Author’s response to reviews

Title: Comparison of The French and CARDS Classifications for Lumbar Degenerative Spondylolisthesis: Reliability and Validity

Authors:

Chao Kong (353397342@qq.com)
Xiangyao Sun (sun.xiang.yao@163.com)
Junzhe Ding (18310693576@qq.com)
Machao Guo (guomachao23@gmail.com)
Xiangyu Li (lxyyxl@sina.com)
Shibao Lu (spinelu@163.com)

Version: 1 Date: 11 Jun 2019

Author’s response to reviews:

Dear editor and reviewers,

Thank you very much for your questions and comments. The questions are responded as following:

Editor comments:

1. Thank you very much for your suggestion. The reference number has been added in the section “Ethics approval and consent to participate”

2. Thank you very much for your suggestion. There are no names or identifiable information in the figures.

Reviewer 1:

1. Thank you very much for your question and suggestions. Reliability refers to consistency of measurement or repeatability. Reproducibility is the closeness of the agreement between the results of measurements. In this article, as you have suggested, reliability is enough. We have revised all sentences that used reproducibility. Thank you again for your suggestion.
2. Thank you very much for your suggestion. One of the major objectives of classification system is to guide surgical treatment. In this study, we wanted to explore whether the two classification systems were associated with preoperative or postoperative pain scores, function and health-related quality of life. In other words, the validity of the two classification systems were compared. As you suggested, we have changed “clinical utility” to “validity” throughout the manuscript. Thank you again for your excellent suggestion.

3. Thank you very much for your suggestion. The minimum clinically important difference (MCID) was first defined as a thresholding value in post-treatment change, and a patient is considered experiencing a clinically meaningful improvement if her/his change exceeds the MCID [1]. Based on receiver operating characteristic (ROC) curve, a few approaches were proposed for its estimation [2-4]. However, different calculation methods may lead to significant variation of MCIDs, and little statistical work has been done for appropriately determining MCID [5]. Some of studies have assessed MCID thresholds in patients undergoing lumbar disc herniation intervention (surgery) but with variable results. Solberg et al. [6] reported MCID thresholds of 20 for ODI, 2.5 for numerical pain scale (NRS) back, and 3.5 for NRS leg and 0.3 for Euroqol (EQ-5D). Another report by Copay et al. [7] had determined lower estimates of 13 points for the ODI, 1.2 points for NRS back pain scale, and 1.6 points for NRS leg pain scale. In this study, the aim of the comparison was to determine whether different classification had different postoperative improvement, rather than to find a cut-off value. So we think the Kruskal–Wallis test is enough to distinguish the difference in this study. However, we find your suggestion very helpful in our future work and thank you again for your valuable suggestion.

4. Thank you for your suggestions. The cases in agreement between the first and second observation have been listed in table 2 and table 3, by which the agreement proportions of different types and different graders can be calculated.

5. Thank you very much for your question and suggestion. We chose 3 weeks’ interval because it was suggested by our statistician. But as you have mentioned, 3 weeks’ interval may be not enough and should be added in a limitation. In this study, Lu SB provided the tutorial to all graders. The tutorial was just like an introduction of the two classification systems to make sure that every grader knew how to grade. During clinical practice, each grader should be familiar with the classification system before grading. But as you mentioned, the tutorial may increase the interobserver agreement and should also be added as a limitation. Thank you very much for your suggestions.

6. Thank you very much for your suggestions. The figures were edited to remove identifiable information which may cause decrease of quality. The two figures were used to show the two classification systems respectively. We have added annotations in the figures.

7. Thank you very much for your suggestion. We have used the term “validity” rather than “clinical utility” as you suggested.
Reviewer 2

1. Thank you very much for your suggestion. Though a variety of surgical approaches have been advocated, there is no consensus among spinal surgeons regarding optimal surgical treatments for patients with lumbar DS. The Meyerding classification, though widely used for describing the degree of spondylolisthesis, is less useful for lumbar DS since almost all cases of lumbar DS would fall into grade I or II category. Also, the Meyerding classification does not take other morphologic parameters such as segmental kyphosis or disc height into consideration, which are related to clinical outcomes. The variability of radiographic features suggests that lumbar DS is a heterogeneous disease and requires a specific grading system.

2. Thank you very much for your question. The final classification was decided by a majority of graders (3/4 or 4/4). If two graders were different with the other two, another grader (Lu SB) will determine the final classification.

3. Thank you very much for your question. The clinical variable adopted in the CARDS classification is leg pain. There are three modifiers in the CARDS classification----modifier 0: patients without leg pain; modifier 1: patients with unilateral leg pain; modifier 2: patients with bilateral leg pain. We have changed the place where “Figure 2” is annotated.

4. Thank you very much for your question. As you mentioned, Table 4 and Table 7 do not contribute to a reliability analysis. Table 4 shows demographics and preoperative clinical scores of patients with different CARDS types. Table 7 shows mean improvements in outcome scores according to French type. The four tables (from Table 4 to Table 7) contribute to the validity of the two classification systems. We have added a subheading (Validity Analysis) in this part.

5. Thank you for your suggestion. We have revised the limitation part as you suggested.

6. We have revised the figures as you suggested.

Thank you again for your time and patience!

Lu Shibao

References:


