# Accepted Manuscript


Title: Quality assessment of systematic reviews for surgical treatment of low back pain: an overview

Author: Delio Eulalio Martins, Nelson Astur, Michel Kanas, Mário Ferretti, Mario Lenza, Marcelo Wajchenberg

PII: S1529-9430(16)00298-9
DOI: http://dx.doi.org/doi: 10.1016/j.spinee.2016.01.185
Reference: SPINEE 56849

To appear in: *The Spine Journal*

Received date: 21-4-2015
Revised date: 16-10-2015
Accepted date: 15-1-2016

Please cite this article as: Delio Eulalio Martins, Nelson Astur, Michel Kanas, Mário Ferretti, Mario Lenza, Marcelo Wajchenberg, Quality assessment of systematic reviews for surgical treatment of low back pain: an overview, *The Spine Journal* (2016), http://dx.doi.org/doi: 10.1016/j.spinee.2016.01.185.

# QUALITY ASSESSMENT OF SYSTEMATIC REVIEWS FOR SURGICAL TREATMENT OF LOW BACK PAIN: AN OVERVIEW

Delio Eulalio Martins (eulalio@me.com) ; Nelson Astur (nelsonan@yahoo.com) ; Michel Kanas (michelkanas@gmail.com) ; Mário Ferretti (Mario.Ferretti@einstein.br) ; Mario Lenza (mario.lenza@einstein.br) ; Marcelo Wajchenberg (marcelo.wajchenberg@einstein.br)

Hospital Israelita Albert Einstein
Av Albert Einstein, 627, office 306, A1 building, 05652-900, Sao Paulo, SP, Brazil

Corresponding author: Delio Eulalio Martins, Av Albert Einstein, 627, office 306, A1 building, 05652-900, Sao Paulo, SP, Brazil phone and fax: #551121519393, eulalio@me.com



**ABSTRACT**

**Background Context**: Low back pain is among the most frequent causes for medical appointments. Surgical treatment is widely controversial and new surgical techniques and treatment modalities have been developed within the last decade. Treatment for Low Back Pain should be evidence-based through Systematic Reviews and Meta Analysis. Thus, the quality of these reviews is sometimes put into question since methodological mistakes are frequently seen.

**Purpose**: The aim of this study is to gather all Systematic Reviews for the surgical treatment of Low Back Pain and analyze their outcomes, quality and conclusion.

**Study Design/Setting:** An overview of Systematic Reviews

**Outcome Measures**: AMSTAR score, PRISMA statement, conclusion supported by descriptive statistics.

**Methods**: A literature search for Systematic Reviews containing Low Back Pain surgical treatment was conducted through different medical databases. Two investigators independently assessed all titles and abstracts for inclusion. Studies should have at least one surgical procedure as an intervention. Diagnoses were categorized as lumbar disc herniation, spondylolisthesis, stenosis, facet joint syndrome, and degenerative disc disease. Quality was assessed through the PRISMA and AMSTAR questionnaires. Study quality related to its PRISMA and/or AMSTAR score percentage was rated as: very poor (<30%), poor (30-50%), fair (50-70%), good (70-90%), and excellent (>90%). Articles were considered conclusive if they had a conclusion for their primary outcome supported by descriptive statistical evidence. This study was funded exclusively by the authors' own resources. None of the authors have any potential conflict of interest to declare.

**Results**: Overall, there were 40 systematic reviews included. According to AMSTAR and PRISMA scores, 5-7.5% of the Systematic Reviews were rated as excellent and most of them were rated as a fair review. AMSTAR indicated that 22.5% of the reviews have very poor quality, while PRISMA stated 7.5% being of very poor quality. For both tools, performing a Meta Analysis made the reviews' quality significantly better. The best-rated diagnosis groups according to PRISMA

were Spondylosis, Lumbar Disc Herniation and Degenerative Disc Disease. Considering the studies' conclusions, 25 (62.5%) out of the 40 Systematic Reviews had a conclusion to their primary outcome and only 11 (27.5%) were supported by descriptive statistical analysis. This means that 44% of the Systematic Reviews with a conclusion were evidence based. There were 15 (37.5%) SRs that did not reach a conclusion to their primary objectives.

**Conclusions**: In conclusion, most SRs for LBP do not reach very good or excellent quality and only 27.5% of them have evidence-based conclusions. Including a meta-analysis is a significant factor to improve quality and evidence for SRs.

**Keywords:** low back pain; systematic review; surgical procedures; outcome assessment; lumbar spine; degenerative disc disease



## INTRODUCTION

The impact of low back pain in healthcare is a major concern as costs are increasing every year and are significantly related to surgical treatment, time to return to work, and work compensation[1,2]. Surgical treatment for LBP is widely controversial and new surgical techniques and treatment modalities have been developed within the last decade.

For best scientific evidence, outcomes for new and standard interventions should be analyzed through randomized clinical trials, [3–5]. However, a majority of treatment modalities for LBP goes through a series of trials with different outcomes amongst themselves that could be influenced by methodology, population or even conflicts with authors' interests or research funding.[6–9] When this happens, the best evidence for treatment outcomes is derived from systematic reviews (SR) with a meta-analysis (MA) of those trials. Unfortunately, there is often a great divide between those outcomes and clinical practice.

Treatment for LBP should be evidence-based through SR and MA. Systematic reviews for LBP are widely available in any medical database but they might lead to different conclusions for the same intervention in the same population. Thus, the quality of these reviews is sometimes put into question since methodological mistakes are frequently seen[10]. Therefore, SR may not be so highly evidenced to guide surgical treatment for the most common degenerative lumbar spine diseases.

The aim of this study is to gather all SR for the surgical treatment of LBP and analyze their outcomes, quality and confidence.

