Systematic Review and Meta-Analysis of 3 Treatment Arms for Vertebral Compression Fractures

A Comparison of Improvement in Pain, Adjacent-Level Fractures, and Quality of Life Between Vertebroplasty, Kyphoplasty, and Nonoperative Management

Sascha Halvachizadeh, MD
Anna-Lea Stalder
David Bellut, MD
Sven Hoppe, MD
Philipp Rossbach, MD
Alessandro Cianfoni, MD
Klaus John Schnake, MD
Ladislav Mica, MD, FEBS
Roman Pfeifer, MD, FEBS
Kai Sprengel, MD, FEBS*
Hans-Christoph Pape, MD, FACS*

Investigation performed at the Department of Trauma, University Hospital Zurich, Zurich, Switzerland

Abstract

Background: Osteoporotic vertebral fractures (OVFs) have become increasingly common, and previous nonrandomized and randomized controlled trials (RCTs) have compared the effects of cement augmentation versus nonoperative management on the clinical outcome. This meta-analysis focuses on RCTs and the calculated differences between cement augmentation techniques and nonsurgical management in outcome (e.g., pain reduction, adjacent-level fractures, and quality of life [QOL]).

Methods: A systematic review was performed according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines, and the following scientific search engines were used: MEDLINE, Embase, Cochrane, Web of Science, and Scopus. The inclusion criteria included RCTs that addressed different treatment strategies for OVF. The primary outcome was pain, which was determined by a visual analog scale (VAS) score; the secondary outcomes were the risk of adjacent-level fractures and QOL (as determined by the EuroQol-5 Dimension [EQ-5D] questionnaire, the Oswestry Disability Index [ODI], the Quality of Life Questionnaire of the European Foundation for Osteoporosis [QUALEFFO], and the Roland-Morris Disability Questionnaire [RDQ]). Patients were assigned to 3 groups according to their treatment: vertebroplasty (VP), kyphoplasty (KP), and nonoperative management (NOM). The short-term (weeks), midterm (months), and long-term (>1 year) effects were compared. A random effects model was used to summarize the treatment effect, including I^2 for assessing heterogeneity and the revised Cochrane risk-of-bias 2 (RoB 2) tool for assessment of ROB. Funnel plots were used to assess risk of publication bias. The log of the odds ratio (OR) between treatments is reported.

Results: After screening of 1,861 references, 53 underwent full-text analysis and 16 trials (30.2%) were included. Eleven trials (68.8%) compared VP and NOM, 1 (6.3%) compared KP and NOM, and 4 (25.0%) compared KP and VP. Improvement of pain was better by 1.31 points (95% confidence interval [CI], 0.41 to 2.21; $p < 0.001$) after VP when compared with NOM in short-term follow-up. Pain effects were similar after VP and KP (midterm difference of

*Kai Sprengel, MD, FEBS, and Hans-Christoph Pape, MD, FACS, contributed equally to this work.

Disclosure: The **Disclosure of Potential Conflicts of Interest** forms are provided with the online version of the article (http://links.lww.com/JBJSREV/A761).

0.0 points; 95% CI, −0.25 to 0.25). The risk of adjacent-level fractures was not increased after any treatment (log OR, −0.16; 95% CI, −0.83 to 0.5; NOM vs. VP or KP). QOL did not differ significantly between the VP or KP and NOM groups except in the short term when measured by the RDQ.

Conclusions: This meta-analysis provides evidence in favor of the surgical treatment of OVFs. Surgery was associated with greater improvement of pain and was unrelated to the development of adjacent-level fractures or QOL. Although improvements in sagittal balance after surgery were poorly documented, surgical treatment may be warranted if pain is a relevant problem.

Level of Evidence: Therapeutic Level I. See Instructions for Authors for a complete description of levels of evidence.

Surgical treatment methods have been developed for osteoporotic vertebral fractures (OVFs), and constant improvements have been made in cementation techniques, which recently have contributed to the active correction of sagittal deformity[1]. Before these techniques were available, it was thought that pain medications and nonoperative treatment with bracing were sufficient to control discomfort without correction of the deformity[2]. However, subsequent worsening of quality of life (QOL) was frequently observed[3]. Consequently, there was a period of enthusiasm for surgical management to improve pain and QOL. One of the first randomized controlled trials (RCTs) that compared surgical management with nonoperative management (NOM) for the treatment of OVF demonstrated favorable outcomes in the group that was treated surgically[4]. However, both surgical techniques—vertebroplasty (VP) and kyphoplasty (KP)—were later questioned because they were thought to lead to overstuffing and increased risk of adjacent-level fractures[5] or secondary loss of sagittal balance[6].

Multiple RCTs have been undertaken to determine if surgical management is truly superior to NOM and, if so, which timing and method might be most beneficial. It became evident that surgical management is usually offered to patients after a short (days) to medium (weeks) period of NOM.

Despite a decade of prospective RCTs, the optimal treatment for an OVF remains a subject of discussion and controversy[7]. Numerous reviews have been published that focus on only 1 treatment arm (VP versus NOM[8]) or have addressed only 1 outcome variable[3].

Therefore, the aim of this study was to include all of the treatment options and to address the following hypotheses: (1) Surgical management of OVF is favorable in terms of long-term pain reduction when compared with NOM. (2) The type of treatment strategy affects the risk of adjacent-level fractures. (3) QOL after an OVF depends on the treatment strategy.

Materials and Methods

This study was conducted following the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines[9].

Search Strategy and Definitions

A systematic literature search was performed, which included the MEDLINE, Embase, Cochrane, Web of Science, and Scopus databases. The inclusion criteria included prospective RCTs assessing treatment modalities for OVF that had been published in the English or German language. The exclusion criteria included other study methodologies and articles without full-text availability. The full search formula is provided in the Appendix.

Data Management

The exports of deduplicated publications were saved in an EndNote (Clarivate) library. Two authors (S.H. and K.S.) received the same library. Blinded independent screening was performed using Rayyan software[10].

Study Selection

After screening of the titles and abstracts, the full text was analyzed. Data were extracted and stored, and qualitative and quantitative synthesis was performed. The articles were selected independently by 2 authors (S.H. and K.S.), and quantitative and qualitative analysis was performed in collaboration. Discrepancies were resolved by consensus or, if necessary, by third-party arbitration (L.M.).

Data Extraction and Group Stratification

Patients were categorized in 3 different treatment groups: (1) VP (surgical treatment), (2) KP (surgical treatment), and (3) NOM.

Trials that resulted in >1 publication were combined into 1 entity. The following data were extracted by 2 authors (S.H. and A.-L.S.): (1) general study information: first author, year, and country; (2) patient characteristics: sample size, age, duration of clinical symptoms prior to treatment, and follow-up; and (3) outcome measures: pain, QOL, and rate of adjacent-level fractures.

Main Outcome Variables

The primary outcome was change in pain with each treatment modality. Pain was measured by a visual analog scale (VAS) score (0 to 10 points [no pain to worst pain ever])[11].

The secondary outcomes included the rate of adjacent-level fracture and QOL, which was assessed by the preferred tool of the trials. The studies used the Oswestry Disability Index (ODI)[12],

the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO)[13], the EuroQol-5 Dimension (EQ-5D) questionnaire[14], or the Roland-Morris Disability Questionnaire (RDQ)[15].

The outcome variables were stratified according to the duration of follow-up: short-term (weeks), midterm (months), and long-term (>1 year). Data were recorded independently and in duplicate by 2 authors (S.H. and A.-L.S.) on separate copies of a spreadsheet. The data were compared, and any discrepancies were resolved by consensus.

Risk-of-Bias (RoB) Assessment

The revised Cochrane risk-of-bias tool (RoB 2) for randomized trials was used to assess RoB[16]. Two authors (S.H. and A.-L.S.) conducted the RoB assessment independently. Discrepancies were resolved by consensus or by third-party arbitration (R.P.). The RoB assessment strictly followed the recommendations provided by the RoB 2 tool and included 5 key domains: (D1) bias arising due to the randomization process, (D2) bias due to deviation from the intended intervention, (D3) bias due to missing outcome data, (D4) bias in the measurement of the outcome, and (D5) bias in the selection of the reported result. These results were visualized with the robvis visualization tool[17]. Additional results from the meta-analysis, including publication bias as demonstrated with a funnel plot, are provided in the Appendix.

