
Article:
Rodriguez, A.J., Fink, H.A. orcid.org/0000-0002-5985-5447, Mirigian, L. et al. (5 more authors) (2017) Pain, quality of life and safety outcomes of kyphoplasty for vertebral compression fractures: report of a task force of the American Society for Bone and Mineral Research. Journal of the Peripheral Nervous System, 32 (9). pp. 1935-1944. ISSN 0884-0431

https://doi.org/10.1002/jbmr.3170

Original Article

Pain, quality of life and safety outcomes of kyphoplasty for vertebral compression fractures: report of a task force of the American Society for Bone and Mineral Research†

Rodriguez AJ[1#], Fink HA[2, 3*#], Mirigian L[4], Guañabens N[5], Eastell R[6], Akesson K[7], Bauer DC[8], Ebeling PR[1, 9, 10]

[1] Bone and Muscle Health Research Group, Department of Medicine, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Australia
[2] Division of Epidemiology and Community Health, School of Public Health, and Department of Medicine, University of Minnesota, Minneapolis USA
[3] Geriatric Research Education and Clinical Center, Veterans Affairs Healthcare System, Minneapolis, USA
[4] American Society for Bone and Mineral Research, Washington DC, USA
[5] Rheumatology Department, Hospital Clínic. IDIBAPS. CIBERehd. University of Barcelona, Spain
[6] Academic Unit of Bone Metabolism, the Mellanby Centre for Bone Research, Northern General Hospital, the University of Sheffield, Sheffield, UK
[7] Clinical and Molecular Osteoporosis Research Unit, Department of Clinical Science, University of Lund, Malmo, Sweden
[8] School of Medicine, University of California San Francisco, San Francisco, USA
[9] Melbourne Medical School (Western Campus), University of Melbourne, St Albans, Australia
[10] Australian Institute for Musculoskeletal Science, St Albans, Australia
[#]Joint first authors

†This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/jbmr.3170]

Additional Supporting Information may be found in the online version of this article.

Initial Date Submitted December 18, 2016; Date Revision Submitted March 16, 2017; Date Final Disposition Set March 28, 2017

Journal of Bone and Mineral Research

DOI 10.1002/jbmr.3170

***Corresponding Author**

Professor Peter R. Ebeling,

Bone and Muscle Health Research Group, Department of Medicine, School of Clinical Sciences at Monash Health, Faculty of Medicine, Nursing and Health Sciences, Monash University

Level 5, Block E, Monash Medical Centre, 246 Clayton Road, Clayton, Victoria, Australia 3168
Ph: + 61 3 8572 2574 Fax: + 61 3 9594 6437
Email: peter.ebeling@monash.edu

Abstract

The relative efficacy and harms of balloon kyphoplasty (BK) for treating vertebral compression fractures (VCF) are uncertain. We searched multiple electronic databases to March 2016 for randomised and quasi-randomised controlled trials comparing BK with control treatment (non-surgical management [NSM], percutaneous vertebroplasty [PV], KIVA®, vertebral body stenting, or other) in adults with VCF. Outcomes included back pain, back disability, quality of life (QoL), new VCF and adverse events (AE). One reviewer extracted data, a second checked accuracy, and two rated risk of bias (ROB). Mean differences and 95% confidence intervals were calculated using inverse-variance models. Risk ratios of new VCF and AE were calculated using Mantel-Haenszel models. Ten unique trials enrolled 1,837 participants (age range: 61-76 years, 74% female), all rated as having high or uncertain ROB. Versus NSM, BK was associated with greater reductions in pain, back-related disability, and better QoL (k=1 trial) that appeared to lessen over time, but were less than minimally clinically important differences. Risk of new VCF at 3 and 12 months was not significantly different (k=2 trials). Risk of any AE was increased at 1 month (RR=1.73 [1.36, 2.21]). There were no significant differences between BK and PV in back pain, back disability, QoL, risk of new VCF or any AE (k=1 to 3 trials). Limitations included lack of a BK versus sham comparison, availability of only one RCT of BK versus NSM, and lack of study blinding. Individuals with painful VCF experienced symptomatic improvement compared with baseline with all interventions. The clinical importance of the greater improvements with BK versus NSM is unclear, may be due to placebo effect, and may not counterbalance short-term AE risks. Outcomes appeared similar between BK and other surgical interventions. Well-conducted randomized trials comparing BK with sham would help resolve remaining uncertainty about the relative benefits and harms of BK.

Keywords: kyphoplasty; vertebral compression fracture; osteoporosis; pain; quality of life;

ageing

Disclosure: The authors declare no competing interests

Acknowledgements: the authors would like to acknowledge Professor Mary Bouxsein (Harvard Medical School) for her assistance in the original taskforce and Ms. Catherine Shore-Lorenti (Monash University) for her efforts in assisting with the literature search. Monash University and the University of Minnesota supported this work. This study was performed by a working party from the American Society for Bone and Mineral Research Task Force charged with reporting outcomes of vertebral augmentation techniques, including balloon kyphoplasty. AJR performed literature searches, extracted data, tabulated data, performed risk of bias assessment, performed meta-analysis, drafted tables, and wrote the first draft of the manuscript. HAF identified grey literature, performed risk of bias assessment, checked data accuracy, drafted tables, edited the draft manuscript, provided clinical feedback in results interpretation and revised the final draft. LM and AJR performed literature searches and extracted data and co-ordinated the group. NG, RE, KA, and DB all assisted as part of the original taskforce. PRE oversaw the project and coordinated the group, performed risk of bias assessment, edited the draft manuscript, provided clinical feedback in results interpretation and revised the final draft. All authors approved the final draft.

Introduction

In the United States, vertebral compression fractures (VCF) account for nearly half of the approximately 1.5 million osteoporotic fractures every year (1). VCF can cause acute and chronic pain, physical impairment and disability, adversely impact quality of life, and are associated with an increased risk of future vertebral and non-vertebral fractures (2).

