</tr>
<tr>
<td>L5–S</td>
<td>7 (6.4)</td>
</tr>
<tr>
<td>Schizas grade of stenosis (n = 109)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0 (0)</td>
</tr>
<tr>
<td>B</td>
<td>24 (22.0)</td>
</tr>
<tr>
<td>C</td>
<td>23 (21.1)</td>
</tr>
<tr>
<td>D</td>
<td>62 (56.9)</td>
</tr>
<tr>
<td>Operation time (min)</td>
<td>52.2 ± 20.3 (40–90)</td>
</tr>
<tr>
<td>Drain tube usage (n)</td>
<td>12 (17.6)</td>
</tr>
<tr>
<td>SICU stay (n)</td>
<td>4 (5.9)</td>
</tr>
<tr>
<td>Hospital stay (day)</td>
<td>3.1 ± 1.0 (2–9)</td>
</tr>
<tr>
<td>Complications (n = 68)</td>
<td></td>
</tr>
<tr>
<td>Root injury</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Incomplete decompression</td>
<td>1 (1.5)</td>
</tr>
<tr>
<td>Transient radicular pain</td>
<td>1 (1.5)</td>
</tr>
</tbody>
</table>

Values are presented as number (%) or mean ± standard deviation (range).
DLSS, degenerative lumbar spinal stenosis; SICU, surgical intensive care unit.

https://doi.org/10.14245/ns.2448376.188
and grade B in 24 segments (22.0%). The operation time was 52.2 ± 20.3 minutes. We could not precisely measure the intraoperative blood loss because the normal saline diluted all the bleeding. No patients required blood transfusion. Drain tubes were used in 12 patients (17.6%). The anesthesiologist recommended overnight intensive care in the surgical intensive care unit for 4 patients (5.9%). The average hospital stay was 3.1 ± 1.0 days (2–9 days). We observed 3 surgical complications: 1 root injury by the radiofrequency wand, 1 incomplete decompression, and 1 transient radicular leg pain. There was no dural tear in the 68 patients. All the complications were managed conservatively with no sequel, and no patient required secondary surgery. There was no progression of spondylolisthesis or scoliosis in the follow-up period. The data are summarized in Table 1.

The preoperative average VAS scores for back pain and leg pain were 4.6 ± 3.2 and 7.0 ± 2.0, which improved to 0.8 ± 1.2 and 1.2 ± 1.5 after the surgery. The JOA scores improved from 31.1 ± 6.1 to 26.8 ± 2.4, with an average improvement rate of 82.0% ± 19.0%. The ODI improved from 50.1 ± 18.7 to 37.6 ± 17.2, corresponding to 87.1 ± 12.1% improvement. All these improvements were statistically significant, with p-values < 0.001. The results are summarized in Table 2.

Postoperative MRI was available in 49 patients with 83 segments of decompression. The average cross-sectional dural area (CSDA) significantly increased from 61.1 ± 26.4 mm² to 151.3 ± 39.5 mm² after the surgery (p < 0.001), with an average increase of 90.2 ± 35.2 mm², corresponding to 205.3% ± 206.1% increase. The ipsilateral facet joints were better preserved than the contralateral facet joints (87.1% ± 12.1% vs. 84.7% ± 11.9%). However, the difference did not reach a statistically significant level. Of all the decompression segments, ipsilateral facet joints were better preserved than the contralateral facet joints in 61%. The data are summarized in Table 3.

### DISCUSSION

Like in other surgical fields, minimal invasiveness is undoubtedly the trend in spine surgery. Minimally invasive spine surgery is usually done through a tubular retractor system with a microscope. Recently, most of these minimally invasive procedures can be accomplished using endoscopic techniques with compatible treatment results. The UBE technique has proved revolutionary and gained worldwide popularity among these endoscopic techniques, starting in South Korea.
The learning curve is always a critical issue for minimally invasive surgery. The learning curve is even steeper for the endoscopic technique. In the early learning curve, the operation time is prolonged, with a higher rate of surgical complications, secondary surgery, and prolonged hospitalization. Most studies claimed that complications in endoscopic spine surgery can be managed conservatively with no severe sequelae. However, potential sequelae of these complications did occur and should not be taken for granted.

The most common complication in UBE surgery is dural tears, with an estimated incidence of 4.1% to 4.5%. UBE surgery is performed under continuous irrigation of normal saline, and the neural tissues are subject to potential injury from the hydrostatic pressure. When a dural tear occurs, normal saline or air bubbles may enter the cerebrospinal fluid space to cause direct injury to the brain. Small tears can be repaired using the fibrin sealant or the patch compression technique. A more than 10-mm tear should be repaired using nonpenetrating hemostatic clips or direct sutures. However, direct repair under the endoscope is highly technically demanding and can only be feasible in an expert's hands.

There are several injury mechanisms for dural tears in UBE surgery. The hydrostatic pressure of normal saline provides a clear surgical field by suppressing bleeding from muscles and epidural vessels. When saline flows into the epidural space, the dura is also suppressed bilaterally and forms a central fold just beneath the central slit of the ligamentum flavum. The surgeon tends to insert the Kerrison punch into the central slit to bite the ligamentum flavum. This movement is generally safe if fatty tissue between the dura and ligamentum flavum exists. However, epidural ligaments may tether the dura, or the fatty tissue may be replaced by epidural adhesion in case of severe stenosis. Using the Kerrison punch may cause a dural tear in this way. Similarly, nerve root injury will occur if we insert the Kerrison punch blindly into the lateral recesses without clearing the epidural space or releasing the epidural adhesion.