Statistical Analysis

Reported means and standard deviations (SDs) were used for calculations of pooled results. For trials that reported means with standard errors (SEs), the SD was computed using the Cochrane Collaboration formula[18]: $SD = SE \times \sqrt{N}$. For trials that reported values as the median with a range or an interquartile range, we estimated the mean and SD according to the formulas described by Wan et al.[19]. To confirm the reliability of these estimations, we performed them in duplicate with the formulas described by Luo et al.[20]. We compared the results of both methods; each demonstrated good reliability for these estimations, even in the presence of deviation from the normal distribution[21]. The results are shown in forest plots of pooled mean differences (MDs) or the log of the odds ratio (OR) with 95% confidence intervals (CIs). To estimate heterogeneity, we used the Cochran Q test (total between-study variation) and calculated the I^2 statistic (the proportion of total variation due to between-study variation) and H^2 statistic (the ratio of the total amount of variability and the amount of between-study variability) among the trials. A random effects (RE) model was utilized in the analysis of the pooled treatment outcomes. All statistical analyses were performed using R version 3.6.1 (R Foundation for Statistical Computing)[22] and the metafor package, which is a free and open-source add-on for conducting meta-analyses using the R statistical software environment[23]. The metafor package consists of a collection of functions that allow the calculation of pooled effect size and fit RE models and meta-regression analyses. Significance was defined as $p < 0.05$.

Results

Study Selection and Characteristics

Of 1,861 articles, the structured screening process revealed 16 eligible trials that addressed the 3 treatment arms: 11 articles (68.8%) investigated

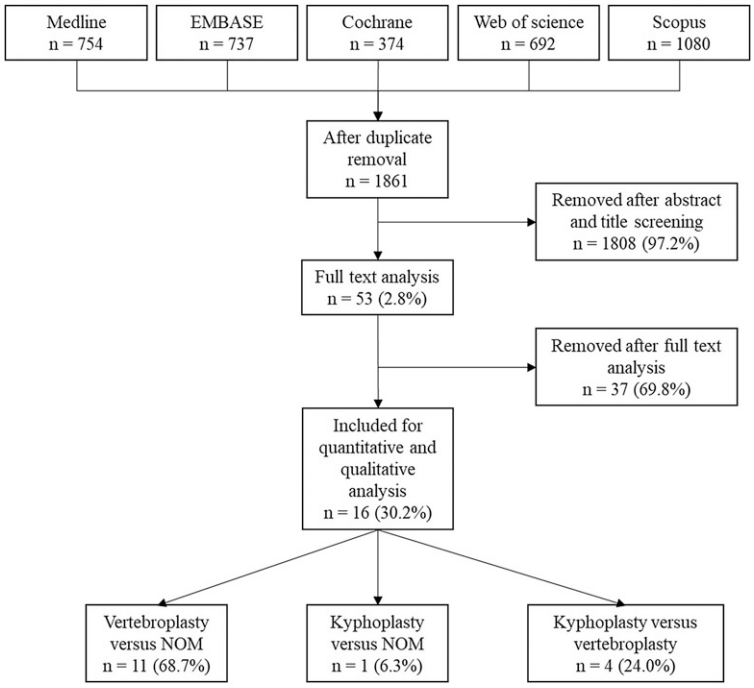

Fig. 1
PRISMA flowchart of the screening process.

TABLE I Study Characteristics and Main Conclusion

Author	Year	Country	Surgical Intervention	Control	Duration of Symptoms *(wk)*	Surgical Intervention Group *(no.)*	Control Group *(no.)*	Mean Age *(yr)*	Main Conclusion
Voormolen et al.[4]	2006	Netherlands	Vertebroplasty	Nonsurgical	11.5	18	16	73.0	Favors intervention
Buchbinder et al.[24]	2009	Australia	Vertebroplasty	Nonsurgical	9.1	38	40	76.6	No benefit from intervention
Kallmes et al.[25]	2009	U.S.	Vertebroplasty	Nonsurgical	18	68	63	73.9	No benefit from intervention
Rousing et al.[34]	2009	Denmark	Vertebroplasty	Nonsurgical	1.8	25	24	80.0	No benefit from intervention
Klazen et al.[28]	2010	Netherlands	Vertebroplasty	Nonsurgical	4	101	101	75.3	Favors intervention
Farrokhi et al.[29]	2011	Iran	Vertebroplasty	Nonsurgical	28.5	40	42	73.0	Favors intervention
Blasco et al.[35]	2012	Spain	Vertebroplasty	Nonsurgical	18.2	64	61	73.3	Favors intervention
Chen et al.[30]	2014	People's Republic of China	Vertebroplasty	Nonsurgical	19.4	46	43	65.6	Favors intervention
Wang et al.[26]	2016	People's Republic of China	Vertebroplasty	Nonsurgical	8	108	109	63.1	Favors intervention
Clark et al.[31]	2016	Australia	Vertebroplasty	Nonsurgical	2.1	61	59	80.5	Favors intervention
Firanescu et al.[32]	2018	Netherlands	Vertebroplasty	Nonsurgical	5.6	91	89	75.8	Favors intervention
Wardlaw et al.[33]	2009	U.K.	Kyphoplasty	Nonsurgical	12	149	151	73.2	Favors intervention
Liu et al.[44]	2010	Taiwan	Kyphoplasty	Vertebroplasty	2.3	50	50	73.3	No benefit from intervention
Korovessis et al.[45]	2013	Greece	Kyphoplasty	Vertebroplasty	12	86	82	71.0	No benefit from intervention
Dohm et al.[46]	2014	U.S.	Kyphoplasty	Vertebroplasty	12	191	190	75.6	Favors intervention
Evans et al.[47]	2016	U.S.	Kyphoplasty	Vertebroplasty	9.4	59	56	75.6	No benefit from intervention

VP versus NOM, 1 article (6.3%) investigated KP versus NOM, and 4 articles (25.0%) investigated KP versus VP (Fig. 1).

In total, 2,371 patients were included: 1,038 (43.8%) in the VP group, 535 (22.6%) in the KP group, and 798 (33.7%) in the NOM group. The mean age of the patients was 73.7 years (SD, 4.3 years). The duration of symptoms prior to intervention was 10.9 weeks (SD, 7.1 weeks) (Table I).

Improvement of Pain

VP or KP Versus NOM

In the short-term follow-up, the VAS score improved by 1.31 (95% CI, 0.41 to 2.21) more following VP or KP when compared with NOM. Operative treatment was not associated with any significant improvement of the VAS score in 3 of 10 trials[24-26], while the remaining 7 trials reported more favorable outcomes following VP or KP[27-33]. The I^2 statistic (99.8%), the significant Cochran Q test ($p < 0.0001$), and the H^2 statistic indicated considerable heterogeneity (Fig. 2-A).

In the midterm follow-up, the VAS score improved by 0.90 (95% CI, 0.25 to 1.54) more in the VP and KP groups compared with the NOM group. Five trials (45.5%) reported comparable outcome measures[24,26,32,34,35], while 6 trials (54.5%) reported favorable outcomes for the VP and KP groups[25,28-31,33]. These results were subject to a large degree of heterogeneity, as shown by the I^2 value of 98.9% and a significant Cochran Q test ($p < 0.0001$) (Fig. 2-B).

In the long-term follow-up, the VAS improvement was 0.89 (95% CI, 0.16 to 1.62) greater in the VP and KP groups compared with the NOM group. Four trials (50%) reported comparable outcomes[24,26,32,34], and 4 trials (50%) reported more favorable outcomes for the VP and KP groups[28-30,33]. Considerable heterogeneity in these results was noted, as shown by the I^2 statistic (99.2%), the H^2 statistic, and a significant Cochran Q test ($p < 0.0001$) (Fig. 2-C).

KP Versus VP

Although the VAS pain score improved following both KP and VP treatment, the magnitude of improvement did not differ between these surgical interventions; the improvement also did not differ between follow-up periods (Figs. 3-A, 3-B, and 3-C).