## MATERIALS AND METHODS

*Search strategy*

Institutional review board approval (number 1942-14) was obtained. A literature search for systematic reviews that involves only LBP treatment was conducted up to January 2014 through different medical databases: Medline

(Pubmed); EMBASE (Ovid); Cochrane database of systematic reviews (CDSR); the Database of Abstracts of Reviews of Effectiveness (DARE); and International Prospective Register of Systematic Reviews (Prospero). No restriction to language or date was applied. To minimize risk of missing relevant reviews, a handsearch of the reference lists of reviews captured by the initial search was performed as well. The search strategy used for Medline is shown on Appendix 1. Other databases followed the same search strategy with minimal adjustments. Two investigators (DM and NA) independently assessed all titles and abstracts to exclude duplicate articles and select potential articles to be included while inconsistencies were resolved through discussion with a third author (ML). When more than one SR with the same or similar interventions from the same author were found, only the most current one was included and was considered an update of previous work.

*Study eligibility criteria*

After a list of studies was gathered from all database searches, systematic reviews of studies involving patients of all ages were included when the study discussed the following diseases: lumbar disc herniation, spondylolisthesis, spinal canal stenosis, facet joint syndrome and degenerative disc diseases. SRs should have at least one surgical procedure as an intervention such as injections of any kind, fusion, or decompression such as laminectomy or discectomy. Either open or minimally invasive techniques were considered. Systematic reviews comparing two non-surgical treatments were excluded as well as those involving cervical or thoracic spine degenerative diseases.

*Data extraction*

Data were extracted independently by four reviewers (DM, NA, MK and MW) that are board certified in spine surgery with a standardized form. To minimize evaluation bias, all reviewers were primarily trained for each item of both questionnaires by one of the senior authors (ML) with expertise in systematic reviews. All four reviewers assessed the first 5 papers together so there would be homogeneity on the interpretation of data. The reviewers assessed the following 35 papers independently. Any disagreement that might have arisen was discussed and resolved by consensus and with an opinion of a fifth reviewer (ML) with expertise in systematic reviews. The following items were included in our form and collected for every SR: study design, searched databases, last date of search strategy, presence of a protocol before conduction of the study, funding sources, number of studies included, number of patients assigned, number of patients assessed at the end of the study, inclusion and exclusion criteria, age of participants, diagnoses enrolled, interventions, primary and secondary outcomes, timing of outcome measures, and presence of positive conclusions.

*Study quality analysis*

Quality of the studies included in the current analysis was assessed through the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)[11] and AMSTAR[12] questionnaires by the same four reviewers. Both forms are validated measurement tools to assess the methodological quality of systematic reviews (Appendices 2 and 3). Each item of the PRISMA form was graded as *yes*, *incomplete* or *no* and respectively scored as

1, 0.5, or 0 points for statistical analysis purposes. Similarly, the AMSTAR tool had each item graded as *yes* or *no* and scored as 1 or 0, respectively. The sum of all items scored for each questionnaire was divided by its maximum possible score to assess study quality as a percentage. Study quality related to its PRISMA and/or AMSTAR score percentage was rated as: very poor (<30%), poor (30-50%), fair (50-70%), good (70-90%), and excellent (>90%).

*Diseases, interventions and outcomes analysis*

For each SR, data were extracted for population, intervention, as well as primary and secondary outcomes. Papers were grouped according to the disease investigated (Degenerative Disk Disease (DDD), Spondylolisthesis (SL), Lumbar Stenosis (LS), Lumbar Disk Herniation (LDH), Spondylosis (S), and Facet Joint Syndrome (FJS)) and compared. Interventions and outcomes were also compared independently. Outcomes considered were: Oswestry Disability Index (ODI)[13], Roland Morris score[14], Short-Form 36 (SF-36)[15], visual analogic scale (VAS), estimated blood loss (EBL), fusion rate, infection, operation time, complications, reoperation, and return to work. Every article was reviewed by two authors to analyze the validity of the conclusion. We considered conclusive SRs those with a conclusion for their respective primary outcome. If articles were conclusive, they were further assessed for description of supportive statistical evidence for this conclusion. Those that met all these criteria were considered as a SR with an evidenced based conclusion.

*Statistical analysis*

Statistical analysis consisted of descriptive statistics including relative and absolute frequencies. In addition, Fisher exact tests were performed to compare categorical variables, and the Mann-Whitney test to compare numeric variables. For all tests, a *P* value level less than 0.05 was considered significant. All the tests have been performed using the software R 3.0.3.

## RESULTS

After a full electronic search, a total of 851 references were identified (Figure 1). A title and abstract screening excluded 755 references and 49 went through a detailed eligibility process in which the full-text article was analyzed. Nine studies were excluded with reasons (Figure 1)[16–24]. Overall, there were 40 systematic reviews included[6–9,25–60]. We also accessed six ongoing studies[61–66] and five studies awaiting classification[67–71] due to non-English written language that were not included.

### *Descriptive analysis*

Out of the 40 included SR of surgical treatment for LBP, half of them were published within the last four years[6,8,9,25,28,30,33,34,36,40,48–51,54,56,58,72–74] with the oldest study being from 1992[43]. The SR with the greatest number of included references had 74citations[43] while there was one SR with a single included study[74] (mean 17.7 studies/SR). More than half of the SR (58%) had no research funding. The majority (57.5%) of the SRs did not include a MA in their statistical analysis. When the diagnosis investigated was considered, there were 13 SRs for degenerative disk disease, 11 for spondylosis,

six for spondylolisthesis, five for lumbar stenosis, three for lumbar disk herniation, and two for facet joint syndrome.

There was a very heterogeneous group of surgical interventions, including injections, direct repair of the pars, arthroplasty, decompression, nucleoplasty, endoscopic discectomy, "surgery" and fusion. There were different fusion techniques studied by the SRs, including posterolateral fusion, posterior lumbar interbody fusion, transforaminal lumbar interbody fusion, and anterior lumbar interbody fusion.

When outcomes were analyzed, SRs that performed a MA had most significant outcomes as well as the major number of items analyzed (Table 1). This comparison between outcomes from SR with a MA and without MA was significant for blood loss, complications, fusion, ODI, operation time, and reoperation, meaning that SR with a MA had more significant results.