Our “no-punch” decompression technique effectively reduces the risk of dural tear and nerve root injury by avoiding the above-mentioned injury mechanisms. A retrospective study conducted by Kim et al. reported that 56% of the dural tears occurred while using a Kerrison punch for laminotomy. We have the same experience that the Kerrison punch is the most offending instrument for dural tears. Therefore, we develop a set of osteotomes with varying curves for laminotomy. The curved osteotomes are designed to undercut the facet joint from its medial aspect. Instead of using the Kerrison punch to piecemeal resect the ligamentum flavum, we detach the ligamentum flavum as a whole piece from its peripheral attachment along with the laminotomy chips (Fig. 10). Under the endoscope, we can check and release the epidural adhesion before taking out the ligamentum flavum. Following these principles, we have no neural injury except one nerve root injury by the radiofrequency wand.

Effective decompression is the primary concern in the surgical treatment of DLSS. Our UBE “no-punch” decompression technique can achieve good decompression with a significant increase in CSDA after the surgery. The postoperative CSDA increased to 151.3 mm² and 205.3% compared to the preoperative CSDA. The patient-report treatment outcomes evaluated by VAS for back and leg pain, JOA scores, and ODI all show significant improvement and excellent results at the final follow-up.

ULBD is the most frequently adopted concept in minimally invasive spine surgery. It can be accomplished using the traditional open technique, minimal access technique with the tubular retractor system, minimally invasive technique with the microendoscopic system, or endoscopic technique with the uniportal or biportal endoscopic system. No matter what technique is adopted, medial facetectomy is required to decompress the lateral recess and release the nerve roots. However, excessive facet joint destruction may result in iatrogenic segmental instability, which may necessitate a secondary spinal fusion surgery. Our “no-punch” decompression technique preserves more than 80% of the facet joints, either ipsilateral or contralateral. This result suggests that this novel decompression technique not only provides effective decompression but also preserves the facet joints well.

Based on the cadaver study for ULBD, the decompression for the ipsilateral lateral recess should be as effective as the contralateral one. However, because of limited visualization and technically demanding access for the contralateral side, the ipsilateral side is usually better decompressed but more destructed. Curved instruments, such as punch and high-speed bur, are also recommended to minimize the invasion and better decompression for the ipsilateral lateral recess. Most studies suggest the ULBD approach from the dominant symptoms side to ensure adequate decompression. However, in our study, the “no-punch” decompression provides comparably good preservation for either ipsilateral or contralateral facet joints. In 61% of the decompression segments, the ipsilateral facet joints were preserved better than the contralateral facet joints. Therefore, the surgeon can choose an approach from the patient’s right or left side, depending on his preference or the patient’s specific

https://doi.org/10.14245/ns.2448376.188
anatomical features.

There are some limitations in this study. This study is a retrospective case series study, and the patient selection is subjected to a bias based on the surgeon’s preference. The surgeries are performed by an experienced surgeon with a particular interest in minimally invasive and endoscopic spine surgery. The treatment results cannot be extrapolated to surgeons with less experience or novice surgeons. The specialized curved osteotomes are designed by and custom-made for the author and are currently only available in Taiwan and his institute. However, surgeons inspired by the “no-punch” concept can request the local manufacturer to produce the osteotomes or chisels according to their preferences and design ideas. The cohort in our study is small, with a limited follow-up period. This study aims to propose a new surgical technique and use the preliminary treatment results to evaluate its potential advantages. Comparative studies with longer-term follow-up are necessary to evaluate the benefits of the UBE “no-punch” decompression technique.

**CONCLUSION**

Our study introduces a novel decompression technique for UBE surgery by not using the Kerrison punch. The “no-punch” decompression technique provides good treatment results, adequate neural decompression, excellent facet joint preservation, and a low complication rate.

**NOTES**

**Conflict of Interest:** The author has nothing to disclose.

**Funding/Support:** The Far Eastern Memorial Hospital supported the study with grant No. FEMH-2024 C-046.

**Acknowledgments:** Special thanks to Ji Woon William Chun for his help in manufacturing the custom-made Pao’s osteotomes.

**Author Contribution:** Single author.

**ORCID**

Jwo-Luen Pao: 0000-0002-2000-5146

**REFERENCES**

16. Schizas C, Theumann N, Burn A, et al. Qualitative grading of severity of lumbar spinal stenosis based on the morphol-


https://doi.org/10.14245/ns.2448376.188
Erratum: Correction of Author Information and Funding

Erratum: Exploring lncRNA Expression Patterns in Patients With Hypertrophied Ligamentum Flavum

Junling Chen¹*, Guibin Zhong¹²*, Manle Qiu², Wei Ke¹, Jingsong Xue¹, Jianwei Chen¹²

¹Department of Orthopedics, Baoshan Branch, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China
²Department of Spine Surgery, Department of Orthopedics, Ren Ji Hospital, School of Medicine Shanghai Jiao Tong University, Shanghai, China

To the editor,

After careful review of the published article, we have identified inaccuracies in the author information and funding details. This is of great significance to the conclusion of my project. The fund information is particularly crucial as if it is not corrected, the financial sponsorship and the results of this article will not be recognized, and all our efforts will be in vain.

