Washington State Health Care Authority

Vertebroplasty, Kyphoplasty, Sacroplasty – Rereview

Draft Evidence Report

August 30, 2024

Health Technology Assessment Program (HTA)
Washington State Health Care Authority
PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126
www.hca.wa.gov/about-hca/hta
shtap@hca.wa.gov

Vertebroplasty, Kyphoplasty, Sacroplasty – Rereview

Provided by:

Aggregate Analytics, Inc.

Prepared by:

Andrea C. Skelly, PhD, MPH
Erika D. Brodt, BS
Rongwei (Rochelle) Fu, PhD
Yun Yu, MS
Shay Stabler-Morris, MSc
Dakota Riopelle, MPH
Asmaa Watson, PhD

Internal clinical, methods review: Roger Chou, MD

August 30, 2024

This technology assessment report is based on research conducted by a contracted technology assessment center, with updates as contracted by the Washington State Health Care Authority. This report is an independent assessment of the technology question(s) described based on accepted methodological principles. The findings and conclusions contained herein are those of the investigators and authors who are responsible for the content. These findings and conclusions may not necessarily represent the views of the HCA/Agency and thus, no statement in this report shall be construed as an official position or policy of the HCA/Agency.

The information in this assessment is intended to assist health care decision-makers, clinicians, patients, and policy makers in making sound evidence-based decisions that may improve the quality and cost-effectiveness of health care services. Information in this report is not a substitute for sound clinical judgment. Those making decisions regarding the provision of health care services should consider this report in a manner similar to any other medical reference, integrating the information with all other pertinent information to make decisions within the context of individual patient circumstances and resource availability.

Aggregate Analytics, Inc. is a contract research organization whose team has over fifteen years of experience in performing health technology assessments, comparative effectiveness reviews, and systematic reviews for a variety of clients based on accepted methodologic standards for such research. AAI's mission is to assist healthcare professionals and organizations in the objective synthesis and generation of evidence to improve future healthcare delivery by providing timely, methodologically rigorous, transparent services and quality evidence synthesis products.

Contents

Tables

Figures

ABBREVIATIONS

AE = adverse events
CI = confidence interval
EQ-5D = EuroQol 5-dimension
HR = hazard ratio
JOA = Japanese Orthopedic Association
KP = kyphoplasty
MD = mean difference
NEADL = Nottingham Extended Activities of Daily Living
NRS = numerical rating scale
ODI = Oswestry Disability Index
OQOLS = Osteoporosis Quality of Life Scale
OR = odds ratio
PMMA = polymethylmethacrylate
QALY = quality-adjusted life year
QHES = The Quality of Health Economic Studies
QoL = quality of life
QUALEFFO = Quality of Life Questionnaire of the European Foundation for Osteoporosis
RCT = randomized controlled trial
RDQ = Roland Morris Disability Questionnaire
ROB = risk of Bias
RR = risk ratio
SAE = serious adverse event
SB = single blinded
SD = standard deviation
SDF-ADL = Study of Osteoporotic Fractures Activities of Daily Living
SF-36 = Short form-36
SF-36 MCS = Short form-36 Mental component score
SF-36 PCS = Short form-36 Physical component score
SF-MPQ = Osteoporosis Quality of Life Scale
SIF = sacral insufficiency fracture
SMD = standardized mean difference
SOE = strength of evidence
SP = sacroplasty
SR = systematic review
SSED = summary of safety and effectiveness data
U.S. = United States
US-FDA = United States Food and Drug Administration
VA/DoD = Veterans Affairs/Department of Defense
VAS = visual analog scales
VCF = vertebral compression fracture
VP = vertebroplasty
WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index
WTP = willingness to pay

Executive Summary

Introduction

Vertebral compression fractures (VCFs) and sacral insufficiency fractures (SIF) often result in considerable pain, loss of function, and decreased quality of life. Patients with osteopenic vertebral or sacral fractures are at greater risk of morbidity and mortality, yet operative intervention (e.g., fusion with instrumentation) may be problematic in this elderly population, making less invasive methods more attractive. VCFs can also occur due to metastatic bone disease leading to disability and morbidity and again, operative interventions may not be feasible.

Vertebroplasty, kyphoplasty and sacroplasty (collectively, percutaneous vertebral and sacral surgery) are minimally invasive surgical procedures used to treat spinal pain believed to be caused by fractures in the vertebra or sacrum. These are all cementoplasty (augmentation) techniques intended to stabilize the fractured bone(s), but the mechanism of pain relief is not clear. Osteoporosis, vertebral metastasis, and multiple myeloma are the most frequently reported indications for these procedures. Cementoplasty may reduce pain and improve stability of the bone.

Vertebroplasty involves injection of bone cement into a partially collapsed vertebral body under computed tomography (CT) or fluoroscopic guidance. Kyphoplasty is a modification of vertebroplasty that expands the partially collapsed vertebral body with an inflatable balloon or other mechanical device before the injection of bone cement. Sacroplasty is an extension of vertebroplasty, involving the injection of bone cement into the sacrum to repair sacral insufficiency fractures.

These surgical procedures are less invasive than other spinal surgical procedures, but more invasive than conservative medical therapy. Vertebroplasty, kyphoplasty and sacroplasty are surgical procedures and are not subject to FDA approval, however materials and devices used as part of these procedures are subject to FDA approval.

Policy Context/Reason for Selection

A Health Technology Assessment titled: ***Vertebroplasty, Kyphoplasty, Sacroplasty***, was published on November 5, 2010 by the Health Care Authority. New evidence has been published subsequent to the 2010 review and additional devices have been FDA approved. The Committee's Coverage Decision is summarized below.

HTCC Coverage Determination
Vertebroplasty, Kyphoplasty and Sacroplasty are not covered benefits.
HTCC Reimbursement Determination
Vertebroplasty, Kyphoplasty and Sacroplasty are not covered benefits.

Objectives

The aim of this report is to systematically review, critically appraise, analyze and synthesize research evidence evaluating the effectiveness and safety of vertebroplasty, kyphoplasty, and sacroplasty. The differential effectiveness and safety of these therapies for subpopulations were also evaluated, as was the cost effectiveness.

Key Questions and Scope

When used in patients with spinal pain ***due to vertebral fracture***:

1. What is the evidence of efficacy and effectiveness of vertebroplasty, kyphoplasty or sacroplasty? Including consideration of:
    a. Short-term and long-term outcomes
    b. Impact on function, pain, quality of life
    c. Other reported measures including: use of pain medications and opioids, return to work
2. What is the evidence of the safety of vertebroplasty, kyphoplasty or sacroplasty? Including consideration of:
    a. Adverse events type and frequency (mortality, major morbidity, other)
    b. Revision/re-operation rates (if not addressed in efficacy)
3. What is the evidence that vertebroplasty, kyphoplasty or sacroplasty has differential efficacy or safety issues in sub populations? Including consideration of:
    a. Gender
    b. Age
    c. Psychological or psychosocial co-morbidities
    d. Diagnosis or time elapsed from fracture
    e. Other patient characteristics or evidence-based patient selection criteria
    f. Provider type, setting or other provider characteristics
    g. Payer/beneficiary type: including worker's compensation, Medicaid, state employees
4. What is the evidence of cost implications and cost-effectiveness of vertebroplasty, kyphoplasty and sacroplasty? Including consideration of:
    a. Costs (direct and indirect) in the short term and over expected duration of use
    b. Revision/re-operation (if not addressed in efficacy)

PICOTS/Scope:

Study Component	Inclusion	Exclusion
Population	Patients with spinal pain due to ***vertebral fracture*** secondary to • Osteoporosis • Malignancy Subgroups, special populations: • Gender • Age • Psychological or psychosocial co-morbidities • Diagnosis or time elapsed from fracture • Other patient characteristics or evidence-based patient selection criteria • Provider type, setting or other provider characteristics • Payer/beneficiary type: including worker's compensation, Medicaid, state employees	• Fractures due to high energy trauma

Study Component	Inclusion	Exclusion
Intervention	• Vertebroplasty • Kyphoplasty • Sacroplasty	• Cements, devices that are not FDA approved unless being studied in a Phase III trial • Spineoplasty graft consisting of mesh filled with bone chips instead of the traditional cement • Percutaneous cement discoplasty (PCD) - intervertebral disc is filled with percutaneously injected acrylic cement; may be used as prep or with vertebroplasty • Studies of exercise/rehab post augmentation • Stentoplasty, vertebral body stenting • Vesselplasty
Comparator	• Sham procedure or placebo • Conservative care, conventional care • Other minimally invasive procedures (e.g., facet joint block, nerve block) • Surgical procedures • Vertebroplasty vs. kyphoplasty	• Comparisons of different cement types • Comparisons of surgical approaches or techniques • Comparison of different vertebroplasty techniques with each other or different forms of kyphoplasty with each other • Use of vertebroplasty, kyphoplasty or sacroplasty as an adjunct to other procedures (e.g., ablation) • Augmentation combined with zoledronic acid (ZOL) versus augmentation alone • Types of imaging guidance, other guidance, e.g., Robotic assisted vs. fluoroscopy • Stentoplasty/vertebral body stenting, • Vesselplasty
Outcomes	**Primary outcomes** • Functional outcomes (e.g., ODI) • Pain relief • Harms/Complications (e.g., procedure related, leakage, new fracture, medical complications, mortality, revision/re-operation) **Secondary outcomes** • Quality of life • Measures of disability (e.g., work lost) • Opioid use • Return to work/return to normal activity	• Measures that are not validated • Intermediate outcomes measures (e.g., radiographic measures of disc height)
Study design	• **Key Question 1**: Comparative clinical studies with a focus on studies with least potential for bias (RCTs); NRSI with concurrent controls that control for confounding will be considered if RCT evidence is not available for KQ 1.	• Case reports • Case series, single arm studies, pre-post studies with fewer than 5 patients (for sacroplasty) • NRSIs for effectiveness or benefit for osteoporotic fractures (KQ1)

Study Component	Inclusion	Exclusion
	• **Key Question 2**, safety, RCTs, NRSI with ≥250 patients that are specifically designed to evaluate safety that control for confounding will be considered; case series will be considered if adequate information is not available from comparative NRSIs and RCTs or for rare or long-term adverse events; systematic reviews may be considered for safety • **Key Question 3:** RCTs only • **Key Question 4:** Full formal economic studies	• NRSI that do not control for confounding (exception for sacroplasty)
Publication	• Full-length studies published in English in peer reviewed journals, published HTAs or publicly available FDA reports • Full formal economic analyses (e.g., cost-utility studies) published in English in HTAs or in a peer-reviewed journal published after those represented in previous HTAs	• Abstracts, editorials, letters • Duplicate publications of the same study which do not report on different outcomes • Single reports from multicenter trials • Studies reporting on the technical aspects of these procedures • White papers • Narrative reviews • Articles identified as preliminary reports when results are published in later versions • Incomplete economic evaluations such as costing studies

FDA = Food and Drug Administration; HTA = health technology assessment; NRSI = nonrandomized studies of interventions; ODI = Oswestry Disability Index; QOL = quality of life.

Methods

The scope of this report and final key questions were refined based on input from clinical experts. Clinical expert input was sought to confirm critical outcomes on which to focus. Draft Key Questions (KQs) and PICOTS scope were published on the HCA website for public comment. Comments were reviewed and considered for the finalization of the KQs and scope and citations were evaluated for inclusion based on the final KQs and scope. Comments from clinical experts and peer-reviewers as well as public comments will be considered for finalization of this report.

A formal, structured systematic search of the peer-reviewed literature was performed across multiple databases including PubMed and EMBASE to identify relevant peer reviewed literature as well as other sources (e.g., ECRI Guideline Trust) to identify pertinent clinical guidelines and previously performed assessments. We hand searched the reference lists of relevant studies and the bibliographies of systematic reviews. Studies were selected for inclusion based on pre-specified criteria detailed in the full report.

All records were screened by two independent reviewers; discrepancies were resolved by consensus. Selection criteria included a focus on studies with the least potential for bias that were written in English and published in the peer-reviewed literature. Included studies reporting on primary outcomes of

interest were critically appraised independently by two reviewers evaluating the methodological quality, study limitations and potential for bias based on study design as well as factors which may bias studies using defined templates and pre-specified criteria.

The method used by Aggregate Analytics, Inc. (AAI) for assessing the quality of evidence of individual studies as well as the overall strength of evidence (SOE) are based on established methods for systematic reviews. Included studies reporting on primary outcomes of interest were critically appraised independently by two reviewers evaluating the methodological quality, study limitations and potential for bias based on study design as well as factors which may bias studies using defined templates and pre-specified criteria. Assessment of RCTs followed appropriate criteria[87] based on methods described in *the Cochrane Handbook for Systematic Reviews of Interventions*[47] and guidance from the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews.*[72] In keeping with the AHRQ methods, each study was given a final rating of "good", "fair", or "poor" quality as described below. Discrepancies in ratings between reviewers were resolved through discussion and consensus. Economic studies were evaluated according to The Quality of Health Economic Studies (QHES) instrument developed by Ofman et al.[67] in conjunction with consideration of epidemiologic principles that may impact findings.

SOE was assessed by two researchers following the principles for adapting GRADE (Grading of Recommendations Assessment, Development and Evaluation) as outlined by the Agency for Healthcare Research and Quality (AHRQ).[2,43,44,72] The SOE was based on the highest quality evidence available for the primary outcomes. In determining the strength of body of evidence regarding a given outcome, the following domains were considered:

- **Risk of bias**: the extent to which the included studies have protection against bias.
- **Consistency:** the degree to which the included studies report results that are similar in terms of effect sizes, range and variability.
- **Directness**: describes whether the evidence is directly related to patient health outcomes or comparisons of interventions are direct (head-to-head).
- **Precision:** describes the level of certainty surrounding the effect estimates.
- **Publication or reporting bias:** is considered when there is concern of selective publishing or selective reporting. This is difficult to assess particularly for nonrandomized studies.

Bodies of evidence consisting of RCTs are initially considered as High SOE. In general, the GRADE and AHRQ methodologies initially consider nonrandomized studies of interventions (NRSIs) as Low SOE as such studies typically are at higher risk of bias due to lack of randomization and inability of investigators to control for critical confounding factors. The SOE could be downgraded based on the limitations described above. There are also situations where studies (particularly observational studies) could be upgraded if the study had large magnitude of effect (strength of association) or if a dose-response relationship is identified and there are no downgrades for the primary domains listed above and confounding is not a concern. Publication and reporting bias are difficult to assess, particularly with fewer than 10 RCTs and for observational studies.[9,77] Publication bias was unknown in all studies and thus this domain was eliminated from the SOE tables. The final SOE was assigned an overall grade of high, moderate, low, or insufficient, which are defined as follows:

- High - Very confident that effect size estimates lie close to the true effect for this outcome; there are few or no deficiencies in the body of evidence; we believe the findings are stable.
- Moderate – Moderately confident that effect size estimates lie close to the true effect for this outcome; some deficiencies in the body of evidence; we believe the findings are likely to be stable but some doubt remains.
- Low – Limited confidence that effect size estimates lie close to the true effect for this outcome; major or numerous deficiencies in the body of evidence; we believe that additional evidence is needed before concluding that findings are stable or that the estimate is close to the true effect.
- Insufficient – We have no evidence, are unable to estimate an effect or have no confidence in the effect estimate for this outcome; OR no available evidence or the body of evidence has unacceptable efficiencies precluding judgment.

Methods for quantitative analysis are described in the full report. Briefly, meta-analyses were conducted using profile likelihood methods and focused on the primary outcomes. To determine the appropriateness of meta-analysis, we considered clinical and methodological diversity and assessed statistical heterogeneity. Sensitivity analyses were considered excluding poor-quality trials, outlying data and related to clinical heterogeneity.

Results

From 4,456 unique citations identified from electronic database searches, hand searching and bibliography review of included studies, a total of 32 RCTs (in 41 publications)[7,11-13,15,17,22,24,27,29,31,32,35-37,42,45,50,52,56-58,60-62,74,75,79,84,86,88-93,95,100,102] met our inclusion criteria (Figure 1): 31 RCTS (in 40 publications)[11-13,15,17,22,24,27,29,31,32,35-37,42,45,50,52,56-58,60-62,74,75,79,84,86,88-93,95,100,102] on osteoporotic fracture and one RCT[7] on fracture due to malignancy. We additionally included: one SR[16] for effectiveness of sacroplasty; six comparative NRSIs controlling for confounding[1,3,5,39,76,101] for effectiveness; nine additional comparative NRSIs[18,34,38,49,55,68,70,96,97] for safety; and 30 case series[4,6,8,10,14,19,25,26,28,33,41,46,51,53,54,59,64-66,71,73,78,81,85,94,98,99,103-105] for safety. The tables below provide an overview of each of these groups of studies by fracture cause (osteoporosis, malignancy, sacral insufficiency) and treatments compared and provides information on the funding source. The most common comparators for vertebroplasty for efficacy were kyphoplasty (9 RCTs, 1 NRSI), usual care (9 RCTs), sham (6 RCTs), and nerve block (2 RCTs, 1 NRSI); The most common comparators for kyphoplasty for efficacy were, usual care (5 RCTs), and other surgical procedures (1 RCT). Of the included RCTs, roughly one fifth (21%) reported industry funding; Furthermore, nearly one third (30%) were not clear about their funding source.

There is substantial new evidence available for this update report compared with the 2010 HTA (see table below). In addition to new evidence from RCTs, longer term follow-up from previously include trials is now available as are more recent studies of cost effectiveness.

	2010 HTA	New or Updated RCTs for2024 Report
VP vs. Sham	2 RCTs	**6 RCTs**
VP vs. UC	3 RCTs	**9 RCTs**
KP vs. Sham	0	**0**
KP vs. UC	1 RCT	**5 RCTs**
VP vs. KP	1 RCT	**9 RCTs**
VP vs. Nerve Block	0	**2 RCTs**
KP vs. Other Surgical Intervention	0	**1 RCT**
Sacroplasty	0	**0**
TOTAL	7	**32**

Results are organized by indication for augmentation, i.e., osteoporosis, tumor/malignancy and sacroplasty, with results from RCT evidence on effectiveness and safety (KQs 1 and 2) described by type of augmentation and comparator. Evidence for differential efficacy or safety (KQ 3) is presented separately and was only available in patients with osteoporotic VCFs for comparisons of VP with sham or usual care and for VP versus KP. Findings for cost-effectiveness (KQ 4) follow the evidence for KQ 3. Evidence presentation for this executive summary focuses on RCT evidence where is it is available as that from NRSIs was insufficient due to study limitations in addition to uncertainty regarding the precision for some outcomes. For sacroplasty, all evidence was from NRSIs and insufficient. Data from NRSIs are detailed in the full report.

Osteoporosis

Vertebroplasty (VP)

Vertebroplasty versus Sham KQ 1 and 2 Table A

Effectiveness:

- Pain response: VP was associated with a greater likelihood of improving baseline VAS pain (0-10 scale) response at most time points, with the exception of ≥1 to ≤2 weeks. There was a large likelihood of response (large effect size) at the earliest time (<1 week) with the likelihood of response favoring VP decreasing over suggestive time frames.
- Pain improvement based on VAS scores (0-10) was similar for VP and sham at the earliest two time frames (up to ≤2 weeks) and at the longest follow up (≥12 months, SOE low). VP was associated with small pain improvement and intermediate times compared to sham treatment.
- VP was associated with small improvements in function versus sham based on the Roland-Morris Disability Questionnaire (RDQ, 0-24 scale) at two time frames: >2 weeks to ≤1 month, and at ≥6 to <12 months (SOE Low). Scores between groups were similar at other times.

Safety:

- Risk of mortality, new vertebral fractures and serious adverse events (SAE) were similar for VP and sham across RCTs
- Cement leakage was common following VP with a range across RCTs of 40% to 91% of treated levels. Authors do not report on whether symptoms were present or on related complications.