Medical treatments of VCF may reduce pain (3), but many patients still experience severe and sometimes long-lasting symptoms. Alternatively, several "vertebral augmentation" procedures have been developed to treat patients with symptomatic VCF. The simplest of these is percutaneous vertebroplasty (PV), in which bone cement is percutaneously injected inside the fractured vertebral body. Balloon kyphoplasty (BK) involves inflation of a balloon inside the fractured vertebral body and balloon removal before injection of bone cement. With vertebral body stenting (VBS), balloon inflation and removal is followed by insertion of an expandable scaffold before injection of bone cement (4). With KIVA®, a nesting, vertically oriented, cylindrical column is inserted over a coil into the fractured vertebral body before injection of bone cement (5).

Vertebral augmentation procedures may reduce pain and back-related disability by restoring vertebral height and stabilizing fractured vertebrae, but risks may include cement leakage associated nerve-root injury, rarely symptomatic pulmonary embolism, and a possible increase in frequency of subsequent vertebral fractures attributable to procedure-related alterations in spine biomechanics (6). Of approximately 300,000 inpatient vertebral augmentation procedures performed in the U.S. between 2005 and 2010, 73% were BK and 27% were PV (7).

Given that BK is the most commonly performed vertebral augmentation procedure, we undertook this systematic review to compare its efficacy, relative efficacy and harms

versus other treatments for VCF in middle-aged and older adults, including versus non-surgical management (NSM), sham control, PV and other vertebral augmentation techniques.

Methods

Data sources

We followed the PRISMA guidelines (8). We searched MEDLINE (from 1946), Cochrane Library, EMBASE, CINAHL, Web of Science, Clinicaltrials.gov, and the World Health Organization International Clinical Trials Registry Platform Search Portal databases to November 2014 using the title/abstract search terms: [("kyphoplasty") AND ("randomized" OR "controlled") AND NOT "malignancy" AND NOT "review"] with no language restriction. We updated this search in March 2016.

Study selection and inclusion criteria

We included randomised controlled trials (RCTs) and quasi-randomised trials that enrolled adults aged >40 years with non-traumatic vertebral fractures and compared BK versus any treatment for at least one validated measure of pain, disability, physical function, health-related quality of life, participant-reported treatment success, balance, falls, posture, bone mineral density (BMD), incident clinical or radiographic vertebral fractures, or adverse events (AE). Two reviewers independently examined titles, abstracts, and full articles for eligibility and resolved discrepancies by discussion and consensus.

Data extraction and risk of bias assessment

Two reviewers independently extracted data on study design, participant characteristics, intervention characteristics, outcomes and adverse events, and an additional reviewer checked accuracy.

Studies were evaluated for risk of bias using the Cochrane Collaboration Risk of Bias Tool (The Cochrane Collaboration, v5.1.0), with potential sources of bias including random sequence generation, allocation concealment, blinding of participants and study personnel, blinding of outcome assessments, and completeness of reported outcomes data (9). For each domain, risk of bias was rated as high, low or unclear by two independent investigators and agreed to in a consensus meeting. When there were multiple reports for a single study, information from all reports was used to rate risk of bias for each domain for that study. Risk of bias was summarized for individual studies as low (low risk of bias for all domains), unclear (unclear risk of bias for >1 domain) or high (high risk of bias for >1 domain).

Data analysis

For continuous outcomes, mean differences (MD) and 95% confidence intervals (CI) were calculated using inverse-variance models. For dichotomous outcomes, risk ratios (RR) and 95%CI were calculated using Mantel-Haenszel models. Statistical heterogeneity was determined by the I^2 statistic, applying random-effects models when $I^2 > 50\%$ (9). When there were multiple reports of a single study, the most recent publication was used as the primary data source and the other reports were used to provide supplemental information.

Results

Literature search

Together, the initial and updated database searches yielded 2,460 unique references. We excluded 2,406 during title and abstract review and 40 during full-text review, leaving 14 reports of 10 unique studies that met eligibility criteria and were included for analysis (4, 5, 10-21) [Figure 1]. Characteristics of included studies are described in Table 1.

Risk of bias

Among the 10 unique eligible studies, the most common source of bias was lack of blinding, followed by inadequate or uncertain concealment of treatment allocation, and incomplete reporting of outcomes [Table 2], though it was not possible to mask assessors of radiographic outcomes to the vertebral cement in participants assigned to BK or other vertebral augmentation procedures. No trials were rated as having low risk of bias, two were rated as having unclear risk of bias (5, 14, 22), and eight were rated as having high risk of bias (4, 10-13, 16-21).

Kyphoplasty versus Placebo or Sham

We identified no eligible trials that compared BK versus a sham BK procedure.

Kyphoplasty versus Non-surgical management

Study characteristics

Five reports met eligibility criteria, four of which were duplicate publications from the FREE trial (10, 11, 17, 20), so that there were two unique trials. Both were rated as having high risk of bias. The FREE trial (n=300) used computer generated permuted block randomization, after which participants and study staff were unblinded. Follow-up was 24 months. The Yi trial (n=200) stated that participants were randomized, but also that treatment assignment was blindly chosen by a single surgeon (21). Outcome assessors but not participants were blinded to treatment assignment and follow-up was 48 months [Table 2].

Patient characteristics

In the FREE trial, mean age was 72.2 years, and 77% of participants were female (10, 11, 17, 20). Qualifying vertebral fractures were <3 months old (mean 6 weeks) and most commonly located at the thoracolumbar junction. Vertebral fractures were attributed to

osteoporosis (approximately 40% of participants had a spine T-score of <-2.5) (10, 20). At baseline, participants reported severe back pain (mean visual analogue scale [VAS (23)] score 6.8), substantial back-related disability (mean Rolland-Morris Disability Questionnaire [RMDQ (24)]) score 17.5), and poor health-related quality of life (mean Short Form-36 Physical Component Summary Scale [SF-36 PCS (25)] score 25.7 and mean EuroQol [EQ-5D (26)] score 0.18). In Yi *et al.* (21), mean age was 61.3 years and 62% were female. Qualifying vertebral fractures were described as symptomatic. However, no data were reported on the age, location or number of vertebral fractures per participant, prevalence of osteoporosis, or participant baseline pain, disability, or quality of life.