I sincerely hope that you can understand the difficulties and hardships we have encountered in our work and help us make these necessary corrections. I truly appreciate your understanding and assistance.

CORRECTED AUTHOR INFORMATION AND FUNDING

1. The second author Guibin Zhong’s affiliation should be revised as follows:

   Junling Chen¹*, Guibin Zhong¹²*, Manle Qiu², Wei Ke¹, Jingsong Xue¹, Jianwei Chen¹²

   ¹Department of Orthopedics, Baoshan Branch, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China
   ²Department of Spine Surgery, Department of Orthopedics, Ren Ji Hospital, School of Medicine Shanghai Jiao Tong University, Shanghai, China

2. The funding information for this study should be revised as follows:

   Funding/Support: The study was supported by the Shanghai Natural Science Fund (No: 21ZR1447500), Scientific Research Project of Shanghai Municipal Health Commission (No. 202040370) and Clinical Research Innovation Cultivation Fund of Baoshan Branch of Renji Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (2023-rb-cxjj-004).
Instructions for Contributors

Revised: February 13, 2024

I. General Information

Neurospine provides spine clinicians and researchers with peer-reviewed articles on basic and clinical investigation of spine and spinal cord to enhance patient management, education, clinical or experimental research, and professionalism. The journal will consider submissions in areas on cranio cervical to lumbosacral spine including the followings; neuroscience and pain research, bone mineral research, disc and joint research, bio and industrial technology, pathophysiology, risk factors, symptomatology, imaging, treatment, rehabilitation of spine, and spinal cord/peripheral nerve diseases. Specifically, basic and technology researches include the most influential research papers from all fields of science and technology, revolutionizing what physicians and researchers practicing the art of spinal neurosurgery worldwide know. Thus, we welcome valuable basic and translational technology research articles to introduce cutting-edge research of fundamental sciences and technology in clinical spinal neurosurgery. Clinical or Basic Research Articles, Review Articles, and Letters to the Editor written in English will be accepted.

Neurospine, the official journal of ASIA SPINE, the Neurospinal Society of Japan, Taiwan Neurosurgical Spine Society, and the Korean Spinal Neurosurgery Society, is an international peer-reviewed open-access journal which published quarterly (last day of March, June, September, and December). It was first published in March 31, 2004 with Volume 1 and Number 1 with the name “Korean Journal of Spine,” and renamed as “Neurospine” since March 2018. Neurospine is indexed/tracked/covered by Emerging Sources Citation Index (ESCI), PubMed, PubMed Central, KoreaMed, KoMCI, EBSCO host, and Google Scholar.

II. Submission of Manuscript

1. Authors are requested to submit their papers electronically by using online manuscript submission available at http://submit.e-neurospine.org.
2. Corresponding author is responsible for submission and revision of the manuscripts. ID is required for processing and can be generated at the homepage.
3. All authors should sign on the Copyright Release, Author Agreement and Disclosure of Conflict of Interest form to certify that the contents of the manuscript have not been published and are not being considered for publication elsewhere. If any research grant has been given by any private company or group, this information should be described on the form. All authors must sign their autobiography by themselves. The form can be downloaded at the homepage of the Neurospine (https://e-neurospine.org), and should be submitted at the time of paper submitting.
4. Regarding author information, the list of the authors in the manuscript should include only those who were directly involved in the process of the work. Authors can refer to the guideline by Harvard University in 1999 to find details on authorship (https://hms.harvard.edu/sites/default/files/assets/Sites/Ombuds/files/AUTHORSHIP%20GUIDELINES.pdf).
5. Decision for the publication of the submitted manuscript will be made solely by the editorial board.
6. Professional editing in English is recommended for non-native speakers. Editorial office may request an English editing. In cases of accepted manuscripts, we may provide copy editing and English proofreading free of charge.
7. All published papers become the permanent property of the Korean Spinal Neurosurgery Society. Copyrights of all published materials are owned by the Korean Spinal Neurosurgery Society. Permission must be obtained from the Korean Spinal Neurosurgery Society for any commercial use of materials. Every author should sign the copyright transfer agreement forms.

III. Manuscript Preparation

Authors should refer to “Uniform Requirements for Manuscripts Submitted to Biomedical Journals” (http://www.icmje.org/about-icmje/faqs/icmje-recommendations/).

1. Title Page
1) The title pages must be composed of external and internal title pages.
2) The external title page must contain the article title, and full names of all authors with their institutional affiliations both. The type of manuscript (original articles, review articles, letters to the editor, brief communications) should also be addressed. When the work includes multiple authors with different affiliations, the institution where the research was mainly conducted should be spelled out first, and then be followed by footnote in superscript Arabic numerals beside the authors’ names to describe their affiliation in a consecutive order of the numbers. Running head must be included consisting of no more than 65 characters/spaces. The external title page must also contain the address, telephone and facsimile numbers, and e-mail address of the correspond-
ing author at the bottom of the page, as well as information on
the previous presentation of the manuscript in conferences and
funding resources, if necessary.
3) The internal title page should only contain the article title. The
internal title page must not contain any information on the
names and affiliations of the authors.

2. Manuscript Format
1) The manuscript should be composed of no more than 5,000
English words for original and review articles except for refer-
ces, tables, and figures. Letter to the editor should consist of
not more than 600 English words, and the editorial should con-
sist of no more than 350 words.
2) The original article should be organized in the order of title, ab-
stract, introduction, materials and methods, results, discussion,
conclusion, references, tables, and figures or illustrations.
3) There should be no more than 40 references in original articles.
4) Manuscript format may vary in review articles. There should be
no more than 100 references in review articles.
5) Text should be written in 11 point fonts with double line spacing.