Table A. Summary of effectiveness and safety evidence for vertebroplasty versus sham in patients with osteoporotic vertebral compression fractures

Effect/Improvement favors VP unless otherwise indicated

Outcomes*	<1 week	≥1 to ≤2 weeks	>2 weeks to ≤1 month	>1 to <6 months	≥6 to <12 months	≥12 months
Pain Response (≥30% improvement from baseline)	Large likelihood, 1 RCT, N=113 (SOE: Low)	Similar likelihood, 2 RCTs, N=186 (SOE: Moderate)	Moderate likelihood, 3 RCTs, N=313 (SOE: Moderate)	Small likelihood, 2 RCTs, N=176 (SOE: Moderate)	Small likelihood, 2 RCTs, N=171 (SOE: Moderate)	Small likelihood, 3 RCTs, N=339 (SOE: Moderate)
VAS pain scores (0-10)	Similar, 4 RCTs, N=500 (SOE: Low)	Similar, 6 RCTs, N=616 (SOE: Moderate)	Small, 6 RCTs, N=616 (SOE: High)	Small, 6 RCTs, N=605 (SOE: High)	Small, 5 RCTs, N=550 (SOE: High)	Similar, 5 RCTs, N=478 (SOE: Low)
RDQ function scores (0-24)	Similar, 2 RCTs, N=244 (SOE: Low)	Similar, 5 RCTs, N=531 (SOE: Low)	Small, 5 RCTs, N=566 (SOE: Moderate)	Similar, 5 RCTs, N=557 (SOE: Low)	Small, 5 RCTs, N=548 (SOE: Low)	Similar, 4 RCTs, N=432 (SOE: Low)
Mortality	Similar, 5 RCTs, N=589, at last follow-up (12-24 months) (SOE: Moderate)					
Any new vertebral fracture	Similar, 4 RCTs, N=408, at last follow-up (6-24 months) (SOE: Moderate)					
Any new symptomatic fracture with bone edema	Similar, 1 RCT, N=34, 12 months (SOE: Low)					
Any SAE	Similar, 4 RCTs, N=409, at last follow-up (3-12 months) (SOE: Low)					
Cement leakage, any	Common after VP, 3 RCTs, N=232 levels, any time (SOE: Moderate)					

RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; SAE = serious adverse event; SOE = strength of evidence; VAS = visual analog scale; VP = vertebroplasty.
*SOE for adverse events focused on cumulative event to last follow-up; data on earlier timepoints if provided is available in the report.

Vertebroplasty versus Usual Care KQ 1 and 2: Table B

Effectiveness:

- VP was associated with large or moderate pain improvement based on VAS scores (0-10) versus usual care for all time frames except ≥6 to <12 months when it was similar for VP and UC.
- VP was associated with small improvements in function based the RDQ (0-24 scale), Oswestry Disability Index (ODI, 0-100), and Dallas Pain Questionnaire Daily Activities (DPQDA, 0-100) for all time frames ≥1 week. No studies reported on function at the earliest time (1 week)

Safety:

- Similar risks were seen for VP and UC for the following: Mortality, new vertebral fractures in general, new symptomatic vertebral fractures, SAEs and reoperation.
- Asymptomatic cement leak was common (13.0%-72.4%; 49.3%-72.4% across the fair-quality trials) and symptomatic cement leak following VP much less common. (0%-1%) Authors do not on related complications.

Table B. Summary of effectiveness and safety evidence for vertebroplasty versus usual care in patients with osteoporotic vertebral compression fractures

Effect/Improvement favors VP unless otherwise indicated

Outcomes*	<1 week	≥1 to ≤2 weeks	>2 weeks to ≤1 month	>1 to <6 months	≥6 to <12 months	≥12 months
Pain Response (<4 on 0-10 VAS)	No evidence	No evidence	No evidence	No evidence	No evidence	INSUFFICIENT
Pain Response (Complete relief)	No evidence	No evidence	No evidence	No evidence	No evidence	INSUFFICIENT
VAS/NRS pain scores (0-10)	Large, 3 RCTs, N=343 (SOE: Moderate)	Moderate, 4 RCTs, N=432 (SOE: Low)†	Large, 3 RCTs, N=398 (SOE: Low)	Moderate, 5 RCTs, N=569 (SOE: Moderate)	Similar, 4 RCTs, N=523 (SOE: Low)	Moderate, 5 RCTs, N=567 (SOE: Low)
Function scores‡	No evidence	Small, 4 RCTs, N=432 (SOE: Low)	Small, 3 RCTs, N=398 (SOE: Moderate)	Small, 4 RCTs, N=440 (SOE: Moderate)	Small, 3 RCTs, N=398 (SOE: Moderate)	Small, 4 RCTs, N=436 (SOE: Moderate)
Mortality	Similar, 6 RCTs, N=844, at last follow-up (6-12 months) (SOE: Moderate)					
Any new vertebral fracture	Similar, 9 RCTs, N=830, at last follow-up (2 weeks to 49 months) (SOE: Low)					
Any new symptomatic vertebral fracture	Similar, 6 RCTs, N=877, at last follow-up (2 weeks to 12 months) (SOE: Low)					
SAEs	Similar, 4 RCTs, N=408, any time (SOE: Low)					
Reoperation	Similar, 1 RCT, N=211, any time (SOE: Low)					
Cement leak, symptomatic	Rare with VP, 7 RCTs, N=661 levels, any time (SOE: Moderate)					
Cement leak, asymptomatic	Common with VP, 7 RCTs, N=661 levels, any time (SOE: Moderate)					

RCT = randomized controlled trial; RCT = randomized controlled trial; SAE = serious adverse event; SOE = strength of evidence; VAS = visual analog scale; VP = vertebroplasty.

*SOE for adverse events focused on cumulative event to last follow-up; data on earlier timepoints if provided is available in the report.

†After exclusion of potential outlier trial (Blasco 2012)

‡Standardized mean difference across three measures of function: Roland Morris Disability Questionnaire (RDQ, 0-24), Oswestry Disability Index (ODI, 0-100), and Dallas Pain Questionnaire Daily Activities (DPQDA, 0-100).

Vertebroplasty versus Kyphoplasty KQ 1 and 2

Effectiveness:

- Pain improvement (VAS 0-10 scale) was similar for VP and KP at all time frames for which there was sufficient evidence to assess this.
- Improvement in function was also similar between VP and KP all time frames for which there was sufficient evidence to assess this.

Safety:

- Similar risk for mortality and new vertebral fracture were seen for VP and KP
- Both symptomatic cement leakage and cement embolism were rare and similar between VP and KP recipients.

Table C. Summary of effectiveness and safety evidence for vertebroplasty versus kyphoplasty in patients with osteoporotic vertebral compression fractures

Effect/Improvement favors VP unless otherwise indicated

Outcomes*	<1 week	≥1 to ≤2 weeks	>2 weeks to ≤1 month	>1 to <6 months	≥6 to <12 months	≥12 months
Pain Response (total effective rate)†	No evidence	No evidence	No evidence	No evidence	No evidence	INSUFFICIENT
VAS/NRS pain scores (0-10)	Similar, 3 RCTs, N=313 (SOE: Moderate)	INSUFFICIENT	Similar, 2 RCTs, N=460 (SOE: Low)‡	Similar, 2 RCTs, N=419 (SOE: Low)‡	Similar, 3 RCTs, N=248 (SOE: Low)	*12-24 months*: Similar, 5 RCTs (N=673) (SOE: Low) *60 months*: INSUFFICIENT
Function scores§	Similar, 1 RCT, N=106 (SOE: Low)	No evidence	INSUFFICIENT	Similar, 2 RCTs, N=399 (SOE: Low)‡	Similar, 3 RCTs, N=238 (SOE: Moderate)	*12 months*: Similar, 5 RCTs (N=643) (SOE: Low) *24 months*: INSUFFICIENT
Mortality	Similar, 4 RCTs, N=631, at latest follow-up (12-24 months) (SOE: Low)					
Any new vertebral fracture	Similar, 6 RCTs, N=781, at latest follow-up (12-49 months) (SOE: Low)					
Cement leak, symptomatic	Similar and rare, 5 RCTs, N=800, any time (SOE: Low)					
Cement embolism, any	Similar and rare, 2 RCTs, N=381, any time (SOE: Low)					
Any new symptomatic vertebral fracture	INSUFFICIENT					
Refracture or worsening at index level	INSUFFICIENT					
SAEs, any and procedure or device related	INSUFFICIENT					
Reoperation for new fracture	INSUFFICIENT					

NRS = numerical pain rating scale; RCT = randomized controlled trial; RCT = randomized controlled trial; SAE = serious adverse event; SOE = strength of evidence; VAS = visual analog scale; VP = vertebroplasty.

*SOE for adverse events focused on cumulative event to last follow-up; data on earlier timepoints if provided is available in the report

†Complete ("cure"), excellent or effective (not defined) improvement in clinical symptoms

‡After exclusion of potential outlier trial (Wang 2023)

§Standardized mean difference across two measures of function: Roland Morris Disability Questionnaire (RDQ, 0-24) and Oswestry Disability Index (ODI, 0-100).

Vertebroplasty versus Medial Branch Nerve or Facet Blocks

Effectiveness:

- While VP was associated with moderate pain improvement (VAS 0-10 scale) versus medial branch nerve or facet blocks at early times (<1 week, ≥1 to ≤2 weeks) improvement at later time frames was similar between groups.
- VP was associated with moderate improvement in function based on the RDQ (0-24 scale) at <1 week and substantial improvement at ≥1 to ≤2 weeks versus medial branch nerve or facet block, however, improvement was similar between groups at later time frames.

Safety:

- Risk of new vertebral fractures was similar for both treatment groups

Table D. Summary of effectiveness and safety evidence for vertebroplasty versus medial branch nerve or facet blocks in patients with osteoporotic vertebral compression fractures

Effect/Improvement favors VP unless otherwise indicated

Outcomes	<1 week	≥1 to ≤2 weeks	>2 weeks to ≤1 month	>1 to <6 months	≥6 to <12 months	≥12 months
VAS/NRS pain scores (0-10)	Moderate, 1 RCT, N=206 (SOE: Low)	Moderate, 2 RCTs, N=233 (SOE: Low)	Similar, 2 RCTs, N=230 (SOE: Low)	Similar, 1 RCT, N=206 (SOE: Low)*	Similar, 1 RCT, N=206 (SOE: Low)	Similar, 1 RCT, N=206 (SOE: Low
RDQ function scores (0-24)	Moderate, 1 RCT, N=206 (SOE: Low)	Large, 1 RCT, N=206 (SOE: Low)*	Similar, 2 RCTs, N=230 (SOE: Low)	Similar, 2 RCTs, N=227 (SOE: Low)	Similar, 1 RCT, N=206 (SOE: Low)	Similar, 1 RCT, N=206 (SOE: Low)
New vertebral fractures	Similar, 1 RCT, N=206, 12 months (SOE: Low)					
Cement leak, asymptomatic	INSUFFICIENT					

NRS = numerical pain rating scale; RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; SOE = strength of evidence; VAS = visual analog scale; VP = vertebroplasty.
*Based on the large, fair-quality trial (Wang 2016).

Kyphoplasty (KP)

Kyphoplasty versus Usual Care

Effectiveness:

- All results are from a single large RCT
- Compared with UC, KP was associated with substantial pain improvement (VAS 0-10 scale) at ≥1 to ≤2 weeks and diminished between >2 weeks to <12 months to moderate and to a small improvement at ≥ 12 months
- Moderate functional improvement was seen with KP at two intermediate time frames (>2 weeks to ≤1 month) and >1 to <6 months, with a small improvement seen for times ≥6 to ≥12 months compared with usual care. Function was similar between groups at 34 months

Safety:

- Risk of mortality, new vertebral fractures and serious adverse events (SAE)

Table E. Summary of effectiveness and safety evidence for kyphoplasty versus usual care in patients with osteoporotic vertebral compression fractures

Effect/Improvement favors KP unless otherwise indicated

Outcomes*	<1 week	≥1 to ≤2 weeks	>2 weeks to ≤1 month	>1 to <6 months	≥6 to <12 months	≥12 months
VAS/NRS pain scores (0-10)	INSUFFICIENT	Large, 1 RCT, N=300 (SOE: Low)†	Moderate, 1 RCT, N=300 (SOE: Low)†	Moderate, 2 RCTs, N=380 (SOE: Low)	Moderate, 1 RCT, N=300 (SOE: Low)†	*12, 24 months* Small, 1 RCT, N=300 (SOE: Low)
Function scores‡	INSUFFICIENT	INSUFFICIENT	Moderate, 1 RCT, N=300 (SOE: Low)†	Moderate, 1 RCT, N=300 (SOE: Low)†	Small, 1 RCT, N=300 (SOE: Low)	*12 months* Small, 1 RCT, N=300 (SOE: Low) *24 months* Similar, 1 RCT, N=300 (SOE: Low)
Mortality	Similar, 1 RCT, N=300, 24 months (SOE: Low)					
Any SAE	Similar, 2 RCTs, N=500, at last follow-up (24-49 months) (SOE: Low)					
Treatment-related SAEs	Similar, 1 RCT, N=300, 30 days (SOE: Low)					
Withdrawals due to AEs	Similar, 1 RCT, N=300, 24 months (SOE: Low)					
New vertebral fracture	Similar, 1 RCT, N=300, 24 months (SOE: Low)					
New symptomatic vertebral fracture	Similar, 2 RCTs, N=500, at last follow-up (24-49 months) (SOE: Low)					
Cement leak, symptomatic	Not uncommon, 2 RCTs, N=228 KP, at last follow-up (24-49 months) (SOE: Low)					
Reoperation for new symptomatic fracture	Similar, 1 RCT, N=300, 24 months (SOE: Low)					

AE = adverse event; KP = kyphoplasty; NRS = numerical pain rating scale; RCT = randomized controlled trial; RCT = randomized controlled trial; SAE = serious adverse event; SOE = strength of evidence; VAS = visual analog scale.

*SOE for adverse events focused on cumulative event to last follow-up; data on earlier timepoints if provided is available in the report.

†Based on the large, fair-quality trial only (FREE trial; Wardlaw 2009, Van Meirhaeghe 2023).

‡Standardized mean difference across two measures of function: Roland Morris Disability Questionnaire (RDQ, 0-24) and Oswestry Disability Index (ODI, 0-100).

Vertebral Compression fractures due to Tumors or Malignancy

Kyphoplasty versus Usual Care KQ1 and 2

Effectiveness:

- Limited evidence from one RCT showed large improvement in pain with KP versus usual care up to 1 month in patients with pathologic fracture due to malignancy. After 1 month, there was substantial crossover from UC to KP (58%); effectiveness results prior to crossover are reported here.
- KP was associated with a large likelihood of functional response >2 weeks to ≤1 month versus usual care based on RDQ (≥2.5-point improvement 0 -24 scale) and Karnofsky Performance Scores (KPS 0-100 scale) at two thresholds (≥5-point improvement, score of ≥70)
- KP was also associated with improved RDQ and KPS scores >2 weeks to ≤1 month.

Safety:

- There was a very high rate of crossover (58%) after 1 month assessment; Safety outcomes are reported prior to crossover (1 month) and based on author's reported ITT analyses.
- Risks of mortality and SAEs were similar between KP and UC at one month and from 1 to ≤12 months.
- New symptomatic fracture risks were similar between groups at 1 month, KP recipients were at higher risk of this between 1 and 12 months.
- Symptomatic cement leak with KP was rare, occurring in 1.4%.

Table F. Summary of effectiveness and safety evidence for kyphoplasty versus usual care in patients with vertebral compression fractures due to tumors or malignancy

Effect/Improvement favors KP unless otherwise indicated

Outcomes*	<1 week	≥1 to ≤2 weeks	>2 weeks to ≤1 month	>1 to ≥12 months
VAS/NRS pain scores (0-10)	No evidence	Large, 1 RCT, N=117 (SOE: Low)	Large, 1 RCT, N=114 (SOE: Low)	No evidence*
Function Responders (≥2.5-point improvement on RDQ)	No evidence	No evidence	Large, 1 RCT, N=113 (SOE: Low)	No evidence*
Function Responders (≥5-point improvement on KPS)	No evidence	No evidence	Large, 1 RCT, N=112 (SOE: Low)	No evidence*
Function Responders (KPS score ≥70)	No evidence	No evidence	Moderate, 1 RCT, N=112 (SOE: Low)	No evidence*
RDQ function scores (0-24)	No evidence	No evidence	Large, 1 RCT, N=113 (SOE: Low)	No evidence*
KPS function scores (0-100)	No evidence	No evidence	Large, 1 RCT, N=112 (SOE: Low)	No evidence*
Mortality	Similar, 1 RCT, N=134, 1 month (SOE: Low) Similar, 1 RCT, N=96, >1 to ≤12 months, ITT (SOE: Low)			
SAEs	Similar, 1 RCT, N=134, 1 month (SOE: Low) Similar, 1 RCT, N=96, >1 to ≤12 months, ITT (SOE: Low)			
New symptomatic fracture	Similar, 1 RCT, N=134, 1 month (SOE: Low) Risk greater with KP, 1 RCT, N=96, >1 to ≤12 months, ITT (SOE: Low)			
Cement leak, symptomatic	Rare, 1 RCT, N=70 in KP, 1 month (SOE: Low)			

KP = kyphoplasty; KPS = Karnofsky Performance Score; NRS = numerical pain rating scale; RCT = randomized controlled trial; RDQ = Roland Morris Disability Questionnaire; SAE = serious adverse event; SOE = strength of evidence; VAS = visual analog scale.
*Due to very high rate of crossover (58%) after 1 month assessment, only outcomes at 1 month or earlier were included for effectiveness. For safety, SOE focused on ITT analysis (patients as randomized) between 1 and 12 months

Vertebroplasty versus Kyphoplasty KQ 1 and 2

The evidence base comparing VP and KP in patients with malignant vertebral fractures remains sparse and insufficient due to high risk of bias, unknown consistency and imprecision for these studies. Three retrospective comparative NRSIs (2 from the prior report,[38,55] 1 newly identified[5]) evaluated the effectiveness of PV and KP for malignant vertebral fracture are summarized in the full report.

- Across the comparative NRSIs, pain response and pain improvement for VP and KP were similar.
- Adverse events were sparsely reported. No neurological or pulmonary complications or new fractures were observed; one death was reported. Detail of adverse events from case series are found in the full report.

Sacroplasty

Sacroplasty versus Nonsurgical Management (Usual Care) and Surgery KQ1 and 2

The evidence base evaluating the effectiveness and safety of sacroplasty remains sparse and insufficient due to high risk of bias, unknown consistency and imprecision for these studies. Studies are summarized in the full report

- Sacroplasty conferred greater improvement in VAS pain scores (0-10 scale) across most timepoints versus usual care (3 studies).[1,39,76] and function scores (ODI scale 0-100)at all timepoints (2 studies).[1,76] Mortality was less common following sacroplasty (1 study).
- One study reported significantly less improvement in ODI scores after sacroplasty compared with daily percutaneous teriparatide injections.[101]
- One study compared sacroplasty with screw fixation (primarily iliosacral screw fixation with cement augmentation)[1] and found that patients in both groups experienced significant improvement in pain (VAS scores) and function (HBI scores) Data were not well reported. HBI function scores at 2 years were similar. Mortality was similar for the two groups.

KQ 3 Vertebroplasty Differential Effectiveness or Safety

Evidence on differential effectiveness or harms of VP reported in included RCTs for subpopulations defined by gender, age, psychological or psychosocial co-morbidities, provider characteristics, or payer type or by fracture age, pain duration or intervention characteristics is sparse. In addition to data from three trials of VP, [13,27,50] results discussed below include those from an AHRQ review[21] that reported stratified analyses for VP. These analyses included all but one of the RCTs of VP versus sham or usual care that are included in this HTA update. No RCTs of KP reported stratified analyses for subpopulations. One RCT comparing VP with KP briefly described such analyses. [32] Analyses in all trials were likely to have low power for detecting effect modification by factors that were evaluated. Confidence in findings from stratified analyses from included studies is very low.

