Degenerative Cervical Myelopathy: A 7-Letter Coding System That Supports Decision-Making for the Surgical Approach

Luca Papavero¹, Gregor Schmeiser¹, Ralph Kothe¹, Bronek Boszczyk², Oliver Heese³, Yoshiharu Kawaguchi⁴, Anna MacDowall⁵, Claes Olerud⁵, Nikolaos Paidakakos⁶, Anastasios Panagiotou⁶, Tobias Pitzen⁷, Marcus Richter⁸, K. Daniel Riew⁹, Aaron Stevenson¹⁰, Lee Tan⁹, Ryo Ueshima⁴, YH Yau¹⁰, Michael Mayer³

¹Clinic for Spine Surgery, Schoen Clinic Hamburg Eilbek, Hamburg, Germany
²Spine Center, Benedictus Hospital, Tutzing, Germany
³Spine Center, Schoen Clinic Munich Harlaching, München, Germany
⁴Orthopedic Surgery, University of Toyama, Toyama, Japan
⁵Orthopedic Department, Uppsala, Sweden
⁶Spine Center, SRH-Clinic, Karlsbad-Langensteinbach, Germany
⁷Spine Center, St. Josefs-Hospital, Wiesbaden, Germany
⁸Spine Center, St. Josef-Hospital, Wiesbaden, Germany
⁹University Medical Center, Spine Division, Columbia University, New York, NY, USA
¹⁰Spinal Unit, Royal Adelaide Hospital, Adelaide, Australia

Objective: To validate with a prospective study a decision-supporting coding system for the surgical approach for multilevel degenerative cervical myelopathy.

Methods: Ten cases were presented on an internet platform, including clinical and imaging data. A single-approach (G1), a choice between 2 (G2), or 3 approaches (G3) were options. Senior and junior spine surgeons analyzed 7 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 multilevel 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 7-letter code (7LC) suggested the surgical approach. Each surgeon performed 2 reads per case within 8 weeks.

Results: G1: Interrater 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 intrarater 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.

Keywords: Degenerative cervical myelopathy, Anterior approach, Posterior approach, Combined approach, Multicenter study
INTRODUCTION

Several parameters must be considered when deciding the surgical approach for patients with multilevel 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.1,2 Patient’s general comorbidities, including surgically relevant bone diseases, such as osteoporosis, rheumatoid arthritis (RA), and ankylosing spondylitis (AS), should be taken into 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.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.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 “7-letter code” (7LC) for didactical purpose because 7 parameters are analyzed (Fig. 1). Completing the checklist 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 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

1. Patient Data

The anonymized clinical and imaging data of 10 patients who underwent surgery for DCM in our clinic in the past 2 years were presented on an internet platform to 16 raters. The clinical data included age, sex, European Myelopathy Score (EMS),8 American Society of Anesthesiologists (ASA) physical status classification,9 and diseases affecting bone quality, such as relevant osteoporosis (T score < -2.5 or a history of fractures), RA, and AS. The imaging data included sagittal and axial T2-weighted magnetic resonance imaging (MRI) slices, axial and reformatted computed tomography (CT) slices, and biplanar plain films with additional flexion/extension radiographs.

2. The 7LC

The 7LC analyzed 7 imaging and clinical parameters (Fig. 1). The imaging consisted of cervical MRI, 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,
Fig. 2. A case in group 1. (A) Clinical data and T2-weighted magnetic resonance imaging. (B) CT imaging. (C) Biplanar and functional X-rays. The 7-letter code 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. DCM, degenerative cervical myelopathy; EMS, European Myelopathy Score; ASA, American Society of Anesthesiologists.

Fig. 3. A case in group 2. (A) Clinical data and T2-weighted magnetic resonance imaging. (B) CT imaging. (C) Biplanar and functional X-rays. The 7-letter code proposal A-A/A/P-A/P-0-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. DCM, degenerative cervical myelopathy; EMS, European Myelopathy Score; ASA, American Society of Anesthesiologists.

> 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 physical status classification ranging from I to V. 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 2 approaches. Bone diseases considered were: relevant osteoporosis, defined as a 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 often require an additional dorsal fixation following the anterior procedure or solely a dorsal procedure. Special circumstances such as the necessity of a 2- or multilevel 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. Sometimes 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 7-letter chain suggests the approach. If 2 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 points 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-A/P-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-A/P-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.