Adjacent-Level Fractures

Of 1,073 patients, 243 (22.6%) sustained an adjacent-level fracture between 3 and 24 months after surgery or NOM. In the VP and KP groups, 125 of 551 patients (22.7%) sustained adjacent-level fractures, as did 118 of 522 patients

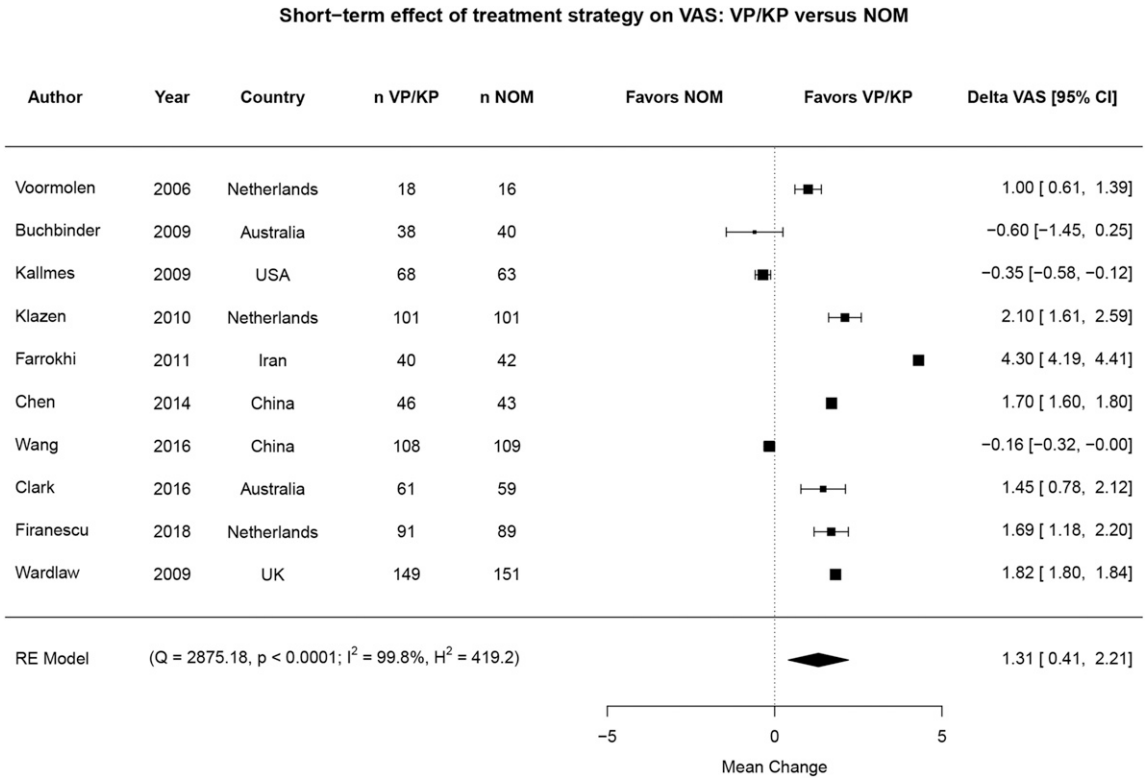

Fig. 2-A

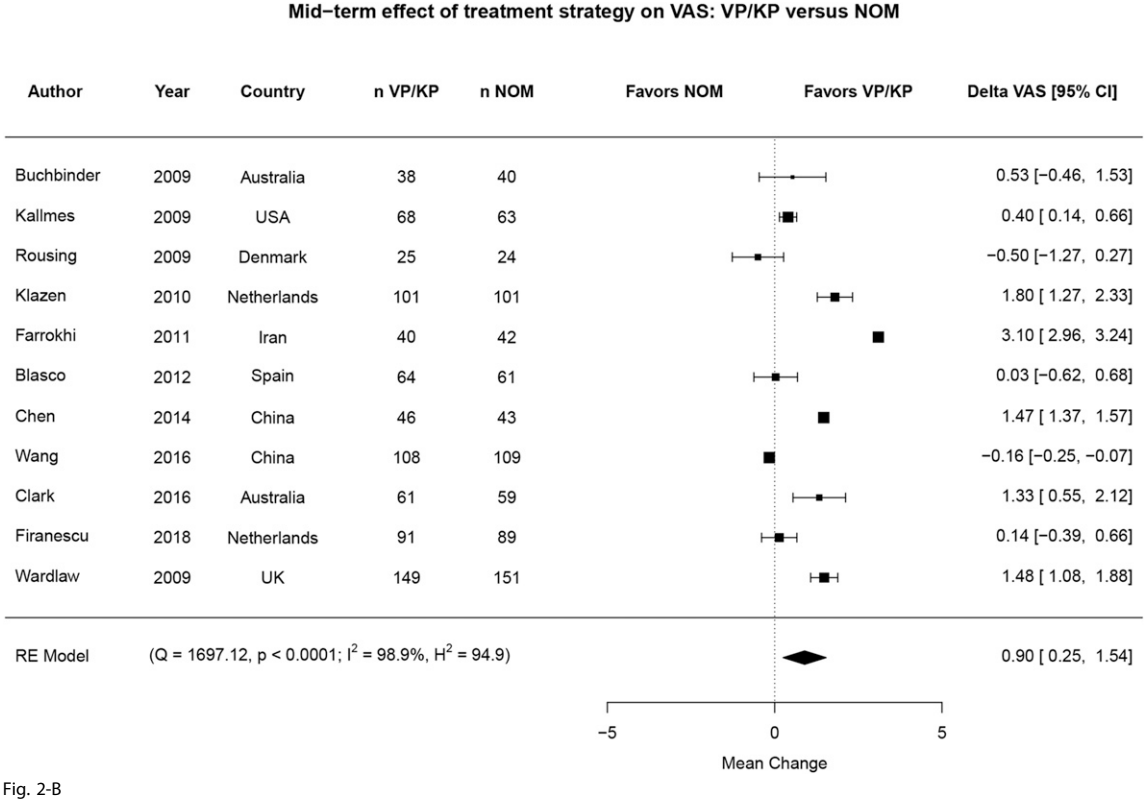

Fig. 2-B

Figs. 2-A, 2-B, and 2-C The effect of the type of management on the change of the VAS pain score when compared with baseline. VP = vertebroplasty, KP = kyphoplasty, and NOM = nonoperative management. **Fig. 2-A** In the short term (within weeks), the VAS pain score improved by 1.31 points more in the VP and KP groups. **Fig. 2-B** In the midterm (within months), the VAS pain score improved by 0.90 points more in the VP and KP groups.

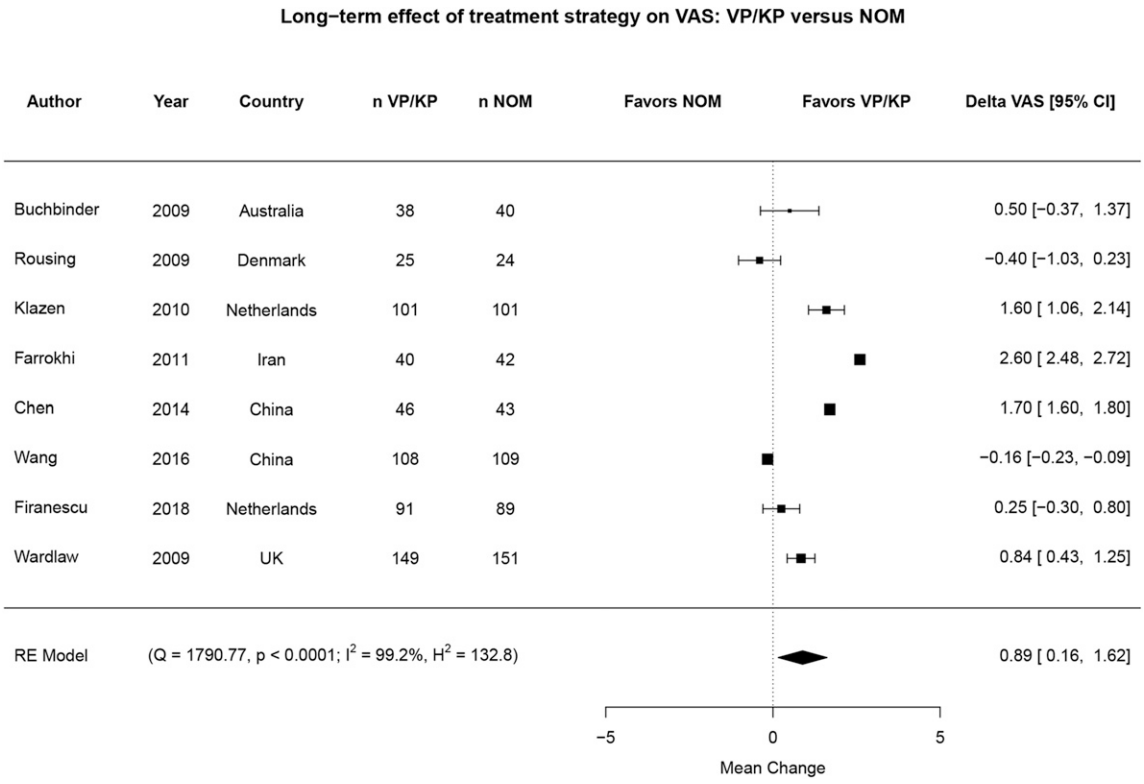

Fig. 2-C

In the long term (after years), the VAS pain score improved by 0.89 points more in the VP and KP groups.

(22.6%) in the NOM group; the difference was not significant. Nine trials (75%) reported data regarding adjacent-level fractures after VP or KP versus NOM. The NOM group demonstrated a comparable risk of adjacent-level fractures to the VP and KP groups (log OR = −0.16; 95% CI, −0.83 to 0.50; heterogeneity, I^2 = 72.5%) (Fig. 4).

QOL

The RDQ revealed significantly better QOL results in the VP or KP group compared with the NOM group (MD, 1.7; 95% CI, 0.01 to 3.47; p = 0.049) in short-term follow-up. A trend toward an improved RDQ can further be seen in midterm follow-up (MD, 1.6; 95% CI, −0.09 to 3.24; p = 0.061) (Table II).