3. Abstract
1) Objective, Methods, Results, and Conclusion sections should
be included in abstract of clinical or laboratory research, but are
not necessary in other types of studies.
2) The abstract should include brief descriptions on the objective,
methods, results, and conclusion as well as a detailed descrip-
tion of the data. An abstract containing 250 words or less is re-
quired for original articles and review articles.
3) Abstract can be revised by the decision of editorial board, and
some sentences can be modified as a result of revision.
4) A list of key words, with a minimum of two items and maxi-
mum of six items, should be included at the end of the abstract.
5) The selection of Key Words should be based on Medical Subject
Heading (MeSH) of Index Medicus and the web site (http://

<table>
<thead>
<tr>
<th>Type</th>
<th>Abstract Max.</th>
<th>Keywords Max.</th>
<th>Main text Max. words</th>
<th>Tables or Figures Max.</th>
<th>References Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Article</td>
<td>250</td>
<td>6</td>
<td>5,000</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>Original Article</td>
<td>250</td>
<td>6</td>
<td>5,000</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>Brief Communication</td>
<td>200</td>
<td>6</td>
<td>1,500</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Letter to the Editor</td>
<td>x</td>
<td>x</td>
<td>600</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Editorial</td>
<td>x</td>
<td>x</td>
<td>350</td>
<td>-</td>
<td>10</td>
</tr>
</tbody>
</table>

Note: Exceptions may be made to the above specifications according to the de-
cision of the editorial committee.

4. Introduction
The introduction should address the purpose of the article concisely,
and include background reports mainly relevant to the purpose of
the paper. Detailed review of the literature should be addressed in
the discussion section.

5. Materials and Methods
1) The article should record research plans, objective, and meth-
ods in order, as well as the data analysis strategies and control of
bias in the study. Enough details should be furnished for the
reader to understand the method(s) without reference to an-
other work in the study described.
2) When reporting experiments with human subjects, the authors
must document the approval received from the local Institu-
tional Review Board. When reporting experiments with animal
subjects, the authors should indicate whether the handling of
the animals was supervised by the research board of the affiliat-
ed institution or such. Approved number of IRB must be noted.
3) Photographs disclosing patients must be accompanied by a
signed release form from the patient or family permitting publi-
cation.
4) Ensure correct use of the terms sex (when reporting biological
factors) and gender (identity, psychosocial or cultural factors),
and, unless inappropriate, report the sex and/or gender of study
participants, the sex of animals or cells, and describe the meth-
ods used to determine sex and gender. If the study was done in-
volving an exclusive population, for example in only one sex,
authors should justify why, except in obvious cases (e.g., pros-
tate cancer). Authors should define how they determined race
or ethnicity and justify their relevance.

6. Results
1) The authors should logically describe their results of observa-
tions and analyses performed using methodology given in the
previous section and provide actual data.
2) For biometric measurements in which considerable amount of
stochastic variation exists, a statistical evaluation is mandatory.
The results must be solely from the findings of the current
study and not refer to any previous reports.
3) While an effort should be made to avoid overlapping descrip-
tions by Tables and by main text, important trends and points
in the Table should be described in the text.

7. Discussion
Discussions about the findings of the research and interpretations in
relation to other studies are made. It is necessary to emphasize the
new and critical findings of the study, not to repeat the results of the
study presented in the previous sections. The meaning and limita-
tion of observed facts should be described, and the conclusion
should be related to the objective of the study only when it is supported by the results of the research.

8. Conclusion
The conclusion section should include a concise statement of the major findings of the study in accordance with the study purpose.

9. References
The authors are responsible for the accuracy of the references. Key the references (double-spaced) at the end of the manuscript. End-Note users can access a direct download of the updated Neurospine Publications style at https://www.e-neurospine.org. References should be numbered consecutively in the order in which they are first mentioned in the text. All references cited in the text must be both listed and cited by the reference number (footnotes are not accepted). Use superscript numerals outside periods and commas, inside colons and semicolons. When more than 2 references are cited at a given place in the manuscript, use hyphens to join the first and last numbers of a closed series; use commas without space to separate other parts of a multiple citation (e.g., As reported previously,\textsuperscript{1,3-8,19}...The derived data were as follows\textsuperscript{3,4,12};)

Do not link the references to the text. Cite unpublished data, such as papers submitted but not yet accepted for publication or personal communications, in parentheses in the text. If there are more than three authors, name only the first three authors and then use et al.

Refer to the List of Journals Indexed in Index Medicus for abbreviations of journal names, or access the list at https://www.nlm.nih.gov/archive/20130415/tsd/serials/lji.html. Sample references are given below:

- **Journal article**

- **Book chapter**

- **Entire book**

- **Software**

- **Online journals**

- **Database**

- **World Wide Web**

10. Tables
1) Tables should be created using the table formatting and editing feature of Microsoft Word. The title of the table must be noted. Tables cannot be submitted in a picture format.
2) Tables should be prepared in detail, in order to understand the contents of the manuscript without further reference.
3) Tables should be submitted separately from manuscript. Do not include vertical lines in table, and refer to the table formats in formal papers in Neurospine.