Outcome measures

Back pain

BK was associated with significantly more reduction in pain than NSM at all time points, though the relative difference between groups in improvement in VAS appeared to diminish over time: MD in decline from baseline at 30 days (-1.82 [-2.37, -1.27]; n= 264 participants); at three months (-1.45 [-2.01, -0.89]; n=246]); at six months (-1.48 [-2.05, -0.91]; n= 241); at 12 months (-0.84 [-1.42, -0.26]; n= 226); and at 24 months (-0.69 [-1.27, -0.11]; n=200) [Supplementary Table 1].

Back-related disability

BK was associated with significantly more reduction in RMDQ than NSM at 30 days (-4.20 [-5.54, -2.86]; n=255), three months (-3.69 [-5.10, -2.28]; n=225), six months (-3.05 [-4.50, -1.60]; n=230) and 12 months (-2.90 [-4.37, -1.43]; n= 204), but not at 24 months (-1.43 [-2.91, 0.05; n=193]). The relative reduction in disability after BK compared with NSM appeared to diminish with time [Supplementary Table 2].

Quality of life

BK was associated with significantly more improvement than NSM on SF-36 PCS at 30 days (5.40 [3.14, 7.66]; n=261), three months (4.00 [1.67, 6.33]; n=241) and six months (3.30 [1.00, 5.60]; n=237), but not at 12 months (1.60 [-0.73, 3.93]; n=225) or 24 months (1.50 [-0.83, 3.83]; n=186) [Supplementary Table 3]. By comparison, BK was associated with significantly more improvement than NSM on the EQ-5D at all time points up to 24 months. For both these outcomes, the difference between groups appeared to diminish over time [Supplementary Table 4].

Incident vertebral fractures

There was no statistically significant difference between BK and NSM participants in risk of new-onset radiographic vertebral fractures occurring at one month (7.4% vs. 4.6%; risk ratio [RR]=1.59 (0.63, 4.00); events=18; n=300), three months (21.9% vs. 27.0%; RR=0.81 [0.51, 1.29]; events=54; n=223), 12 months (33.0% vs. 25.3%; RR=1.31 [0.85, 2.02]; events=62; n=220), or 24 months (47.5% vs. 44.1%; RR=1.08 [0.81, 1.44]; events=101; n=220) [Supplementary Table 5]; or at 24 months in incident adjacent radiographic vertebral fractures (23.7% vs. 16.7%; RR=1.54 [0.89, 2.65]; events= 45, n=220) or incident clinical vertebral fractures (20.8% vs. 17.9%; RR=1.07 [0.69, 1.68]; events=58; n= 300) (10, 17, 20, 21) [Supplementary Table 6].

Adverse events

The FREE study reported a significantly increased risk of any AE occurring within one month following BK compared with NSM (63.1% vs. 36.4%; RR=1.73 [1.36, 2.21]; events= 149; n= 300), with the most common AEs being back pain, new vertebral fracture, nausea or vomiting, and urinary tract infection (17). By comparison, there was no difference in risk of any AE within 24 months (89.9% vs. 88.7%; RR=1.01 [0.94, 1.10]; events= 268; n=

300) (10, 17, 20). There was no significant difference in risk of serious AE, either within one month (16.1% vs. 11.3%; RR=1.43 [0.80, 2.55]; events= 41; n= 300) or 24 months of the intervention (49.7% vs. 48.3%; RR=1.03 [0.82, 1.29]; events= 147; n= 300) (10, 17, 20) [Supplementary Table 7].

Kyphoplasty versus Vertebroplasty

Study characteristics

Six reports (12, 14, 15, 18, 19, 21) met eligibility criteria, including five unique RCTs, (n=857) and one quasi-randomized study (n=112) (13). All six trials were rated as having high risk of bias. Two trials were single-blinded (19, 21), two trials were unblinded (12, 18), and two trials had no blinding specified (13-15). Treatment allocation was performed by computerised block randomisation in two studies (12, 14, 15), assigned by the operating surgeon in two studies (13, 21), and was not specified in two studies (18, 19). Follow-up duration ranged from six to 60 months (14, 15) [Table 2].

Patient characteristics

Mean participant age was 71.6 years and about 75% were female. Qualifying VCF were acute (<6 weeks old) or subacute (6-12 weeks old) (12-15), with some studies requiring or reporting supportive findings on MRI (12, 13, 19, 21). Fractures were most commonly located near the thoracolumbar junction. Three studies limited participation to individuals who had failed several weeks of conservative therapy (13, 18, 19). Three studies were limited to or were mostly comprised of participants with osteopenia or osteoporosis (12, 18, 19). At baseline, participants reported severe back pain (mean VAS range 7.6-8.1) (12-15, 19), substantial back-related disability (mean Oswestry Disability Index [ODI (27)] range 58-66%) (12, 13, 19), and fair to poor quality of life (mean SF-36 PCS and EQ-5D approximately 28 and 0.42, respectively) (12).

Outcome measures

Back pain

In two RCTs, results favored BK over VP at 30 days VAS MD= -0.28 [-0.43, -0.13]; n=107, k=1 trial) (19) and favored VP over BK at 5 years (0.60 [0.09, 1.11]; n=100, k=1 trial) (14), but there were no statistically significant differences at other time points (12, 15, 19) [Figure 2]. In the quasi-randomised study, there were no statistically significant differences ($p<0.05$) in VAS scores between treatment groups at any follow-up time point, but there was a statistically significant difference between groups in change from baseline to two year follow-up (0.60 [0.22, 0.98]; n=86) that was small and not likely to have been clinically meaningful (13).

Back-related disability

In two RCTs (12, 19) and one quasi-randomized study (13), there was no statistically significant difference between treatments in improvement in ODI from baseline to any time points ranging between three months and two years [Supplementary Table 8].

Quality of life

There was no statistically significant difference between treatments in improvement in SF-36 PCS or EQ-5D at any time point (12) [Supplementary Table 9 & 10].

Incident vertebral fractures

There was no statistically significant difference between BK and VP in risk of incident radiographic vertebral fractures occurring within three months of intervention (23.3% vs. 27.4%, RR=0.85 [0.58, 1.26], k=1) (12), 12 months (28.3% vs. 31.5%, RR=0.89 [0.66, 1.19], k=2) (12, 19), or 24 months (49.1% vs. 57.7%, RR=0.85[0.66,1.09], k=1) (12).