*Descriptive study quality analysis*

Percentage measures for AMSTAR, PRISMA and PRISMA domains are shown in Table 2. To assess viability of our metrics tools for SRs quality, we have conducted a comparison for correlation between AMSTAR and PRISMA through a Spearman test resulting in a high level correlation coefficient (0.81; $p<0.001$). Final AMSTAR score according to percentage indicates that most frequent rate for LBP SR is in between 50 to 70%, which is a fair review. Only 7.5% of the reviews are rated as excellent. 22.5% of the SRs are considered as having a very poor quality. PRISMA had less reviews at the extremes of the rating measures, with only 7.5% being very poor and 5% excellent. Most of the reviews were

rated as fair (37.5%). When PRISMA domains are analyzed, most flaws were found in the abstract, methods and results sections, and these reviews tend to be rated as very poor or poor (80% for abstract, 55% for methods, and 60% for results). Best rates (good and excellent) are in title, introduction, discussion, and funding domains (87.5%, 82.5%, 57.5%, and 57.5%, respectively).

Systematic reviews that performed a meta-analysis (n=17) were compared to those without a meta-analysis (n=23) according to their AMSTAR and PRISMA scores. For both tools, performing a MA made the reviews' quality significantly better (Figure 2). According to AMSTAR, 47% of the SR with MA was good or excellent, while 21.7% of those without MA were only good. That was not statistically significant though ($p$=0.089). However, PRISMA had a statistically ($p$=0.008) higher rate of good and excellent reviews for those where a MA was performed (47.1%, versus 8.7% of those without MA). Three out of the seven PRISMA domains were significantly better for SR with MA, they were methods, results, and discussion (Mann-Whitney test $p$<0.001 for all three).

*Quality analysis according to disease studied*

Degenerative disk disease was the most studied group with 13 SR, followed by Spondylosis with 11. There were six SR for Spondylolisthesis, five for Lumbar Stenosis, three for Lumbar Disc Herniation, and two for Facet Joint Syndrome. The best-rated diagnosis groups according to PRISMA were Spondylosis, Lumbar disk herniation and Degenerative disk disease with 36.3%, 33.3%, and 30.7%, respectively, for good and excellent SR (Table 3). Lumbar stenosis and Facet joint disease had no SR scored over 70% according to PRISMA. Furthermore, most of the papers, not considering diagnosis group, were

rated as very poor, poor, or fair. When the AMSTAR tool was performed, results were apparently better, although not statistically significant. Half of the papers within Spondylolisthesis or Facet Joint Syndrome groups were classified as good or excellent, while the group with the greater number of SR rated over 70% by PRISMA (Spondylosis) had the worst performance with AMSTAR (18.8% of good or excellent papers). When quality of the SR was compared between diagnosis groups, there was no statistical difference between either of the questionnaires' scores according to Fisher exact test. Nevertheless, SR for LDH[39,53,59] had the smallest score for both AMSTAR and PRISMA.

*Intervention and outcomes analysis*

There were a wide variety of interventions. Surgery was the intervention for all SRs since it is an inclusion criterion for this study. Specific techniques were composed of Arthroplasty (6 SRs), Fusion (10SRs), Endoscopy (1 SR), DeKompressor®(1 SR), Injections (4 SR), Instrumentation (1 SR), Minimally Invasive Interbody Fusion (2 SRs), Nucleoplasty (4 SRs), Posterolateral Fusion (4 SRs), and Osteosynthesis (1 SR). Clinical outcomes were mostly measured by the ODI (assessed in 20 SRs), followed by the VAS pain score in 14 SRs. There were five SRs analyzing SF-36 and only one used the Roland Morris score. Surgical outcomes were most described as spine fusion (8 SRs), followed by EBL and operation time (5 SRs each). Postoperative outcomes were measured by complications (12 SRs), return to work (7 SRs), and reoperation rate (7 SRs).

Considering the studies' conclusions, 25 (62.5%) out of the 40 SRs had a conclusion to their primary outcome[6–9,27,33,34,36,38,39,42–44,48–51,53–56,58,59,72,73] and only 11 (27.5%) were supported by descriptive statistical

analysis[6–9,39,48–50,56,58,73]. This means that 44% of the SRs with a conclusion were evidence based. There were 15 (37.5%) SRs that did not reach a conclusion to their primary objectives [25,26,28–31,35,37,40,45,52,57,60,75,76] and the most frequent reason for this was the lack of available randomized controlled trials. If we calculate the average PRISMA and AMSTAR scores for these 11 evidence based conclusive SRs, we see 76.6 and 75.21 respectively, meaning that those were very good studies. If we apply the same calculus for the 15 SRs without a conclusion, average PRISMA and AMSTAR were respectively 47.28 and 41.82 meaning fair quality studies.

Among the 11 conclusive SRs, there were four studies for DDD, four for S, one for LDH, one for LS, and one for SL. Neither of the SRs for FJS had evidence based conclusion.

**DISCUSSION**

The pursuit for scientific evidence is a challenge in clinical medical. Treatments and clinical decisions should be based on the best evidence possible but this is often not a reality due to numerous obstacles to achieve this goal. Surgical treatment is often under-supported by the medical literature due to heterogeneity of patients, surgical techniques, implant differences, surgeon skill, and evolving new technology. The presence of statistical heterogeneity poses a challenge to conduct valuable outcomes studies with reliable results that would change clinical practice. Best evidence should come from systematic reviews followed by meta-analysis of the data, which would ideally support all medical decisions. An overload of systematic reviews has been published within the last

decades, with still increasing numbers on the way[77]. In our series of 40 SRs, half of them were published within the last four years. Systematic reviews need frequent updates as well since well-designed randomized clinical trials published after a SR search was conducted may significantly change results. Randomized clinical trials of the same intervention and population may disagree and SR should analyze and compare trials for definitive evidence. Although many SRs for LBP surgical treatment are available, there is still no strong evidence favoring most of surgical procedures from an evidence-based approach[32,36,74,78]. Based on all literature available, spine surgeons still cannot answer most of patients' questions regarding the best treatment for specific spine diagnoses. Searching for answers, surgeons should not blindly trust a SR since the validity of a significant number of them is questioned. Few points should always be questioned before taking SR into practice: search conduction method; bias in the study selection; current/updated SR; quality of the studies assessed; evidence combined and summarized appropriately; publication bias; and, justified conclusions[79].