11. Figures and Illustrations
1) Figures should have resolution of 300 dpi or above and should be submitted individually (Namely, if Figure 1 is divided into A, B, C, and D, do not combine them into one, but submit each of them separately). Allowable file format for figures are JPG or TIF (TIFF) only.
2) Figures should be named according to figure name (example: Fig-1.A.tif). If the quality of the photographs is considered as inappropriate for printing, resubmission of them can be requested by the journal.
3) Line art should have resolution of 1,200 dpi or more in JPG or TIF format.

12. Author Check List
1) Before submitting the manuscript, authors should double-check all requirements noted in the agreement form regarding the registration and copyrights of their manuscript. A manuscript that does not fit the author instructions of the journal regarding format and references will be returned to the authors for further correction.
2) The page numbers in the manuscript should be counted from the page with the abstract, and the name and affiliation of the authors should not appear thereafter.
3) Author check list should be prepared, signed by corresponding
IV. Peer Review Process

All manuscripts are considered confidential. They are peer-reviewed by at least 2 anonymous reviewers selected by the Editor. The corresponding author is notified as soon as possible of the Editor’s decision to accept, reject, or ask for revisions. The average time interval for an initial review process that involves both editorial and peer reviews is approximately 1 month; occasionally, there are unavoidable delays, usually because a manuscript needs multiple reviews or several revisions. When manuscripts are returned for revision, a cover letter from the Editor provides directions that should be followed carefully. When submitting the revised manuscript, authors should include a Response Letter, which describes how the manuscript has been revised. A point-by-point response to the Editor should be included with the revised manuscript. Authors who plan to resubmit but cannot meet this deadline should contact the Editorial Office. Manuscripts held for revision will be retained for a maximum of 90 days. The revised manuscript and the author's comments will be reviewed again. If a manuscript is completely acceptable, according to the criteria set forth in these instructions, it is scheduled for publication in the next available issue.

We neither guarantee the acceptance without review nor very short peer review times for unsolicited manuscripts. Commissioned manuscripts also are reviewed before publication.

We adopt double-blind peer review in which case, not only authors but also reviewers do not know each other.

V. Publication and Charges

1) Once a manuscript is accepted for publication by the journal, it will be sent to the press, and page proofs will be sent to authors. Authors must respond to the page proofs as soon as possible after making necessary corrections of misspellings, and the location of the photographs, figures or tables. Authors can make corrections for only typing errors, and are not allowed to make any author alteration or substantive changes of the text. Proofs must be returned to the press within 48 hours of receipt. No response from the authors within this time frame will lead the publication of the proof read without corrections, and the editorial board will not be responsible for any mistakes or errors occurring in this process.

2) Starting from January 2024, Neurospine will introduce article processing charge (APC) for papers that are submitted and accepted for publication. This APC has been set at $400. Recognizing the potential impact of this change, we are concurrently working on establishing a discount benefit system.

This will be available to members of various cooperative societies and individuals who have shown exceptional dedication to Neurospine. Our aim is to alleviate the financial burden for our most active and committed contributors.

VI. Ethical Guidelines

1. Research Ethics

1) All of the manuscripts should be prepared in strict observation of research and publication ethics guidelines recommended by the Council of Science Editors (CSE), International Committee of Medical Journal Editors (ICMJE), World Association of Medical Editors (WAME), and the Korean Association of Medical Journal Editors (KAMJE).

2) Any study including human subjects or human data must be reviewed and approved by a responsible institutional review board (IRB). Please refer to the principles embodied in the Declaration of Helsinki (https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/) for all investigations involving human materials.

3) Animal experiments also should be reviewed by an appropriate committee (Institutional Animal Care and Use Committee, IACUC) for the care and use of animals. Also studies with pathogens requiring a high degree of biosafety should pass review of a relevant committee (Institutional Biosafety Committee, IBC). The editor of Neurospine always request submission of copies of informed consents from human subjects in clinical studies or IRB approval documents.

2. Conflict of Interest

1) The corresponding author of an article is asked to inform the Editor of the authors’ potential conflicts of interest possibly influencing their interpretation of data. A potential conflict of interest should be disclosed in the cover letter even when the authors are confident that their judgments have not been influenced in preparing the manuscript. Such conflicts may be financial support or private connections to pharmaceutical companies, political pressure from interest groups, or academic problems. Disclosure form shall be same with ICMJE Uniform Disclosure Form for Potential Conflicts of Interest (http://www.icmje.org/coi_disclosure.pdf).

2) The Editor will decide whether the information on the conflict should be included in the published paper. Before publishing such information, the Editor will consult with the corresponding author. In particular, all sources of funding for a study should be explicitly stated. The Neurospine asks referees to let its Editor...
3. Journal Policies on Authorship and Contributorship

1) Authors are required to make clear of their contribution to their manuscript in cover letter. To be listed as an author one should have contributed substantially to all three categories established by the International Committee of Medical Journal Editors (ICMJE): (1) conception and design, or acquisition, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published; and (4) agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

2) When a large, multicenter group has conducted the work, the group should identify the individuals who accept direct responsibility for the manuscript. When submitting a manuscript authored by a group, the corresponding author should clearly indicate the preferred citation and identify all individual authors as well as the group name. Journals generally list other members of the group in the Acknowledgments. Acquisition of funding, collection of data, or general supervision of the research group alone does not constitute authorship. Authors are responsible for replying to all questions asked by reviewers or editors that relate to the accuracy or integrity of any part of the work. All persons who have made a substantial contribution, but who are not eligible as authors, should be named in the acknowledgments. Authors are expected to consider carefully the way authors should be listed and ordered before submitting their manuscripts, and to provide a definitive list of authors with their original submission. Any addition, deletion, or rearrangement of author names in the authorship list should be made before the manuscript has been accepted—and only if approved by the journal Editor. To request such a change, the Editor must receive the following from the corresponding author: (a) the reason for requesting a change in the list of authors; and (b) written confirmation (by e-mail or letter) from all authors to say that they agree with the addition, removal, or rearrangement.