2. The 10 Test Cases
The 10 real cases were divided into 3 groups according to their clinical and imaging characteristics. In the first group of 5 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 5 cases had the same imaging data but different clinical data; these were intended as an “internal control.” In the second group, the 3 cases were expected to show equipoise between the 2 approaches (Fig. 3). This group was intended to highlight the individual preferences and the cultural differences of the surgeons. The third group comprised 2 complex cases that could be treated by 3 approaches, depending on the weight put on the clinical and imaging parameters (Fig. 4).

3. The Raters

We invited 16 surgeons to use the 7LC to rate the 10 cases. The 8 senior raters each had more than 20 years of experience in cervical spine surgery, and the 8 junior raters each had less than 3 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 (3 centers in Germany and 1 center each in Sweden, United Kingdom, United States, Australia, and Japan).

4. 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 8 weeks. For the second read, the cases were presented in a new random sequence.

5. Statistics

Kappa values were computed to determine the inter-RR and intra-RR. To provide an overall index of agreement, Cohen 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 3 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. All statistical analyses were performed using R ver. 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.

6. Ethics

This study was approved by the Ethics Committee of the Federal State of Hamburg (Ethikkommission Ärztekammer Hamburg) (protocol No. PV 5817). According to the deliberation of the Ethics Committee of the Federal State of Hamburg, 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.

RESULTS

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

1. 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 years 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.

2. The Second and Third Group of Cases (Balanced Distribution Between 2 or 3 Approaches)

The correspondence between the approach as determined by the 7LC and the approach as chosen by the surgeons is shown in Fig. 5.
Several studies have addressed the difficulty of deciding upon a surgical approach for multilevel cervical spondylotic myelopathy (CSM), which is now called DCM. In 1996, Naderi et al.\(^{19}\) 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, Ghogawala et al.\(^{20}\) 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.\(^{21}\) 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.\(^{22}\) 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 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 intrasurgeon reliability was good but not perfect. 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 ($\kappa = 0.71$). 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 higher 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 2 or even 3 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 2-level corpectomy with anterior plating requires an additional posterior approach for dorsal fixation. In 2016, Hartmann et al.\(^{23}\) surveyed among 302 spine surgeons in central Europe. One-third of the respondents used additional dorsal fixation to increase stability following a 2-level corpectomy. Some biomechanical studies recommend combined ventral-dorsal support after a 2-level corpectomy,\(^{24,25}\) but others report favorable results after 3-level or even multilevel corpectomies without dorsal fixation.\(^{26,27}\) The 7LC includes 2-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 1, 2, or 3 possible approaches. Junior surgeons disagreed with the approach proposed by the 7LC in 18% of the complex (2-approach) cases and 50% of the very complex (3-approach) cases. Senior surgeons disagreed with the 7LC proposal in 15% of the complex cases and 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 the decision-making of less experienced spine surgeons for the most common DCM cases. If we had increased the number of cases to 5 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 cannot 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 3 years. The substantial inter- and intraobserver reliability suggests that the 7LC includes most of the factors influencing the surgeon’s surgical planning process. Despite testing being 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 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.

CONFLICT OF INTEREST

The authors have nothing to disclose.

ACKNOWLEDGMENTS

This study was supported by a grant from the Deutsche Arthrose-Hilfe e.V. (German Arthrosis Association). Part of the manuscript content has been presented on the Annual Meeting of the German Spine Society, Wiesbaden December 6-7, 2018. The authors thank Florin Sburlea Andoni, MD, for completing the first read of the cases.