The main objective of this study was to assess quality of systematic reviews for surgical treatment of low back pain. After 40 selected systematic reviews were analyzed through the AMSTAR and PRISMA questionnaires, just a few of them (7.5% and 5% respectively) were classified as excellent (scores >90%). However, they were not the same according to each assessment tool. PRISMA rated the works by Yajun et al[7] and Zhou et al[49] as excellent while AMSTAR rated excellent the works by Wang et al[6], Gibson et al[39], and Jacobs et al[8]. All these papers performed a meta-analysis of the data, which makes it a paramount condition for a SR to be considered as excellent. Interestingly, three

out of the five top rated SRs had the same patient-intervention combination although their results may disagree[6–8]. Not all papers elucidate their conflicts of interest, which may lead to risk of bias. Those were all SRs published between 2010-13 comparing fusion versus total disc replacement for the degenerative disc disease of the lumbar spine. Wang concludes that there is evidence that the risk of clinical adjacent segment pathology following fusion is higher when compared with total disc replacement. On the other hand, Yajun states that disk replacement does not show significant superiority for the treatment of lumbar DDD compared with fusion. Jacobs concludes that although there is no clinical relevance, total disc replacement seems to be effective in treating LBP in selected patients and is equivalent to fusion. A fourth SR rated as fair[9] significantly favored total disc replacement for its safety and efficacy. Additionally, some flaws in the SRs may be present but are not always detected by this study metrics. For instance, although Yajuns' SR had a high score for both tools, a subtle flaw was not detect, since two trials[80,81] from the same study group including the same subset of patients were included and considered different papers. Even though there are already three highly qualified SRs and another three lower scored SRs[9,57,75] on the same topic, this is still a highly debated topic in every spine surgery meeting and we do not seem to be getting closer to an answer. Following the conclusions above, most SRs report the need for further evidence and more randomized clinical trials. In addition to conflict of interest and industry role in research, caution should be taken on the interpretation of industry supported or funded research while inappropriate influence of funders is often regarded as an important risk of bias and it is not unusual that the sponsor either owns the data or needs to approve the

manuscript. This is well stated in the *Cochrane Handbook for Systematic Reviews*[82]

Conducting an excellent SR is not an easy task, and sometimes is just not feasible since after applying a well-designed search strategy and study selection, authors are frequently facing insufficient data for a meta-analysis. In order to keep up with the proposed protocol, observational studies are included and sometimes quality decreases. However, Furlan et al[83] stated that observational, non-randomized studies frequently agree with randomized controlled trials for low back pain. Despite all challenges, majority of the SRs for surgical treatment of LBP were considered as "good" in the present study. Those good reviews usually have high scores for domains like title, introduction, and discussion but their biggest flaws are in the abstract, methods, and results. According to PRISMA, this is where most studies loose their quality and authors do not follow ideal patterns to conduct a SR. A significant factor for better evidence in SR is applying MA to results. Studies with MA had significantly higher scores than those without an MA according to both AMSTAR and PRISMA. This difference was more evident through the PRISMA tool where 47% of the SR with a MA was rated as good or excellent against only 8.7% of those without a MA ($p<0.015$). Many SRs for LBP achieve a final conclusion and draw benefits of a technique over another. However, readers that directly access the conclusion of the paper should be careful, for a significant number of SRs analyzed draw conclusions without achieving evidence. Even with fair or good quality scores, some SRs conclude facts that are not supported by their statistical analysis[32,35,36,38,40–42,44,54] or do not even perform a meta-analysis. For this reason, SRs should be carefully read throughout the entire manuscript

before assumptions are taken. For example, in a SR that favors nucleoplasty for contained disc herniation in its conclusion[54], there is no evidence or adequate statistical analysis in its methodology and results to support conclusion. No MA was performed as well.

Analyzing diagnosis groups is challenging since definition of diagnosis is not always clear. If we only take SRs that specifically analyzed LDH[39,53,59], those had the lower mean score for both PRISMA and AMSTAR. This is not significant since one the SRs in this group had the lowest score among all reviews for PRISMA and significantly decreased the average in a group of three. Grouping three SRs for LDH and 13 for DDD may sound conflicting since LDH and DDD sometimes mean the same disease. That may also be true for spondylosis, DDD, and sometimes, even lumbar stenosis. Diseases studied in SRs for low back pain are not easy to define, and trials may gather different pain generators turning the population even more heterogeneous. That happens due to the clinical difficulty to define the source of pain in a patient, being all part of a natural degenerative process of the spine. Two SRs analyzed facet joint syndrome[38,60] and their average rate was fair reviews. Despite one of them[60] was rated as good by the AMSTAR tool, its funding and conflicts of interest of the authors are not very clear, which is extremely important when a review of a commercial product is being made (radiofrequency). The same happens in a SR of a specific percutaneous discectomy device available by the spine industry[53] were funding and conflicts of interest are also questionable. Spondylolisthesis may be the most non-conflicting diagnosis due to its clear imaging findings. There were six SRs[28,33,34,37,50,52] for this topic with average fair ratings. The most studied outcome was fusion rate for

spondylolisthesis and different surgical techniques were compared. An earlier study by Jacobs et al[37] in 2006 compared posterolateral fusion with other interbody fusion techniques and concluded that posterolateral fusion appeared to be the general gold standard for the treatment of adult isthmic spondylolisthesis although there was no scientific evidence. About seven years later, two SRs[33,50] rated as good comparing posterolateral fusion and posterior lumbar interbody fusion for spondylolisthesis, reverted this conclusion favoring interbody fusion, with significant evidence for higher fusion rates in this group. The same intervention-outcome was significantly concluded by Umeta and Avanzi[48], although for a population classified as spondylosis (which included cases with spondylolisthesis).

Although there is a trend toward minimally invasive spine surgery (MISS) techniques, evidence is still weak. There were seven SRs concerning minimally invasive techniques, represented by nucleoplasty[42,44,54,60], MISS transforaminal lumbar interbody fusion[51,76], and endoscopy techniques[59] and none of them had significant or evidence-based conclusions. Those were average, fair-quality reviews but their biggest flaw was on study selection, the lack of good randomized controlled trials for these techniques made a meta-analysis of the data impossible. This is probably due to the late onset of these techniques by spine surgeons and no good clinical trials had been performed yet.