4. Redundant Publication and Plagiarism

1) Redundant publication is defined as “reporting (publishing or attempting to publish) substantially the same work more than once, without attribution of the original source(s)”. Characteristics of reports that are substantially similar include the following: (a) “at least one of the authors must be common to all reports (if there are no common authors, it is more likely plagiarism than redundant publication),” (b) “the subject or study populations are often the same or similar,” (c) “the methodology is typically identical or nearly so,” and (d) “the results and their interpretation generally vary little, if at all.”

2) When submitting a manuscript, authors should include a letter informing the editor of any potential overlap with other already published material or material being evaluated for publication and should also state how the manuscript submitted to Neurospine differs substantially from this other material. If all or part of your patient population was previously reported, this should be mentioned in the Materials and Methods, with citation of the appropriate reference(s).

3) The editorial committee checks the similarity by using the iThenticate (http://www.ithenticate.com/) program for all submitted articles to prevent plagiarism. The editorial committee rejects the article suspected of plagiarism and asks the author to check whether it is plagiarized and make a resubmission.

5. Readership

It is primarily for clinicians and researchers who care patients with spine and spinal cord diseases. They are able to obtain tailored information to adopt for their research and practice. Its readership can be expanded to other positions: • Researchers can get the recent topics of clinical research in spine and spinal cord field and detailed research methods; • Clinicians in the field can get the new information and recent development for care of patients; • Medical teacher can access and adopt a variety of data in medical education; • Allied health professionals including nurses are able to get the recent information for care of patients with spine and spinal cord diseases; • Policy makers are able to reflect the results of the articles to the nation-wide health care policies for patients with spine and spinal cord diseases; • The public, especially family of patients with spine and spinal cord diseases are able to read the advance ment in their family’s diseases so that they have a better knowledge on the diseases and a confidence in the clinicians’ devotion to their family.

6. Obligation to Register Clinical Trial

1) Clinical trial defined as “any research project that prospectively assigns human subjects to intervention and comparison groups to study the cause-and-effect relationship between a medical intervention and a health outcome” should be registered to the primary registry to be prior publication.

7. Process for Identification of and Dealing With Allegations of Research Misconduct
When the Journal faces suspected cases of research and publication misconduct such as a redundant (duplicate) publication, plagiarism, fabricated data, changes in authorship, undisclosed conflicts of interest, an ethical problem discovered with the submitted manuscript, a reviewer who has appropriated an author’s idea or data, complaints against editors, and other issues, the resolving process will follow the flowchart provided by the Committee on Publication Ethics (http://publicationethics.org/resources/flowcharts). The Editorial Board will discuss the suspected cases and reach a decision. We will not hesitate to publish errata, corrigenda, clarifications, retractions, and apologies when needed.

Neurospine adheres to the research and publication ethics policies outlined in International Standards for Editors and Authors (http://publicationethics.org) and the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (http://icmje.org). Any studies involving human subject must comply with the principles of the World Medical Association Declaration of Helsinki. Clinical research should be approved by the Institutional Review Board, as well through patient consent. A patient's personal information cannot be published in any form. However, if it is absolutely necessary to use a patient's personal information, the consent of the patient or his/her guardian will be needed before publishing. Animal studies should be performed in compliance with all relevant guidelines, observing the standards described in the NIH Guide for the Care and Use of Laboratory Animals.

Cases that require editorial expressions of concern or retraction shall follow the COPE flowcharts available from: http://publicationethics.org/resources/flowcharts. If correction is needed, it will follow the ICMJE Recommendation for Corrections, Retractions, Replications and Version Control available from: http://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/corrections-and-version-control.html as follows:

Honest errors are a part of science and publishing and require publication of a correction when they are detected. Corrections are needed for errors of fact. Minimum standards are as follows: First, it shall publish a correction notice as soon as possible, detailing changes from and citing the original publication on both an electronic and numbered print page that is included in an electronic or a print Table of Contents to ensure proper indexing. Second, it shall post a new article version with details of the changes from the original version and the date(s) on which the changes were made through CrossMark; Third, it shall archive all prior versions of the article. This archive can be either directly accessible to readers; and Fourth, previous electronic versions shall prominently note that there are more recent versions of the article via CrossMark.

8. Handling Complaints and Appeals
The policy of the journal is primarily aimed at protecting the authors, reviewers, editors, and the publisher of the journal. If not described below, the process of handling complaints and appeals follows the guidelines of the Committee of Publication Ethics available from: https://publicationethics.org/appeals

Who complains or makes an appeal?
Submitters, authors, reviewers, and readers may register complaints and appeals in a variety of cases as follows: falsification, fabrication, plagiarism, duplicate publication, authorship dispute, conflict of interest, ethical treatment of animals, informed consent, bias or unfair inappropriate competitive acts, copyright, stolen data, defamation, and legal problem. If any individuals or institutions want to inform the cases, they can send a letter to editor through https://www.e-neurospine.org/about/contact.php. For the complaints or appeals, concrete data with answers to all factual questions (who, when, where, what, how, why) should be provided.

