ABSTRACT

**Purpose:** To validate with a prospective study a decision-supporting coding system for the surgical approach for multi-level degenerative cervical myelopathy (mDCM).

**Methods:** Ten cases were presented on an internet platform, including clinical and imaging data. A single approach (G1), a choice between two (G2), or three approaches (G3) were options. Senior and junior spine surgeons analyzed seven parameters: location and extension of the compression of the spinal cord; C-spine alignment and instability; general morbidity and bone diseases; and K-line and multi-level corpectomy. For each parameter, an anterior, posterior, or combined approach was suggested. The most frequent letter or the last letter (if C) of the resulting seven letter code (7LC) suggested the surgical approach. Each surgeon performed two reads per case within eight weeks.

**Results:** G1: Inter-rater reliability between junior surgeons improved from the first read (κ=0.40) to the second (κ=0.76; p<0.001) but did not change between senior surgeons (κ=0.85). The intra-rater reliability was similar for junior (κ=0.78) and senior (κ=0.71) surgeons. G2: Junior/senior surgeons agreed completely (58%/62%), partially (24%/23%), or did not agree (18%/15%) with the 7LC choice. G3: Junior/senior surgeons agreed completely (50%/50%) or partially (50%/50%) with the 7LC choice.

**Conclusion:** The 7LC showed good overall reliability. Junior surgeons went through a learning curve and converged to senior surgeons in the second read. The 7LC helps less experienced surgeons to analyze, in a structured manner, the relevant clinical and imaging parameters influencing the choice of the surgical approach, rather than simply pointing out the only correct one.

**Key words:** degenerative cervical myelopathy, surgical treatment, scores, preoperative validation, multicenter study
Several parameters must be considered when deciding the surgical approach for patients with multi-level degenerative cervical myelopathy (mDCM). Imaging shows the location of the spinal cord compression, number of levels involved, sagittal alignment, and instability of the C-spine and influences decision making.\textsuperscript{1,2} Patient’s general comorbidities including surgically relevant bone diseases, such as osteoporosis, rheumatoid arthritis (RA), and ankylosing spondylitis (AS) should be taken in account. Furthermore, factors such as the need for multi-level corpectomy or “K-line negative” type of compression also play a role. Lastly, the surgeon’s familiarity with specific techniques and his overall surgical experience influence the choice of the surgical approach as well.\textsuperscript{3,4}

The clinical outcomes of anterior, posterior, and combined approaches have been shown to be equivalent, but they differ in terms of complications and cost-effectiveness.\textsuperscript{5-7} Deciding upon the most appropriate surgical approach in mDCM can be challenging for senior surgeons and certainly for junior surgeons. To tailor the surgical approach to the needs of a specific patient, we introduced into our daily practice a coding system, which we termed the “seven-letter code” (7LC) for didactical purpose, because seven parameters are analyzed (Fig. 1). Completing the check-list results in a chain of letters that indicate an approach: A, anterior; P, posterior; and C, combined. The most frequent letter or the last letter (if C) suggests the approach that should be chosen. However, in some cases, the balance of letters may indicate equipoise regarding the surgical route.

We aimed to validate the 7LC with an international panel of senior and junior spine surgeons. The following questions were investigated: Does the 7LC provide substantial inter- and intra-rater reliability (inter-RR and intra-RR)? Do junior and senior surgeons make different decisions? To what extent does the approach suggested by the 7LC correspond to the surgeon’s personal choice?

If an algorithm considers most of the relevant factors that influence the decision-making, one would expect similar surgical plans. On the contrary, if it is not clinically sound, then one would expect low inter-RR and intra-RR.

MATERIALS and METHODS
Patient data

The anonymized clinical and imaging data of ten patients who underwent surgery for DCM in our clinic in the past two years were presented on an internet platform to 16 raters. The clinical data included age, gender, European Myelopathy Score (EMS), American Society of Anesthesiologists scale (ASA), and diseases affecting bone quality, such as relevant osteoporosis (T score < -2.5 or a history of fractures), rheumatoid arthritis (RA), and ankylosing spondylitis (AS). The imaging data included sagittal and axial T2-weighted magnetic resonance imaging (MRI) slices; axial and reformatted computed tomography slices; and biplanar plain films with additional flexion/extension radiographs.

The 7LC

The four steps to generate the code are discussed below.