Similarly, there was no significant difference in risk of incident adjacent radiographic vertebral fracture occurring up to 12 months [6.0% vs. 7.0%; RR=0.91 (0.39, 2.15); n=278; k=3] (14, 15, 18, 19), 24 months [16.0% vs. 14.0%; RR= 1.14 (0.45, 2.91); n=100; k=1] (14) or 60 months [16.0% vs. 14.0%; RR= 1.14 (0.45, 2.91); n=100; k=1] (14). There was no significant difference in risk of incident clinical vertebral fracture at one month (4.7% vs. 8.9%; RR=0.53 [0.24, 1.15]; k=1) (12), or 12 months (16.3% vs. 22.9%; RR=0.77, 0.53, 1.11, k=2) (12, 18) [Supplementary Table 11], or, in one quasi-randomized study, at two years (18.2% vs. 14.3%, RR=1.27 [0.48, 3.36], n=86) (13).

Adverse events

Only one study reported data on adverse events, and found no increased risk of serious AE at 30 days (26.2% vs. 27.4%, p=0.82) (12, 28). This trial reported the most common individual types of AEs within 30 days of surgery as procedural pain 6% for BK vs. 5% for VP, back pain 7% for BK vs. 15% for VP, and new symptomatic fracture 5% for BK vs. 9% for VP. Within 2 years of surgery, there did not appear to be a significant difference between treatment groups in risk of falls, pneumonia, bronchitis, or urinary tract infection.

Kyphoplasty versus Vertebral body stenting

One eligible trial (n=63, 100 treated levels) randomized participants to BK versus VBS, and reported only that there were no neurologic sequelae in the immediate post-operative period (4) [Table 2]. This trial was rated as having high risk of bias for lack of blinding of outcome assessment.

Kyphoplasty versus KIVA

Study characteristics

Two eligible trials randomized participants to BK versus KIVA (5, 16). One trial was blinded to participants, investigators and outcome assessors and (5) was rated as having unclear risk of bias due to unclear allocation of treatment assignment, while the other was blinded only to participants, and then only until after the procedure was completed (16), and was therefore rated as having high risk of bias. Follow-up was 14 and 12 months, respectively [Table 2].

Patient characteristics

Mean participant age was 73.7 years and 72.8% were female. Qualifying vertebral fractures were acute or subacute, and, in the one study that reported location, most commonly around the thoracolumbar junction (16). One trial limited participation to individuals who had failed conservative treatment, and reported a mean spine T-score in the osteopenic range. At baseline, participants reported severe back pain (mean VAS 8.6) and considerable back-related disability (mean ODI 63%) (16).

Outcomes

Back pain, Back-related disability, and Quality of life

There was no difference between treatment groups in the proportion of participants with >5.5 points improvement in back pain (VAS) (43% for BK vs. 54% for KIVA; RR=0.80 [0.58, 1.10] (5)), with >1.5 points improvement in back pain (VAS) (97.6% vs. 95.3% (16)), or with undefined "improved" SF-36 PCS (59% for BK vs. 51% for KIVA (5)). There also was no between-group difference in achievement of a composite endpoint requiring a 1.5 point improvement in VAS, and absence of either a >10 point worsening in ODI or a device-

related AE (16), and no between-group difference in mean improvement from baseline in back pain (VAS) (5, 16), back-related disability (ODI) (5, 16), or quality of life (SF-36 PCS) (5, 16).

Incident vertebral fractures

There was no difference between BK and KIVA participants in risk of incident radiographic vertebral fractures (12.8% vs. 12.2%, p=0.91) or incident adjacent radiographic vertebral fractures (17.1% vs. 15.7%, p=0.64) (5).

Adverse events

There was no difference in serious adverse events between BK and KIVA participants through 12 months (34.6% vs. 28.6%; RR=1.21; 0.84, 1.75; events= 80; n= 253] (16). Individual types of AEs were reported only if judged by trial investigators to be procedure-related, including three participants randomized to KIVA with herpes zoster, post-procedural pain, and pruritus, respectively, and four assigned to BK with an airway complication, back pain, ischemic stroke, and rash, respectively (16).

Discussion

On average, individuals with painful VCF experienced statistically significant symptomatic improvement compared with baseline with all studied treatment interventions, including NSM. Though we found that BK was associated with improved pain, back-related disability and quality of life outcomes compared with NSM, these results were derived almost entirely from a single trial. Further, magnitude of improvement from baseline in these outcome measures after BK relative to NSM appeared to diminish over time, mean between group differences were smaller than previously reported minimally clinically important differences in a population of individuals being treated for chronic low back pain (29), raising

concerns about their clinical significance. Because we identified no eligible trials of BK versus sham BK, it also was not possible to determine to what extent the observed improvements of BK versus NSM were attributable to a placebo effect.

Compared with NSM, BK was not associated with a statistically significantly increased risk of incident VCF, though confidence intervals were wide and results could not exclude a clinically meaningful increase in risk. Further, compared with NSM, BK was associated with a near doubling in risk of any AE within 30 days of intervention. Based on the mean between-group differences in efficacy outcomes of uncertain clinical importance, an increase in early AE, and the high risk of bias of the largest eligible BK versus NSM trial, it is uncertain whether any benefits of BK versus NSM for treatment of VCF outweigh potential harms, both in the VCF population overall and within selected patient subgroups.

We found no significant difference between BK and either PV or KIVA in pain, back-related disability, or quality of life outcomes, or in risk of incident VCF, or risk of any AE or serious AE. These results were limited by the lack of results reporting the proportion of participants in each treatment group that experienced a clinically important difference in each efficacy outcome, wide confidence intervals around the estimates for risk of incident VCF that could not exclude clinically important differences in fracture risk, and limited reporting of AE outcomes. The high risk of bias ratings of all the trials that compared BK versus PV further limits confidence in these findings.