- ***VP versus Sham or Usual Care***
    - Fracture age/pain duration: There does not appear to be modification of the treatment effect for vertebroplasty (versus sham) in patients with acute osteoporotic fractures, compared with those with more chronic fractures, based on reported subgroup analyses from included RCTs or from reported stratified analyses of RCTs comparing VP to a combined UC and sham across RCTs for the outcomes of pain or function.
    - Other factors:
        - No modification of treatment effect based on sex, presence or absence of pervious fractures or treating center was reported by one RCT of VP versus sham
        - For the outcomes of pain and function, there appears to be no modification of treatment effect based for the following subgroups based on stratified analysis of RCTs comparing VP with sham or usual care: PMMA volume, study enrollment requirement of MRI findings of bone marrow edema
        - In stratified analysis of RCTs of VP, control type appeared to modify treatment at 2-4 weeks with a smaller difference in effect size for pain observed in trials with sham control versus usual care as a control. Interaction between control types was not statistically significant for function.
- ***VP versus KP***
    - One RCT comparing VP versus KP reported that no appreciable differences in the magnitude of pain reduction were seen for subgroup analysis on sex, age, preoperative pain scores or preoperative RDQ scores. Authors do not provide data or p-values for interaction.

KQ 4 Cost-effectiveness

Six full economic studies relevant to populations with osteoporotic vertebral compression fractures[30,40,48,80,82,83] and one relevant to cancer-related VCF[69]were identified for this update report. Two studies were U.S. based[30,48]. Both were industry funded (Medtronic). Of the non-US based studies two were reported by government entities one from the UK[80] the other from Canada Health Quality Ontario, 2016 #13}. One[83] was performed in Japan and received no funding. The other two were performed in Sweden[40] and the UK[82] and were industry funded (Medtronic). No economic studies on sacroplasty were identified. Given the differences in healthcare systems and reimbursement policies between the U.S. and other countries, the generalizability of findings from studies from outside of the U.S. is unclear.

In general, most economic studies suggest that vertebral augmentation may be cost effective versus nonoperative conventional management based on conventional willingness to pay (WTP) thresholds.

- The only cost-utility analysis (CUA) in patients with malignant VCF was performed by Health Quality Canada.[69] It concluded that KP and VP may be cost-effective with ICERs of Canadian Dollars (CAD) $33,471/QALY gained for KP and CAD $17,870/QALY gained for VP, both in comparison to nonsurgical management. It was rated as good quality.
- The highest quality, most comprehensive analysis was performed by the UK National Institute for Health Research (NIHR)[80] and included data from both sham controlled trials and unblinded trials comparing VP and KP with usual care in patients with osteoporotic VCF. Based on extensive sensitivity analyses, including consideration of whether sham involving local anesthetic might be considered a more “active” control, authors conclude that ICERs are driven by the clinical scenarios chosen:
    - KP was consistently cost-effective (at WTP below £20,000) if modeling included differential mortality benefit versus UC. When no mortality benefit was assumed, the method for utility determination influenced cost/QALY
    - ICERs for VP and KP were often greater than £20,000 when blinded trials were used
    - PV was constantly cost effective at ICER below £20,000 when a pooled beneficial effect was used.
    - Authors note that while vertebral augmentation may lead to decreased mortality, the data for this is from administrative data (registry) and that causal inference is not possible given lack of detailed information on causes of death.
- Both US based studies (one poor quality, the other good quality), and one good quality UK-based study relied at least in part on Medicare Claims data and similar methods to model mortality based on studies that suggest lower mortality for vertebral augmentation compared to nonoperative management in their base-cases. All reported vertebral augmentation was cost-effective versus non-operative management.
    - From sensitivity analyses is the two good quality studies, cost-effectiveness was influenced by varying the degree of assumed mortality “benefit” from augmentation.
    - Similarly, one good-quality study from Japan found that reducing the assumed mortality “benefit” substantially increased the incremental cost-effectiveness ratio (ICER).

Use of data from or analyses based on Medicare/CMS data for mortality is an important limitation of these studies. Sensitivity analyses in most studies suggest that assumptions regarding mortality had important impacts on cost-effectiveness. Well known limitations of such administrative database studies include selection bias, inability to control confounding, confounding by indication, missing data and misclassified data. Thus, causal inference for mortality benefit is not possible. Some studies modeled a life-time horizon or longer-term horizons (5 years) however long-term RCT data from are sparse. Patient

populations modeled were generally >65 and changes in health status and co-morbidities may impact life years and quality of life. The impact of adverse events and potential for subsequent fractures were infrequently modeled or considered in sensitivity analyses.

Strength of Evidence Summaries

Detailed SOE tables, including reasons for downgrading, are found in Section 5 of the full report.

Considerations

Research published subsequent to the 2010 HTA now includes a much broader evidence base of RCTs of vertebral augmentation in patients with osteoporotic vertebral compression fractures. Most RCTs were considered to be good or fair quality. While additional trials of VP versus sham were identified, there are still no trials of KP versus sham. One RCT comparing KP with usual care in patients with VCF due to malignancy was identified to enhance the prior evidence base, however no RCTs for sacroplasty were identified and evidence remains spares and was considered insufficient.

In trials of VP versus sham, effects for pain improvement with VP were smaller compared with sham at some time frames when an association was seen, and improvement was similar between VP and sham at other times. In contrast, VP was associated with large or moderate improvements in pain compared with usual care, at all but one time. The reason for this is not clear. This observation may be in part due to placebo and nonspecific effects not related to treatment, given the inability to blind patients receiving usual care leading to a potential overestimate of effect. Authors of one included trial[23,27] suggest that fracture age/acuity may impact clinical outcomes, noting that most placebo-controlled trials performed VP later in the natural history of the fracture and this may partially explain findings of no benefit for VP versus placebo. There is inadequate information from subanalyses of included studies for Key Question 3 to effectively evaluate differential effectiveness or harms based on fracture age/pain duration. Stratified analyses from a recent AHRQ review[20] across 10 RCTs (N=1093), comparing VP to sham or usual care control (most of which are included in this HTA update) found no statistically significant interaction at for subgroups of baseline pain duration or by study inclusion criteria based on pain duration for the outcome of pain. Similarly, for the outcome of function, no statistically significant interaction for either of these pain duration subgroups was observed. Authors note that estimates based on small numbers of trials and are imprecise. The use of local anesthetic might be considered a more “active” control and partially explain the smaller or no effect seen between VP and sham control.[15,37] Periosteal infiltration of local anesthetic was done for patients randomized to vertebroplasty in four RCTs[13,15,36,50] was injected into the vertebral body in a another trial [45] and was confined to subcutaneous infiltration without periosteal numbing and a 4 mm skin incision was made in the sixth trial.[22] Effect sizes for pain from this last trial, were greater than those found in the other trials. However this trial also enrolled patients with shorter mean fracture duration (≤3 weeks in 79% of patients), used a higher PMMA volume (7.5 ml vs. range of 1.4 ml to 5 ml) compared with the other trials reporting pain scores.

Across RCTs, adverse events were variably and sparsely reported. Serious adverse events were variably defined and trials report that most were not procedure related. Comparative NRSIs reporting safety were included and summarized in the full report but evidence from these was considered insufficient due to study limitations primarily. Evidence from large administrative database studies suggested during

public comment to the key question posting reported lower mortality with vertebral augmentation compared with non-operative management of osteoporotic vertebral fractures. Mortality is a rare event. The effect sizes for mortality VP versus UC from the included RCTs and those from the administrative data studies are reasonably consistent, however, the RCTs show no statistical difference between VP and sham or VP and usual care. While the RCTs are less biased and allow for causal inference, some may have been underpowered to detect differences in mortality and had shorter follow-up. Causal inference, however, is not possible from such studies, and their results should be considered within the context of the general limitations of administrative database studies (claims data). These include the potential for selection bias, inability to control some important confounding or prognostic factors that cannot be measured in administrative data, the potential for coding-related misclassification of variables and missing data. Although included database studies described methods to control selection bias, confounding and other biases (e.g.. via propensity matching), residual confounding and unmeasured selection bias are possible. These may lead to an overestimate procedure benefits.[63] It is interesting to note that some of the large database studies that reported statistically significant mortality benefit were industry funded.

Executive Summary References

1. Andresen JR, Radmer S, Andresen R, et al. Comparative outcome of different treatment options for fragility fractures of the sacrum. BMC Musculoskelet Disord 2022;23:1106.
2. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. Bmj 2004;328:1490.
3. Bae IS, Chun HJ, Bak KH, Yi HJ, Choi KS, Kim KD. Medial Branch Block Versus Vertebroplasty for 1-Level Osteoporotic Vertebral Compression Fracture: 2-Year Retrospective Study. World Neurosurg 2019;122:e1599-e605.
4. Bae JS, Park JH, Kim KJ, Kim HS, Jang IT. Analysis of Risk Factors for Secondary New Vertebral Compression Fracture Following Percutaneous Vertebroplasty in Patients with Osteoporosis. World Neurosurg 2017;99:387-94.
5. Bae JW, Gwak HS, Kim S, et al. Percutaneous vertebroplasty for patients with metastatic compression fractures of the thoracolumbar spine: clinical and radiological factors affecting functional outcomes. Spine J 2016;16:355-64.
6. Beall DP, Shonnard NH, Shonnard MC, et al. An Interim Analysis of the First 102 Patients Treated in the Prospective Vertebral Augmentation Sacroplasty Fracture Registry. J Vasc Interv Radiol 2023;34:1477-84.
7. Berenson J, Pflugmacher R, Jarzem P, et al. Balloon kyphoplasty versus non-surgical fracture management for treatment of painful vertebral body compression fractures in patients with cancer: a multicentre, randomised controlled trial. Lancet Oncol 2011;12:225-35.
8. Bergmann M, Oberkircher L, Bliemel C, Frangen TM, Ruchholtz S, Krüger A. Early clinical outcome and complications related to balloon kyphoplasty. Orthop Rev (Pavia) 2012;4:e25.
9. Berkman ND, Lohr KN, Ansari M, et al. AHRQ Methods for Effective Health Care Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008.
10. Bian F, Bian G, Zhao L, Huang S, Fang J, An Y. Risk factors for recollapse of new vertebral compression fractures after percutaneous kyphoplasty in geriatric patients: establishment of a nomogram. BMC Musculoskelet Disord 2022;23:458.

11. Blasco J, Martinez-Ferrer A, Macho J, et al. 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;27:1159-66.
12. Boonen S, Van Meirhaeghe J, Bastian L, et al. Balloon kyphoplasty for the treatment of acute vertebral compression fractures: 2-year results from a randomized trial. J Bone Miner Res 2011;26:1627-37.
13. Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. N Engl J Med 2009;361:557-68.
14. Burton AW, Mendoza T, Gebhardt R, et al. Vertebral compression fracture treatment with vertebroplasty and kyphoplasty: experience in 407 patients with 1,156 fractures in a tertiary cancer center. Pain Med 2011;12:1750-7.
15. Carli D, Venmans A, Lodder P, et al. Vertebroplasty versus Active Control Intervention for Chronic Osteoporotic Vertebral Compression Fractures: The VERTOS V Randomized Controlled Trial. Radiology 2023;308:e222535.
16. Chandra V, Wajswol E, Shukla P, Contractor S, Kumar A. Safety and Efficacy of Sacroplasty for Sacral Fractures: A Systematic Review and Meta-Analysis. J Vasc Interv Radiol 2019;30:1845-54.
17. 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;21:473-7.
18. Cheng J, Muheremu A, Zeng X, Liu L, Liu Y, Chen Y. Percutaneous vertebroplasty vs balloon kyphoplasty in the treatment of newly onset osteoporotic vertebral compression fractures: A retrospective cohort study. Medicine (Baltimore) 2019;98:e14793.
19. Choo S, Malik AT, Jain N, Yu E, Kim J, Khan SN. 30-day adverse outcomes, re-admissions and mortality following vertebroplasty/kyphoplasty. Clin Neurol Neurosurg 2018;174:129-33.
20. Chou R, Fu R, Dana T, Pappas M, Hart E, Mauer KM. Interventional Treatments for Acute and Chronic Pain: Systematic Review. Comparative Effectiveness Review No. 247. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 75Q80120D00006.) AHRQ Publication No. 21-EHC030. Rockville, MD: Agency for Healthcare Research and Quality; September 2021. DOI: https://doi.org/10.23970/AHRQEPCCER247. Posted final reports are located on the Effective Health Care Program search page.
21. Chou R, Fu R, Dana T, Pappas M, Hart E, Mauer KM. AHRQ Comparative Effectiveness Reviews. Interventional Treatments for Acute and Chronic Pain: Systematic Review. Rockville (MD): Agency for Healthcare Research and Quality (US); 2021.
22. Clark W, Bird P, Gonski P, et al. Safety and efficacy of vertebroplasty for acute painful osteoporotic fractures (VAPOUR): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet 2016;388:1408-16.
23. Clark W, Diamond T. Early Vertebroplasty for Severely Painful Acute Osteoporotic Compression Fractures: A Critical Review of the Literature. Cardiovasc Intervent Radiol 2023;46:1439-46.
24. Comstock BA, Sitlani CM, Jarvik JG, Heagerty PJ, Turner JA, Kallmes DF. Investigational vertebroplasty safety and efficacy trial (INVEST): patient-reported outcomes through 1 year. Radiology 2013;269:224-31.
25. Cui Y, Pan Y, Lin Y, Mi C, Wang B, Shi X. Risk factors for predicting cement leakage in percutaneous vertebroplasty for spinal metastases. J Orthop Sci 2022;27:79-83.
26. Deibert CP, Gandhoke GS, Paschel EE, Gerszten PC. A Longitudinal Cohort Investigation of the Development of Symptomatic Adjacent Level Compression Fractures Following Balloon-assisted Kyphoplasty in a Series of 726 Patients. Pain Physician 2016;19:E1167-e72.

27. Diamond T, Clark W, Bird P, Gonski P, Barnes E, Gebski V. Early vertebroplasty within 3 weeks of fracture for acute painful vertebral osteoporotic fractures: subgroup analysis of the VAPOUR trial and review of the literature. Eur Spine J 2020;29:1606-13.
28. Ding J, Zhang Q, Zhu J, et al. Risk factors for predicting cement leakage following percutaneous vertebroplasty for osteoporotic vertebral compression fractures. Eur Spine J 2016;25:3411-7.
29. Dohm M, Black CM, Dacre A, Tillman JB, Fueredi G. A randomized trial comparing balloon kyphoplasty and vertebroplasty for vertebral compression fractures due to osteoporosis. AJNR Am J Neuroradiol 2014;35:2227-36.
30. Edidin AA, Ong KL, Lau E, Schmier JK, Kemner JE, Kurtz SM. Cost-effectiveness analysis of treatments for vertebral compression fractures. Appl Health Econ Health Policy 2012;10:273-84.
31. Endres S, Badura A. Shield kyphoplasty through a unipedicular approach compared to vertebroplasty and balloon kyphoplasty in osteoporotic thoracolumbar fracture: a prospective randomized study. Orthop Traumatol Surg Res 2012;98:334-40.
32. Evans AJ, Kip KE, Brinjikji W, et al. Randomized controlled trial of vertebroplasty versus kyphoplasty in the treatment of vertebral compression fractures. J Neurointerv Surg 2016;8:756-63.
33. Fadili Hassani S, Cormier E, Shotar E, et al. Intracardiac cement embolism during percutaneous vertebroplasty: incidence, risk factors and clinical management. Eur Radiol 2019;29:663-73.
34. Faloon MJ, Ruoff M, Deshpande C, et al. Risk Factors Associated with Adjacent and Remote- Level Pathologic Vertebral Compression Fracture Following Balloon Kyphoplasty: 2-Year Follow-Up Comparison Versus Conservative Treatment. J Long Term Eff Med Implants 2015;25:313-9.
35. 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;14:561-9.
36. Firanescu CE, de Vries J, Lodder P, et al. Percutaneous Vertebroplasty is no Risk Factor for New Vertebral Fractures and Protects Against Further Height Loss (VERTOS IV). Cardiovasc Intervent Radiol 2019;42:991-1000.
37. Firanescu CE, de Vries J, Lodder P, et al. Vertebroplasty versus sham procedure for painful acute osteoporotic vertebral compression fractures (VERTOS IV): randomised sham controlled clinical trial. Bmj 2018;361:k1551.
38. Fourney DR, Schomer DF, Nader R, et al. Percutaneous vertebroplasty and kyphoplasty for painful vertebral body fractures in cancer patients. J Neurosurg 2003;98:21-30.
39. Frey ME, Warner C, Thomas SM, et al. Sacroplasty: A Ten-Year Analysis of Prospective Patients Treated with Percutaneous Sacroplasty: Literature Review and Technical Considerations. Pain Physician 2017;20:E1063-e72.
40. Fritzell P, Ohlin A, Borgström F. Cost-effectiveness of balloon kyphoplasty versus standard medical treatment in patients with osteoporotic vertebral compression fracture: a Swedish multicenter randomized controlled trial with 2-year follow-up. Spine (Phila Pa 1976) 2011;36:2243-51.
41. García-Maroto R, García-Coiradas J, Milano G, Cebrián JL, Marco F, López-Durán L. [Safety and efficacy of kyphoplasty in the treatment of tumoral disease of the spine]. Rev Esp Cir Ortop Traumatol 2015;59:406-12.
42. Griffoni C, Lukassen JNM, Babbi L, et al. Percutaneous vertebroplasty and balloon kyphoplasty in the treatment of osteoporotic vertebral fractures: a prospective randomized comparison. Eur Spine J 2020;29:1614-20.
43. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383-94.
44. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64:380-2.

45. Hansen EJ, Rousing, R., Tropp, H., Carreon, L.Y., Osterhaden Andersen, M. Vertebroplasty vs. SHAM for treating osteoporotic vertebral compression fractures: a double blind RCT (VOPE). Integrative Journal of Orthopaedics ad Traumatology 2019;2:2-6.
46. Heo DH, Park CK. Percutaneous Sacroplasty for Non-neoplastic Osteoporotic Sacral Insufficiency Fractures. Pain Physician 2017;20:89-94.
47. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. 2011.
48. Hopkins TJ, Eggington S, Quinn M, Nichols-Ricker CI. Cost-effectiveness of balloon kyphoplasty and vertebroplasty versus conservative medical management in the USA. Osteoporos Int 2020;31:2461-71.
49. Huang KY, Lee SC, Liu WL, Wang JD. Long Term Survival of Pathological Thoracolumbar Fractures Treated with Vertebroplasty: Analysis Using a Nationwide Insurance Claim Database. J Clin Med 2019;9.
50. Kallmes DF, Comstock BA, Heagerty PJ, et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. N Engl J Med 2009;361:569-79.
51. Kim HJ, Zuckerman SL, Cerpa M, Yeom JS, Lehman RA, Jr., Lenke LG. Incidence and Risk Factors for Complications and Mortality After Vertebroplasty or Kyphoplasty in the Osteoporotic Vertebral Compression Fracture-Analysis of 1,932 Cases From the American College of Surgeons National Surgical Quality Improvement. Global Spine J 2022;12:1125-34.
52. Klazen CA, Lohle PN, de Vries J, et al. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. Lancet 2010;376:1085-92.
53. Kobayashi N, Noguchi T, Kobayashi D, et al. Safety and Efficacy of Percutaneous Vertebroplasty for Osteoporotic Vertebral Compression Fractures: A Multicenter Retrospective Study in Japan. Interv Radiol (Higashimatsuyama) 2021;6:21-8.
54. Kortman K, Ortiz O, Miller T, et al. Multicenter study to assess the efficacy and safety of sacroplasty in patients with osteoporotic sacral insufficiency fractures or pathologic sacral lesions. J Neurointerv Surg 2013;5:461-6.
55. Kose KC, Cebesoy O, Akan B, Altinel L, Dincer D, Yazar T. Functional results of vertebral augmentation techniques in pathological vertebral fractures of myelomatous patients. J Natl Med Assoc 2006;98:1654-8.
56. Kroon F, Staples M, Ebeling PR, et al. Two-year results of a randomized placebo-controlled trial of vertebroplasty for acute osteoporotic vertebral fractures. J Bone Miner Res 2014;29:1346-55.
57. Leali PT, Solla F, Maestretti G, Balsano M, Doria C. Safety and efficacy of vertebroplasty in the treatment of osteoporotic vertebral compression fractures: a prospective multicenter international randomized controlled study. Clin Cases Miner Bone Metab 2016;13:234-6.
58. Li Y, Zhu, J, Xie, C. A comparative study of percutanous kyphoplasty and conservative therapy on vertebral osteoporotic compression fractures in elderly patients. Int J Clin Exp Med 2017;10:8139-45.
59. Lin D, Hao J, Li L, et al. Effect of Bone Cement Volume Fraction on Adjacent Vertebral Fractures After Unilateral Percutaneous Kyphoplasty. Clin Spine Surg 2017;30:E270-e5.
60. Liu JT, Li CS, Chang CS, Liao WJ. Long-term follow-up study of osteoporotic vertebral compression fracture treated using balloon kyphoplasty and vertebroplasty. J Neurosurg Spine 2015;23:94-8.
61. Liu JT, Liao WJ, Tan WC, et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. Osteoporos Int 2010;21:359-64.
62. Liu Q, Cao J, Kong JJ. Clinical effect of balloon kyphoplasty in elderly patients with multiple osteoporotic vertebral fracture. Niger J Clin Pract 2019;22:289-92.