REFERENCES

1. Epstein JA. The surgical management of cervical spinal stenosis, spondylisis, and myeloradiculopathy by means of the posterior approach. Spine (Phila Pa 1976) 1988;13:864-9.
2. Nouri A, Martin AR, Nater A, et al. Influence of magnetic resonance imaging features on surgical decision-making in degenerative cervical myelopathy: results from a global survey of AOSpine International Members. World Neurosurg 2017;105:864-74.
3. Irwin ZN, Hilibrand A, Gustavel M, et al. Variation in surgical decision making for degenerative spinal disorders. Part II: cervical spine. Spine (Phila Pa 1976) 2005;30:2214-9.
4. Lawrence BD, Jacobs WB, Norvell DC, et al. Anterior versus posterior approach for treatment of cervical spondylotic myelopathy: a systematic review. Spine (Phila Pa 1976) 2013;38(22 Suppl 1):S173-82.
5. Fehlings MG, Barry S, Kopjar B, et al. Anterior versus posterior surgical approaches to treat cervical spondylotic myelopathy: outcomes of the prospective multicenter AOSpine North America CSM study in 264 patients. Spine (Phila Pa 1976) 2013;38:2247-52.
6. Fehlings MG, Jha NK, Hewson SM, et al. Is surgery for cervical spondylotic myelopathy cost-effective? A cost-utility analysis based on data from the AOSpine North America prospective CSM study. J Neurosurg Spine 2012;17(1 Suppl):89-93.
7. Kato S, Nouri A, Wu D, et al. Comparison of anterior and posterior surgery for degenerative cervical myelopathy: an MRI-based propensity-score-matched analysis using data from the prospective multicenter AOSpine CSM North America and International Studies. J Bone Joint Surg Am 2017;99:1013-21.

8. Herdmann J, Linzbach M, Krzan M, et al. The European myelopathy score. In: Baucher BL, Brock M, Klinger M, editors. Advances in neurosurgery. Berlin: Springer; 1994. p. 266-8.

9. Hackett NJ, De Oliveira GS, Jain UK, et al. ASA class is a reliable independent predictor of medical complications and mortality following surgery. Int J Surg 2015;18:184-90.

10. Luo J, Cao K, Huang S, et al. Comparison of anterior approach versus posterior approach for the treatment of multilevel cervical spondylotic myelopathy. Eur Spine J 2015;24:1621-30.

11. Leven D, Cho SK. Pseudarthrosis of the cervical spine: risk factors, diagnosis and management. Asian Spine J 2016;10:776-86.

12. Fujiyoshi T, Yamazaki M, Kawabe J, et al. A new concept for making decisions regarding the surgical approach for cervical ossification of the posterior longitudinal ligament: the K-line. Spine (Phila Pa 1976) 2008;33:E990-3.

13. Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 1968;70:213-20.

14. Finn RH. A note on estimating the reliability of categorical data. Educ Psychol Meas 1970;30:71-6.

15. Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bull 1971;76:365-77.

16. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8:23-34.

17. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.

18. Gamer M. Irr: various coefficients of interrater reliability and agreement. R package ver. 0.5. Vienna (Austria); R Foundation for Statistical Computing; 2005.

19. Naderi S, Benzel EC, Baldwin NG. Cervical spondylotic myelopathy: surgical decision making. Neurosurg Focus 1996;1:e1.

20. Ghogawala Z, Martin B, Benzel EC, et al. Comparative effectiveness of ventral vs dorsal surgery for cervical spondylotic myelopathy. Neurosurgery 2011;68:622-30.

21. Baksheshian J, Mehta VA, Liu JC. Current diagnosis and management of cervical spondylotic myelopathy. Global Spine J 2017;7:572-86.

22. Hartmann S, Tschugg A, Obernauer J, et al. Cervical corpectomies: results of a survey and review of the literature on diagnosis, indications, and surgical technique. Acta Neurochir (Wien) 2016;158:1859-67.

23. Aryan HE, Sanchez-Mejia RO, Ben-Haim S, et al. Successful treatment of cervical myelopathy with minimal morbidity by circumferential decompression and fusion. Eur Spine J 2007;16:1401-9.

24. Sevki K, Mehmet T, Ufuk T, et al. Results of surgical treatment for degenerative cervical myelopathy: anterior cervical corpectomy and stabilization. Spine (Phila Pa 1976) 2004;29:2493-500.

25. Uribe JS, Sangala JR, Duckworth EA, et al. Comparison between anterior cervical discectomy fusion and cervical corpectomy fusion using titanium cages for reconstruction: analysis of outcome and long-term follow-up. Eur Spine J 2009;18:654-62.

26. Singh K, Vaccaro AR, Kim J, et al. Biomechanical comparison of cervical spine reconstructive techniques after a multilevel corpectomy of the cervical spine. Spine (Phila Pa 1976) 2003;28:2352-8.

27. Koller H, Hempfing A, Ferraris L, et al. 4- and 5-level anterior fusions of the cervical spine: review of literature and clinical results. Eur Spine J 2007;16:2055-71.