Uncertainty about the benefits and harms of BK relative to NSM, sham BK, PV or other treatments is further complicated by uncertainty about the effect of PV. A prior systematic review found no statistically significant difference between PV and sham PV in mean change from baseline to 1 month in either pain or back-related disability (30), both overall and in two participant subgroups postulated to be more likely to benefit from PV--

those with recent VCF (<6 weeks) or with severe baseline back pain (score on 0-10 scale >8). Further analysis suggested that compared with participants randomized to sham PV, those assigned to PV may have been slightly more likely to experience clinically meaningful improvements in pain at 1 month (reduction in pain score of >3: RR=1.3 [0.8, 1.9]; reduction in pain score of >30%: RR=1.3 [1.0, 1.8]). A more recent RCT of PV versus sham PV for treatment of acute, severely painful VCF reported that participants assigned PV were significantly more likely than those in the sham PV group to have a clinically meaningful reduction in pain score between baseline and follow-up time points through 6 months (31). It is uncertain whether the apparent differences in outcomes between the earlier and more recent PV versus sham PV trials are attributable to methodological differences, including the lack of numbing of the periosteum and the greater volume of bone cement used in the recent trial. Combined with the differences in participant characteristics and in the PV intervention groups between the recent PV versus sham PV trial and those in prior BK versus PV trials, it seems unwise to make indirect comparisons between these sets of studies to draw inferences about the unstudied comparison of BK versus sham BK.

The current review was limited by available evidence. Though ten unique trials met eligibility criteria, after considering the different BK treatment comparisons, outcome measures and time points, only relatively few participants provided information about the efficacy and safety of BK versus other interventions. Second, because all but two trials reported results for efficacy outcomes only as overall group means (5, 16), it was difficult to determine how many and which types of participants achieved clinically meaningful improvements with treatment. Third, AEs and incident vertebral fractures were rarely systematically reported and often were not reported at all, particularly for individual types of AEs. This hampered our ability to weigh the relative harms of BK against any potential benefits. Fourth, most trials were rated as having high risk of bias, most commonly due to

lack of blinding of participants and/or outcome assessors, and less often due to a lack of allocation concealment, both which could have led to overestimation of the true effect of interventions..

In conclusion, we found that in middle-aged and older adults with VCF, based on only a single RCT, BK was associated with greater improvement in pain, disability and quality of life and an increase in risk of early AE compared with NSM. However, the magnitude of treatment differences may have been too small to be clinically meaningful, diminished over time, and likely was attributable at least in part to a placebo effect. Based on a small number of heterogeneous (and high risk of bias) studies, there was no difference in these outcomes between BK and either the PV or KIVA vertebral augmentation techniques. Risks of subsequent fracture were not statistically significantly different between BK and other treatments, but results could not rule out important differences. These remaining areas of uncertainty should be addressed by future trials that randomize participants with acute and subacute painful VCF to BK versus sham BK, PV or KIVA®, mask study participants, investigators and outcome assessors, follow participants for at least one year, include adequately powered responder analyses for efficacy outcomes (e.g. proportion achieving a clinically important improvement in back pain), systematically collect results for clinical and radiographic vertebral fracture and AE, and include *a priori* subgroup analyses for patient groups with characteristics suspected to modify the effects of treatment.

References

1. Riggs BL, Melton LJ, 3rd. The worldwide problem of osteoporosis: insights afforded by epidemiology. Bone. 1995;17(5 Suppl):505s-11s.
2. Ross PD. Clinical consequences of vertebral fractures. The American journal of medicine. 1997;103(2a):30S-42S; discussion S-3S.
3. Prather H, Watson JO, Gilula LA. Nonoperative management of osteoporotic vertebral compression fractures. Injury. 2007;38(3, Supplement):40-8.
4. Werner CM, Osterhoff G, Schlickeiser J, Jenni R, Wanner GA, Ossendorf C, et al. Vertebral body stenting versus kyphoplasty for the treatment of osteoporotic vertebral compression fractures: a randomized trial. J Bone Joint Surg Am. 2013;95(7):577-84.
5. 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;38(4):292-9.
6. Ensrud KE, Schousboe JT. Clinical practice. Vertebral fractures. The New England journal of medicine. 2011;364(17):1634-42.
7. Goz V, Errico TJ, Weinreb JH, Koehler SM, Hecht AC, Lafage V, et al. Vertebroplasty and kyphoplasty: national outcomes and trends in utilization from 2005 through 2010. The spine journal : official journal of the North American Spine Society. 2015;15(5):959-65.
8. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009;151(4):264-9, w64.
9. JPT Higgins, Green S. Cochrane Handbook for Systematic Reviews of Interventions The Cochrane Collaboration; 2011. Available from: www.cochrane-handbook.org.
10. Boonen S, Van Meirhaeghe J, Bastian L, Cummings SR, Ranstam J, Tillman JB, et al. Balloon kyphoplasty for the treatment of acute vertebral compression fractures: 2-year results from a randomized trial. J Bone Miner Res. 2011;26(7):1627-37.
11. Borgstrom F, Olafsson G, Strom O, Tillman JB, Wardlaw D, Boonen S, et al. The impact of different health dimensions on overall quality of life related to kyphoplasty and non-surgical management. Osteoporosis International. 2013;24(7):1991-9.
12. Dohm M, Black CM, Dacre A, Tillman JB, Fueredi G, investigators K. A randomized trial comparing balloon kyphoplasty and vertebroplasty for vertebral compression fractures due to osteoporosis. Ajnr: American Journal of Neuroradiology. 2014;35(12):2227-36.
13. Du J, Li X, Lin X. Kyphoplasty versus vertebroplasty in the treatment of painful osteoporotic vertebral compression fractures: two-year follow-up in a prospective controlled study. Acta Orthopaedica Belgica. 2014;80(4):477-86.
14. Liu JT, Li C, Chang CS, Liao WJ. Long-term follow-up study of osteoporotic vertebral compression fracture treated using balloon kyphoplasty and vertebroplasty. Journal of neurosurgery. 2015;Spine. 23(1):94-8.
15. Liu JT, Liao WJ, Tan WC, Lee JK, Liu CH, Chen YH, et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. Osteoporosis International. 2010;21(2):359-64.
16. Tutton SM, Pflugmacher R, Davidian M, Beall DP, Facchini FR, Garfin SR. KAST Study: The Kiva System As a Vertebral Augmentation Treatment-A Safety and Effectiveness Trial: A Randomized, Noninferiority Trial Comparing the Kiva System With Balloon Kyphoplasty in Treatment of Osteoporotic Vertebral Compression Fractures. Spine. 2015;40(12):865-75.
17. Van Meirhaeghe J, Bastian L, Boonen S, Ranstam J, Tillman JB, Wardlaw D. A randomized trial of balloon kyphoplasty and nonsurgical management for treating acute vertebral compression fractures: Vertebral body kyphosis correction and surgical parameters. Spine. 2013;38(12):971-83.