The motivation to perform this study was the impression that most SRs for LBP surgical treatment were lacking concrete conclusions and that seemed to be a huge obstacle for SRs to have an impact on surgeons clinical routine. Our analysis proved that point when only 27.5% of the SRs for LBP surgical treatment had a conclusion for their primary objective supported by statistical

significance. Furthermore, 66% of the SRs relating a conclusion to their objectives had no statistics to support it. If most of the SRs that could not even reach a conclusion to their primary objectives put the blame on the lack of good randomized clinical trials, why should we keep doing SRs if our flaw is still on the primary studies? Although a good number of SRs reaches good quality scores, many of them cannot reach a conclusion due to the lack of primary studies in the literature. An obvious increase in quality scores was observed in the few SRs with a conclusion based on significant statistics while those that did not determine a conclusion had the lower scores. One of the main limitations of this study was the fact that the quality tools for SRs were independently applied by four different reviewers and inter and intra-reliability of both questionnaires was not assessed. However, we have conducted a comparison for correlation between AMSTAR and PRISMA through a Spearman test resulting in a high level correlation coefficient (0.81; $p<0.001$). Furthermore, previous authors have successfully used PRISMA and AMSTAR for quality assessment of SRs of other topics.[84–87] Another limitation is the extremely population-intervention heterogeneity of reviews which limited statistical analysis of outcomes between SRs. The unclear definition of disease group by SRs authors may also have affected analysis. An important limitation was the usage of non-validated tools to determine study quality although there is none available in the literature for this purpose. Future investigation of SRs for LBP is mandatory and must be constant. While evidence is paramount to create treatment guidelines for LBP surgery, quality analysis of the evidence provided by SRs should be assessed before changing routine treatment decisions. Alongside clinical evidence, there should be noted that “absence of evidence is not evidence of absence” and individual

interpretation should be applied to all trials and systematic reviews. Challenges are always related to producing good quality randomized clinical trials, but those are time consuming and bonded to high budgets. While it is easy to criticize the lack of prospective, randomized studies, and those that are industry funded, particularly the rigorously executed Governmental-regulated trials, funding is almost an indispensable condition for a large well-performed randomized clinical trial. In our opinion, we should not ban funding, but prioritize well-design studies proposals following strict methodological validated standards even before the trial starts. Protocols should be registered and published, before started, in an international peer-reviewed periodic. A clear well-performed methodology would minimize funding bias.

In conclusion, most systematic reviews for low back pain do not reach very good or excellent quality and only 27.5% of them have evidence-based conclusions. Therefore, authors suggest that researchers should concentrate efforts in performing randomized clinical trials in surgical treatment for low back pain before attempting secondary studies. Including a meta-analysis is a significant factor to improve quality and evidence for systematic reviews.

**Acknowledgments**: Thanks to Elivane Victor and Ana Carolina Mafra for the technical support with statistics. Thanks to Aaron T. Creek, MD for his support in English Review.

**REFERENCES**

[1] Gou HR, Tanaka S, Halperin WE, Cameron LL. Back pain prevalence in US industry and estimates of lost workdays. Am J Public Health 1999;89:1029–35.

[2] Frymoyer JW, Cats-Baril WL. An overview of the incidences and costs of low back pain. Orthop Clin North Am 1991;22:263–71.

[3] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332.

[4] Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.

[5] Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609–13.

[6] Wang JC, Arnold PM, Hermsmeyer JT, Norvell DC. Do lumbar motion preserving devices reduce the risk of adjacent segment pathology compared with fusion surgery? A systematic review. Spine (Phila Pa 1976) 2012;37:S133–43.

[7] Yajun W, Yue Z, Xiuxin H, Cui C. A meta-analysis of artificial total disc replacement versus fusion for lumbar degenerative disc disease. Eur Spine J 2010;19:1250–61.

[8] Jacobs WCH, van der Gaag NA, Kruyt MC, Tuschel A, de Kleuver M, Peul WC, et al. Total disc replacement for chronic discogenic low back pain: a Cochrane review. Spine (Phila Pa 1976) 2013;38:24–36.

[9] Wei J, Song Y, Sun L, Lv C. Comparison of artificial total disc replacement versus fusion for lumbar degenerative disc disease: a meta-analysis of randomized controlled trials. Int Orthop 2013;37:1315–25.

[10] Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. PLoS Med 2007;4:447–55.

[11] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. Ann Intern Med 2009;151.

[12] Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.

[13] Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH. The North American spine society lumbar spine outcome assessment Instrument: reliability and validity tests. Spine (Phila Pa 1976) 1996;21:741–9.

[14] Dunn KM, Cherkin DC. The Roland-Morris Disability Questionnaire. Spine (Phila Pa 1976) 2007;32:287.

[15] McHorney CA, Ware JE, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32:40–66.

[16] Menzin J, Nichols C, Mwamburi M, Federico V, Kemner JE. PSU30 A Systematic Review of the Use of Patient-Centered Outcome Measures in Studies of Spinal Fusion Surgery for Degenerative Disc Disease: Implications for Comparative Effectiveness Research. Value Heal 2012;15:A79.

[17] Menzin J, Nichols C, Mwamburi M, Federico V, Kemner JE. Pain and Functional Outcomes Following Spinal Fusion Surgery for Degenerative Disc Disease: A Systematic Literature Review and Meta-Analysis. Spine J 2012;12:S8–9.

[18] Pull ter Gunne AF, Hosman AJF, Cohen DB, Schuetz M, Habil D, van Laarhoven CJHM, et al. A methodological systematic review on surgical site infections following spinal surgery: part 1: risk factors. Spine (Phila Pa 1976) 2012;37:2017–33.

[19] Van Middendorp JJ, Pull ter Gunne AF, Schuetz M, Habil D, Cohen DB, Hosman AJF, et al. A methodological systematic review on surgical site infections following spinal surgery: part 2: prophylactic treatments. Spine (Phila Pa 1976) 2012;37:2034–45.

[20] Longo UG, Loppini M, Romeo G, Maffulli N, Denaro V. Errors of level in spinal surgery: an evidence-based systematic review. J Bone Joint Surg Br 2012;94:1546–50.