63. McCullough BJ, Comstock BA, Deyo RA, Kreuter W, Jarvik JG. Major medical outcomes with spinal augmentation vs conservative therapy. JAMA Intern Med 2013;173:1514-21.
64. Molloy S, Sewell MD, Platinum J, et al. Is balloon kyphoplasty safe and effective for cancer-related vertebral compression fractures with posterior vertebral body wall defects? J Surg Oncol 2016;113:835-42.
65. Moulin B, Tselikas L, Gravel G, et al. Safety and Efficacy of Multilevel Thoracolumbar Vertebroplasty in the Simultaneous Treatment of Six or More Pathologic Compression Fractures. J Vasc Interv Radiol 2020;31:1683-9.e1.
66. Ning L, Zhu J, Tian S, et al. Correlation Analysis Between Basic Diseases and Subsequent Vertebral Fractures After Percutaneous Kyphoplasty (PKP) for Osteoporotic Vertebral Compression Fractures. Pain Physician 2021;24:E803-e10.
67. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. J Manag Care Pharm 2003;9:53-61.
68. Ong KL, Beall DP, Frohbergh M, Lau E, Hirsch JA. Were VCF patients at higher risk of mortality following the 2009 publication of the vertebroplasty "sham" trials? Osteoporos Int 2018;29:375-83.
69. Ontario HQ. Vertebral Augmentation Involving Vertebroplasty or Kyphoplasty for Cancer-Related Vertebral Compression Fractures: An Economic Analysis. Ont Health Technol Assess Ser 2016;16:1-34.
70. Purvis TE, De la Garza Ramos R, Sankey EW, Karikari IO, Goodwin CR, Sciubba DM. In-hospital medical complications after non-operative and operative treatment of thoracolumbar fractures in patients over 75 years of age. J Clin Neurosci 2018;50:83-7.
71. Qi J, Hu Y, Yang Z, et al. Incidence, Risk Factors, and Outcomes of Symptomatic Bone Cement Displacement following Percutaneous Kyphoplasty for Osteoporotic Vertebral Compression Fracture: A Single Center Study. J Clin Med 2022;11.
72. Quality AfHRa. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. 2014.
73. Rocha Romero A, Hernández-Porras BC, Plancarte-Sanchez R, et al. Risk of New Fractures in Vertebroplasty for Multiple Myeloma. A Retrospective Study. Pain Med 2020;21:3018-23.
74. 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 (Phila Pa 1976) 2009;34:1349-54.
75. Rousing R, Hansen KL, Andersen MO, Jespersen SM, Thomsen K, Lauritsen JM. Twelve-months follow-up in forty-nine patients with acute/semiacute osteoporotic vertebral fractures treated conservatively or with percutaneous vertebroplasty: a clinical randomized study. Spine (Phila Pa 1976) 2010;35:478-82.
76. Sarigul B, Ogrenci A, Yilmaz M, Koban O, Mammadov M, Dalbayrak S. Sacral insufficiency fracture: a single-center experience of 185 patients with a minimum 5-year follow-up. Eur Spine J 2023.
77. Schünemann HJ, Tugwell P, Reeves BC, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Res Synth Methods 2013;4:49-62.
78. Spross C, Aghayev E, Kocher R, Röder C, Forster T, Kuelling FA. Incidence and risk factors for early adjacent vertebral fractures after balloon kyphoplasty for osteoporotic fractures: analysis of the SWISSspine registry. Eur Spine J 2014;23:1332-8.
79. Staples MP, Howe BM, Ringler MD, et al. New vertebral fractures after vertebroplasty: 2-year results from a randomised controlled trial. Arch Osteoporos 2015;10:229.
80. Stevenson M, Gomersall T, Lloyd Jones M, et al. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: a systematic review and cost-effectiveness analysis. Health Technol Assess 2014;18:1-290.

81. Sun X, Deng M, Xu W, et al. CT features and risk factors of pulmonary cement embolism after vertebroplasty or kyphoplasty in patients with vertebral compression fracture: a retrospective cohort study. Quant Imaging Med Surg 2023;13:2397-407.
82. Svedbom A, Alvares L, Cooper C, Marsh D, Ström O. Balloon kyphoplasty compared to vertebroplasty and nonsurgical management in patients hospitalised with acute osteoporotic vertebral compression fracture: a UK cost-effectiveness analysis. Osteoporos Int 2013;24:355-67.
83. Takahashi S, Hoshino M, Yasuda H, et al. Cost-effectiveness of Balloon Kyphoplasty for Patients With Acute/Subacute Osteoporotic Vertebral Fractures in the Super-Aging Japanese Society. Spine (Phila Pa 1976) 2019;44:E298-e305.
84. Tan CW, Arlachov Y, Czernicki M, et al. Spinal medial branch nerve root block (MBNB) intervention compared to standard care-vertebroplasty (VP) for the treatment of painful osteoporotic vertebral fractures in frail, older hospitalised patients: a feasibility study. Arch Osteoporos 2023;18:126.
85. Tang B, Cui L, Chen X, Liu Y. Risk Factors for Cement Leakage in Percutaneous Vertebroplasty for Osteoporotic Vertebral Compression Fractures: An Analysis of 1456 Vertebrae Augmented by Low-Viscosity Bone Cement. Spine (Phila Pa 1976) 2021;46:216-22.
86. 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 (Phila Pa 1976) 2013;38:971-83.
87. Viswanathan M, Ansari MT, Berkman ND, et al. Chapter 9: Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. 2014.
88. Vogl TJ, Pflugmacher R, Hierholzer J, et al. Cement directed kyphoplasty reduces cement leakage as compared with vertebroplasty: results of a controlled, randomized trial. Spine (Phila Pa 1976) 2013;38:1730-6.
89. Voormolen MH, Mali WP, Lohle PN, et al. Percutaneous vertebroplasty compared with optimal pain medication treatment: short-term clinical outcome of patients with subacute or chronic painful osteoporotic vertebral compression fractures. The VERTOS study. AJNR Am J Neuroradiol 2007;28:555-60.
90. 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;25:3486-94.
91. 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:E187-94.
92. Wang S, Wang H, Niu L. Clinical efficacy of PVP and PKP in the treatment of OVCFs after bilateral resection of ovarian cancer. Oncol Lett 2018;16:151-6.
93. Wang W, Liu Y, Wan H, et al. Effectiveness and prognostic factors of different minimally invasive surgeries for vertebral compression fractures. BMC Musculoskelet Disord 2023;24:11.
94. Wang YT, Wu XT, Chen H, Wang C, Mao ZB. Adjacent-level symptomatic fracture after percutaneous vertebral augmentation of osteoporotic vertebral compression fracture: a retrospective analysis. J Orthop Sci 2014;19:868-76.
95. Wardlaw D, Cummings SR, Van Meirhaeghe J, et al. Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet 2009;373:1016-24.
96. Wen Z, Mo X, Zhao S, et al. Comparison of Percutaneous Kyphoplasty and Pedicle Screw Fixation for Treatment of Thoracolumbar Severe Osteoporotic Vertebral Compression Fracture with Kyphosis. World Neurosurg 2021;152:e589-e96.

97. Wu CL, Wu JC, Huang WC, et al. The risk of stroke after percutaneous vertebroplasty for osteoporosis: a population-based cohort study. PLoS One 2012;7:e31405.
98. Wu W, Zhang X, Li X, et al. Comparison of the clinical outcomes of percutaneous kyphoplasty for the management of osteolytic and osteoblastic-related metastatic vertebral lesions. J Neurointerv Surg 2022;14:938-41.
99. Wu W, Zhang X, Li X, Yu S. Clinical analysis of percutaneous kyphoplasty for spinal metastases in older adults with comorbidities. J Geriatr Oncol 2023;14:101410.
100. Yang EZ, Xu JG, Huang GZ, et al. Percutaneous Vertebroplasty Versus Conservative Treatment in Aged Patients With Acute Osteoporotic Vertebral Compression Fractures: A Prospective Randomized Controlled Clinical Study. Spine (Phila Pa 1976) 2016;41:653-60.
101. Yang YC, Hsieh MH, Chien JT, Liu KC, Yang CC. Teriparatide treatment shows faster healing than sacroplasty for postmenopausal women with sacral insufficiency fracture. Osteoporos Sarcopenia 2023;9:27-31.
102. Yi X, Lu H, Tian F, et al. Recompression in new levels after percutaneous vertebroplasty and kyphoplasty compared with conservative treatment. Arch Orthop Trauma Surg 2014;134:21-30.
103. Zhang TY, Zhang PX, Xue F, Zhang DY, Jiang BG. Risk factors for cement leakage and nomogram for predicting the intradiscal cement leakage after the vertebra augmented surgery. BMC Musculoskelet Disord 2020;21:792.
104. Zhao C, Liu X, Wang Y, et al. The effects of biomechanical factors on adjacent vertebral compression fractures after percutaneous kyphoplasty: a propensity score matching analysis. Osteoporos Int 2022;33:1795-806.
105. Zou J, Mei X, Gan M, Yang H. Kyphoplasty for spinal fractures from multiple myeloma. J Surg Oncol 2010;102:43-7.

1 Appraisal

1.1 Background and Rationale

Vertebral compression fractures (VCFs) and sacral insufficiency fractures (SIF) often result in considerable pain, loss of function, and decreased quality of life. Patients with osteopenic vertebral or sacral fractures are at greater risk of morbidity and mortality, yet operative intervention (e.g. fusion with instrumentation) may be problematic in this elderly population, making less invasive methods more attractive. VCFs can also occur due to metastatic bone disease leading to disability and morbidity and again, operative interventions may not be feasible.

Vertebroplasty, kyphoplasty and sacroplasty (collectively, percutaneous vertebral and sacral surgery) are minimally invasive surgical procedures used to treat spinal pain believed to be caused by fractures in the vertebra or sacrum. These are all cementoplasty (augmentation) techniques intended to stabilize the fractured bone(s), but the mechanism of pain relief is not clear. Osteoporosis, vertebral metastasis and multiple myeloma are the most frequently reported indications for these procedures. Cementoplasty may reduce pain and improve stability of the bone.

Vertebroplasty involves injection of bone cement into a partially collapsed vertebral body under computed tomography (CT) or fluoroscopic guidance. Kyphoplasty is a modification of vertebroplasty that expands the partially collapsed vertebral body with an inflatable balloon or other mechanical device before the injection of bone cement. Sacroplasty is an extension of vertebroplasty, involving the injection of bone cement into the sacrum to repair sacral insufficiency fractures.

These surgical procedures are less invasive than other spinal surgical procedures, but more invasive than conservative medical therapy. Vertebroplasty, kyphoplasty and sacroplasty are surgical procedures and are not subject to FDA approval, however materials and devices used as part of these procedures are subject to FDA approval.

1.2 Policy Context

A Health Technology Assessment titled: ***Vertebroplasty, Kyphoplasty, Sacroplasty***, was published on November 5, 2010 by the Health Care Authority. New evidence has been published subsequent to the 2010 review and additional devices have been FDA approved. The Committee's Coverage Decision is summarized below.

HTCC Coverage Determination

Vertebroplasty, Kyphoplasty and Sacroplasty are not covered benefits.

HTCC Reimbursement Determination

Vertebroplasty, Kyphoplasty and Sacroplasty are not covered benefits.

1.3 Objectives

The aim of this report is to systematically review, critically appraise, analyze and synthesize research evidence evaluating the effectiveness and safety of vertebroplasty, kyphoplasty and sacroplasty for primary treatment of vertebral or sacral fracture due to osteoporosis or tumor/malignancy compared with placebo/sham, no treatment, surgery or common conventional treatment options to reflect evidence published subsequent to the 2010 report. Vertebroplasty and kyphoplasty will be compared

with each other. The differential effectiveness and safety of these therapies for subpopulations will be evaluated, as will the cost effectiveness.

1.4 Key Questions

Public comments related to the topic nomination and selection posting did not result in changes to the DRAFT key questions or PICOTS. Suggested citations from the comments will be evaluated for inclusion against the final KQ and PICOTS. Public comments from the public posting of the DRAFT key questions were evaluated and clinical expert perspectives have been sought and were used to inform finalization of the KQ and PICOTs. The assessment update will be restricted to devices approved by the FDA where applicable.

When used in patients with spinal pain ***due to vertebral fracture***:

1. What is the evidence of efficacy and effectiveness of vertebroplasty, kyphoplasty or sacroplasty? Including consideration of:
    a. Short-term and long-term outcomes
    b. Impact on function, pain, quality of life
    c. Other reported measures including: use of pain medications and opioids, return to work
2. What is the evidence of the safety of vertebroplasty, kyphoplasty or sacroplasty? Including consideration of:
    a. Adverse events type and frequency (mortality, major morbidity, other)
    b. Revision/re-operation rates (if not addressed in efficacy)
3. What is the evidence that vertebroplasty, kyphoplasty or sacroplasty has differential efficacy or safety issues in sub populations? Including consideration of:
    a. Gender
    b. Age
    c. Psychological or psychosocial co-morbidities
    d. Diagnosis or time elapsed from fracture
    e. Other patient characteristics or evidence-based patient selection criteria
    f. Provider type, setting or other provider characteristics
    g. Payer/beneficiary type: including worker's compensation, Medicaid, state employees
4. What is the evidence of cost implications and cost-effectiveness of vertebroplasty, kyphoplasty and sacroplasty? Including consideration of:
    a. Costs (direct and indirect) in the short term and over expected duration of use
    b. Revision/re-operation (if not addressed in efficacy)

Scope:

Summary of inclusion and exclusion criteria

PICO inclusion/exclusion criteria below will be finalized following consultation with agency and after review of public comment on key questions and clinical expert input.

PICOTS/Scope:

Study Component	Inclusion	Exclusion
Population	Patients with spinal pain due to ***vertebral fracture*** secondary to • Osteoporosis • Malignancy Subgroups, special populations: • Gender • Age • Psychological or psychosocial co-morbidities • Diagnosis or time elapsed from fracture • Other patient characteristics or evidence-based patient selection criteria • Provider type, setting or other provider characteristics • Payer/beneficiary type: including worker's compensation, Medicaid, state employees	• Fractures due to high energy trauma
Intervention	• Vertebroplasty • Kyphoplasty • Sacroplasty	• Cements, devices that are not FDA approved unless being studied in a Phase III trial • Spineoplasty graft consisting of mesh filled with bone chips instead of the traditional cement • Percutaneous cement discoplasty (PCD) - intervertebral disc is filled with percutaneously injected acrylic cement; may be used as prep or with vertebroplasty • Studies of exercise/rehab post augmentation • Stentoplasty, vertebral body stenting • Vesselplasty
Comparator	• Sham procedure or placebo • Conservative care, conventional care • Other minimally invasive procedures (e.g., facet joint block, nerve block) • Surgical procedures • Vertebroplasty vs. kyphoplasty	• Comparisons of different cement types • Comparisons of surgical approaches or techniques • Comparison of different vertebroplasty techniques with each other or different forms of kyphoplasty with each other • Use of vertebroplasty, kyphoplasty or sacroplasty as an adjunct to other procedures (e.g., ablation) • Augmentation combined with zoledronic acid (ZOL) versus augmentation alone • Types of imaging guidance, other guidance, e.g., Robotic assisted vs. fluoroscopy • Stentoplasty/vertebral body stenting, • Vesselplasty

Study Component	Inclusion	Exclusion
Outcomes	**Primary outcomes** • Functional outcomes (e.g., ODI) • Pain relief • Harms/Complications (e.g., procedure related, leakage, new fracture, medical complications, mortality, revision/re-operation) **Secondary outcomes** • Quality of life • Measures of disability (e.g., work lost) • Opioid use • Return to work/return to normal activity	• Measures that are not validated • Intermediate outcomes measures (e.g., radiographic measures of disc height)
Study design	• **Key Question 1**: Comparative clinical studies with a focus on studies with least potential for bias (RCTs); NRSI with concurrent controls that control for confounding will be considered if RCT evidence is not available for KQ 1. • **Key Question 2**, safety, RCTs, NRSI with ≥250 patients that are specifically designed to evaluate safety that control for confounding will be considered; case series will be considered if adequate information is not available from comparative NRSIs and RCTs or for rare or long-term adverse events; systematic reviews may be considered for safety • **Key Question 3:** RCTs only • **Key Question 4:** Full formal economic studies	• Case reports • Case series, single arm studies, pre-post studies with fewer than 5 patients (for sacroplasty) • NRSIs for effectiveness or benefit for osteoporotic fractures (KQ1) • NRSI that do not control for confounding (exception for sacroplasty)
Publication	• Full-length studies published in English in peer reviewed journals, published HTAs or publicly available FDA reports • Full formal economic analyses (e.g., cost-utility studies) published in English in HTAs or in a peer-reviewed journal published after those represented in previous HTAs	• Abstracts, editorials, letters • Duplicate publications of the same study which do not report on different outcomes • Single reports from multicenter trials • Studies reporting on the technical aspects of these procedures • White papers • Narrative reviews • Articles identified as preliminary reports when results are published in later versions • Incomplete economic evaluations such as costing studies

FDA = Food and Drug Administration; HTA = Health Technology Assessment; ODI = Oswestry Disability Index; QOL = quality of life; QALY = Quality adjusted life years.

1.5 Outcomes Assessed

This review focuses on the following primary effectiveness outcomes: validated measures of pain and function and opioid use. We focus on serious treatment-related adverse events, i.e., treatment-related events that may be life-threatening or required medical intervention. Clinical input on prioritization of harms and adverse events was obtained and reflected in the reporting of these. We also report on cost-effectiveness measures from full economic analyses. Table 1 provides a list of validated primary outcomes measures used in this review. We used definitions for the magnitude of effect size consistent with prior AHRQ reviews for treatment of pain, Appendix Q.