18. Vogl TJ, Pflugmacher R, Hierholzer J, Stender G, Gounis M, Wakhloo A, et al. Cement directed kyphoplasty reduces cement leakage as compared with vertebroplasty: results of a controlled, randomized trial. Spine. 2013;38(20):1730-6.
19. Wang CH, Ma JZ, Zhang CC, Nie L. Comparison of high-viscosity cement vertebroplasty and balloon kyphoplasty for the treatment of osteoporotic vertebral compression fractures. Pain physician. 2015;18(2):E187-94.
20. Wardlaw D, Cummings S, Van Meirhaeghe J, Bastian L, Tillman J, Ranstam J, et al. A randomized trial of balloon kyphoplasty and nonsurgical care for acute vertebral compression fracture. Osteoporosis International. 2009;20:S273-S4.
21. Yi XD, Lu HL, Tian F, Wang Y, Li CD, Liu H, et al. Recompression in new levels after percutaneous vertebroplasty and kyphoplasty compared with conservative treatment. Archives of Orthopaedic and Trauma Surgery. 2014;134(1):21-30.
22. Liu JT, Liao WJ, Tan WC, Lee JK, Liu CH, Chen YH, et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. Osteoporos Int. 2010;21(2):359-64.
23. DeLoach LJ, Higgins MS, Caplan AB, Stiff JL. The visual analog scale in the immediate postoperative period: intrasubject variability and correlation with a numeric scale. Anesthesia and analgesia. 1998;86(1):102-6.
24. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine (Phila Pa 1976). 1983;8(2):141-4.
25. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care. 1992;30(6):473-83.
26. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Annals of medicine. 2001;33(5):337-43.
27. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy. 1980;66(8):271-3.
28. KAVIAR Study - Kyphoplasty And Vertebroplasty In the Augmentation and Restoration of Vertebral Body Compression Fractures. In: ClinicalTrials.gov, editor. 2013.
29. Maughan EF, Lewis JS. Outcome measures in chronic low back pain. European Spine Journal. 2010;19(9):1484-94.
30. Staples MP, Kallmes DF, Comstock BA, Jarvik JG, Osborne RH, Heagerty PJ, et al. Effectiveness of vertebroplasty using individual patient data from two randomised placebo controlled trials: meta-analysis. BMJ (Clinical research ed). 2011;343:d3952.
31. Clark W, Bird P, Gonski P, Diamond TH, Smerdely P, McNeil HP, et al. Safety and efficacy of vertebroplasty for acute painful osteoporotic fractures (VAPOUR): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet (London, England). 2016.

Figure Legends

Figure 1.

Study selection flow diagram.

Figure 2.

Mean changes in back pain scores from baseline at 1, 3, 6, 12 and >12 months in kyphoplasty versus percutaneous vertebroplasty study groups.

Figure 3.

Incident radiographic vertebral fractures at 3, 12, and >12 months (number of events through to time point) in kyphoplasty versus percutaneous vertebroplasty study groups.

Figure 4.

Incident adjacent radiographic vertebral fractures at >12 months (number of events through to time point) in kyphoplasty versus percutaneous vertebroplasty study groups.