[21] Ahn UM, Ahn NU, Buchowski JM, Garrett ES, Sieber AN, Kostuik JP. Cauda equina syndrome secondary to lumbar disc herniation: a meta-analysis of surgical outcomes. Spine (Phila Pa 1976) 2000;25:1515–22.

[22] Beynon R, Hawkins J, Laing R, Higgins N, Whiting P, Jameson C, et al. The diagnostic utility and cost-effectiveness of selective nerve root blocks in patients considered for lumbar decompression surgery: a systematic review and economic model. Health Technol Assess 2013;17:1–88, v – vi.

[23] Charles Malveaux WMS, Sharan AD. Adjacent Segment Disease After Lumbar Spinal Fusion: A Systematic Review of the Current Literature. Semin Spine Surg 2011;23:266–74.

[24] Levy RM, Deer TR. Systematic safety review and meta-analysis of procedural experience using percutaneous access to treat symptomatic lumbar spinal stenosis. Pain Med 2012;13:1554–61.

[25] Malik KM, Cohen SP, Walega DR, Benzon HT. Diagnostic criteria and treatment of discogenic pain: a systematic review of recent clinical literature. Spine J 2013;13:1675–89.

[26] Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. Spine (Phila Pa 1976) 2007;32:816–23.

[27] Conn A, Buenaventura RM, Datta S, Abdi S, Diwan S. Systematic review of caudal epidural injections in the management of chronic low back pain. Pain Physician 2009;12:109–35.

[28] Westacott DJ, Cooke SJ. Functional outcome following direct repair or intervertebral fusion for adolescent spondylolysis: a systematic review. J Pediatr Orthop B 2012;21:596–601.

[29] Turner J a, Herron L, Deyo R a. Meta-analysis of the results of lumbar spine fusion. Acta Orthop 1993;64:120–2.

[30] Friedman JH, Dighe G. Systematic review of caudal epidural injections in the management of chronic back pain. R I Med J (2013) 2013;96:12–6.

[31] Niggemeyer O, Strauss JM, Schulitz KP. Comparison of surgical procedures for degenerative lumbar spinal stenosis: a meta-analysis of the literature from 1975 to 1995. Eur Spine J 1997;6:423–9.

[32] Jarrett MS, Orlando JF, Grimmer-Somers K. The effectiveness of land based exercise compared to decompressive surgery in the management of lumbar spinal-canal stenosis: a systematic review. BMC Musculoskelet Disord 2012;13:30.

[33] Ye Y-P, Xu H, Chen D. Comparison between posterior lumbar interbody fusion and posterolateral fusion with transpedicular screw fixation for isthmic spondylolithesis: a meta-analysis. Arch Orthop Trauma Surg 2013;133:1649–55.

[34] Wood KB, Fritzell P, Dettori JR, Hashimoto R, Lund T, Shaffrey C. Effectiveness of spinal fusion versus structured rehabilitation in chronic low back pain patients with and without isthmic spondylolisthesis: a systematic review. Spine (Phila Pa 1976) 2011;36:S110–9.

[35] Ibrahim T, Tleyjeh IM, Gabbar O. Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials. Int Orthop 2008;32:107–13.

[36] Phillips FM, Slosar PJ, Youssef J a, Andersson G, Papatheofanis F. Lumbar spine fusion for chronic low back pain due to degenerative disc disease: a systematic review. Spine (Phila Pa 1976) 2013;38:E409–22.

[37] Jacobs WCH, Vreeling A, De Kleuver M. Fusion for low-grade adult isthmic spondylolisthesis: a systematic review of the literature. Eur Spine J 2006;15:391–402.

[38] Abdi S, Datta S, Trescot AM, Schultz DM, Atluri SL, Smith HS, et al. Epidural Steroids in the Management of Chronic Spinal Pain: A Systematic Review 2007:185–212.

[39] Jna G, Waddell G. Surgical interventions for lumbar disc prolapse ( Review ) 2008.

[40] May S, Comer C. Is surgery more effective than non-surgical treatment for spinal stenosis, and which non-surgical treatment is more effective? A systematic review. Physiotherapy 2013;99:12–20.

[41] Van den Eerenbeemt KD, Ostelo RW, van Royen BJ, Peul WC, van Tulder MW. Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature. Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc 2010;19:1262–80.

[42] Andersson GBJ, Mekhail NA, Block JE. Treatment Of Intractable Discogenic Low Back Pain . A Systematic Review Of Spinal Fusion And Intradiscal Electrothermal Therapy ( Idet ) 2006:237–48.

[43] Turner JA, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. Spine (Phila Pa 1976) 1992;17:1–8.

[44] Manchikanti L, Derby R, Benyamin RM, Helm S, Hirsch JA. A systematic review of mechanical lumbar disc decompression with nucleoplasty. Pain Physician n.d.;12:561–72.

[45] Carreon LY, Glassman SD, Howard J. Fusion and nonsurgical treatment for symptomatic lumbar degenerative disease: a systematic review of Oswestry Disability Index and MOS Short Form-36 outcomes. Spine J 2008;8:747–55.

[46] Kovacs FM, Urrútia G, Alarcón JD. Surgery versus conservative treatment for symptomatic lumbar spinal stenosis: a systematic review of randomized controlled trials. Spine (Phila Pa 1976) 2011;36:E1335–51.

[47] Wu RH, Fraser JF, Ha R. Minimal Access Versus Open Transforaminal Lumbar Interbody Fusion Meta-Analysis of Fusion Rates 2010;35:2273–81.

[48] Umeta RSG, Avanzi O. Techniques of lumbar-sacral spine fusion in spondylosis: systematic literature review and meta-analysis of randomized clinical trials. Spine J 2011;11:668–76.

[49] Zhou Z-J, Zhao F-D, Fang X-Q, Zhao X, Fan S-W. Meta-analysis of instrumented posterior interbody fusion versus instrumented posterolateral fusion in the lumbar spine. J Neurosurg Spine 2011;15:295–310.

[50] Liu X, Wang Y, Qiu G, Weng X, Yu B. A systematic review with meta-analysis of posterior interbody fusion versus posterolateral fusion in lumbar spondylolisthesis. Eur Spine J 2014;23:43–56.