RoB Assessment

All of the trials reported an intention-to-treat analysis; however, they did not also report a per-protocol analysis. Fourteen trials (87.5%) provided some concern for bias, and the remaining 2 (12.5%) provided a high concern for bias. None of the studies reported a low concern for bias (Fig. 5).

Discussion

In the U.S., the proportion of individuals over the age of 65 years has been projected increase by 8.2% from 2016 to 2060 (from 15.2% to 23.4%[36]), and the European Union projects an increase of approximately 1.5-fold by 2050[37]. This demographic change is expected to be associated with an increase in the incidence of osteoporosis[38] and associated OVFs[39].

Among the treatment strategies for OVF, NOM continues to play a relevant role, although surgical intervention with VP or KP has demonstrated constant improvement in cement augmentation options and surgical techniques[40-42].

Nevertheless, there is heterogeneity among results regarding management of OVFs, as well as rather short-term follow-up, inconclusive evidence of the best timing of intervention, and a lack of guidelines for pain control. This was confirmed in this meta-analysis, which focused on studies between 2006 and 2019; most technical improvements, such as cement augmentation techniques and overstuffing, did not appear to play a relevant role[5].

Interestingly, most of the available studies focused on the comparison of surgical management versus NOM, rather than comparing the indications and/or surgical techniques. We aimed to account for this issue by comparing 3 treatment groups. As discussed below, we found high variability.

The current meta-analysis of RCTs compared 3 routinely performed treatment strategies for OVF, and the main results were as follows:

1. There was greater improvement of pain in both of the patient groups that were treated surgically (with KP and VP) compared with those treated with NOM.
2. Neither surgical management nor NOM demonstrated a greater risk of adjacent-level fractures.
3. QOL, although assessed with various tools, was not associated

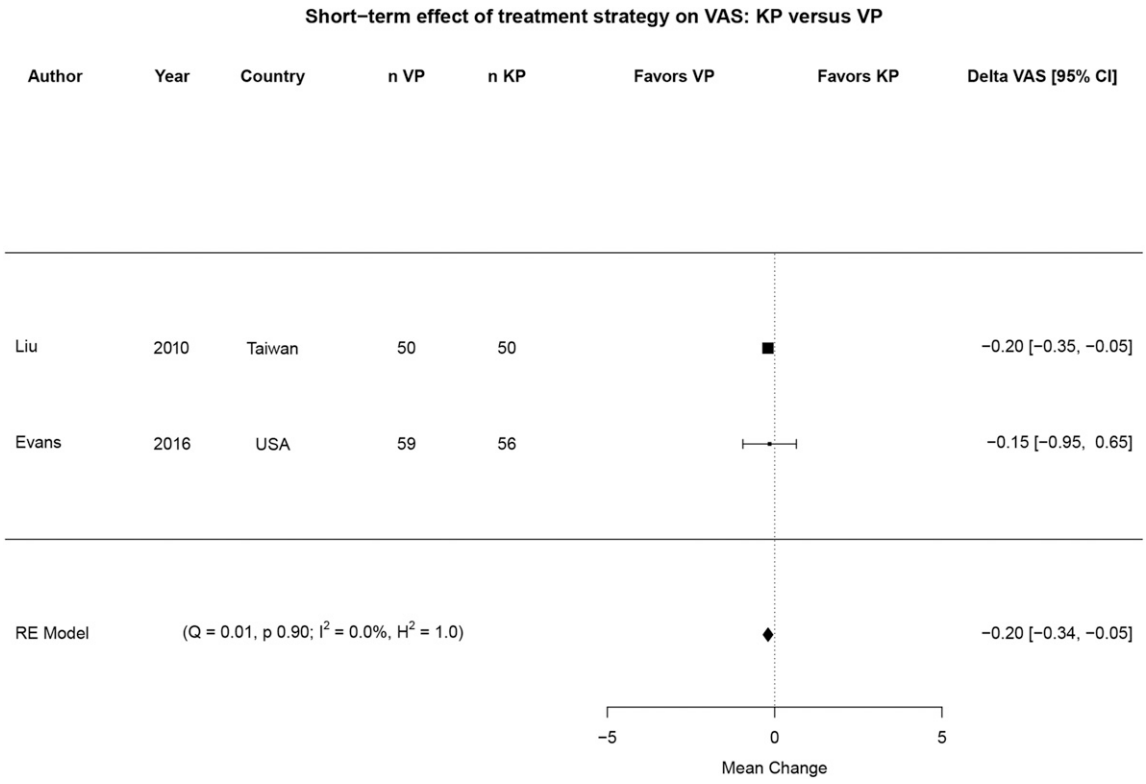

Fig. 3-A

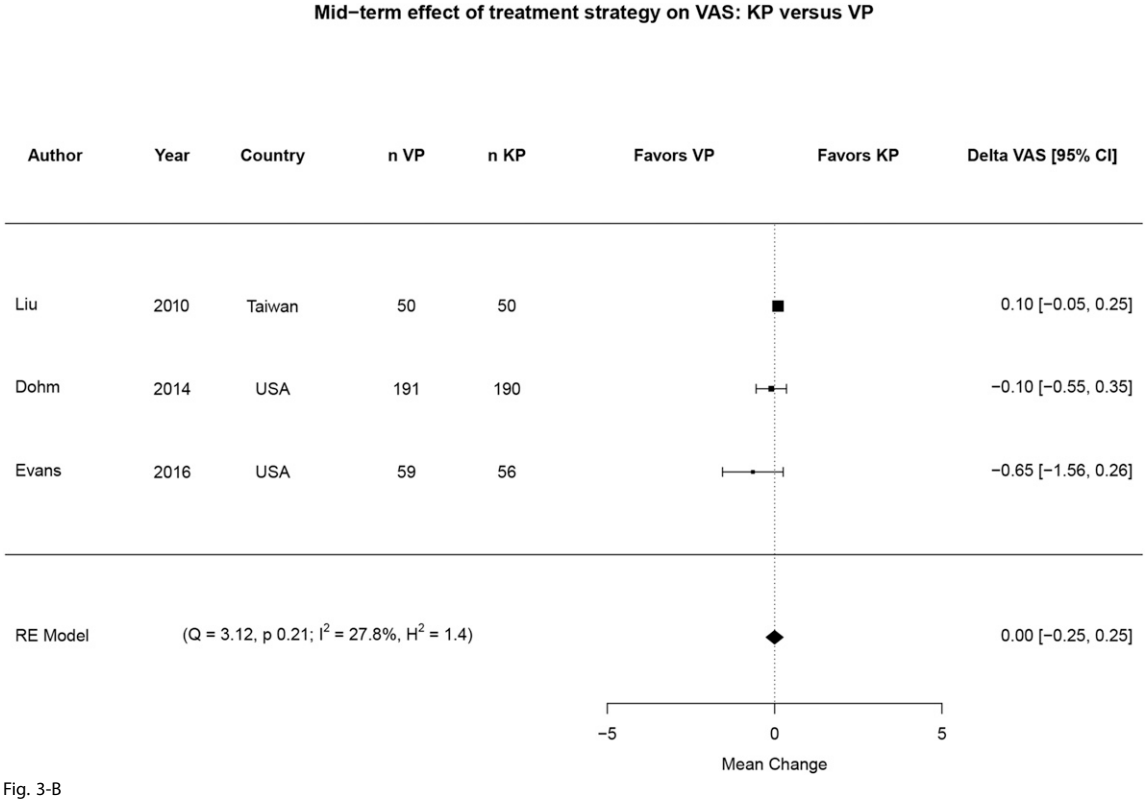

Fig. 3-B

Figs. 3-A, 3-B, and 3-C The effect of the type of surgical intervention on the change of the VAS pain score when compared with baseline. VP = vertebroplasty and KP = kyphoplasty. In the short term (within weeks, **Fig. 3-A**), midterm (within months, **Fig. 3-B**), and long term (after years, **Fig. 3-C**), the improvement of the VAS pain score was comparable in both groups.