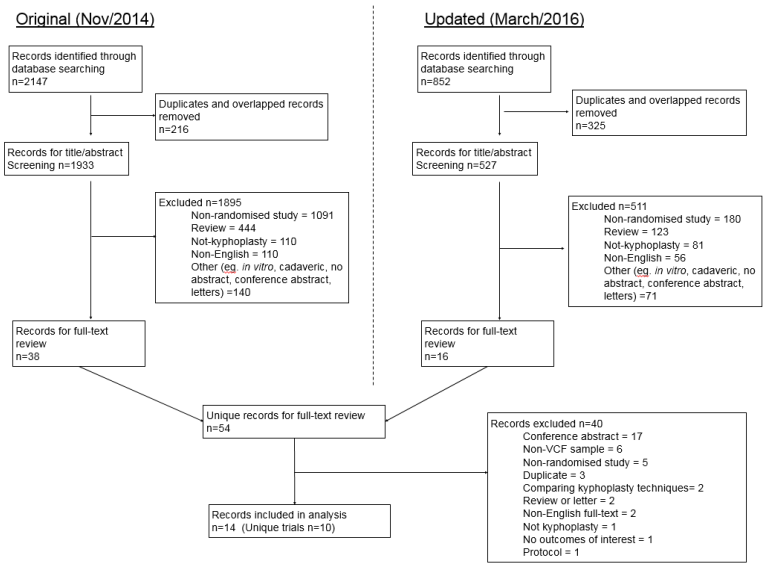

Figure 1

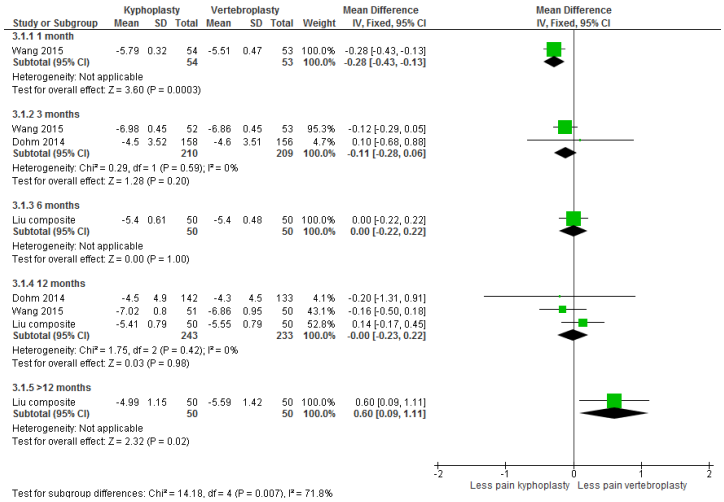

Figure 2

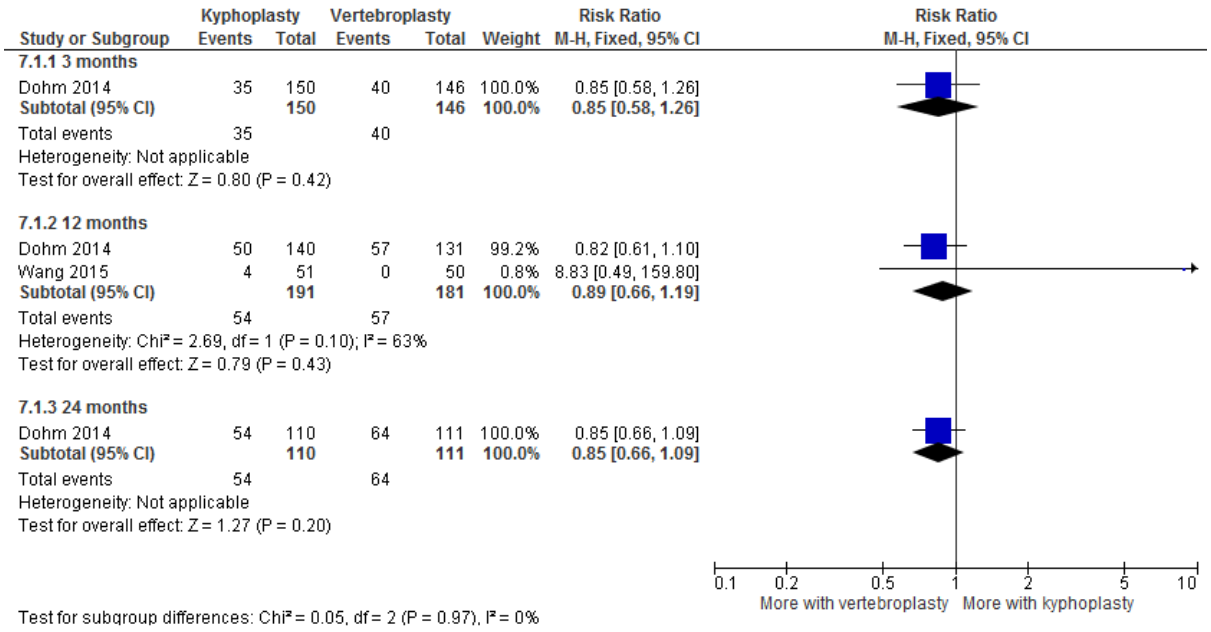

Figure 3

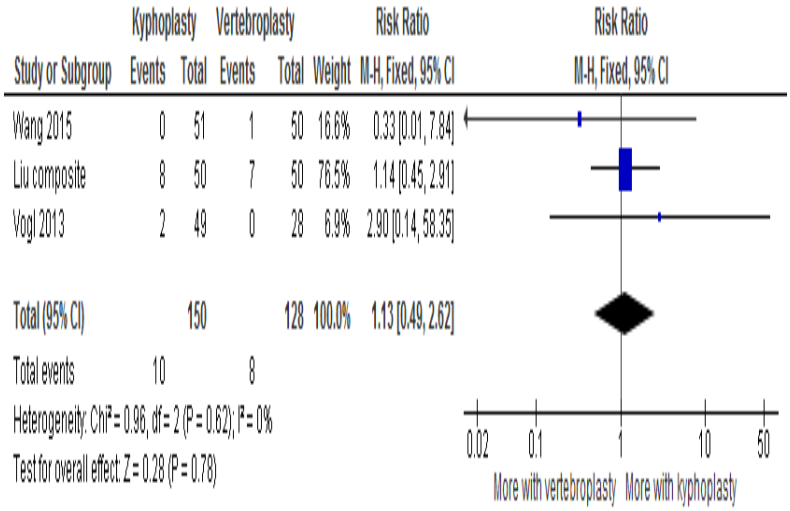

Figure 4

Table 1. Study Characteristics

Study	Blinding	Follow-up (mo)	Centres	Countries	Participants	Randomisation	Outcome Measures*	Funding
BK vs. NSM								
Wardlaw 2009, Boonen 2011, Borgstrom 2013, Van Meirhaeghe 2013 (FREE Trial)	Unblinded	24	Multiple	Belgium, Germany, Sweden, Scotland	300	Computer-generated permuted block	SF-36 PCS*, EQ5D, VAS, analgesic use RMDQ, patient satisfaction, TUG test, new VCF, kyphosis angle, RAD, QALY, vertebral height, AE	Medtronic Spine LLC
Yi 2014	Single blind (outcome assessor)	48	Single	China	200	Both stated "randomized" and "surgeon…blindly chose…treatment"	New VCF*, AE	No disclosures
BK vs. PV								
Du 2006	Not specified	24	Single	China	86	Quasi-randomized (divided according to the surgeon)	VAS, ODI, vertebral height, participant satisfaction, kyphotic angle, new VCF	No disclosures
Liu 2010, 2015	Single blind (radiology techs)	60	Single	Taiwan	100	Permuted block	Vertebral height, kyphotic angle, VAS, new VCF	University/ Hospital grant
Vogl 2013	Single blind (participants)	12	Multiple	Germany, USA	77	Stated only as "randomized"	Cement leakage*, vertebral height, new VCF	Soteira Inc.
Dohm 2014 (KAVIAR Trial)	Single blind (radiology techs)	24	Single	USA	404	Computer generated dynamic minimisation	New VCF*, SF-36 PCS, EQ-5D, ODI, VAS, kyphosis correction, AE	Medtronic Spine LLC
Wang 2015	Blinded (participants, radiologists)	12	Single	China	107	Stated only as "randomized"	VAS, ODI, cement leakage, vertebral height restoration rate, new VCF, AE	No disclosures
BK vs. VBS								
Werner 2013	Unblinded	Post-op	Single	Switzerland	100	Computer generated block	Kyphotic angle*, cement leakage, material-related complications	No disclosures
BK vs. KIVA								
Korovessis 2013	Blinded (participants, investigators, outcome assessors)	14	Single	Greece	163	Stated only as "randomized"	Cement leakage, vertebral height, kyphotic angle, VAS, SF-36, ODI	No disclosures
Tutton 2015 (KAST Trial)	Single blind (participants only until after procedure)	12	Multiple	USA, EU	300	Computer generated block	Composite (reduced VAS by 15mm on 100mm VAS, improved or ≤10 point worsening ODI, absence of device-related SAE)*, VAS, ODI, new VCF	Benvenue Medical, Inc