Primary Outcome Measures Used in Included Studies

Outcome Measure	Assessed By	Components	Score Range	Interpretation	MCID
Pain Visual Analog Scale (VAS-pain) / Numeric pain scale (NPS) / Numeric Pain Rating Scale (NPRS)	Patient	Patients are asked to indicate on a scale line (100 mm in length) where they rate their pain level of the day. One variation of this measure includes changing the length of the line.	0 to variable maximum of 10 or 100 (total score)	The higher the score, the greater the pain. No pain: 0 to 4 mm Mild pain: 5 to 44 mm Moderate pain: 45 to 74 mm Severe pain: 74 to 100 mm	For CLBP, FBSS, PDN or CRPS: NR
Oswestry Disability Scale (ODI)	Patient	Questionnaire examines perceived level of disability in 10 everyday activities of daily living. The 6 statements are scored from 0 to 5 and the final score is calculated as a percentage of the total points possible.	0%-100%	The higher the score, the greater the disability 0% to 20%: minimal disability 21%-40%: moderate disability 41%-60%: severe disability 61%-80%: crippled 81%-100%: bed bound	In patients with low back pain (various pathologies): Range, 9.5 to 12.9 points
Roland Morris Disability Questionnaire (RDQ)	Patient	Questionnaire assess self-rated physical disability caused by low back pain. Answers are yes	0-24	The higher the score, the greater the disability	3-point change, or 30% improvement from baseline

Outcome Measure	Assessed By	Components	Score Range	Interpretation	MCID
		(1)/no(0) format. The 24 questions are summed to get the total score			

CLBP = Chronic low back pain; CRPS = Complex regional pain syndrome; DN4 = Douleur Neuropathique 4 Questionnaire; EQ5D = EuroQol 5-Dimension Questionnaire; EQ-VAS = EuroQol Visual Analog Scale; FBSS = Failed back surgery syndrome; MCID = Minimal clinically important difference; MCS = Mental Component Score; NPS = Numerical Pain Scale; NPRS = Numeric Pain Rating Scale; ODI = Oswestry Disability Index; PCS = Physical Component Score; PDN = Painful diabetic neuropathy; SF-12 = Short Form 12; SF-36 = Short Form 36; VAS = Visual analogue scale.
*MCIDs were only found if an outcome was significant in any of the results of this report. Those that are significant in the results, but not found searching the literature, then the MCID is reported as NR.

1.6 Washington State Utilization Data

2 Background

2.1 Epidemiology and Burden of Disease

As the population of elderly adults increases rapidly, the incidence of conditions related to age is expected to rise. Osteoporosis is a bone disease defined by low bone mass and an increased risk of fragility. Already, an estimated 61 million elderly patients are living with osteoporotic conditions in the United States, with increasing rates seen particularly in women.[88,146] By 2050, the number of Americans aged 65 and older is expected to increase 47% to 68%,[25,26,104] further increasing the number of individuals living with this condition.

Vertebral compression fractures (VCFs) and sacral insufficiency fractures (SIFs) are a common consequence of living with osteoporosis, often resulting in considerable pain, loss of function, and decreased quality of life.[125] VCFs are characterized by low back pain varying in severity.[172] At least 1 million VCFs are reported each year in the US,[53] making it the third most frequent fragility related fracture.[166] According to a report published for the World Health Organization, the highest prevalence is seen amongst post-menopausal women.[172] In men, there is a differential prevalence of fractures with data from a large cohort[151] reporting non-Hispanic white American 17.1%; Afro-Caribbean, 5.5%; African American, 15.1%; Hispanic-American, 13.7%; Asian American, 10.5%; Hong Kong Chinese, 5.6%, and Korean, 5.1%.

VCFs can also occur due to metastatic bone disease leading to disability and morbidity. Incidence varies widely depending on cancer diagnosis, but can range from 2% to 28%,[45,140] though older studies suggest even higher rates.[6] There are additional risks in patients following radiotherapy.[122]

Patients with osteopenic vertebral or sacral fractures are at greater risk of morbidity and mortality,[30] yet operative intervention (e.g. fusion with instrumentation) may be problematic in this elderly population, making less invasive methods more attractive. Non-invasive Management of VCFs generally include control through medication, bracing, and physical therapy.[53] Guidelines for fracture management are inconsistent[65] and based on weak evidence.[133] However, elderly patients often require extensive bed rest – which can lead to immobility related complications – seldom tolerate pain related to bracing,[2] and must be mindful of adverse events from medications.[117]

2.2 Technologies & Interventions

Vertebroplasty, kyphoplasty and sacroplasty (collectively, percutaneous vertebral and sacral surgery) are minimally invasive surgical procedures used to treat spinal pain believed to be caused by fractures in the vertebra or sacrum. First described in 1987,[73] these technologies have been previously reviewed by the Washington Health Care Authority.[188] The mechanism of pain relief is not described well in the literature. A recent analysis of a national database found that up to 9.2% of patients with VCFs receive vertebral augmentation.[126] Meta-analyses suggest pain relief compared to other non-invasive treatments.[23,170,204] Vertebral augmentation may also be beneficial in cancer-related fractures.[130]

Fluoroscopy is used in percutaneous vertebroplasty (VP) to guide the injection of bone cement, such as polymethylmethacrylate (PMMA), through the pedicle into the collapsed vertebral bodies. The aim is to stabilize the fractures and effectively reduce pain.[54,112] Kyphoplasty (KP) is a modification of vertebroplasty first introduced a decade after VP. This technique expands the partially collapsed vertebral body with an inflatable balloon or other mechanical device before the injection of bone cement under low pressure. KP was designed for fracture reduction, height restoration, and kyphosis correction,[46,54] with the additional aim to decrease the risk of cement leakage.[112] Sacroplasty is an

extension of vertebroplasty, involving the injection of bone cement into the sacrum to repair sacral insufficiency fractures. These surgical procedures are less invasive than other spinal surgical procedures, but more invasive than conservative medical therapy. Because they are surgical procedures, they are not subject to FDA approval; however materials and devices used as part of these procedures – including polymethylmethacrylate (PMMA) based bone cements - are subject to FDA approval.[68]

Recent developments to vertebral augmentation include the following FDA approved devices:

- The SpineJack System, which is used to restore the height of compressed vertebra before the use of balloon kyphoplasty and cement injection.[60,112,176]
- The OsseoFix System, an expandable titanium mesh cage implanted and slowly expanded, following by cement injection.[58,112,175]
- The Kiva System, a flexible implant that holds cement and is used to restore vertebral body height.[112,128,155]

Vertebroplasty and sacroplasty are minimally invasive procedures typically performed with only local anesthesia or without conscious sedation, though general anesthesia may also be used. Kyphoplasty, on the other hand, almost always requires general anesthesia and necessitates at least one overnight hospital stay. During all three procedures, the patient must lie prone, and multiple levels can be treated in a single session. Despite higher costs, balloon kyphoplasty is performed three times as frequently as vertebroplasty in the United States.[54]

The most common indications for these procedures include[3,112,123,142,150,171]

- Osteoporotic VCFs causing non-radicular and intractable pain despite conservative treatment.
- Painful VCFs that fail to improve with time and non-surgical management.
- Symptomatic vertebral body microfracture.
- Rapidly progressive fracture preceding kyphosis.
- Severe kyphosis restricting pulmonary compliance.
- Recurrent or adjacent fracture.
- VCFs associated with osteonecrosis, nonunion, or cystic degeneration.
- Primary osteolytic diseases causing refractory pain or restricting activities of daily living.
- Painful primary bone tumors.
- Osteolytic metastases causing pain or restricting activities of daily living.
- Metastatic bone tumors preceding pathological fracture or pending fracture.
- Fractures due to osteogenesis imperfecta.
- Pseudoarthrosis following avascular necrosis of the vertebral body.
- Patients hospitalized for pain and functional impairment following VCFs.
- Fractures following Kümmell Disease

Absolute contraindications to vertebral augmentation include:[142,148,150]

- Asymptomatic fractures.
- Clinical improvement during non-surgical care.
- History of osteomyelitis or spinal infection.
- Allergy to bone fillers, bone cement, or opacification agents.
- Uncorrected coagulopathy.
- Systemic infection.
- Fracture that breaches the posterior vertebral wall.

- Burst fracture.
- Retropulsed bone fragments.

Additional relative contraindications include a loss of vertebral body height ≥75%, damaged pedicles and facets, and tumors invading the spinal canal. The time between fracture and augmentation varies, with experts suggesting at least three weeks of non-surgical care first.[148]

2.3 Published Clinical Guidelines

The ECRI Guideline Trust (based on the former National Guideline Clearing House), PubMed, Google, Google Scholar, professional societies, references in other publications, were searched for evidence-based clinical guidelines related to the use of vertebral augmentation for treating osteoporotic and malignant/tumor fractures. Nineteen evidence-based clinical guidelines were identified via the ECRI Guidelines Trust. A summary of the identified clinical guidelines and their associated TRUST score (when available) and strength of recommendations are provided in Table 1 below.

Table 1. Summary of Guidelines and Consensus Statements

Guideline	Evidence Base	Recommendation/Consensus	TRUST Score Strength of Recommendation
American Academy of Orthopaedic Surgeons (AAOS), 2010 (McGuire, 2011), updated 2023	5 RCTs: -2 RCTs of grade level I (i.e., defined as reliable) -3 RCTs of grade level II (i.e., defined as moderately reliable) Inconclusive evidence comparing the procedure with conservative care and vertebroplasty	**Vertebroplasty:** Not recommended for osteoporotic spinal compression fractures without neurological impairment.	Strong
		Kyphoplasty: Option for osteoporotic spinal fractures; benefits in pain and function up to 6 months.	Limited
		Calcitonin: Suggested for acute fractures (0-5 days) for 4 weeks.	Moderate
		Ibandronate/Strontium Ranelate: Options to prevent additional symptomatic fractures.	Limited
		L2 Nerve Root Block: Option for acute L3/L4 fractures with neurological intactness.	Limited
		Bed Rest/Alternative Medicine/Analgesics: Options for managing osteoporotic spinal fractures.	Inconclusive
		Bracing: Option for osteoporotic spinal fractures with correlating symptoms.	Inconclusive
		Exercise Program: Supervised or unsupervised for managing osteoporotic spinal fractures.	Inconclusive
		Electrical Stimulation: Option for managing osteoporotic spinal fractures with correlating symptoms.	Inconclusive
American College of Radiology (ACR), 2022	ACR Appropriateness Criteria® Management of Vertebral Compression Fractures: Variants 1 to 9 https://acsearch.acr.org/list	**Vertebroplasty:** Recommended for osteoporotic compression fractures with spinal deformity, worsening symptoms, or pulmonary dysfunction; no active management for asymptomatic VCFs without pain or activity restriction.	NR
		MRI Evaluation: Suggested before vertebral augmentation in patients with malignancy history or atypical features; helps differentiate recent from chronic fractures.	NR
American College of Radiology (ACR), American Society of Neuroradiology (ASNR), Society of	NR	Vertebral augmentation is recognized as safe and established by ACR, ASN, ASSR, SIR, and SNIS, with guidelines for patient selection and procedure. Indications include symptomatic osteoporotic fractures, insufficiency fractures unresponsive to therapy, weakened vertebrae from osteoporosis or neoplasia, symptomatic microfractures,	NR

Guideline	Evidence Base	Recommendation/Consensus	TRUST Score Strength of Recommendation
Neurointerventional Surgery (SNIS), American Society of Spine Radiology (ASSR), and the Society of Interventional Radiology (SIR), 2017 (updated 2022)		benign painful lesions, progressive fractures, and severe kyphosis. Not recommended for prophylactic use against future fractures.	
American Society of Interventional and Therapeutic Neuroradiology, Society of Interventional Radiology, American Association of Neurological Surgeons/Congress of Neurological Surgeons, and the American Society of Spine Radiology, 2007	NR	In 2007, a position statement affirmed that percutaneous vertebral augmentation (vertebroplasty and kyphoplasty) is safe, effective, and durable for symptomatic osteoporotic and neoplastic fractures, recommended when traditional therapy fails to relieve pain or significantly impacts the patient's lifestyle.	NR
International Society for the Advancement of Spine Surgery (ISASS), 2019	NR	The 2019 policy statement (Lamlice et al.) deems vertebral augmentation eligible for patients with severe pain-related functional limitations, history of VCFs, physical exam consistent with VCFs, and confirmed fracture by imaging. Contraindications include blood-borne infection, surgical site infection, or osteomyelitis. ISASS 2019 supports vertebral augmentation (preferably kyphoplasty) as safe, effective, and beneficial over conservative management, emphasizing early treatment to reduce mortality and morbidity.	NR
North American Spine Society (NASS), 2023	Studies, RCTs (Chandra et al. (2014), NICE's key conclusions, meta-analyses, RCTs, retrospective	**Coverage Recommendations (March 2023):** NASS recommends vertebral augmentation for vertebral body fractures due to osteoporosis, avascular necrosis, or neoplasm with severe pain unresponsive to conservative treatment, impaired daily activities, and confirmed acute fracture on imaging. No specific tools or products recommended; not applicable to traumatic fractures or primary vertebral tumors.	NR

Guideline	Evidence Base	Recommendation/Consensus	TRUST Score Strength of Recommendation
	multicenter studies, prospective cohort studies, SRs, VAPOUR study)	**Absolute Contraindications:** Vertebral augmentation is contraindicated for chronic fractures without active imaging evidence, active systemic or local infection, and during pregnancy.	NR
		Relative Contraindications: Caution is advised in cases of allergy to fill material, coagulopathy, spinal instability, myelopathy, neurologic deficit, or neural impingement.	NR
National Institute for Health and Care Excellence (NICE) (United Kingdom), 2013	Technology appraisal guidance 9 RCTs, 5 open-label trials Risk assessment, diagnosis and management (CG75)	**Vertebroplasty/Kyphoplasty (NICE 2013 & 2008):** Recommended for severe, ongoing pain from recent vertebral fractures unresponsive to pain management, and in cases of vertebral metastases without spinal cord compression or instability, following specialist agreement. Guidance last reviewed in 2014, next review in 5 years.	NR
American Academy of Family Physicians (AAFP), 2016	NR	**AAFP 2016 Recommendations:** Offer conservative therapy for vertebral compression fractures. Consider percutaneous vertebral augmentation if nonsurgical care fails to relieve pain or if pain significantly impacts quality of life. Evaluate patients for osteoporosis and initiate preventive therapy if needed.	NR
American Association of Clinical Endocrinologists (AACE) and American College of Endocrinology (ACE) (Camacho et al., 2016; Updated 2020)	NR	**Recommendation:** Vertebroplasty and kyphoplasty are not recommended as first-line treatments for vertebral fractures due to unclear pain relief benefits and potential increased risk of adjacent vertebral fractures (Grade A, BEL 1; downgraded).	NR
American Association of Neurological Surgeons (AANS)	NR	**AANS 2023 Guideline:** Candidates for vertebroplasty or kyphoplasty include patients with osteoporotic VCFs (present >2 weeks, moderate to severe pain, unresponsive to conservative therapy), painful metastases or multiple myelomas, painful vertebral hemangiomas, vertebral osteonecrosis, and for reinforcement of a weak vertebral body before surgical stabilization.	NR
		AANS 2023 Contraindications: Vertebroplasty or kyphoplasty should not be performed in patients with fully healed or conservatively managed VCFs, VCFs older than one year, vertebral body collapse >80-90%, non-osteoporotic spinal curvature,	NR

Guideline	Evidence Base	Recommendation/Consensus	TRUST Score Strength of Recommendation
		spinal stenosis or herniated discs unrelated to VCF, untreated coagulopathy, osteomyelitis, discitis, or significant spinal canal compromise from bone fragments or tumors.	
Society of Interventional Radiology (SIR), American Association of Neurological Surgeons (AANS) and the Congress of Neurological Surgeons (CNS), American College of Radiology (ACR), American Society of Neuroradiology (ASNR), American Society of Spine Radiology (ASSR), Canadian Interventional Radiology Association (CIRA), and Society of NeuroInterventional Surgery (SNIS), 2014	NR	**2014 Consensus Statement:** Vertebroplasty and kyphoplasty are considered safe, effective, and durable for symptomatic osteoporotic and neoplastic fractures when non-operative therapy fails to relieve pain or significantly affects quality of life. No current indication exists for prophylactic use to prevent future fractures; recommendations may evolve with future research.	NR
Society of NeuroInterventional Surgery (SNIS), 2014	SR	**2014 SNIS Report:** Kyphoplasty is superior to conservative therapy in reducing pain, disability, and improving quality of life in cancer patients with vertebral fractures (Class IIA, Level B). Vertebroplasty and kyphoplasty are reasonable options for severe, refractory back pain from cancer or osteoporotic vertebral fractures (Class IIA, Level B).	NR
German Society for Orthopaedics and Trauma (DGOU), 2018	Review of literature and case series (i.e., 707 clinical cases from 16 hospitals were evaluated)	**Management of Osteoporotic Vertebral Fractures:** Conservative management is recommended for OF type 1 and 2 fractures (and those scoring <6 on the OF scale). Vertebral augmentation with instrumentation is indicated for OF type 3, 4, and 5 fractures. Consider intraoperative complications of cement augmentation, including neurological injuries, cement leakage, embolization, vertebral body perforation, hematoma, pneumothorax, and contrast fluid incompatibilities.	

Guideline	Evidence Base	Recommendation/Consensus	TRUST Score Strength of Recommendation
WFNS Spine Committee, 2022	Literature search (2010 to 2021) (i.e., RCTs, prospective non-randomized studies, retrospective studies, SRs)	**Cement Augmentation for Osteoporotic Compression Fractures:** - Conflicting studies on efficacy; meta-analyses are inconclusive regarding pain reduction. - Insufficient evidence to determine optimal timing for vertebral augmentation. - No significant difference between unilateral and bilateral approaches in pain control, quality of life, or mobilization. - Complications: cement leakage common in vertebroplasty; progressive vertebral height loss, adjacent fractures, and cardiac issues more frequent in kyphoplasty. **Recommendation:** Further high-quality, well-designed randomized controlled studies are needed to establish the role of vertebral augmentation in osteoporotic compression fractures.	NR
American Society of Anesthesiologist (ASA), American Society of Regional Anesthesia and Pain Medicine (ASRA), 2010	RCTs	**Consensus:** Consultants, ASA members, and ASRA members strongly agree that minimally invasive spinal procedures should be performed for pain related to vertebral compression fractures.	NR
Society of Interventional Radiology (SIR), 2014	NR	**2014 SIR Guideline:** Vertebral augmentation is recommended for compression fractures unresponsive to medical therapy, including cases where patients are nonambulatory due to pain, unable to tolerate physical therapy despite analgesics, or experience unacceptable side effects (e.g., sedation, confusion, constipation) from necessary pain medication.	NR
American Society of Pain and Neuroscience (ASPN), 2021	NR	**Recommendation:** Vertebral augmentation is strongly recommended for symptomatic vertebral compression fractures from spinal metastases (Level 1-A). However, ASPN notes limited data on the superiority of vertebroplasty versus kyphoplasty in treating malignant fractures.	NR
International Myeloma Working Group (IMWG), 2013	NR	**Guideline Summary:** Vertebroplasty and kyphoplasty are effective for pain relief and functional improvement in neoplastic spinal fractures, but the role of vertebroplasty in myeloma patients remains unclear due to a lack of randomized trials. Two randomized studies showed no benefit of vertebroplasty over conservative therapy for osteoporotic fractures.	NR

Guideline	Evidence Base	Recommendation/Consensus	TRUST Score Strength of Recommendation
Cardiovascular and Interventional Radiological Society of Europe (CIRSE), 2017	NR	**Vertebroplasty Indications:** Painful osteoporotic VCFs, benign bone tumors, malignant osteolysis, osteonecrosis, vertebrae plana, acute/chronic fractures, or for reinforcement before surgery. **Absolute Contraindications:** Asymptomatic/improving VCFs, unstable fractures, infections, severe coagulopathy, or allergies to materials. Not for prophylaxis in osteoporosis. **Relative Contraindications:** Radicular pain, tumor extension, posterior column fractures, sclerotic metastasis, or multiple metastases. **Percutaneous Kyphoplasty Indications:** Best for acute traumatic VCFs with kyphosis; similar indications to VP. **Recommendation:** CIRSE does not find strong evidence for preferring KP over VP in routine cases. KP may be preferred when height restoration is crucial, e.g., acute kyphotic fractures in younger patients.	NR

AANS = American Association of Neurological Surgeons; ACR = American College of Radiology; ASA = American Society of Anesthesiologists; ASNR = American Society of Neuroradiology; ASRA = American Society of Regional Anesthesia and Pain Medicine; ASSR = American Society of Spine Radiology; BEL = Best Evidence Level; CIRA = Canadian Interventional Radiology Association; CIRSE = Cardiovascular and Interventional Radiological Society of Europe; CNS = Congress of Neurological Surgeons; IMWG = International Myeloma Working Group; KP = kyphoplasty; VP = vertebroplasty.