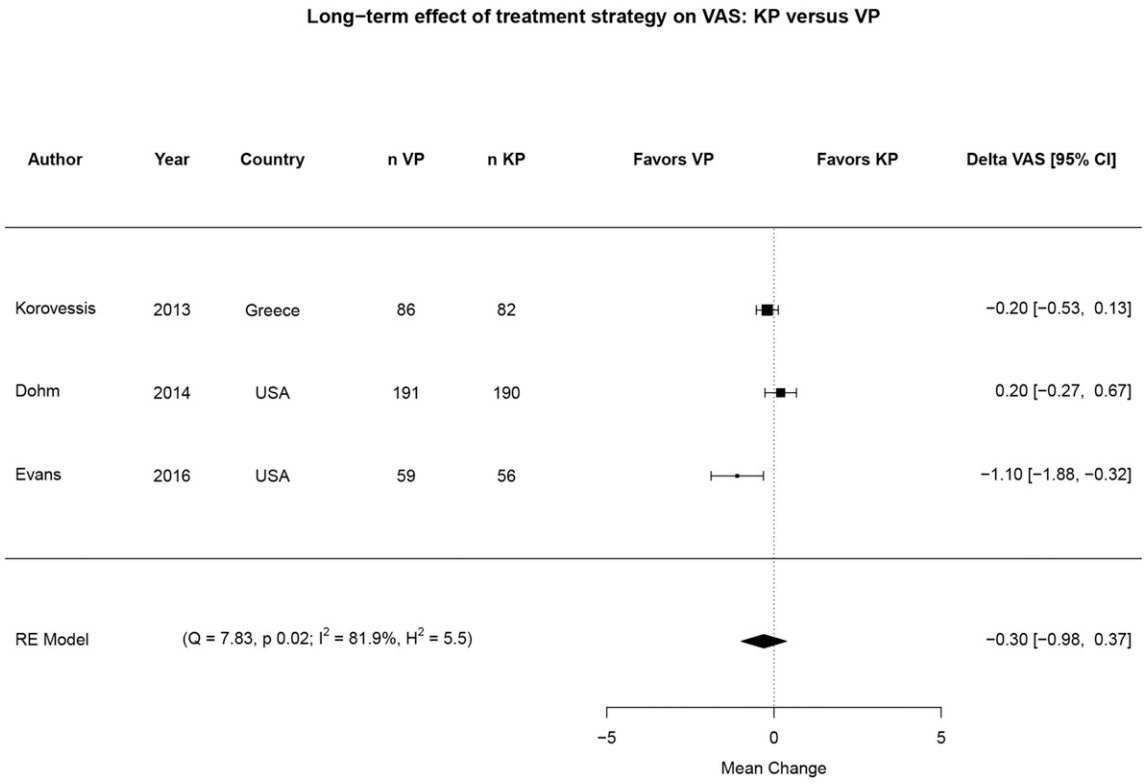

Fig. 3-C

with improvement in pain and overall function.

The consistent superiority of surgical techniques across most of the studies during multiple follow-up time periods was striking. Overall, improvement of the VAS pain score was greater after VP and KP compared with NOM. Moreover, the majority of trials (9 of 12) demonstrated more favorable pain relief when the fracture was treated surgically compared with NOM, independent of the surgical technique. Beall et al. have provided additional evidence favoring surgical intervention over NOM[43]. Their meta-analysis included RCTs and non-RCTs in their calculation, which increased the heterogeneity of the studies and decreased the level of evidence[18].

The present meta-analysis included 4 RCTs that investigated the effects of KP versus VP[44-47]; they found no significant difference in outcome measures. This is in line with similar studies that included both RCTs and non-RCTS. Papanastassiou et al. concluded that VP and KP provide greater pain relief and fewer subsequent fractures than NOM of OVFs[48]. The superiority of VP and KP compared with NOM was assessed in numerous trials, providing evidence for similar effectiveness and safety[49] and improved outcome measures[48,50,51].

Overall, the improvement in pain was not surprising. The duration of pain from the onset of clinical symptoms until the beginning of surgical treatment was reported to be around 10 weeks. Thus, our results are in keeping with previous RCTs.

We have tried to account for the fact that pain medication strategies have helped in surgical treatment as well as NOM. However, we were surprised that none of the studies described a uniform pain management strategy for their patients. In contrast, all of the studies described an individualized approach in detail but did not specify the types of pain medications that were used with VP or KP, which was widely accepted as a successful procedure for treating OVF.

At first glance, the lack of evidence for a greater risk of adjacent-level fracture in surgically treated patients is surprising because a certain loss of sagittal balance has been described over time[32,52]. It is unclear whether the adjacent-level vertebral disc was the primary reason for this loss of balance or whether there was additional loss of reduction due to further osseous destruction around the cement.

Our meta-analysis tried to account for the issue of sagittal balance. However, we did not arrive at a meaningful conclusion because of the non-standardized reporting of the studies. One may argue that sagittal balance may become more important if adjacent-level fractures develop. However, the long-term changes in sagittal balance were not a focus of the prospective RCTs, which may explain why no changes in sagittal balance were typically mentioned.

In other studies, the rate of adjacent-level fracture has also appeared to not be affected by the treatment modality[53-55]. However, an additional

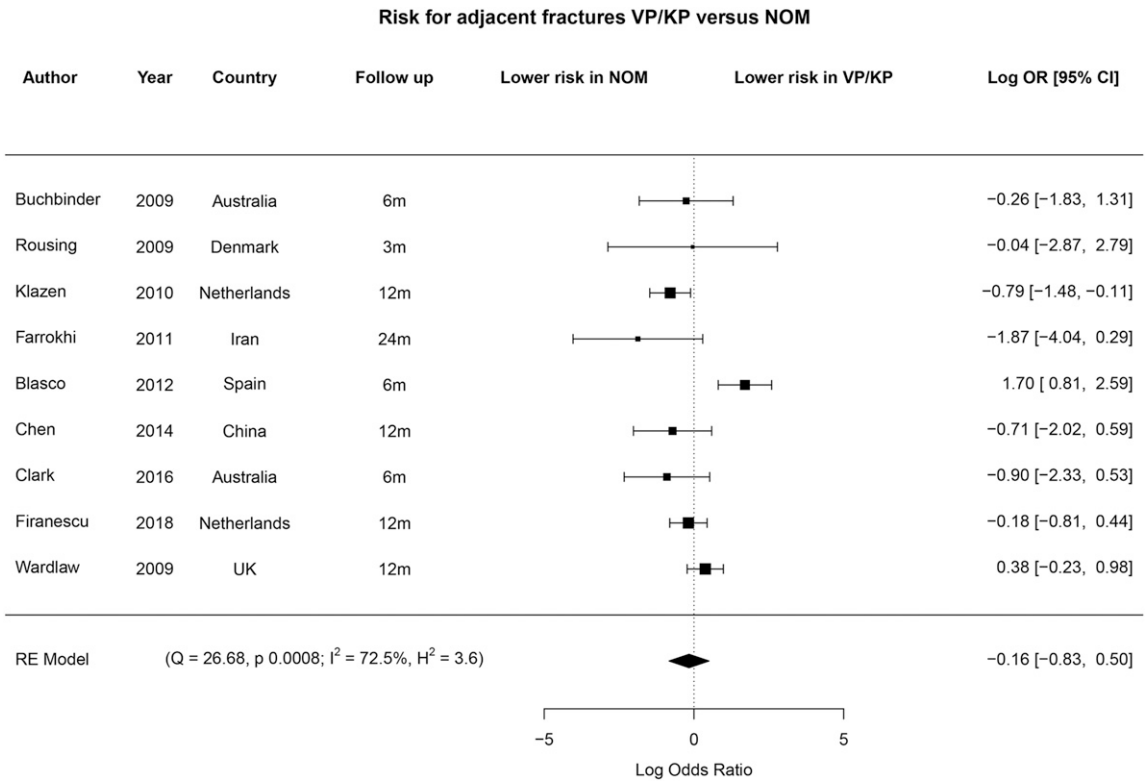

Fig. 4

VP and KP did not result in higher rates of adjacent-level fractures, and NOM did not reduce the risk of having adjacent-level fractures. (For instance, a value of 0.3 would be approximately a 2-fold increase, while a value of –0.3 would be a 2-fold decrease.) VP = vertebroplasty, KP = kyphoplasty, NOM = nonoperative management, OR = odds ratio, and CI = confidence interval.

study showed that adjacent-level fractures may occur earlier after surgical treatment even though the rate of adjacent-level fractures within 1 year remained independent of the treatment modality[56].

We were surprised to find no association between improvement of pain, or other outcome parameters, and

TABLE II Mean Difference of QOL Assessment in Short-, Mid-, and Long-Term Follow-up Following Operative Treatment Compared with NOM*

Time and Assessment Tool	Articles *(no.)*	Mean Difference, VP or KP vs. NOM	95% CI	P Value	Interpretation
Short-term: weeks					
EQ-5D	4	0.01	−0.04 to 0.06	0.497	Comparable
ODI	3	11.05	−19.75 to 41.84	0.286	Comparable
QUALEFFO	5	2.3	−7.49 to 12.17	0.596	Comparable
RDQ	9	1.7	0.01 to 3.47	0.049	Favoring VP
Midterm: months					
EQ-5D	3	−0.02	−0.38 to 0.34	0.89	Comparable
ODI	1	10.08	−14.13 to 34.29	0.224	Comparable
QUALEFFO	4	3.4	−13.36 to 20.17	0.636	Comparable
RDQ	8	1.6	−0.09 to 3.24	0.061	Trend favoring VP
Long-term: years					
EQ-5D	2	−0.08	−1.21 to 1.04	0.664	Comparable
ODI	1	10.9	−12.15 to 33.95	0.192	Comparable
QUALEFFO	4	1.9	−8.33 to 12.19	0.660	Comparable
RDQ	8	1.9	−0.68 to 4.60	0.130	Comparable

*In order to improve comparability, a positive mean difference implies a better score in the operative treatment (VP and KP) groups when compared with NOM. A negative mean difference implies a better score in the NOM group.