AE= adverse event; BK= balloon kyphoplasty; EQ5D= EuroQol 5 dimensions; NSM= non-surgical management; ODI= Oswestry disability index; PV= percutaneous vertebroplasty; QALY= quality adjusted life year; RAD= restricted activity days; RMDQ= Rolland Morris Disability questionnaire; SAE= serious adverse event; SF-36 PCS= short form 36 physical component summary score; TUG= timed up and go; VAS= visual analogue scale; VBS= vertebral body stenting; VCF= vertebral compression fracture
*Specified as primary outcome measure.

Table 2. Study characteristics of included literature

Report (Year)	Trial ID	Blinding	Follow-up (mo)	Centred	Country(ies)	n	Randomisation	Intervention	Approach	Comparator	Primary Outcome	Secondary Outcomes	Funding
Yi	-	Single blind	48	Single	China	200	Surgeon decision	BK	Bipedicular	NSM	New VCF	AE	No disclosures
Werner	-	Unblind	Post-op	Single	Switzerland	100	5xBlock of 20	BK	Bipedicular	VBS	Post-op change Cobb's angle	Complications	No disclosures
Wardlaw	FREE (NCT00211211)	Unblind	12	Multiple	Belgium, Germany, Sweden, Scotland	300	Permuted block 1:1	BK	Bipedicular	NSM	Difference in change in 1 month SF36-PCS	EQ5D, VAS, analgesic use RMQD, Patient satisfaction, TUG test, fracture, Cobb's angle, RAD, vertebral body height, AE	Medtronic Spine LLC
Wang	-	Single blind	12	Single	China	107	n/sp	HVCV	Unipedicular	BK	VAS, ODI	Compression ratio, leakage	No disclosures
Vogl	-	Unblind	12	Multiple	Germany, USA	77	n/sp	BK		PV	Cement leakage	Vertebral height, new Fx, other complications	Soteira Inc. (Natick, MA)
Van Meirhaeghe	FREE (NCT00211211)	Unblind	24	Multiple	Belgium, Germany, Sweden, Scotland	300	Permuted block 1:1	BK	Bipedicular	NSM	1-month SF36-PCS	EQ5D, VAS, RMQD, Patient satisfaction, TUG test, fracture, Cobb's angle, vertebral body height, AE	Medtronic Spine LLC
Tutton	KAST (NCT01123512)	Single blind	12	Multiple	USA (n=15); EU (n=6)	253/300	Stratified by site, number of intended treatment levels level; Block allocation	KIVA	Bipedicular	BK	Composite (reduce VAS by 15mm on 100mm VAS; maintain/improve ODI; absence of SAE)	Volume bone cement used, VAS score, ODI, adjacent level fracture rates	Benvenue Medical, Inc
Liu 2015	-	n/a	60	Single	Taiwan	100	"Divided equally"	BK	Bipedicular	PV	Vertebral body height, Cobb's angle and VAS	n/r	Uni/Hospital grant
Liu 2010	-	n/a	6	Single	Taiwan	100	Permuted block 1:1	BK	Bipedicular	PV	Vertebral body height, Cobb's angle and VAS	n/r	Uni/Hospital grant
Korovessis	-	Single blind	14	Single	Greece	163	Block	KIVA	Bipedicular	BK	Anterior vertebral body height ratio (AVBHr), midline vertebral body height ratio segmental kyphotic angle (Gardner angle)(MVBHr), posterior vertebral body height ratio (PVBHr), and	VAS, SF36, ODI	No disclosures
Du	-	n/a	24	Single	China	86	Surgeon decision	BK	Bipedicular	PV	VAS, ODI	Vertebral height, patient satisfaction, Cobb's angle, new VCF	No disclosures
Dohm	KAVIAR (NCT00323609)	Unblind	24	Single	USA	404	Computer-generated dynamic minimisation	BK	Bipedicular	PV	New radiographic VCF	SF-36, EQ5D, RMQD, ODI, VAS,	Medtronic Spine LLC
Borgstrom	FREE (NCT00211211)	Unblind	24	Multiple	Belgium, Germany, Sweden, Scotland	300	Permuted block 1:1	BK	Bipedicular	NSM	Difference in change in 1 month SF36-PCS	EQ5D, VAS, analgesic use RMQD, Patient satisfaction, TUG test, fracture, Cobb's angle, RAD, vertebral body height, AE, QALY	Medtronic Spine LLC
Boonen	FREE (NCT00211211)	Unblind	24	Multiple	Belgium, Germany, Sweden, Scotland	300	Permuted block 1:1	BK	Bipedicular	NSM	Difference in change in 1 month SF36-PCS	EQ5D, VAS, analgesic use RMQD, Patient satisfaction, TUG test, fracture, Cobb's angle, RAD, vertebral body height, AE	Medtronic Spine LLC

AE= adverse event; BK= balloon kyphoplasty; EQ5D= EuroQol 5 dimensions; EU=European Union; HVCV= high viscosity cement vertebroplasty; Inc.= incorporated; LLC= limited liability company; n/a = not applicable; n/r= not reported; NSM= non-surgical management; n/sp= not specified; ODI= Oswestry disability index; PV= percutaneous kyphoplasty; QALY= quality adjusted life year; RAD; restricted activity days; RMQD; Rolland Morris Disability questionnaire; SAE= serious adverse event; SF36(PCS)= short form 36 physical component summary; TUG= timed up and go; USA= United States of America; VAS= visual analogue scale; VBS= vertebral body stenting; VCF= vertebral compression fracture