2.4 Previous Systematic Reviews & Health Technology Assessments

Systematic reviews (SRs) and health technology assessments (HTAs) were found by searching PubMed, EMBASE, the Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews, from database inception to January 3,, 2024. Reference lists of relevant studies and the bibliographies of SRs were hand searched. See Appendix B for search terms and full search strategy.

We chose the most recent and complete systematic reviews (SRs) to summarize. They needed to include recent RCTs and, where non-available, NRSI, and to have been methodologically sound. We summarized SRs that looked at comparative studies to ascertain effect sizes. Where no SRs looking at RCTs are available, we opted for SRs looking only at NRSI and case series for completeness.

17 previous SRs were identified. Amongst SRs looking at osteoporotic VCFs, two[138,200] looked at VP or KP compared to sham or usual care, one[38] compared VP or KP to a combined sham or usual care, two[23,107] looked at VP compared to sham or usual care, one[23] looked at VP compared to facet joint injection, four[23,44,111,143] looked at VP compared to KP, and one[111] looked at KP versus usual care. One SR[85] focused on mortality compared vertebral augmentation to VP, KP, or both to conventional medical management. One large SR[160] within a HTA in the UK compared VP, KP, and other combined controls (Sham, usual care, and conservative management). Two SRs looked at cancer-related VCFs, with one[154] looking at VP or KP (non-comparative), and the other[116] looking KP versus non-surgical management. The latter study also looked at various studies of combined treatments, including VP, KP, and non-surgical management, but that is out of the scope of the present report. Two studies looked at sacroplasty in patients with sacral insufficiency fractures, with one[22] comparing to usual care and screw fixation, and the other[113] only looking at case series. Three SRs[20,134,135] looked at economic studies in VP and KP compared to usual care. One SR[199] looked at VP versus KP in patients with Kummel's Disease. Two economic SRs include all of the same studies with the exception of one additional study in the newer SR, so they are combined. One SR[111] was a network meta-analysis. See Table 2 for details.

Table 2. Summary of Selected Systematic Reviews for Efficacy and Harms in Vertebroplasty, Kyphoplasty, and Sacroplasty

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Systematic Reviews						
Chou, 2021[38] 1990 to April 2021 Ovid, Medline, PsychINFO, Cochrane Chou, 2021 1990 to April 2021 Ovid, Medline, PsychINFO, Cochrane	VP vs. Sham or usual care Osteoporotic VCFs	<u>Pain</u> VAS NRS Likelihood of pain response <u>Function</u> RDQ <u>QoL</u> EQ-5D QUALEFFO SF-36 PCS SF-36 MCS <u>Opiate use</u> Likelihood of continued use <u>Harms</u> Incident vertebral fracture, morality	5 RCTs (Sham) 8 RCTs (Usual Care)	Yes (Cochrane)	Yes	**Pain:** Moderate quality evidence shows VP improves pain at 2-4 weeks (MD -1.05, 95% CI -1.80 to -0.32, I^2 = 64.2%), 1-6 months (MD -0.76, 95% CI -1.17 to -0.38, I^2 = 5.5%), 6-12 months (MD -0.73, 95% CI -1.33 to -0.15, I^2 = 42.9%), and ≥12 months (MD -0.87, 95% CI -1.43 to -0.31, I^2 = 41.9%), with no difference at 1-2 weeks (MD -0.53, 95% CI -1.36 to 0.24, I^2 = 74.6%). **Function:** VP improves function from 2-4 weeks (SMD -0.27, 95% CI -0.42 to -0.12, I^2 = 0%) to ≥12 months (SMD -0.23, 95% CI -0.39 to -0.06, I^2 = 0%), but no difference at 1-2 weeks (SMD -0.21, 95% CI -0.48 to 0.04, I^2 = 49.3%). **Quality of Life (QoL):** Small improvement in EQ-5D at 2-4 weeks (MD 0.05, 95% CI 0.02 to 0.09, I^2 = 0%) and 6-12 months (MD 0.06, 95% CI 0.02 to 0.11, I^2 = 0%); no difference in QUALEFFO or SF-36 PCS/MCS. **Opiate Use:** Data inconclusive, one trial found similar rates (Data NR). **Harms:** No difference in vertebral fractures (RR 1.02, 95% CI 0.66 to 1.62, I^2 = 9.6%) or mortality (RR 0.88, 95% CI 0.50 to 1.53, I^2 = 0%) between VP and sham/usual care.

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
	KP vs. sham or usual care Osteoporotic VCFs	<u>Pain</u> VAS <u>Function</u> RDQ <u>QoL</u> SF-36 PCS SF-36 MCS EQ-5D <u>Opiate use</u> Likelihood of use Strong opioid use Analgesic use <u>Harms</u> Mortality Serious AEs Any AEs New or worsening fracture	2 RCTs	Yes (Cochrane)	Yes	**Pain:** One trial found decreased pain with KP at 1 week (MD -2.2, 95% CI -2.8 to -1.6) and continued at 1-2 years (ANOVA MD -0.8 to -0.9). Another trial showed a large effect at 1 week and 1 month (MD from baseline -3.5 to -3.3, 95% CI NR). **Function:** One trial showed increased function at 1 month (ANOVA MD -4.0, 95% CI -5.5 to -2.6), at 1 year (ANOVA MD -2.6, 95% CI -4.1 to -2.0), and a small difference at 2 years (ANOVA MD -1.4, 95% CI NR). Another trial found a large improvement at 1 month (MD -8.4, 95% CI -7.6 to -9.2). **Quality of Life (QoL):** One trial found a small improvement in SF-36 PCS at 1 month (ANOVA MD 5.2, 95% CI 2.9 to 7.4), but no improvement at 1 or 2 years (Data NR). Another trial found a moderate improvement in SF-36 PCS (MD from baseline 11.1, 95% CI 10.7 to 11.5). SF-36 MCS showed a small improvement at 1 month in one trial (MD from baseline 8.4, 95% CI 7.7 to 9.1). EQ-5D improvement was noted at 1 month in one trial (ANOVA MD 0.18, 95% CI 0.08 to 0.28), but not at 1 or 2 years. **Opiate Use:** One trial found a difference in likelihood of use at 1 month (29.8% vs. 42.9%, p=0.40), but not at 1 year (28.0% vs. 33.7%, p=1.00). Strong opioid use showed no difference at 1 month or 1 year (Data

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
						NR). KP was associated with reduced analgesic use at 1 month (RR 0.64, 95% CI 0.49 to 0.83). **Harms:** One trial found no difference in mortality between KP and usual care (Data NR), while another found an increased likelihood of mortality with KP (32.9% vs. 18.8%). Serious adverse events showed no difference (Data NR), and no difference in the likelihood of any adverse events was found (RR 1.27, 95% CI 0.78 to 2.06). One trial found more new or worsening fractures in KP (RD 7.7%, 95% CI -4.5 to 20.0).
Buchbinder, 2018[23] Up to November 2017 CENTRAL, Medline, Embase	VP vs. sham Osteoporotic VCFs	<u>Pain</u> VAS Global assessment of success <u>Function</u> RDQ <u>QoL</u> QUALEFFO EQ-5D <u>Harms</u>	5 RCTs	Yes (Cochrane)	Yes	**Pain:** High quality evidence shows VP improves pain at 1 month (SMD -0.27, 95% CI -0.44 to -0.10, I^2 = 0.0%), but not at other time points up to 24 months. Pooled analyses indicate a higher likelihood of pain success in VP patients at 3 months (RR 1.60, 95% CI 1.12 to 2.30, I^2 = 0.0%) and 6 months (RR 1.38, 95% CI 1.02 to 1.87, I^2 = 0.0%), with no differences at other time points up to 12 months. **Function:** High quality evidence shows improvement in RDQ at 1 month (MD 1.8, 95% CI 0.3 to 3.1, I^2 = 0.0%), but not at other time points.

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
		Symptomatic fractures, Serious AEs				**Quality of Life (QoL):** Moderate quality evidence indicates improvement in QUALEFFO at 1 to 2 weeks (MD 4.76, 95% CI 1.68 to 7.38, I^2 = 0.0%) and 1 month (MD 2.75, 95% CI 3.53 to 9.02, I^2 = 67%), with no differences at other time points. Improvement in EQ-5D was observed at 1 month (MD 0.05, 95% CI 0.01 to 0.09, I^2 = 0.0%), 3 months (MD 0.04, 95% CI 0.0 to 0.08, I^2 = 0.0%), and 6 months (MD 0.06, 95% CI 0.01 to 0.10, I^2 = 0.0%), but not at other time points up to 24 months. **Harms:** Moderate quality evidence shows no difference in symptomatic vertebral fractures at 12 months (RR 1.08, 95% CI 0.62 to 1.87) and no difference in serious adverse events at 12-24 months (RR 0.64, 95% CI 0.36 to 1.12).
	VP. Usual care Osteoporotic VCFs	<u>Pain</u> VAS Global assessment of success <u>Function</u> RDQ <u>QoL</u> QUALEFFO EQ-5D <u>Harms</u>	8 RCTs	No (only VP vs sham)	Yes	**Pain:** Pooled analyses showed improvement in pain favoring VP at all time points up to 12 months (SMD -1.02 to -2.06, I^2 = 94% to 96%), but not at 24 months (1 trial, SMD -0.45, 95% CI -0.90 to 0.01). **Function:** Pooled analyses showed improvement in disability (ODI or RDQ) at all time points up to 24 months (SMD -1.52 to -5.65, I^2 = 97% to 98%). **Quality of Life (QoL):** Pooled analyses showed no difference in QUALEFFO

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
		Symptomatic fractures, Serious Aes				between VP and usual care at any time point up to 12 months. EQ-5D showed improvement in QOL at 1 to 2 weeks up to 3 months (1 to 3 trials, MD 0.08 to 0.10, I^2 = 27%), but not at 6 or 12 months.
	VP vs. facet joint injection Osteoporotic VCFs	Pain VAS Function RDQ QoL SF-36	1 RCT	No (only VP vs sham)	Yes	**Pain:** One trial showed a difference favoring VP at 1 to 2 weeks (MD -1.61, 95% CI -1.84 to -1.38), but not at other time points up to 12 months. **Function:** One trial showed a difference favoring VP at 1 to 2 weeks (MD -3.42, 95% CI -3.72 to -3.12), but not at other time points up to 12 months. **Quality of Life (QoL):** One trial showed no difference between VP and facet joint injection at any time points up to 12 months.
	VP vs. KP Osteoporotic VCFs	Pain VAS Function ODI QoL EQ-5D Harms Serious AEs	7 RCTs	No (only VP vs sham)	Yes	**Pain:** Pooled analyses showed no difference between VP and KP at any time point up to 24 months. **Function:** Pooled analyses showed no difference between VP and KP in ODI at any time point up to 24 months. **Quality of Life (QoL):** Pooled analyses showed no difference between VP and KP in EQ-5D up to 24 months.

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
						Harms: One trial showed no difference in serious adverse events between VP and KP at 12 months (RR 0.20, 95% CI 0.01 to 4.14) or 24 months (RR 0.91, 95% CI 0.42 to 1.97).
Hinde, 2020[85] Between 2006 and 2018 PubMed, Embase, Cochrane	VP, KP, or both vs. Conventional medical management Osteoporotic VCFs	Harms Mortality	16 NRSI (5 databases)	Yes (Newcastle)	Yes	**Mortality:** Pooled analysis (7 studies, Newcastle 6 to 9) showed that vertebral augmentation (VP or KP) reduced mortality risk compared to non-surgical management overall (HR 0.78, 95% CI 0.66 to 0.92, I^2 = 68%), at 24 months (HR 0.70, 95% CI 0.69 to 0.71, I^2 = 0%), and approached the null at 60 months with high heterogeneity (HR 0.79, 95% CI 0.62 to 1.00, I^2 = 88%).
Daher, 2023[44] Up to June 2022 PubMed, Cochrane, Google Scholar	VP vs. KP Osteoporotic VCFs	Function ODI Pain VAS Harms Cement leakage, adjacent level fractures	2 RCTs 6 NRSI	Yes (Cochrane)	Yes	**Pain:** Pooled analysis showed no difference on VAS between VP and KP (MD from baseline -0.10, 95% CI -0.36 to 0.16, I^2 = 97%). **Function:** Pooled analysis showed no difference on ODI between VP and KP (MD from baseline -0.40, 95% CI -1.70 to 2.51, I^2 = 88%). **Harms:** Pooled analysis showed a difference in the risk of cement leakage favoring KP (RR 0.44, 95% CI 0.20 to 0.95, I^2 = 36%). Pooled analysis showed no difference in the risk of adjacent level fractures between VP and KP (RR 1.41, 95% CI 0.65 to 3.08, I^2 = 0%).

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Qui, 2023[138] Up to September 2022 PubMed, Embase, Cochrane Library	VP or KP* vs. sham or usual care* Osteoporotic VCFs	Harms New vertebral fractures Adjacent fractures	2 RCTs 5 NRSI†	No	Yes	**Harms:** Pooled analyses showed that vertebral augmentation is associated with an increased likelihood of new vertebral fractures (OR 2.10, 95% CI 1.35 to 3.28, I^2 = 0.0%) and adjacent level fractures (OR 2.17, 95% CI 1.23 to 3.82, I^2 = 0.0%).
Rose, 2023[143] 2014 to 2024 Embase, Medline	VP vs. KP Osteoporotic VCFs	Harms Cement leakage	3 RCTs 3 NRSI†	No	No	**Harms:** Cement leakage occurred in 39.3% of patients with VP compared to 28.9% with KP, showing VP was associated with an increased leak rate (p=0.0005). No pulmonary embolisms or nerve injuries were recorded. One case of decompressive surgery was reported in the KP group (0.2%).
Zhang, 2024[200] Up to May 2023 Cochrane, Embase, Medline, PubMed, Web of Science	VP or KP vs. non-surgical management Osteoporotic VCFs	Harms Mortality	5 national database studies representing 3 databases	No	Yes	**Harms:** Overall pooled analyses showed an association between vertebral augmentation and a reduced risk of mortality (HR = 0.82, 95% CI 0.78 to 0.85, I^2 = 75%). This reduction was consistent in short-term (HR = 0.29, 95% CI 0.26 to 0.32), mid-term (HR = 0.78, 95% CI 0.76 to 0.81), and long-term (HR = 0.70, 95% CI 0.50 to 1.00) follow-up. KP was associated with reduced mortality (HR 0.82, 95% CI 0.79 to 0.86) compared to non-surgical management, but VP was not (HR 0.84, 95% CI 0.66 to 1.07).

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Zhang, 2022[199] Up to September 2020 PubMed, Cochrane, Embase, Web of Science, CNKI, Wanfang Data	VP vs. KP Osteoporotic Kummel Disease	<u>Pain</u> VAS <u>Function</u> ODI <u>Harms</u> Cement leakage, new fractures, total complications	8 NRSI	Yes (NOQAS)	Yes	**Pain:** High quality evidence (NOQAS) from pooled analysis showed no difference in pain between VP and KP at post-operative (SMD 0.03, 95% CI -0.16 to 0.22, I^2 = 33.7%) or final follow-up (timing NR; SMD -0.06, 95% CI -0.25 to 0.13, I^2 = 37.2%). **Function:** High quality evidence (NOQAS) from pooled analysis showed no difference in function between VP and KP at post-operative (SMD -0.20, 95% CI -0.43 to 0.04, I^2 = 32.8%) or final follow-up (timing NR; SMD -0.14, 95% CI -0.36 to 0.08, I^2 = 31.1%). **Harms:** High quality evidence (NOQAS) from pooled analysis showed KP was associated with a lower likelihood of cement leakage (OR 0.50, 95% CI 0.31 to 0.81, I^2 = 0.0%). There was no difference between VP and KP in the likelihood of new vertebral fractures (OR 0.79, 95% CI 0.36 to 1.73, I^2 = 0.0%). KP was associated with a lower likelihood of total complications (OR 0.63, 95% CI 0.39 to 1.00, I^2 = 0.0%).
Stevenson, 2014 (NICE)[160] Up to November 2011	VP vs. KP vs. Controls (Sham, Usual care, conservative management) Osteoporotic VCFs	<u>Pain</u> VAS/NRS <u>Function</u> RDQ Barthel Index <u>QoL</u>	9 RCTs	Yes (Cochrane)	No	**Pain:** Nine studies (unpooled) used VAS to measure pain. Four reported significant short- and medium-term reductions with VP or KP, while three found no significant differences. One study showed greater pain improvement with VP at 1 month, and a meta-analysis (two trials) found no significant difference at 1 month. One

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Medline, CINAHL, Embase, EconLit, Cochrane, DARE		DPQ EQ-5D QUALEFFO <u>Analgesic use</u> Opioid use <u>Harms</u> Mortality Cement leakage New Fractures Complications Other AEs <u>Economic</u>				study indicated that VP provided earlier pain relief compared to conservative treatment. **Function:** Five trials assessed RDQ scores. Short-term results favored VP. Medium-term results showed no differences, except for one trial favoring KP at 12 months and another favoring VP at 1 year. A meta-analysis found no differences in RDQ improvements. Regarding the Barthel Index, one trial initially favored VP at 12 months, but the result was not significant after adjusting for baseline differences. **Quality of Life (QoL):** One RCT reported no difference between VP and control groups on AQoL. Another RCT found no difference between VP and sham on DPQ, except for work and leisure favoring VP at 3 months. Two unpooled RCTs found no difference between VP and conservative treatments on EQ-5D. A third RCT favored conservative treatment, while a fourth favored KP over non-surgical management at 1, 12, and 24 months. Four unpooled RCTs showed no difference between VP and conservative treatment on QUALEFFO at any time point, except for 1 week. One study showed initial benefits for KP on SF-36 PCS, but no difference after 6 months, while two others found no significant differences. Three unpooled studies found

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
						no significant differences in SF-36 MCS scores. **Analgesic Use:** Six studies reported on analgesic use. One study found no significant differences between groups, though more control patients required rescue therapy (25% vs. 5%). Two studies showed decreased opioid use in both VP and control groups over time, with no differences between groups. KP patients were less likely to use opioids at 1 and 6 months in one study. A meta-analysis of two studies found VP patients more likely to use opioids at 1 month (RR 1.25, 95% CI 1.14 to 1.36). Another study found reduced analgesic use and pain scores favoring VP at early follow-up, with significant differences at 1 day, 2 weeks, and 1 month, but not at later follow-ups. **Harms:** Six studies reported all-cause mortality with no differences between treatment groups. A meta-analysis of three studies at 12 months also found no difference. Seven studies reported cement leakage rates, with VP showing a 44% incidence and KP 27%. Leakage ranged from 0% to 72%, with the highest rates in studies using CT and higher cement volumes. Most leaks were asymptomatic, but some led to complications. One study found a significantly higher proportion of clinical

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
						fractures in the VP group (71%) compared to the control group (9%). Three other studies found no differences. In the KP group, 7.4% of patients experienced fractures possibly related to the intervention at 24 months. Perioperative issues included thecal sac injury (1 patient) and hospitalization for tachycardia (1 patient). Postoperative complications included infections (3 studies), osteomyelitis (1 patient), and pulmonary embolisms (3 KP patients). Other complications were rare and mostly non-severe. Adverse events varied. One study reported various events within 6 months, without specifying patient details. Another study provided detailed data on serious adverse events, with a few linked to the procedure, including a hematoma and UTI exacerbation. Deaths occurred in both treatment and control groups, but none were related to the procedure. **Economic:** If differential mortality effect chosen, KP had a cost-per-QALY-gained ratio below £20,000. If pooled beneficial effect assumed, VP had a cost-per-QALY-gained ratio below $10,000. VP typically was the dominant intervention or had a cost-per-QALY-gained ratio below £15,000, except when several parameters were altered unfavorably. Exploratory analyses indicated that using high-

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
						viscosity cement in all patients was unlikely to have a cost-per-QALY-gained value below £20,000.
Sorensen, 2019[154] January 2000 to January 2018 Medline, Embase, Cochrane[‡]	VP vs. KP[§] Cancer-related VCFs	<u>Pain</u> Vas <u>Function</u> ODI KPS <u>Harms</u> Cement leakage, symptomatic complications	2 RCTs 60 NRSI 25 case series/reports[†]	No	No	**Pain:** Patients treated with VP experienced pain relief with VAS improving from 7.48 preoperatively to 3.00 postoperatively. KP patients improved from 7.05 preoperatively to 2.96 postoperatively. All improvements persisted during follow-ups. **Function:** VP patients improved in ODI from 74.68 preoperatively to 17.74 at <4 weeks post-op, while KP patients improved from 66.02 to 43.73 at <4 weeks. These improvements plateaued during follow-ups. KPS for combined VP and KP patients improved from 66.99 preoperatively to 80.28 postoperatively, persisting through follow-ups. **Harms:** Cement leakage occurred in 37.9% of VP patients compared to 13.6% of KP patients. There were 43 cases of symptomatic complications related to the procedure (VP=35, KP=8), but details were not reported.