1. Select the parameters

The 7LC analyzed seven imaging and clinical parameters (Fig. 1). The imaging consisted of cervical MRI, computed tomography (CT), and biplanar X-Ray with flexion/extension images. MRI and CT showed if the prevalent impingement of the spinal cord is anterior, posterior, or combined. Furthermore, the number of levels involved (< 3, 3, > 3) can be defined (Figs. 2a-b, 3a-b, 4a-b). The plain X-Ray films depicted the (mal)alignment and the (in)stability of the C-spine (Figs. 2c, 3c, 4c). Comorbidities also influenced the decision about the surgical approach (Figs. 2a, 3a, 4a). The general ones are summarized by ASA ranging from 1 to 5. In a recent meta-analysis, the anterior versus the posterior approach for the treatment of mDCM was compared; intraoperative blood loss, operation time, and reoperation rate were significantly higher in the anterior surgery with no apparent difference in the neural function recovery rate between the two approaches. Bone diseases considered were: relevant osteoporosis, defined as history of spontaneous fractures or T-score worse than -2.5, AS, and RA. The reduced bone density of the vertebral body and the altered biomechanics common to these diseases require often an additional dorsal fixation following the anterior procedure or solely a dorsal procedure. Special circumstances such as the necessity of a two- or multi-level vertebral body replacement or as a K-line negative type of compression were more of an exception rather than the
rule. Therefore, they were the last step of the analysis or the last letter of the 7LC. If these circumstances apply, the combined approach is always suggested.

2. Fulfilled imaging/clinical parameter results in a letter

Each parameter can be categorized in a way that suggests a surgical approach (Fig. 1): A (anterior); P (posterior); C (combined); or 0 (zero) when the parameter does not apply, e.g. no instability or bone diseases. Sometime an equipoise as A/P or P/C may result and mirrors the clinical reality.

3. Count the letters

The most frequent letter or the last letter (if C) of the resulting seven-letter chain suggests the approach. If two letters showed an equal frequency, then the surgeon may opt for his favorite approach.

4. The majority wins

The 7LC-algorithm weights a single letter (e.g. A) with 1 point and a combined letter (e.g. A/P) with 0.5 point each. The 7LC can be downloaded for personal use on the smartphone (https://7LC.org).

The analysis of the case shown in Fig. 2 generates the code A-A-A/P-0-A/P-0-0, indicating that an anterior approach (A:3 vs P:1) should be considered.

The analysis of the case shown in Fig. 3 generates the code A-A/P-A/P-0-A/P-0-0, which could lead as well to an anterior approach as to a posterior one (A:2.5 vs P:1.5).

The intent of the 7LC is primarily to support surgeons to analyze, in a structured manner, the relevant clinical and imaging parameters influencing the choice of the surgical approach, rather than simply pointing out the only correct one.

The ten test cases

The ten real cases were divided into three groups according to their clinical and imaging characteristics. In the first group of five cases, each case was expected to result in the choice of a single surgical approach (Fig. 2). This group was intended to familiarize the raters with the 7LC system. Two of the five cases had the same imaging data but different clinical data; these were intended as an “internal control.” In the second group, the three cases were expected to show equipoise
between the two approaches (Fig. 3). This group was intended to highlight the individual preferences and the cultural differences of the surgeons. The third group comprised two complex cases that could be treated by three approaches, depending on the weight put on the clinical and imaging parameters (Fig. 4).

The raters

We invited sixteen surgeons to use the 7LC to rate the ten cases. The eight senior raters each had more than 20 years of experience in cervical spine surgery, and the eight junior raters each had less than three years of experience. To ensure a heterogeneous group, we considered whether each surgeon had a known preference for anterior or posterior procedures or no known preference. We invited surgeons from centers around the world to represent different cultures (three centers in Germany and one center each in Sweden, United Kingdom, United States, Australia, and Japan).

The internet platform

Each rater received a personal link to a professional online platform (www.evidat.com, Luebeck, Germany) that presented in English an introductory lecture about the use of the 7LC. After the successful completion of a demonstration case, the raters could then score the individual cases in a random sequence. The scoring could be completed in several sessions if desired. The procedure was repeated within eight weeks. For the second read, the cases were presented in a new random sequence.