Who is responsible to resolve and handle complaints and appeals?
The Editor, Editorial Board, or Editorial Office is responsible for them.

What may be the consequence of remedy?
It depends on the type or degree of misconduct. The consequence of resolution will follow the guidelines of the Committee of Publication Ethics (COPE).

9. Postpublication Discussions and Corrections
The postpublication discussion is available through letter to the editor. If any readers have a concern on any articles published, they can submit letter to the editor on the articles. If there founds any errors or mistakes in the article, it can be corrected through errata, corrigenda, or retraction.

10. Policies on data sharing and reproducibility
Until 2020, authors will be encouraged to share their data openly, but starting in 2021, they will be mandated to do so. The related regulation follows the open data sharing policy outlined below.

1) Open data sharing policy
For clarification on result accuracy and reproducibility of the results, raw data or analysis data will be deposited to a public repository, for example, Harvard Dataverse (https://dataverse.harvard.edu/) after acceptance of the manuscript. Therefore, submission of the raw data or analysis data is mandatory. If the data is already a public one, its URL site or sources should be disclosed. If data cannot be publicized, it can be negotiated with the editor. If there are any inquiries on depositing data, authors should contact the editorial office.

2) Clinical data sharing policy
This journal follows the data sharing policy described in "Data Sharing Statements for Clinical Trials: A Requirement of the International Committee of Medical Journal Editors” (https://doi.org/10.3346/jkms.2017.32.7.1051). As of July 1, 2018 manuscripts submitted to ICMJE journals that report the results of interventional clinical trials must contain a data sharing state-
Table. Examples of Data Sharing Statements That Fulfill These ICMJE Requirements*

<table>
<thead>
<tr>
<th>Will individual participant data be available (including data dictionaries)?</th>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
<th>Example 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>What data in particular will be shared?</td>
<td>All of the individual participant data collected during the trial, after deidentification.</td>
<td>Individual participant data that underlie the results reported in this article, after deidentification (text, tables, figures, and appendices).</td>
<td>Individual participant data that underlie the results reported in this article, after deidentification (text, tables, figures, and appendices).</td>
<td>Not available</td>
</tr>
<tr>
<td>When will data be available (start and end dates)?</td>
<td>Immediately following publication. No end date.</td>
<td>Beginning 3 months and ending 5 years following article publication.</td>
<td>Beginning 9 months and ending 36 months following article publication.</td>
<td>Not applicable</td>
</tr>
<tr>
<td>With whom?</td>
<td>Anyone who wishes to access the data.</td>
<td>Researchers who provide a methodologically sound proposal.</td>
<td>Investigators whose proposed use of the data has been approved by an independent review committee (learned intermediary) identified for this purpose.</td>
<td>Not applicable</td>
</tr>
<tr>
<td>For what types of analyses?</td>
<td>Any purpose.</td>
<td>To achieve aims in the approved proposal.</td>
<td>For individual participant data meta-analysis.</td>
<td>Not applicable</td>
</tr>
<tr>
<td>By what mechanism will data be made available?</td>
<td>Data are available indefinitely at (Link to be included).</td>
<td>Proposals should be directed to xxx@yyy. To gain access, data requestors will need to sign a data access agreement. Data are available for 5 years at a third party website (Link to be included).</td>
<td>Proposals may be submitted up to 36 months following article publication. After 36 months the data will be available in our University’s data warehouse but without investigator support other than deposited metadata. Information regarding submitting proposals and accessing data may be found at (Link to be provided).</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

* These examples are meant to illustrate a range of, but not all, data sharing options.

11. Artificial intelligence guideline

The emergence of artificial intelligence (AI), exemplified by Chat GPT’s announcement in 2022, has brought about significant transformations in the academic landscape. AI has proven its capacity to perform tasks at an exceptionally high level in scholarly paper writing, challenging the traditional belief that academic writing is solely within the realm of human intellectual abilities. This shift is positive as it brings us closer to uncovering truths that may have been overlooked. It is also helpful to enhance the overall readability of academic papers and overcome the language barrier.

However, it is crucial to acknowledge that AI still grapples with unresolved issues, with hallucination being a prominent concern. The reliability of AI is not yet definitive. Copyright-related concerns arising from creative processes facilitated by AI necessitate societal consensus. It is imperative to evaluate this technology not solely for its convenience but also considering its broader implications. We highly respect the creative process of researchers and acknowledge its intrinsic value. Therefore, the editorial board of Neospire earnestly requests authors to adhere to the following guidelines when
submitting papers to the journal:
1) Artificial intelligence tools cannot be listed or cited as one of the authors due to its inability to take responsibility for errors.
2) Authors should exert every effort to ensure the reliability of their papers when utilizing artificial intelligence, holding responsibility for any plagiarism or false information generated using AI.
3) Authors must provide detailed information, including prompts, AI tools used, and their versions, in the Materials and Methods or Acknowledgment section when employing AI tools.
4) Images or videos created using AI, without societal consensus on copyright, cannot be included in papers at the moment.
5) Reviewers are cautioned against sharing the manuscript outside during the peer-review process. Even simple uploading papers to external AI tools can break confidentiality.
6) The editor may refuse to proceed with review of the paper if inappropriate use of AI is detected.