[51] Tian N-F, Wu Y-S, Zhang X-L, Xu H-Z, Chi Y-L, Mao F-M. Minimally invasive versus open transforaminal lumbar interbody fusion: a meta-analysis based on the current evidence. Eur Spine J 2013;22:1741–9.

[52] Martin CR, Gruszczynski AT, Braunsfurth HA, Fallatah SM, O'Neil J, Wai EK. The surgical management of degenerative lumbar spondylolisthesis: a systematic review. Spine (Phila Pa 1976) 2007;32:1791–8.

[53] Singh V, Benyamin RM, Datta S, Falco FJ, Helm S, Manchikanti L. Systematic review of percutaneous lumbar mechanical disc decompression utilizing Dekompressor. Pain Physician 2009;12:589–99.

[54] Manchikanti L, Falco FJE, Benyamin RM, Caraway DL, Deer TR, Singh V, et al. An Update of the Systematic Assessment of Mechanical Lumbar Disc Decompression with Nucleoplasty 2013:25–54.

[55] Daubs MD, Norvell DC, McGuire R, Molinari R, Hermsmeyer JT, Fourney DR, et al. Fusion versus nonoperative care for chronic low back pain: do psychological factors affect outcomes? Spine (Phila Pa 1976) 2011;36:S96–109.

[56] Choma TJ, Schuster JM, Norvell DC, Dettori JR, Chutkan NB. Fusion versus nonoperative management for chronic low back pain: do comorbid diseases or general health factors affect outcome? Spine (Phila Pa 1976) 2011;36:S87–95.

[57] De Kleuver M, Oner FC, Jacobs WCH. Total disc replacement for chronic low back pain: background and a systematic review of the literature. Eur Spine J 2003;12:108–16.

[58] Saltychev M, Eskola M, Laimi K. Lumbar fusion compared with conservative treatment in patients with chronic low back pain: a meta-analysis. Int J Rehabil Res 2014;37:2–8.

[59] Gotfryd A, Avanzi O. A systematic review of randomised clinical trials using posterior discectomy to treat lumbar disc herniations. Int Orthop 2009;33:11–7.

[60] Niemisto L, Ea K, Malmivaara A, Seitsalo S, Hurri H. Radiofrequency denervation for neck and back pain ( Review ) 2010.

[61] Scheepers M, Dissertations P. PROSPERO International prospective register of systematic reviews Effectiveness of surgical versus conservative treatment for symptomatic unilateral spondylolysis of the lumbar spine in athletes : a systematic review protocol 2013:1–5.

[62] Shamliyan T, Goldmann D, Sands-lincoln M. PROSPERO International prospective register of systematic reviews Epidural steroid injections for radicular lumbosacral pain 2014:1–3.

[63] Poetscher A, Lenza M, Gentil A, Ferretti M. PROSPERO International prospective register of systematic reviews Facet joint radiofrequency rizotomy for low back pain : a systematic review 2013:1–4.

[64] Fernandez M, Ferreira P, Ferreira M, Refshauge K, Hartvigsen J, Silva I. PROSPERO International prospective register of systematic reviews Surgery or physical activity in the management of sciatica : a systematic review 2013:1–4.

[65] Jacobs W, Rubinstein S, Willems P, Moojen W, Oner C, Tulder M Van. PROSPERO International prospective register of systematic reviews The evidence on surgical interventions for low back disorders , an overview of systematic reviews 2011:1–3.

[66] Niemisto L, Jousimaa J, Hurri H, Ea K, Malmivaara A. Radiofrequency denervation for chronic low-back pain ( Protocol ) 2010.

[67] Surgery for lumbar isthmic spondylolisthesis in adults.pdf n.d.

[68] Posterior approach discectomy with or without fusion for lumbar disc herniation a systematic review.pdf n.d.

[69] Decompression with and without fusion in the treatment of degenerative lumbar disease a systematic review.pdf n.d.

[70] bri140336@140323-135.pdf n.d.

[71] Artificial Total disc replacement for lumbar degenerative disc disease a systematic review.pdf n.d.

[72] Jarrett MS, Orlando JF, Grimmer-Somers K. The effectiveness of land based exercise compared to decompressive surgery in the management of lumbar spinal-canal stenosis: a systematic review. BMC Musculoskelet Disord 2012;13:30.

[73] Kovacs FM, Urrútia G, Alarcón JD. Surgery versus conservative treatment for symptomatic lumbar spinal stenosis: a systematic review of randomized controlled trials. Spine (Phila Pa 1976) 2011;36:E1335–51.

[74] Daubs MD, Norvell DC, McGuire R, Molinari R, Hermsmeyer JT, Fourney DR, et al. Fusion versus nonoperative care for chronic low back pain: do psychological factors affect outcomes? Spine (Phila Pa 1976) 2011;36:S96–109.

[75] Van den Eerenbeemt KD, Ostelo RW, van Royen BJ, Peul WC, van Tulder MW. Total disc replacement surgery for symptomatic degenerative lumbar

disc disease: a systematic review of the literature. Eur Spine J 2010;19:1262–80.

[76] Wu RH, Fraser JF, Härtl R. Minimal access versus open transforaminal lumbar interbody fusion: meta-analysis of fusion rates. Spine (Phila Pa 1976) 2010;35:2273–81.

[77] Egger M, Smith GD. Meta-Analysis. Potentials and promise. BMJ 1997;315:1371–4.

[78] Delgado-López PS, Rodríguez-Salazar A, Castilla-Díez JM, Martín-Velasco V, Fernández-Arconada O. Papel de la cirugía en la enfermedad degenerativa espinal. Análisis de revisiones sistemáticas sobre tratamientos quirúrgicos y conservadores desde el punto de vista de la medicina basada en la evidencia. Neurocirugia 2005;16:142–57.

[79] Chou R. Using evidence in pain practice: Part I: Assessing quality of systematic reviews and clinical practice guidelines. Pain Med 2008;9:518–30.

[80] Delamarter RB, Bae HW, Pradhan BB. Clinical results of ProDisc-II lumbar total disc replacement: report from the United States clinical trial. Orthop Clin North Am 2005;36:301–13.