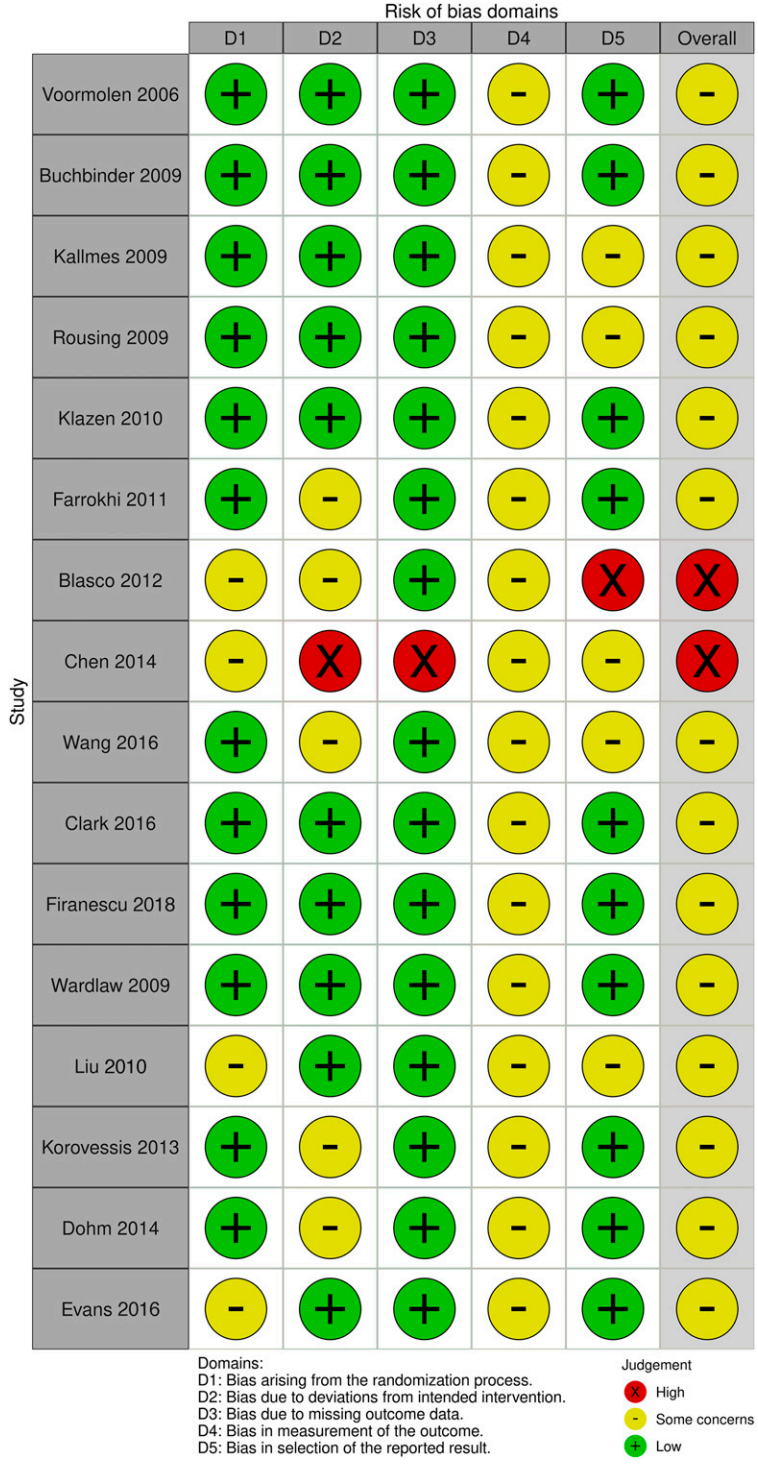

Fig. 5

Risk-of-bias assessment with use of the RoB 2 tool, visualized with use of robvis. All of the studies showed at least some concern for bias.

QOL. Because surgical intervention yields improved stability, one might have expected improvement of pain to affect QOL as well, as shown in previous studies examining early mobility and structural damage[57]. Some previous studies have indicated improvement in QOL in the short term or midterm following successful treatment of a fracture[58] (although it remained impaired in the long term in patients in whom the fracture treatment resulted in problems such as nonunion, instability, or increasing kyphosis[59]). VP has been shown to be associated with a higher

level of mobility and lower pain levels[60]. Our meta-analysis generally did not confirm a higher level of mobility. One may argue that this is a result of our statistical analyses: because the included studies provided different measures of QOL, condensing the data would lead to a loss of information. However, the summary of our data does clearly demonstrate that QOL either remains similar or improves after surgical intervention; none of the included trials found that surgical intervention was associated with a decrease in QOL or that NOM results in a higher QOL compared with surgical intervention. Because of the lack of association between QOL and improvements in other outcomes, one may hypothesize that factors other than pain alone have been more important in determining QOL, especially in a geriatric population. This has been shown convincingly with other types of osteoporotic fractures and in orthogeriatric comanagement[61]. The issue of a minimal clinically important difference (MCID) may also play a role. Our results indicated that in the short term, VP and KP are associated with a greater reduction in the VAS pain score (by 1.3 points) compared with NOM. In the medium and long-term analysis, the reduction of the VAS pain score was 0.9 points greater in the VP and KP groups compared with the NOM group. While these results were significant, some might challenge their clinical relevance[62,63]. Some of the published reports do define an improvement in the VAS pain score of 0.9 points as clinically relevant[64], but some set the threshold at 1.6 points[65] and others recommend a change of 2.5 points[66]. Similarly, the recommended threshold for the MCID for the ODI ranges between 12.8[65] and 20.0[66] to 24.0 points[67]. These discrepancies and the lack of standardized quantification methods for the MCID may have prevented the clear demonstration of clinically relevant differences.

Our study shares the general limitations of other systematic reviews and meta-analyses (e.g., the heterogeneity of the included articles). As discussed above, certain important outcome measures (e.g., sagittal balance) are inconclusively documented, thus preventing an adequate analysis of those aspects. NOM was not standardized and, in most studies, was determined by the preference of the treating physician (see Appendix). This may have contributed to the heterogeneity of the studies and, therefore, the results should be interpreted with caution[68]. Furthermore, although our primary outcome measure (the VAS pain score) represents a widely accepted tool, there remains a certain degree of bias in measurement of this outcome because the VAS score is a nonlinear parameter. These limitations have previously been reported by other authors[11]. Finally, the included studies indicated that the treated fractures were OVFs; however, only 6 trials (37.5%) provided a dual x-ray absorptiometry score, while the others failed to provide evidence of osteoporosis. The medical treatment for osteoporosis was not standardized and was thus subject to the individual treatment strategy of the treating physician (see Appendix). These limitations may have impacted the interpretation of the presented results.

Overview

We believe that the present meta-analysis provides some evidence in favor of surgical treatment for OVFs. Although we found no evidence for an increased risk of adjacent-level fractures and the reporting of the effect of treatment strategies on sagittal balance was poorly documented, most of the studies documented an improvement in pain compared with nonoperative treatment. Surgical treatment appears to be warranted if pain is a relevant problem.

Source of Funding

No external funding sources were utilized for this study.

Appendix

Supporting material provided by the authors is posted with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJSREV/A762).