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Briggs, 2023[22] Medline, Embase Up to January 2022	Sacroplasty vs. conservative care, screw fixation only, augmented screw fixation, or all screw fixation (±augmentation) Sacral insufficiency fractures	Pain VAS Harms Cement leakage	11 NRSI 24 case series**	No	No	**Pain:** Sacroplasty patients experienced a higher difference in VAS (MD 5.83, SD 1.14) compared to conservative care (MD 3.7, SD 2.71), screw fixation (MD 3.63, SD 1.36), augmented screw fixation (MD 4.38, SD 0.79), and all screw fixation ± augmentation (MD 4.1, SD 1.12). **Harms:** Cement leakage occurred in 3.3% of sacroplasty patients, with 0.4% experiencing symptomatic cement leakage.
Mahmood, 2019[113] PubMed, SCOPUS, Ovid	Sacroplasty (not comparative) Sacral insufficiency fractures	Pain VAS Harms Cement leakage Repeat procedures	19 NRSI 12 case series	No	No	**Pain:** The mean reduction in VAS at the latest follow-up was 5.8 ± 1.3, with most studies reporting up to 12 months. **Harms:** Nine studies reported cement leakage in at least one patient, but most events were clinically insignificant. One study reported 6 repeat procedures.
Network Meta-analysis						
Liu, 2023[111††] Up to September 2023 PubMed, Web of Science, Embase, Cochrane	VP vs. KP Osteoporotic VCFs	Pain VAS Function ODI Harms Cement leakage, new fracture	5 RCTs	Yes (RoB 2)	Yes	**Pain:** Pooled analysis showed no difference in pain between VP and KP at short term (MD -0.17, 95% CI -1.01 to 0.66) or long term (MD -0.09, 95% CI -0.53 to 0.34). **Function:** Pooled analysis showed no difference in ODI between VP and KP at short term (MD -1.80, 95% CI -4.89 to 1.29) or long term (MD -1.78, 95% CI -7.18 to 3.63).

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
						Harms: Pooled analysis showed no difference in the risk of cement leakage between VP and KP (RR 0.90, 95% CI 0.62 to 1.32) or in the risk of new fractures (RR 1.10, 95% CI 0.68 to 1.79).
	VP vs. Sham Osteoporotic VCFs	Pain VAS Harms Cement leakage, new fractures	6 RCTs	Yes (RoB 2)	Yes	**Pain:** Pooled analysis showed no difference in pain between VP and the sham procedure at short term (MD 0.17, 95% CI -0.86 to 1.19). **Harms:** Pooled analysis showed no difference in the risk of cement leakage between VP and the sham procedure (OR 57.00, 95% CI 3.45 to 942.90) and no difference in the risk of new fractures (RR 1.18, 95% CI 0.53 to 2.62).
	VP vs. Usual care Osteoporotic VCFs	Pain VAS Function ODI Harms New fractures	6 RCTs	Yes (RoB 2)	Yes	**Pain:** Pooled analysis showed a difference in VAS favoring VP compared to usual care at short term (MD 3.14, 95% CI 2.31 to 3.98) but not at long term (MD 1.08, 95% CI 0.62 to 1.55). **Function:** Pooled analysis showed a difference in ODI favoring VP compared to usual care at both short term (MD 14.13, 95% CI 11.50 to 16.76) and long term (MD 8.69, 95% CI 3.16 to 14.21). **Harms:** Pooled analysis showed no difference in the risk of new fractures between VP and usual care (RR 1.28, 95% CI 0.80 to 2.03).

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
	KP vs. Usual care Osteoporotic VCFs	<u>Pain</u> VAS <u>Function</u> ODI <u>Harms</u> New fractures	1 RCT	Yes (RoB 2)	Yes	**Pain:** One RCT showed a difference in VAS favoring KP compared to usual care at short term (MD 3.32, 95% CI 2.32 to 4.31) but not at long term (MD 1.17, 95% CI 0.63 to 1.72). **Function:** One RCT showed a difference in ODI favoring KP compared to usual care at both short term (MD 15.93, 95% CI 12.32 to 19.54) and long term (MD 10.46, 95% CI 3.52 to 17.40). **Harms:** One RCT showed no difference in the risk of new fractures between KP and usual care (RR 1.16, 95% CI 0.73 to 1.82).

AE = adverse event; AQoL = Assessment of Quality of Life; CI = confidence interval; DPQ = Dallas Pain Questionnaire; EQ-5D = EuroQoL 5D; HR = hazard ratio; HTA = Health Technology Assessment; KP = kyphoplasty; MCS = mental component score; MD = mean difference; NICE = National Institute for Health and Care Excellence; NOQAS = Newcastle-Ottowa Quality Assessment Scale; NR = not reported; NRSI = non-randomized study of intervention; ODI = Oswestry Disability Index; PCS = physical component score; QALY = quality adjusted life year; QoL = quality of life; QUALEFFO = Quality of Life Questionnaire; RCT = randomized control trial; RD = risk difference; RDQ = Roland Morris Disability Questionnaire; RoB = Risk of Bias; RR = risk ratio; SMD = standardized mean difference; SF-36 = 36-item Short-Form Survey; SR = systematic review; VAS = visual analogue scale; VCF = vertebral compression fracture; VP = vertebroplasty.

* VP and KP were combined; sham and usual care were pooled.

† Analyses pooled RCTs and NRSI

‡ Sorensen 2019 also included articles identified by the HTA performed by Health Quality Ontario.

§ Included studies do not necessary compare VP to KP, but Sorensen 2019 only reports on and compares these two groups aggregated across all studies.

** Briggs 2023 does not report which studies report on which treatments, so all are reported together.

†† Liu 2023 also performed a network meta-analysis.

2.5 Medicare and Representative Private Insurer Coverage Policies

For the purposes of this report, we obtained and summarized payer policies from three bellwether payers and any relevant information on National Coverage Determinations (NCDs) and/or Local Coverage Determinations (LCDs) from the Centers for Medicare and Medicaid Services (CMS). Coverage decisions are summarized briefly below (Table 5).

- **Centers for Medicare and Medicaid Services (CMS) National Coverage Determination**
- **Cigna**
- **Aetna**
- **United HealthCare**

Table 3. Overview of CMS and Payer Policies

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
Centers for Medicare and Medicaid Services (CMS) (2023) Original Effective Date: For services performed on or after 10/01/2015 Revision Effective Date: For services performed on or after 01/10/2021	RCTs, meta-analysis	According to the "Percutaneous Vertebral Augmentation (PVA) for Osteoporotic Vertebral Compression Fracture (VCF)" LCD with an original effective date on or after 10/01/2015 and a revision effective date on or after 01/10/2021, Percutaneous Vertebral Augmentation (PVA) for Osteoporotic Vertebral Compression Fracture (VCF) is covered in patients with both the following: **Inclusion:** Acute (< 6 weeks) or subacute (6-12 weeks) osteoporotic VCF (T1 – L5) with documented imaging; Symptomatic: severe pain (NRS/VAS ≥ 8) if hospitalized, or moderate to severe pain (NRS/VAS ≥ 5) if non-hospitalized, unresponsive to optimal non-surgical management; Continuum of care: referral for BMD evaluation and osteoporosis education, plus participation in an osteoporosis prevention/treatment program. **Exclusion:** - **Absolute:** Pain not due to VCF, infection, pregnancy. - **Relative:** More than three fractures per procedure, allergy to materials, uncorrected coagulopathy, spinal instability, myelopathy, neurologic deficit, neural impingement, fracture retropulsion/canal compromise.	Ensure continuum of care

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
Cigna (2023) Effective Date: 8/15/2023 Next Review Date: 6/15/2024	Meta-analyses, HTAs, SRs, RCTs, Prospective case series	**Covered:** Percutaneous vertebroplasty or kyphoplasty is medically necessary if standard therapy fails and one of the following is present: Osteoporotic, osteolytic, osteonecrotic (Kummel disease), or steroid-induced VCF with persistent pain unresponsive to 6+ weeks of conservative treatment; Severe back pain due to osteolytic metastasis or multiple myeloma; Painful/aggressive hemangioma or eosinophilic granuloma of the spine. **Not Covered:** Percutaneous vertebroplasty, kyphoplasty, and sacroplasty are considered experimental or investigational for all other indications.	Percutaneous vertebroplasty and kyphoplasty are widely accepted as safe and effective for pain relief, increased mobility, and improved quality of life in patients with painful osteolytic lesions and osteoporotic compression fractures unresponsive to conservative treatment. However, more clinical trials are needed to assess their long-term safety and efficacy. There is currently insufficient evidence on the safety, efficacy, and long-term outcomes of sacroplasty.
Aetna (2023) Effective: 07/31/1995 Next Review: 01/11/2024	Primary studies, Systematic reviews, previous HTAs, guidelines	**Medically Necessary**: Percutaneous vertebroplasty or kyphoplasty for persistent, debilitating pain in cervical, thoracic, or lumbar vertebrae due to: Primary malignant bone or marrow neoplasm; Secondary osteolytic metastasis (excluding sacrum/coccyx); Steroid-induced fractures; Multiple myeloma; Painful/aggressive hemangiomas; Painful vertebral eosinophilic granuloma; Painful, debilitating osteoporotic acute/subacute fractures (e.g., Kummell's disease) **For Osteoporotic or Steroid-Induced Fractures:** Pain localized to the treated level; Severe pain or loss of mobility unrelieved by 6+ weeks of optimal medical therapy; Other pain causes ruled out by CT/MRI; Affected vertebra at least one-third of original height with intact posterior cortex; Max 3 vertebral fractures per procedure; Documentation of osteoporosis care and education **Not Covered:** Sacroplasty for osteoporotic sacral insufficiency fractures and other indications due to insufficient evidence.	No rationale for policy given CPT codes if selection criteria is met: 22510, 22511, 22512, 22513, 22514, 22515
United HealthCare (2023)	Retrospective studies, previous HTAs, RCTs,	**Medically Necessary:** Percutaneous vertebroplasty and kyphoplasty for pain causing functional/physical impairment in	Percutaneous vertebroplasty and kyphoplasty are proven and medically necessary for treating

Payer (year)	Evidence Base Available	Policy	Rationale/Comments
Effective Date: November 1, 2023	systematic reviews, Meta-analyses, Cochrane review	cervical, thoracic, or lumbar vertebrae within 4 months of onset, unresponsive to optimal medical therapy for: - Osteoporotic VCF - Steroid-induced vertebral fracture - Osteolytic metastatic disease - Multiple myeloma - Aggressive vertebral hemangioma - Unstable fractures due to osteonecrosis (e.g., Kummel disease) **Required Imaging Exclusions:** CT/MRI ruling out other causes like foraminal stenosis, facet arthropathy, herniated disk, or other spinal conditions. **Contraindications:** Spinal cord compression, significant vertebral collapse (<1/3 height), healed VCF, sacral/coccygeal lesions, asymptomatic VCFs, or fractures responding to conservative therapy.	pain causing Functional or Physical Impairment in cervical, thoracic, or lumbar vertebral bodies within 4 months of pain onset that has failed to respond to Optimal Medical Therapy for some indications (see list of criteria in the above section)

BMD = bone mineral density; CMS = Centers for Medicare and Medicaid Services; Cochrane = Cochrane Review; CPT = Current Procedural Terminology; CT = computed tomography; HTA = Health Technology Assessment; Kummel disease = Osteonecrosis; LCD = Local Coverage Determination; MRI = magnetic resonance imaging; NRS = Numeric Rating Scale; PVA = Percutaneous Vertebral Augmentation; RCT = randomized controlled trial; SR = systematic review; VAS = Visual Analog Scale; VCF = vertebral compression fracture.

3 The Evidence

3.1 Methods of the Systematic Literature Review

3.1.1 Objectives

The aim of this report is to systematically review, critically appraise, analyze and synthesize research evidence evaluating the effectiveness and safety of vertebroplasty, kyphoplasty, and sacroplasty. The differential effectiveness and safety of these therapies for subpopulations will be evaluated, as will the cost effectiveness.

3.1.2 Key Questions

When used in patients with spinal pain ***due to vertebral fracture***:

1. What is the evidence of efficacy and effectiveness of vertebroplasty, kyphoplasty or sacroplasty? Including consideration of:
    a. Short-term and long-term outcomes
    b. Impact on function, pain, quality of life
    c. Other reported measures including: use of pain medications and opioids, return to work
2. What is the evidence of the safety of vertebroplasty, kyphoplasty or sacroplasty? Including consideration of:
    a. Adverse events type and frequency (mortality, major morbidity, other)
    b. Revision/re-operation rates (if not addressed in efficacy)
3. What is the evidence that vertebroplasty, kyphoplasty or sacroplasty has differential efficacy or safety issues in sub populations? Including consideration of:
    a. Gender
    b. Age
    c. Psychological or psychosocial co-morbidities
    d. Diagnosis or time elapsed from fracture
    e. Other patient characteristics or evidence-based patient selection criteria
    f. Provider type, setting or other provider characteristics
    g. Payer/beneficiary type: including worker's compensation, Medicaid, state employees
4. What is the evidence of cost implications and cost-effectiveness of vertebroplasty, kyphoplasty and sacroplasty? Including consideration of:
    a. Costs (direct and indirect) in the short term and over expected duration of use
    b. Revision/re-operation (if not addressed in efficacy)

3.1.3 Inclusion/Exclusion Criteria

The scope of this report and final key questions were refined based on input from clinical experts. Clinical expert input was sought to confirm critical outcomes on which to focus. Draft Key Questions and PICOTS scope were published on the HCA website for public comment. Four were received. Public comments as well as those from clinical experts and peer-reviewers were considered for finalization of this report. See Table 4 below for inclusion and exclusion criteria.

Table 4. Summary of inclusion and exclusion criteria

Study Component	Inclusion	Exclusion
Population	Patients with spinal pain due to ***vertebral fracture*** secondary to 1. Osteoporosis 2. Malignancy	• Fractures due to high energy trauma
Intervention	• Vertebroplasty • Kyphoplasty • Sacroplasty	• Cements, devices that are not FDA approved
Comparator	• Sham procedure or placebo • Conservative care • Surgical procedures • Vertebroplasty vs. kyphoplasty	• Comparisons of different cement types • Comparisons of surgical approaches or techniques • Use of vertebroplasty, kyphoplasty or sacroplasty as an adjunct to other procedures (e.g. ablation) • Combined with zoledronic acid (ZOL) versus augmentation alone • Types of imaging guidance, other guidance
Outcomes	**Primary** • Functional outcomes (e.g. ODI) • Pain relief • Quality of life outcomes • Measures of disability (e.g., work lost) • Complications (e.g. procedure related, leakage, new fracture, medical complications, death. Revision/re-operation) Return to work **Economic** • Cost-effectiveness (e.g., cost per improved outcome), cost-utility (e.g., cost per quality adjusted life year (QALY), incremental cost effectiveness ratio (ICER) outcome	• Non-clinical outcomes • Non-validated measures (e.g., for pain, function, QOL)
Timing	Review will focus on persistence of relief 1 or more months post-treatment	
Study design	**Key Questions 1, 3 and 4:** • High quality systematic reviews of RCTs will be considered if available and they address the key questions • Randomized controlled trials (RCTs) • In the absence of RCTs, high quality non-randomized comparative studies will be	• Indirect comparisons • Comparisons with historical cohorts • Nonrandomized studies which do not control for confounding • Incomplete economic evaluations such as costing studies • Case series with fewer than 5 patients (for sacroplasty) • Case reports

Study Component	Inclusion	Exclusion
	considered in the absence of RCTs with a focus on comparative prospective studies **Key Question 2:** • In the absence of RCTs, high-quality non-randomized studies designed specifically to evaluate harms/adverse events that are rare or occur long-term • Case series will be considered if adequate information not available from comparative studies **Key Question 4:** • Only full, formal economic studies (i.e., cost-effectiveness, cost-utility, cost-minimization, and cost-benefit studies) will be considered.	• Studies in which <80% of patients have a condition of interest • Studies that do not report on primary outcomes or harms
Publication	• Full-length studies published in English in peer reviewed journals, published HTAs or publicly available FDA reports • Full formal economic analyses (e.g. cost-utility studies) published in English in HTAs or in a peer-reviewed journal published after those represented in previous HTAs.	• Abstracts, conference proceedings, editorials, letters • Duplicate publications of the same study which do not report on different outcomes • Single reports from multicenter trials • White papers • Narrative reviews • Articles identified as preliminary reports when results are published in later versions • Incomplete economic evaluations such as costing studies