Statistics

Kappa values were computed to determine the inter-RR and intra-RR. To provide an overall index of agreement, Cohen’s weighted kappa coefficients or Finn coefficients were first computed for all rater pairs, and then the arithmetic mean of these estimates was calculated as suggested by Light for fully crossed designs with three or more raters. As an additional index of agreement, the median of the paired kappa coefficients was calculated. The level of significance was set to 5%.

Kappa values can range from -1 (less agreement than expected by chance) to 0 (agreement expected by chance) and to 1 (perfect agreement). The interpretation of the kappa values was based on the recommendations of Landis and Koch, which were as follows: kappa <0, poor agreement; kappa 0–
0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement.\textsuperscript{17}

All statistical analyses were performed using R Version 3.4.0 (R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing; Vienna, Austria; URL: www.R-project.org) and the R package irr.\textsuperscript{18}

\textbf{Ethics}

According to the deliberation of the Ethics Committee of the \textbf{Federal State of Hamburg} (PV5817) informed consent can be waived for retrospective database-based studies, whenever the data was acquired, saved and treated anonymously. This applies to the present study.

\textbf{RESULTS}

Fifteen raters completed both readings, and one rater performed the first reading, resulting in a total of 31 data sets.

\textit{The first group of cases (majority of single-approach ratings)}

For the junior surgeons, the inter-RR improved significantly from the first read (\(\kappa=0.40\)) to the second read (\(\kappa=0.76; p<0.001\)). For the senior surgeons, the inter-RR was the same for both reads (\(\kappa=0.85\)). The intra-RR was similar for junior surgeons (\(\kappa=0.78\)) and senior surgeons (\(\kappa=0.71\)).

Two cases (nos. 1 and 8) for which an anterior approach was most likely had the same imaging but different ages (29 vs. 70 years), sex (female vs. male), and findings (EMS 14, BMI 42 vs. EMS 8, diabetes mellitus, and treatment with warfarin for cardiovascular disease). A single junior surgeon changed his mind on the second reading and switched from the anterior to the posterior approach in case no. 1. Two raters chose a posterior approach for case no. 8. Seven senior surgeons confirmed the anterior approach in both cases; one changed his mind and favored a posterior approach in the elderly patient.

\textit{The second and third group of cases (balanced distribution between two or three approaches)}
The correspondence between the approach as determined by the 7LC and the approach as chosen by the surgeons is shown in Figure 5.

DISCUSSION

Several studies have addressed the difficulty of deciding upon a surgical approach for multi-level cervical spondylotic myelopathy (CSM), which is now called DCM. In 1996, Naderi et al. concluded that although there are many proposed surgical approaches for patients with CSM, there is no standard method to determine which approach is preferable. In 2011, Goghawala et al. noted the comparative effectiveness of anterior and dorsal surgery for CSM and stated that there is most likely not a single “best” approach to CSM. In 2015, Luo et al. performed a meta-analysis of studies that compared the anterior and the posterior approach for the treatment of multilevel CSM and found that a definitive conclusion could not be reached regarding which surgical approach is more effective for the treatment of multilevel CSM. In 2017, Bakhsheshian et al. reviewed the current management of CSM and concluded that “further high-quality randomized clinical studies with long-term follow-up are still needed to help predict the ideal surgical strategy”. Given the multitude of clinical and imaging parameters that should be considered, the surgeon’s personal experience during training and in subsequent practice, and even cultural differences, it is not surprising that it is difficult to compare decision-making among different surgeons and in different settings.

We evaluated a coding system, the 7LC, that is based on the relevant clinical and imaging parameters of patients with DCM. As expected with any clinical decision-making scheme, we found that inter- and intra-surgeon reliability was good but not perfect. For the senior surgeons, the inter-RR was the same for both reads (κ=0.85). The intra-RR was similar for junior surgeons (κ=0.78) and senior surgeons (κ=0.71). Therefore, we do not recommend that surgeons rely only on the coding system for choosing the surgical approach. However, the value of the 7LC is based on how it lists, in a structured manner, the factors that influence this complex decision, rather than simply pointing out the only correct surgical approach. In other words, the 7LC uses a checklist to help promote the kind of mental
process that experienced senior spine surgeons might utilize. Such a system may be of benefit to their younger colleagues.

We asked whether the 7LC provides substantial inter-RR and intra-RR in cases for which a single surgical approach is most likely appropriate, and the answer was yes for both junior and senior surgeons. The greater consistency of the senior surgeons was probably due to the confluence between their mental algorithm (based on experience) and the checklist. The consistent scores of all raters for case nos. 1 and 8, which had different clinical data but the same imaging data, confirmed the importance of imaging in decision-making.