Sascha Halvachizadeh, MD[1,2],
Anna-Lea Stalder[3],
David Bellut, MD[4],
Sven Hoppe, MD[5],
Philipp Rossbach, MD[6],
Alessandro Cianfoni, MD[7,8],
Klaus John Schnake, MD[9,10],
Ladislav Mica, MD, FEBS[1,2],
Roman Pfeifer, MD, FEBS[1,2],
Kai Sprengel, MD, FEBS[1,2],
Hans-Christoph Pape, MD, FACS[1,2]

[1]Department of Trauma, University Hospital Zurich, Zurich, Switzerland

[2]Harald-Tscherne Laboratory for Orthopedic and Trauma Research, University of Zurich, Zurich, Switzerland

[3]Faculty of Medicine, University of Basel, Basel, Switzerland

[4]Department of Neurosurgery, University Hospital Zurich, Zurich, Switzerland

[5]Department of Orthopedic Surgery, Inselspital University Hospital of Bern, Bern, Switzerland

[6]Department of Rheumatology, University Hospital Zurich, Zurich, Switzerland

[7]Department of Neuroradiology, Neurocenter of Southern Switzerland, Ospedale Regionaledi Lugano, Lugano, Switzerland

[8]Department of Interventional and Diagnostic Neuroradiology, Inselspital University Hospital of Bern, Bern, Switzerland

[9]Center for Spinal and Scoliosis Surgery, Malteser Waldkrankenhaus St. Marien, Erlangen, Germany

[10]Department of Orthopedics and Traumatology, Paracelsus Private Medical University Nuremberg, Nuremberg, Germany

Email for corresponding author: Sascha.Halvachizadeh@usz.ch

References

1. Renaud C. Treatment of vertebral compression fractures with the cranio-caudal expandable implant SpineJack®: Technical note and outcomes in 77 consecutive patients. Orthop Traumatol Surg Res. 2015 Nov;101(7):857-9.

2. Watts NB; GLOW investigators. Insights from the Global Longitudinal Study of Osteoporosis in Women (GLOW). Nat Rev Endocrinol. 2014 Jul;10(7):412-22.

3. Al-Sari UA, Tobias J, Clark E. Health-related quality of life in older people with osteoporotic vertebral fractures: a systematic review and meta-analysis. Osteoporos Int. 2016 Oct;27(10): 2891-900.

4. Voormolen MHJ, Lohle PN, Lampmann LE, van den Wildenberg W, Juttmann JR, Diekerhof CH, de Waal Malefijt J. Prospective clinical follow-up after percutaneous vertebroplasty in patients with painful osteoporotic vertebral compression fractures. J Vasc Interv Radiol. 2006 Aug;17(8):1313-20.

5. Liebschner MA, Rosenberg WS, Keaveny TM. Effects of bone cement volume and distribution on vertebral stiffness after vertebroplasty. Spine. 2001 Jul 15;26(14):1547-54.

6. Pateder DB, Khanna AJ, Lieberman IH. Vertebroplasty and kyphoplasty for the management of osteoporotic vertebral compression fractures. Orthop Clin North Am. 2007 Jul;38(3):409-18; abstract vii.

7. Sanli I, van Kuijk SMJ, de Bie RA, van Rhijn LW, Willems PC. Percutaneous cement augmentation in the treatment of osteoporotic vertebral fractures (OVFs) in the elderly: a systematic review. Eur Spine J. 2020 Jul;29(7): 1553-72.

8. Buchbinder R, Golmohammadi K, Johnston RV, Owen RJ, Homik J, Jones A, Dhillon SS, Kallmes DF, Lambert RG. Percutaneous vertebroplasty for osteoporotic vertebral compression fracture. Cochrane Database Syst Rev. 2015 Apr 30;(4):CD006349.

9. Salameh JP, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, Deeks JJ, Leeflang M, Korevaar DA, Whiting P, Takwoingi Y, Reitsma JB, Cohen JF, Frank RA, Hunt HA, Hooft L, Rutjes AWS, Willis BH, Gatsonis C, Levis B, Moher D, McInnes MDF. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. BMJ. 2020 Aug 14;370:m2632.

10. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016 Dec 5; 5(1):210.

11. Langley GB, Sheppeard H. The visual analogue scale: its use in pain measurement. Rheumatol Int. 1985;5(4):145-8.

12. Fairbank JC, Pynsent PB. The Oswestry disability index. Spine. 2000 Nov 15;25(22): 2940-52, discussion 2952.

13. Lips P, Cooper C, Agnusdei D, Caulin F, Egger P, Johnell O, Kanis JA, Kellingray S, Leplege A, Liberman UA, McCloskey E, Minne H, Reeve J, Reginster JY, Scholz M, Todd C, de Vernejoul MC, Wiklund I; Working Party for Quality of Life of the European Foundation for Osteoporosis. Quality of life in patients with vertebral fractures: validation of the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO). Osteoporos Int. 1999;10(2):150-60.

14. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. Appl Health Econ Health Policy. 2017 Apr;15(2):127-13.

15. Trout AT, Kallmes DF, Gray LA, Goodnature BA, Everson SL, Comstock BA, Jarvik JG. Evaluation of vertebroplasty with a validated outcome measure: the Roland-Morris Disability Questionnaire. AJNR Am J Neuroradiol. 2005 Nov-Dec;26(10):2652-7.

16. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernán MA, Hopewell S, Hróbjartsson A, Junqueira DR, Jüni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019 Aug 28;366:l4898.

17. McGuinness LA, Higgins JPT. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. Res Synth Methods. 2021 Jan;12(1): 55-61.

18. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA. Cochrane handbook for systematic reviews of interventions. Wiley-Blackwell; 2019.

19. Wan X, Wang W, Liu J. Tong. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014 Dec 19;14: 135.

20. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res. 2018 Jun;27(6): 1785-805.

21. Weir CJ, Butcher I, Assi V, Lewis SC, Murray GD, Langhorne P, Brady MC. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. BMC Med Res Methodol. 2018 Mar 7;18(1):25.

22. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2013.

23. Viechtbauer W, Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010 Aug;36(3):1-48..

24. Buchbinder R, Osborne RH, Ebeling PR, Wark JD, Mitchell P, Wriedt C, Graves S, Staples MP, Murphy B. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. N Engl J Med. 2009 Aug 6; 361(6):557-68.

25. Kallmes DF, Comstock BA, Heagerty PJ, Turner JA, Wilson DJ, Diamond TH, Edwards R, Gray LA, Stout L, Owen S, Hollingworth W, Ghdoke B, Annesley-Williams DJ, Ralston SH, Jarvik JG. A randomized trial of vertebroplasty for osteoporotic spinal fractures. N Engl J Med. 2009 Aug 6;361(6):569-79.

26. Wang B, Guo H, Yuan L, Huang D, Zhang H, Hao D. A prospective randomized controlled study comparing the pain relief in patients with osteoporotic vertebral compression fractures with the use of vertebroplasty or facet blocking. Eur Spine J. 2016 Nov;25(11):3486-94.

27. Voormolen MH, van Rooij WJ, Sluzewski M, van der Graaf Y, Lampmann LE, Lohle PN, Juttmann JR. Pain response in the first trimester after percutaneous vertebroplasty in patients with osteoporotic vertebral compression fractures with or without bone marrow edema. AJNR Am J Neuroradiol. 2006 Aug;27(7):1579-85.

28. Klazen CA, Lohle PN, de Vries J, Jansen FH, Tielbeek AV, Blonk MC, Venmans A, van Rooij WJ, Schoemaker MC, Juttmann JR, Lo TH, Verhaar HJ, van der Graaf Y, van Everdingen KJ, Muller AF, Elgersma OE, Halkema DR, Fransen H, Janssens X, Buskens E, Mali WP. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. Lancet. 2010 Sep 25;376(9746):1085-92.

29. Farrokhi MR, Alibai E, Maghami Z. Randomized controlled trial of percutaneous vertebroplasty versus optimal medical management for the relief of pain and disability in acute osteoporotic vertebral compression fractures. J Neurosurg Spine. 2011 May;14(5): 561-9.

30. Chen D, An ZQ, Song S, Tang JF, Qin H. Percutaneous vertebroplasty compared with conservative treatment in patients with chronic painful osteoporotic spinal fractures. J Clin Neurosci. 2014 Mar;21(3):473-7.

31. Clark W, Bird P, Gonski P, Diamond TH, Smerdely P, McNeil HP, Schlaphoff G, Bryant C, Barnes E, Gebski V. Safety and efficacy of vertebroplasty for acute painful osteoporotic fractures (VAPOUR): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet. 2016 Oct 1;388(10052):1408-16.

32. Firanescu CE, de Vries J, Lodder P, Venmans A, Schoemaker MC, Smeets AJ, Donga E, Juttmann JR, Klazen CAH, Elgersma OEH, Jansen FH, Tielbeek AV, Boukrab I, Schonenberg K, van Rooij WJJ, Hirsch JA, Lohle PNM. Vertebroplasty versus sham procedure for painful acute osteoporotic vertebral compression fractures (VERTOS IV): randomised sham controlled clinical trial. BMJ. 2018 May 9; 361:k1551.

33. Wardlaw D, Cummings SR, Van Meirhaeghe J, Bastian L, Tillman JB, Ranstam J, Eastell R, Shabe P, Talmadge K, Boonen S. Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet. 2009 Mar 21;373(9668):1016-24.

34. Rousing R, Andersen MO, Jespersen SM, Thomsen K, Lauritsen J. Percutaneous vertebroplasty compared to conservative treatment in patients with painful acute or subacute osteoporotic vertebral fractures: three-months follow-up in a clinical randomized study. Spine. 2009 Jun 1;34(13):1349-54.