We recognize that technology's rapid evolution continually shapes the academic writing process, and the points mentioned above may evolve based on future societal agreements.

All correspondences, business communications and manuscripts should be mailed to:

**Editor-in-Chief:** Inbo Han

**Editorial Office**
CHA University, CHA School of Medicine Bundang Medical Center, 59 Yatap-ro, Bundang-gu, Seongnam-si, Gyeonggi-do 13496, Korea
Tel: +82-31-780-1924, Fax: +82-31-780-5269
Email: support@e-neurospine.org
Copyright Release, Author Agreement, Human and Animal Right, and Disclosure of Conflict of Interest

The author(s) submit my/our manuscript with the following title

__________________________________________________________________________________________

In consideration of the Editorial Board reviewing for the Neurospine publishing.

1. Copyright Release and Author Agreement
I/we undersigned hereby transfer all rights, interest, copyright and digital copyright related to the journal to the Neurospine upon acceptance of the manuscript for publication. I/we have all rights, such as right to apply patents and right to use part or all of the contents of the manuscript, except copyright. I/we can use materials from the manuscript under written agreement of the Neurospine, and in this case, I/we will clarify the reference. All authors have made a concrete and intellectual contribution to the content of the manuscript, and will take public responsibility for its content.
The author(s) certify that the manuscript was prepared in strict observation of research and publication ethics guidelines recommended by the editorial committee of the Neurospine.
The author(s) certify that the contents of the manuscript have not been published and are not being considered for publication elsewhere.

2. Human and Animal Right
In case of experimenting on human, the author(s) have certified that the process of the research is in accordance with ethical standards of Helsinki declaration, domestic and foreign committees that preside over human experiment.
If any doubts are raised whether the research was proceeded in accordance with the declaration, the author(s) would explain it.
In case of experimenting on animals, the author(s) have certified that the author(s) had followed the domestic and foreign guideline related to experiment of animals in a laboratory.

3. Disclosure of Conflict of Interest
The author(s) of the journal have clarified everything that interest may arise such as research expenses, consultant expenses, stock, particularly concerned person of the judges on the document of disclosure of conflict of interest.
If there are conflicts of interest, authors should state their content on the title page of the manuscript.

<table>
<thead>
<tr>
<th>Author's name</th>
<th>Author's Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1. Mandatory components of a manuscript
   1) Formats and contents of the manuscripts are checked by corresponding author. □ Yes / □ No
   2) All manuscripts should be written in English. Manuscripts may be no longer than 5,000 English words
      for original articles except for references, tables, and figures. □ Yes / □ No
   3) Manuscripts should be prepared in the following orders.
      Original article: external title page, internal title page, abstract, key words, introduction, materials, and
      methods, results, discussion, conclusion, references, table, and figure legends. □ Yes / □ No
   4) "Editing in English is done prior to submission of a manuscript." □ Yes / □ No

2. External title page
   The external title page should be a separate file, and must contain names and affiliations of all authors and
   contact information of the corresponding author. □ Yes / □ No

3. Internal title page
   Only the English title of the manuscript is listed. Any information on the names and affiliations of the
   authors is not shown on the internal title page. □ Yes / □ No

4. Abstract
   1) Abstract should have no longer than 250 words for original articles and review articles. □ Yes / □ No
   2) Abstract includes Objective, Methods, Results, and Conclusion in clinical or laboratory research. □ Yes / □ No
   3) The selection of Key Words is based on medical subject headings (MeSH) terms. □ Yes / □ No

5. Manuscript
   1) Text is written in 11-point fonts with double line spacing. □ Yes / □ No
   2) Figures and tables are cited in numerical order in the order they are mentioned in the text. □ Yes / □ No

6. References
   1) References should be numbered consecutively in Arabic numeric order in which they are first men-
      tioned in the text. □ Yes / □ No
   2) All references cited in the text must be both listed and cited by the reference number (footnotes are not
      accepted). □ Yes / □ No
   3) When more than 2 references are cited at a given place in the manuscript, use hyphens to join the first
      and last numbers of a closed series; use commas without space to separate other parts of a multiple ci-
      tation (e.g., As reported previously,[1,3,8,9]...The derived data were as follows[3,4,12]:) □ Yes / □ No
   4) If there are more than 3 authors in end-reference list, name only the first 3 authors and then use et al. □ Yes / □ No
   5) Use superscript numerals outside periods and commas, inside colons and semicolons. □ Yes / □ No

7. Tables, Figures and Illustrations
   1) Tables and figures are prepared in separate files. □ Yes / □ No
   2) Figures are submitted individually not incorporated into one file. □ Yes / □ No
   3) Figures and illustrations are saved in JPG or TIF file format and have a resolution of 300 DPI or more.
      (Line art should have resolution of 1,200 dpi or more) □ Yes / □ No
8. Copyright Release, Author Agreement, Human and Animal Right, and Disclosure of Conflict of Interest
All authors signed on the Copyright Release, Author Agreement, Human and Animal Right, and Disclosure of Conflict of Interest form to verify that the purpose of the research is not related to personal interests and the manuscript was prepared in strict observation of research and publication ethics guidelines.
The above form is submitted with the manuscript.
☐ Yes / ☐ No

9. Ethical approval of studies and informed consent
For all manuscripts reporting data from studies involving human participants or animals, formal review and approval, or formal review and waiver, by an appropriate institutional review board or ethics committee is required and should be described in the Materials and Methods section.
☐ Yes / ☐ No