The correspondence between the approach based on the 7LC results and the surgeon’s personal choice in complex cases in which two or even three approaches might be suitable was influenced by several factors that could affect the final decision. The role of the combined approach, especially the single-staged approach, has increased during the last decade because of the availability of better implants and improved fixation devices. On the other hand, a combined approach is typically less common at the beginning of the surgical learning curve. Furthermore, it is still unclear whether a two-level corpectomy with anterior plating requires an additional posterior approach for dorsal fixation. In 2016, Hartmann et al. conducted a survey among 302 spine surgeons in central Europe. One-third of the respondents used additional dorsal fixation to increase stability following a two-level corpectomy.\textsuperscript{22}

Some biomechanical studies recommend combined ventral-dorsal support after a two-level corpectomy,\textsuperscript{23,24} but others report favorable results after three-level or even multi-level corpectomies without dorsal fixation.\textsuperscript{25,27} The 7LC includes two-level vertebral body replacement as a criterion for additional posterior fixation, but a given surgeon may not make this choice. This could explain why there is a substantial but not complete correspondence between the 7LC and the choices of the surgeons.

There are general limits for establishing surgical guidelines. Therefore, the 7LC should be considered a tool for structuring and supporting the surgical approach decision rather than a tool that indicates a single best approach. Furthermore, the individual experience of the surgeon and cultural familiarity
with one specific approach may shorten the surgical time and improve the results, which in turn reinforces the surgeon’s preference for that approach.

One limitation of the study was the different number of cases in each group, i.e. in the groups with cases that had one, two, or three possible approaches. Junior surgeons disagreed with the approach proposed by the 7LC in 18% of the complex (two-approach) cases and in 50% of the very complex (three-approach) cases. Senior surgeons disagreed with the 7LC proposal in 15% of the complex cases and in 50% of the very complex cases. Given the limited number of cases, we could not draw any statistically valid conclusions. One reason for the unequal distribution of cases was that the 7LC was designed as a tool for supporting decision-making of less experienced spine surgeons for the most common DCM cases. If we had increased the number of cases to five for each group, scoring would have taken a lot of time. A further limitation of the 7LC is that the parameters related to cervical sagittal balance and the influence of smoking on bone quality are not considered. This is because the restoration of the lordotic curve depends more on the surgical technique, and especially on the fixation, than on the surgical approach. Furthermore, the influence of smoking can not be quantified clear-cut.

Another limitation is that one cannot base a surgical strategy on a coding system that was tested on only 10 cases. However, in the unit of the first authors, the 7LC has been used during the last three years. The substantial inter- and intra-observer reliability suggests that the 7LC includes most of the factors influencing the surgeon’s surgical planning process. Despite testing being been performed by a heterogeneous international group of surgeons, the satisfactory agreement in choosing the surgical approach confirms the clinical usefulness of the 7LC.

In summary, the results of the present study showed a substantial correspondence between surgical approach decisions based on the 7LC and the personal choices of the surgeons. The differences between senior and junior surgeons were not striking. The senior surgeons showed a shorter learning curve and a somewhat higher inter-RR.

CONCLUSION
An international heterogeneous group of spine surgeons tested a decision-supporting coding system for the surgical approach for DCM. It showed substantial overall reliability, but as expected, it was not 100%. Junior surgeons went through a learning curve and converged to senior surgeons in the second read. The value of the 7LC for younger surgeons is that it helps to analyze, in a structured manner, the relevant clinical and imaging parameters influencing the choice of the surgical approach, rather than simply pointing out the only correct one.
REFERENCES


Cohen J. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 1968; 70: 213-220.


Gamer M. Irr: various coefficients of interrater reliability and agreement. R package 2005; version 0.5


FIGURE LEGENDS

Fig. 1
K-line negative compression: The osteophytes exceed the K-line (a straight line that connects the midpoints of the spinal canal at C2 and C7 on the T2-weighted sagittal MRI image). The analysis of 7 parameters, i.e. $C_{S2}S_{M2}C_{M1}$ (Compression-Spine-Morbidity-C-modifier) is used to create a seven-letter-code, for example, A-A/P-A-0-P-A-0. The most frequent letter, which was ‘A’ in this case, suggests the use of an anterior approach. A, anterior; P, posterior; C, combined; 0, not applicable; RA, rheumatoid arthritis; AS, ankylosing spondylitis; relevant osteoporosis: a history of fractures or T score worse than -2.5.