35. Blasco J, Martinez-Ferrer A, Macho J, San Roman L, Pomés J, Carrasco J, Monegal A, Guañabens N, Peris P. Effect of vertebroplasty on pain relief, quality of life, and the incidence of new vertebral fractures: a 12-month randomized follow-up, controlled trial. J Bone Miner Res. 2012 May;27(5):1159-66.

36. United States Census Bureau. An aging nation: projected number of children and older adults. 2019 Oct 8. https://www.census.gov/library/visualizations/2018/comm/historic-first.html

37. Kluge FA, Goldstein JR. Transfers in an aging European Union. J Econ Ageing. 2019 May;13: 45-54.

38. Wagner SC, Formby PM, Helgeson MD, Kang DG. Diagnosing the undiagnosed: osteoporosis in patients undergoing lumbar fusion. Spine. 2016 Nov 1;41(21):E1279-83.

39. Diebo BG, Sheikh B, Freilich M, Shah NV, Redfern JAI, Tarabichi S, Shepherd EM, Lafage R, Passias PG, Najjar S, Schwab FJ, Lafage V, Paulino CB. Osteoporosis and Spine Surgery: A Critical Analysis Review. JBJS Rev. 2020 Jun;8(6):e0160.

40. Nicoletti G, Furnari M, Giuffrida M, Ponzo G, Iacopino DG, Cammarata G, Scalia G, Graziano F. A new tool to improve pedicle screw placement accuracy in navigated spine surgery: a monocentric study. J Neurosurg Sci. 2021 Jun;65(3):348-53.

41. Carl B, Bopp M, Saß B, Voellger B, Nimsky C. Implementation of augmented reality support in spine surgery. Eur Spine J. 2019 Jul;28(7):1697-711.

42. Epstein NE. A comparison of kyphoplasty, vertebroplasty, or non-surgical treatment of traumatic/atraumatic osteoporotic vertebral compression fractures: a short review. Surg Neurol Int. 2019 Apr 24;10:54.

43. Beall D, Lorio MP, Yun BM, Runa MJ, Ong KL, Warner CB. Review of Vertebral Augmentation: An Updated Meta-analysis of the Effectiveness. Int J Spine Surg. 2018 Aug 15;12(3):295-321.

44. Liu JT, Liao WJ, Tan WC, Lee JK, Liu CH, Chen YH, Lin TB. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. Osteoporos Int. 2010 Feb; 21(2):359-64.

45. Korovessis P, Vardakastanis K, Repantis T, Vitsas V. Balloon kyphoplasty versus KIVA vertebral augmentation—comparison of 2 techniques for osteoporotic vertebral body fractures: a prospective randomized study. Spine. 2013 Feb 15;38(4):292-9.

46. Dohm M, Black CM, Dacre A, Tillman JB, Fueredi G; KAVIAR investigators. A randomized trial comparing balloon kyphoplasty and vertebroplasty for vertebral compression fractures due to osteoporosis. AJNR Am J Neuroradiol. 2014 Dec;35(12):2227-36.

47. Evans AJ, Kip KE, Brinjikji W, Layton KF, Jensen ML, Gaughen JR, Kallmes DF. Randomized controlled trial of vertebroplasty versus kyphoplasty in the treatment of vertebral compression fractures. J Neurointerv Surg. 2016 Jul;8(7):756-63.

48. Papanastassiou ID, Phillips FM, Van Meirhaeghe J, Berenson JR, Andersson GBJ, Chung G, Small BJ, Aghayev K, Vrionis FD. Comparing effects of kyphoplasty, vertebroplasty, and non-surgical management in a systematic review of randomized and non-randomized controlled studies. Eur Spine J. 2012 Sep;21(9):1826-43.

49. Taylor RS, Taylor RJ, Fritzell P. Balloon kyphoplasty and vertebroplasty for vertebral compression fractures: a comparative systematic review of efficacy and safety. Spine. 2006 Nov 1;31(23):2747-55.

50. Beall DP, Chambers MR, Thomas S, Amburgy J, Webb JR Jr, Goodman BS, Datta DK, Easton RW, Linville D 2nd, Talati S, Tillman JB. Prospective and multicenter evaluation of outcomes for quality of life and activities of daily living for balloon kyphoplasty in the treatment of vertebral compression fractures: the EVOLVE trial. Neurosurgery. 2019 Jan 1;84(1):169-78.

51. Gu CN, Brinjikji W, Evans AJ, Murad MH, Kallmes DF. Outcomes of vertebroplasty compared with kyphoplasty: a systematic review and meta-analysis. J Neurointerv Surg. 2016 Jun;8(6):636-42.

52. Wang Z, Wang G, Yang H. Comparison of unilateral versus bilateral balloon kyphoplasty for the treatment of osteoporotic vertebral compression fractures. J Clin Neurosci. 2012 May;19(5):723-6.

53. Mudano AS, Bian J, Cope JU, Curtis JR, Gross TP, Allison JJ, Kim Y, Briggs D, Melton ME, Xi J, Saag KG. Vertebroplasty and kyphoplasty are associated with an increased risk of secondary vertebral compression fractures: a population-based cohort study. Osteoporos Int. 2009 May; 20(5):819-26.

54. Harrop JS, Prpa B, Reinhardt MK, Lieberman I. Primary and secondary osteoporosis' incidence of subsequent vertebral compression fractures after kyphoplasty. Spine. 2004 Oct 1;29(19):2120-5.

55. Lavelle WF, Cheney R. Recurrent fracture after vertebral kyphoplasty. Spine J. 2006 Sep-Oct;6(5):488-93.

56. Teuber H, Tiziani S, Halvachizadeh S, Frey D, Sprengel K, Pape HC, Osterhoff G. Single-level vertebral kyphoplasty is not associated with an increased risk of symptomatic secondary adjacent osteoporotic vertebral compression fractures: a matched case-control analysis. Arch Osteoporos. 2018 Jul 27;13(1):82.

57. López-Medina C, Garrido-Castro JL, Castro-Jiménez J, González-Navas C, Calvo-Gutiérrez J, Castro-Villegas MC, Ortega-Castro R, Escudero-Contreras A, Font-Ugalde P, Collantes-Estévez E. Evaluation of quality of life in patients with axial spondyloarthritis and its association with disease activity, functionality, mobility, and structural damage. Clin Rheumatol. 2018 Jun;37(6):1581-8.

58. DeVine J, Norvell DC, Ecker E, Fourney DR, Vaccaro A, Wang J, Andersson G. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. Spine. 2011 Oct 1; 36(21)(Suppl):S69-74.

59. Lau D, Funao H, Clark AJ, Nicholls F, Smith J, Bess S, Shaffrey C, Schwab FJ, Lafage V, Deviren V, Hart R, Kebaish KM, Ames CP; International Spine Study Group. The clinical correlation of the Hart-ISSG Proximal Junctional Kyphosis Severity Scale with health-related quality-of-life outcomes and need for revision surgery. Spine. 2016 Feb;41(3):213-23.

60. Stanghelle B, Bentzen H, Giangregorio L, Pripp AH, Bergland A. Associations between health-related quality of life, physical function and pain in older women with osteoporosis and vertebral fracture. BMC Geriatr. 2019 Nov 4;19(1):298.

61. Halvachizadeh S, Gröbli L, Berk T, Jensen KO, Hierholzer C, Bischoff-Ferrari HA, Pfeifer R, Pape HC. The effect of geriatric comanagement (GC) in geriatric trauma patients treated in a level 1 trauma setting: A comparison of data before and after the implementation of a certified geriatric trauma center. PLoS One. 2021 Jan 11;16(1):e0244554.

62. van Rijn MHC, Bech A, Bouyer J, van den Brand JAJG. Statistical significance versus clinical relevance. Nephrol Dial Transplant. 2017 Apr 1;32(suppl_2):ii6-12.

63. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007 Sep-Oct; 7(5):541-6.

64. Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? Acad Emerg Med. 1998 Nov;5(11):1086-90.

65. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. Spine J. 2008 Nov-Dec;8(6): 968-74.

66. Solberg T, Johnsen LG, Nygaard ØP, Grotle M. Can we define success criteria for lumbar disc surgery?: estimates for a substantial amount of improvement in core outcome measures. Acta Orthop. 2013 Apr;84(2):196-201.

67. Hung M, Saltzman CL, Kendall R, Bounsanga J, Voss MW, Lawrence B, Spiker R, Brodke D. What are the MCIDs for PROMIS, NDI, and ODI instruments among patients with spinal conditions? Clin Orthop Relat Res. 2018 Oct; 476(10):2027-36.

68. Longo UG, Loppini M, Denaro L, Maffulli N, Denaro V. Conservative management of patients with an osteoporotic vertebral fracture: a review of the literature. J Bone Joint Surg Br. 2012 Feb;94(2):152-7.