[81] Zigler J, Delamarter R, Spivak JM, Linovitz RJ, Danielson GO, Haider TT, et al. Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-

level degenerative disc disease. Spine (Phila Pa 1976) 2007;32:1155–62; discussion 1163.

[82] Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011.

[83] Furlan AD, Tomlinson G, Jadad AAR, Bombardier C. Examining heterogeneity in meta-analysis: comparing results of randomized trials and nonrandomized studies of interventions for low back pain. Spine (Phila Pa 1976) 2008;33:339–48.

[84] Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PLoS One 2007;2.

[85] Gagnier JJ, Kellam PJ. Reporting and methodological quality of systematic reviews in the orthopaedic literature. J Bone Jt Surg Am 2013;95:e771–7.

[86] Ho RST, Wu X, Yuan J, Liu S, Lai X, Wong SYS, et al. Methodological quality of meta-analyses on treatments for chronic obstructive pulmonary disease: a cross-sectional study using the AMSTAR (Assessing the Methodological Quality of Systematic Reviews) tool. NPJ Prim Care Respir Med 2015;25:14102.

[87] Tunis AS, McInnes MDF, Hanna R, Esmail K. Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology

journals changed since publication of the PRISMA statement? Radiology 2013;269:413–26.

Figure Legends:

Fig. 1 - Flow chart for study screening and eligibility.

Fig. 2 - Boxplots for AMSTAR and PRISMA total percentage according to systematic reviews (SR) with or without meta-analysis (MA). Mann-Whitney test.

| *Outcomes* | *Results* | *SR with MA (n=17)* | *SR without MA (n=23)* | *p-value* |
|---|---|---|---|---|
| | | n (%) | n (%) | |
| Estimated blood loss | NA | 12 (70.6) | 23 (100) | 0.009 |
| | NS | 3 (17.6) | 0 (0) | |
| | SR | 2 (11.8) | 0 (0) | |
| Complications | NA | 8 (47.1) | 21 (91.3) | 0.004 |
| | NS | 6 (35.3) | 2 (8.7) | |
| | SR | 3 (17.6) | 0 (0) | |
| Fusion | NA | 9 (52.9) | 23 (100) | <0.001 |
| | NS | 3 (17.6) | 0 (0) | |
| | SR | 5 (29.4) | 0 (0) | |
| Infection | NA | 16 (94.1) | 23 (100) | 0.425 |
| | NS | 1 (5.9) | 0 (0) | |
| ODI | NA | 5 (29.4) | 15 (65.2) | 0.035 |
| | NS | 7 (41.2) | 7 (30.4) | |
| | SR | 5 (29.4) | 1 (4.3) | |
| Operation time | NA | 12 (70.6) | 23 (100) | 0.009 |
| | NS | 2 (11.8) | 0 (0) | |
| | SR | 3 (17.6) | 0 (0) | |
| Pain VAS | NA | 9 (52.9) | 17 (73.9) | 0.385 |
| | NS | 6 (35.3) | 4 (17.4) | |
| | SR | 2 (11.8) | 2 (8.7) | |
| Reoperation | NA | 10 (58.8) | 23 (100) | 0.001 |
| | NS | 5 (29.4) | 0 (0) | |
| | SR | 2 (11.8) | 0 (0) | |
| Return to work | NA | 14 (82.4) | 19 (82.6) | 1 |
| | NS | 2 (11.8) | 3 (13) | |
| | SR | 1 (5.9) | 1 (4.3) | |

| | | | | |
|---|---|---|---|---|
| Roland-Morris | NA | 17 (100) | 22 (95.7) | 1 |
| | NS | 0 (0) | 1 (4.3) | |
| SF-36 | NA | 14 (82.4) | 21 (91.3) | 0.634 |
| | NS | 3 (17.6) | 2 (8.7) | |

**Table 1 – Analysis of outcomes considering systematic reviews with or without a meta-analysis. SR = systematic review; MA =meta-analysis; NA = not analyzed in the systematic review; NS = non-significant; SR = significant result; ODI = Oswestry disability index; VAS = visual analogic scale; SF-36 = Short-Form 36. *p* values are from Fisher exact test.**

| *Quality tool* | *Categories* | | | | |
|---|---|---|---|---|---|
| | Very Poor | Poor | Fair | Good | Excelent |
| | % | % | % | % | % |
| Amstar total | 22.5 | 15 | 30 | 25 | 7.5 |
| Prisma total | 7.5 | 30 | 37.5 | 20 | 5 |
| Prisma - Title | 12.5 | 0 | 0 | 0 | 87.5 |
| Prisma - Abstract | 17.5 | 62,5 | 0 | 0 | 20 |
| Prisma - Introduction | 0 | 17.5 | 0 | 37.5 | 45 |
| Prisma - Methods | 22.5 | 32.5 | 20 | 20 | 5 |
| Prisma - Results | 35 | 25 | 12.5 | 15 | 12.5 |
| Prisma - Discussion | 5 | 17.5 | 20 | 7.5 | 50 |
| Prisma - Funding | 37.5 | 5 | 0 | 0 | 57.5 |

**Table 2 - AMSTAR and PRISMA quality distribution by percentage of final score. PRISMA is also subdivided by domains percentages. Very poor (<30%); Poor (30-50%); Fair (50-70%); Good (70-90%); Excellent (>90%)**

| | *n* | *PRISMA >70* | *AMSTAR >70* |
|---|---|---|---|
| DDD | 13 | 4 (30.7%) | 5 (38.4%) |
| SL | 6 | 1 (16.6%) | 3 (50%) |
| LS | 5 | 0 (0) | 1 (20%) |
| LDH | 3 | 1 (33.3%) | 1 (33.3%) |
| S | 11 | 4 (36.3%) | 2 (18.8%) |
| FJS | 2 | 0 (0) | 1 (50%) |

**Table 3 – Absolute and relative frequencies of very good and excellent systematic reviews according to disease group in PRISMA and AMSTAR tools. DDD = Degenerative disc disease, SL = Spondylolisthesis, LS = Lumbar stenosis, LDH = Lumbar disc herniation, S = Spondylosis, FJS = Facet joint syndrome. P=0.713 (PRISMA) and p=0.747 (AMSTAR). Fisher exact test.**