Fig. 2
A case in group 1. a Clinical data and T2-weighted MRI. b CT imaging. c Biplanar and functional X-rays. The 7LC proposal A-A/A/P-0-A/P-0-0 suggested an anterior approach (A:3 vs P:1), and 27 of 31 ratings chose this option.

Fig. 3
A case in group 2. a Clinical data and T2-weighted MRI. b CT imaging. c Biplanar and functional X-rays. The 7LC proposal A-A/P-A/A/P-A/P-0-0 showed a slight prevalence A: 2.5 vs P: 1.5 suggesting as well an anterior approach (14 ratings), as a posterior approach (15 ratings); 2 ratings favored a combined approach.

Fig. 4
A case in group 3. a Clinical data and T2-weighted MRI. b CT imaging. c Biplanar and functional X-rays. The 7LC proposal P/C-A-A/P-A/P-A/P-0-0 suggested slightly more (A:2.5 vs P:2 vs C:0.5) an anterior approach (13 ratings) than a posterior approach (9 ratings). But a combined approach, considering the instability C3/C4, was also a viable option (9 ratings).

Fig. 5
The correspondence between the approach suggested by the 7LC code and the surgical choice by each surgeon was substantial and was somewhat higher for senior surgeons than for junior ones. G1, patient group 1; G2, patient group 2; G3, patient group 3.
<table>
<thead>
<tr>
<th>C&lt;sub&gt;2&lt;/sub&gt;S&lt;sub&gt;2&lt;/sub&gt;M&lt;sub&gt;2&lt;/sub&gt;Cm&lt;sub&gt;1&lt;/sub&gt;</th>
<th>7 Letter-code</th>
<th>Appr</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compression</strong> 1 Location</td>
<td>Ant → hill type or flat 1-2L</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>→ flat ≥ 3L</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>Comb</td>
<td>P/C</td>
</tr>
<tr>
<td><strong>2 Extension</strong></td>
<td>&lt; 3</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>A/P</td>
</tr>
<tr>
<td></td>
<td>&gt; 3</td>
<td>P</td>
</tr>
<tr>
<td><strong>Spine</strong> 3 Alignment</td>
<td>Lordotic</td>
<td>A/P</td>
</tr>
<tr>
<td></td>
<td>Kyphotic → flex. or &lt; 10°</td>
<td>A/P</td>
</tr>
<tr>
<td></td>
<td>→ fixed or &gt; 10°</td>
<td>A/C</td>
</tr>
<tr>
<td><strong>4 Instability</strong> (subaxial)</td>
<td>&lt; 6 mm</td>
<td>A/P</td>
</tr>
<tr>
<td></td>
<td>&gt; 6 mm</td>
<td>C</td>
</tr>
<tr>
<td><strong>Morbidity</strong> 5</td>
<td>&lt; ASA 3</td>
<td>A/P</td>
</tr>
<tr>
<td></td>
<td>≥ ASA 3</td>
<td>P</td>
</tr>
<tr>
<td><strong>6 bone diseases</strong></td>
<td>Relevant osteoporosis, RA, AS</td>
<td>P/C</td>
</tr>
<tr>
<td><strong>C – modifier</strong> 7</td>
<td>• „K-line negative“ type of compression</td>
<td>C/0</td>
</tr>
<tr>
<td></td>
<td>• Corpectomy ≥ 2 levels</td>
<td></td>
</tr>
</tbody>
</table>
Case Group 1: F, 59 yrs
8 months ataxia + paresthesias in hands
DCM: EMS score: 15, grade I

Comorbidities:
- Stroke in the past (platelet inhibitors)
- ASA 2
- No bone diseases
Case Group 2: M, 55 yrs
DCM: EMS score: 14, grade I
Comorbidities:
- hypertension
- MI 8 years ago, stenting
- ASA 3
- No bone diseases
Case Group 3: F, 77 yrs
DCM: EMS score: 10, grade II
Comorbidities:
- Cardiovascular risks
- PNP (vit. B12 substitution)
- Diabetes mellitus (oral drugs)
- Non-relevant osteoporosis
- ASA 3
- No bone diseases