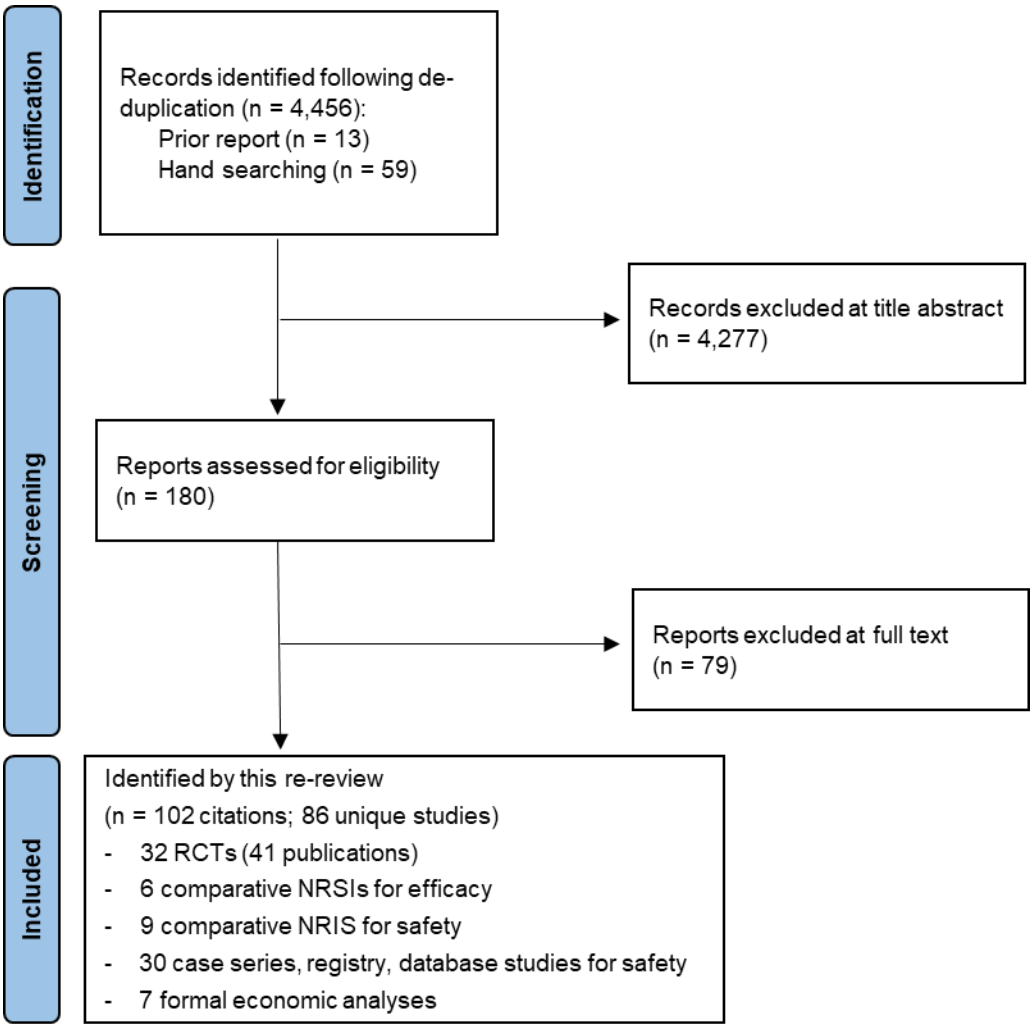
FDA = U.S. Food and Drug Administration; ODI = Owestry Disability Index; QOL = Quality of Life; RCT = Randomized Controlled Trial; HTA = Health Technology Assessment

3.1.4 Data Sources and Search Strategy

We searched electronic databases from January 1, 2010 to January 3, 2024 for trials related to vertebral augmentation to identify publications evaluating these treatments for osteoporotic vertebral compression fracture and malignancy-related fractures that had been published since the prior reports. The start dates of our searches overlapped by a few months with the end date of the searches in the prior reports. A formal, structured systematic search of the peer-reviewed literature was performed across a number of databases including PubMed, EMBASE, the Cochrane Central Register of Controlled Trials, and the Cochrane Database of Systematic Reviews (see Appendix B for full search strategy) to identify relevant peer reviewed literature as well as other sources (ClinicalTrials.gov, ECRI Guidelines Trust, Center for Reviews and Dissemination Database) to identify pertinent clinical guidelines and previously performed assessments. We conducted a comprehensive search on clinicaltrials.gov to identify relevant ongoing research trials. However, no conclusive findings were obtained from the search. We also hand searched the reference lists of relevant studies and the bibliographies of systematic reviews.

The clinical studies included in this report were identified using the algorithm shown in Appendix A. The process involves four stages. The first stage of the study selection process consisted of the comprehensive electronic search and bibliography review. We then screened all possible relevant articles using titles and abstracts in stage two. This was done by two individuals independently. Those articles that met a set of *a priori* retrieval criteria were included for full-text review. We excluded conference abstracts, non-English-language articles, duplicate publications that did not report different data or follow-up times, white papers, narrative reviews, preliminary reports, and incomplete economic evaluations. Any disagreement between screeners that were unresolved resulted in the article being included for the next stage. Stage three involved retrieval of the full text articles remaining. The final stage of the study selection algorithm consisted of the review and selection of those studies using a set of *a priori* inclusion criteria, again, by two independent investigators. Discrepancies were resolved through discussion and if necessary, adjudicated by a third investigator. See Figure 1 below for a flow diagram of the search results. A list of excluded articles along with the reason for exclusion is available in Appendix C. The remaining articles form the evidence base for this report.

Figure 1. Flow of studies diagram



NRSI = non-randomized study of intervention; RCT = Randomized controlled trial

3.1.5 Data Extraction

Reviewers extracted the following data from the clinical studies: study design, setting, country, source of funding, sample size, inclusion and exclusion criteria, study population characteristics, follow-up time, device details, PMMA volume, study outcomes and adverse events. Data from figures were estimated using Web Plot Digitizer v5.[1] For economic studies, data related to sources used, economic parameters and perspectives, results, and sensitivity analyses were abstracted. An attempt was made to reconcile conflicting information among multiple reports presenting data from the same study. Detailed study and patient characteristics and results are available in Appendix F to L.

3.1.6 Quality Assessment: Overall Strength of Evidence (SOE), Risk of Bias, & QHES evaluation

The method used by Aggregate Analytics, Inc. (AAI) for assessing the quality of evidence of individual studies as well as the overall strength of evidence (SOE) are based on established methods for systematic reviews. Included studies reporting on primary outcomes of interest were critically appraised independently by two reviewers evaluating the methodological quality, study limitations and potential for bias based on study design as well as factors which may bias studies using defined templates and pre-specified criteria. Assessment of RCTs followed appropriate criteria[178] based on methods described in *the Cochrane Handbook for Systematic Reviews of Interventions*[84] and guidance from the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*[139] In keeping with the AHRQ methods, each study was given a final rating of "good", "fair", or "poor" quality as described below. Discrepancies in ratings between reviewers were resolved through discussion and consensus. Criteria are detailed in Appendix D.

Table 5. Criteria for grading the quality of individual studies

Rating	Description and Criteria
Good	• Low risk of bias; study results generally considered valid • Employed valid methods for selection, inclusion, and allocation of patients to treatment; report similar baseline characteristics/key risk factors for testing groups being compared; clearly describe attrition and have low attrition; use appropriate means for preventing bias (e.g., blinded outcomes assessment); and use appropriate analytic methods (e.g., intention-to-treat analysis); full reporting on pre-specified outcomes. • For studies of testing, pre-specification of thresholds for a positive test,
Fair	• Study is susceptible to some bias but not enough to necessarily invalidate results • May not meet all criteria for good quality, but no flaw is likely to cause major bias; the study may be missing information making it difficult to assess limitations and potential problems • This category is broad; studies with this rating will vary in strengths and weaknesses; some fair-quality studies are likely to be valid, while others may be only possibly valid
Poor	• Significant flaws that imply biases of various kinds that may invalidate results; the study contains "fatal flaws" in design, analysis or reporting; large amounts of missing information; discrepancies in reporting or serious problems with intervention or test delivery • Study results are at least as likely to reflect flaws in the study design or execution as the true difference between the compared interventions • Considered to be less reliable than higher quality studies when synthesizing the evidence, particularly if discrepancies between studies are present

Economic studies were evaluated according to The Quality of Health Economic Studies (QHES) instrument developed by Ofman et al. in conjunction with consideration of epidemiologic principles that may impact findings.[127]

Based on these quality criteria, each comparative study chosen for inclusion for a Key Question was given a risk of bias (RoB) (or QHES) rating; details of each rating are available in Appendix E.

SOE was assessed by two researchers following the principles for adapting GRADE (Grades of Recommendation Assessment, Development and Evaluation) as outlined by the Agency for Healthcare Research and Quality (AHRQ).[7,78,79,139] The strength of evidence was based on the highest quality evidence available for the primary outcomes.

In determining the strength of body of evidence regarding a given outcome, the following domains were considered:

- **Risk of bias**: the extent to which the included studies have protection against bias
- **Consistency:** the degree to which the included studies report results that are similar in terms of effect sizes, range and variability.
- **Directness**: describes whether the evidence is directly related to patient health outcomes or comparisons of interventions are direct (head-to-head).
- **Precision:** describes the level of certainty surrounding the effect estimates.
- **Publication or reporting bias:** is considered when there is concern of selective publishing or selective reporting. This is difficult to assess particularly for nonrandomized studies.

Bodies of evidence consisting of RCTs are initially considered as High SOE. In general, the GRADE and AHRQ methodologies initially consider nonrandomized studies as Low SOE as such studies typically are at higher risk of bias due to lack of randomization and inability of investigators to control for critical confounding factors. The SOE could be downgraded based on the limitations described above. There are also situations where studies (particularly observational studies) could be upgraded if the study had large magnitude of effect (strength of association) or if a dose-response relationship is identified and there are no downgrades for the primary domains listed above and confounding is not a concern. Publication and reporting bias are difficult to assess, particularly with fewer than 10 RCTs and for observational studies.[16,149] Publication bias was unknown in all studies and thus this domain was eliminated from the strength of evidence tables. The final SOE was assigned an overall grade of high, moderate, low, or insufficient, which are defined as follows:

- High - Very confident that effect size estimates lie close to the true effect for this outcome; there are few or no deficiencies in the body of evidence; we believe the findings are stable.
- Moderate – Moderately confident that effect size estimates lie close to the true effect for this outcome; some deficiencies in the body of evidence; we believe the findings are likely to be stable but some doubt remains.
- Low – Limited confidence that effect size estimates lie close to the true effect for this outcome; major or numerous deficiencies in the body of evidence; we believe that additional evidence is needed before concluding that findings are stable or that the estimate is close to the true effect.

- Insufficient – We have no evidence, are unable to estimate an effect or have no confidence in the effect estimate for this outcome; OR no available evidence or the body of evidence has unacceptable efficiencies precluding judgment.

Assessing the SOE for studies performing subgroup analysis for evaluation of differential effectiveness or safety requires additional considerations discussed below. Methods for determining the overall quality (strength) of evidence related to economic studies have not been reported, thus the overall strength of evidence for outcomes reported in Key Question 4 was not assessed.

3.1.7 Analysis

Evidence was summarized qualitatively and quantitatively. Risk ratio (RR) and 95% confidence intervals (CI) were used for dichotomous outcomes to evaluate the presence of an association between testing and the outcome. In the absence of adjusted effect size estimates, for dichotomous outcomes, crude risk ratios (RR) and 95% confidence intervals were calculated using either STATA 14.0[158] or spreadsheets based on Rothman Episheet and GraphPad.[76] For instances with fewer than five observations per cell, exact methods were employed. Where effect estimates that were adjusted for confounding were reported by study authors, they were preferred and reported. For continuous variables, mean differences (MD) and associated 95% CIs were calculated if the outcomes were reported using the same scale.

Meta-analyses were conducted as appropriate in order to summarize primary outcome data from multiple studies and to obtain more precise and accurate estimates using STATA 14.0.[158] To determine the appropriateness of meta-analysis, we considered clinical and methodological diversity and assessed statistical heterogeneity. Statistical heterogeneity among the studies was assessed using Cochran's χ^2 test and the I^2 statistic.[83] To combine trials, we used a random effects model based on the profile likelihood method which provides a more conservative effect estimate; in the case of non-convergence with profile likelihood, the Der Simonian and Laird estimates were reported.[81] For continuous variables, differences in mean follow-up scores between treatments were analyzed to determine mean differences as an effect size. Methods for calculating the standard deviations and for imputing missing standard deviations followed the recommendations given in The Cochrane Handbook 7.7.[84] Where no events occurred in one arm of a study, a value of 0.50 was used for that arm in accordance with Cochrane methods. Studies in which no events occurred in either study arm did not contribute to effect estimates (0% weight) but were retained in some plots for visual effect and completeness. Sensitivity analyses were conducted excluding poor-quality studies, outlying data and clinically heterogeneous trials where there were sufficient data. We classified the magnitude of effects for continuous measures of pain and function using the same system as in prior AHRQ reviews on pain.[36,39,40,152,153] Effects below the threshold for small were categorized as no effect. Outcomes are detailed in the evidence tables in the appendices and/or the body of the report.

We did not conduct analyses to evaluate potential markers for publication bias given the substantial heterogeneity in patient populations and small number of trials available for some analyses.[159]

To evaluate differential efficacy and safety (heterogeneity of effect, interaction), we focused on RCTs as they have the least potential for bias and confounding thus allowing for causal inference. Further, only RCTs that formally tested for interaction between subgroups were considered for Key Question 3. SOE for these studies is based on consideration of the overall study risk of bias (study quality) as well as whether subgroup variables and analyses were specified a priori, the hypothesized impact of a subgroup

on the outcome/effect and sample size as evaluation of interaction requires greater sample size. Such analyses should be interpreted cautiously and consider the biologic plausibility of differential efficacy or safety. Such analyses are generally considered hypothesis generating, and additional confirmatory evidence should be sought.[132,164,186]

4 Results

4.1 Number of Studies Retained & Overall Quality of Studies

From 4,456 unique citations identified from electronic database searches, hand searching and bibliography review of included studies, a total of 32 RCTs (in 41 publications)[14,18,19,24,29,33,41,42,50,52,59,61,64,66,67,77,80,89,93,99,101,105,108-110,144,145,157,168,174,179-184,187,195,197] met our inclusion criteria (Figure 1): 31 RCTS (in 40 publications)[18,19,24,29,33,41,42,50,52,59,61,64,66,67,77,80,89,93,99,101,105,108-110,144,145,157,168,174,179-184,187,195,197] on osteoporotic fracture and one RCT[14] on fracture due to malignancy. We additionally included: one SR[31] for effectiveness of sacroplasty; six comparative NRSIs controlling for confounding[4,8,10,71,147,196] for effectiveness; 9 additional comparative NRSIs[34,63,69,87,98,129,136,189,192] for safety; and 30 case series[9,13,15,17,27,35,43,47,51,62,74,82,92,96,97,106,120,121,124,137,141,156,163,169,185,193,194,201-203] for safety. The tables below provide an overview of each of these groups of studies by fracture cause (osteoporosis, malignancy, sacral insufficiency) and treatments compared and provides information on the funding source. The most common comparators for vertebroplasty for efficacy were kyphoplasty (9 RCTs, 1 NRSI), usual care (9 RCTs), sham (6 RCTs), and nerve block (2 RCTs, 1 NRSI); The most common comparators for kyphoplasty for efficacy were, usual care (5 RCTs), and other surgical procedures (1 RCT). Of the included RCTs, roughly one fifth (21%) reported industry funding; Furthermore, nearly one third (30%) were not clear about their funding source. Additionally, seven formal economic analyses were included: 2 in the United States[57,86], one in the United Kingdom and Sweden[165], and one each in Canada[131], the United Kingdom[160], Sweden[72], and Japan[167]

Number of RCTs included for each comparison of efficacy of included interventions for treatment of vertebral fracture

Comparisons	RCTs (publications)	Funding : No. RCTs (Publications)			
		Industry	Other*	None	NR
OVCF					
VP vs. Sham	6 (11)[24,29,41,42,50,66,67,80,89,99,157]	2[41,50,66,67]	3[24,42,80,89,99,157]	---	1[29]
VP vs. UC	9 (10)[18,33,64,93,101,144,145,180,195,197]	---	4[18,64,93,131,144,145]	2[33,195]	3[101,180,197]
VP vs. Nerve Block	2[168,181]	---	1[168]	1[181]	---
VP vs. KP	9 (10)[52,59,61,77,108,109,179,182-184,197]	3[52,61,179]	2[108,109,184]	2[182,183]	3[59,77,197]

Comparisons	RCTs (publications)	Funding : No. RCTs (Publications)			
		Industry	Other*	None	NR
KP vs. UC	4 (6)[19,105,110,174,187,197]	1[19,174,187]	---	---	3[105,110,197]
KP vs. Other Surgical Procedures	1[190]	---	---	1[190]	---
Total: OVCF	**31 (40)**[18,19,24,29,33,41,42,50,52,59,61,64,66,67,77,80,89,93,99,101,105,108-110,144,145,157,168,174,179-184,187,195,197]				
Malignancy					
KP vs. UC	1[14]	1[14]	---	---	---
Total: Malignancy	**1**[14]				
	32 (41)[14,18,19,24,29,33,41,42,50,52,59,61,64,66,67,77,80,89,93,99,101,105,108-110,144,145,157,168,174,179-184,187,195,197]				

RCT = Randomized Controlled Trial; NR = Not Reported; OVCF = Osteoporotic Vertebral Compression Fracture; VP = Vertebroplasty; UC = Usual Care; KP = Kyphoplasty.

Number of comparative NRSIs included for effectiveness for each comparison of efficacy of included interventions for treatment of vertebral fracture

Comparisons	NRSIs (publications)	Funding : No. NRSIs (Publications)			
		Industry	Other*	None	NR
OVCF					
VP vs. Medial Branch Block	1[8]	---	---	1[8]	---
Total: OVCF	**1**[8]				
Malignancy					
VP vs. KP	1[10]	---	---	---	1[10]
Total: Malignancy	**1**[10]				
SIF					
SP vs. UC	2[4,147]	---	---	2[4,147]	---
SP vs. Other Surgical Intervention	1[4]	---	---	1[4]	---
SP vs. Non-surgical Management	2[71,196]	---	---	---	2[71,196]
Total: SP	**4**[4,71,147,196]				

Comparisons	NRSIs (publications)	Funding : No. NRSIs (Publications)			
		Industry	Other*	None	NR
Total Overall	**6**[4,8,10,71,147,196]				

NRSI = Non-Randomized Study of Interventions; NR = Not Reported; OVCF = Osteoporotic Vertebral Compression Fracture; VP = Vertebroplasty; UC = Usual Care; KP = Kyphoplasty; SIF = Sacral Insufficiency Syndrome; SP = Sacroplasty.

Number of comparative NRSIs included for harms only for each comparison of efficacy of included interventions for treatment of vertebral fracture

Comparisons	NRSIs (publications)	Funding : No. NRSIs (Publications)			
		Industry	Other*	None	NR
OVCF					
Mixed VP/KP vs. Other Surgical Management or UC	1[136]	---	---	1[136]	---
VP vs. Other Surgical Management	2[87,192]	---	1[87]	1[192]	---
VP vs. Non-operative Management	1[129]	1[129]	---	---	---
VP vs. KP	2[34,129]	1[129]	---	1[34]	---
KP vs. UC	1[63]	---	---	---	1[63]
KP vs. Non-operative Management	1[129]	1[129]	---	---	---
KP vs. Other Surgeries	1[189]	---	---	1[189]	---
Total: OVCF	**7**[34,63,87,129,136,189,192]				
Malignancy					
VP vs. KP	2[69,98]	---	---	---	2[69,98]
Total: Malignancy	**2**[69,98]				
Total Overall	**9**[34,63,69,87,98,129,136,189,192]				

NRSI = Non-Randomized Study of Interventions; NR = Not Reported; OVCF = Osteoporotic Vertebral Compression Fracture; VP = Vertebroplasty; UC = Usual Care; KP = Kyphoplasty

Number of case series included for harms only for each comparison of efficacy of included interventions for treatment of vertebral fracture

Comparisons	NRSIs (publications)	Funding : No. NRSIs (Publications)			
		Industry	Other*	None	NR
OVCF					
VP	5[9,51,62,96,169]	---	---	3[62,96,169]	2[9,51]
KP	8[15,17,47,106,124,137,156,202]	---	4[106,124,137,202]	2[17,47]	2[15,156]
Mixed VP/KP	5[35,92,163,185,201]	---	---	2[35,92]	3[163,185,201]
SP	3[13,82,97]	---	1[13]	2[82,97]	---
Total: OVCF	**21**[9,13,15,17,35,47,51,62,82,92,96,97,106,124,137,156,163,169,185,201,202]				
Malignancy					
VP	3[43,121,141]	---	---	2[43,141]	1[121]
KP	5[74,120,193,194,203]	---	1[194]	1[74]	3[120,193,203]
Mixed VP/KP	1[27]	1[27]	---	---	---
Total: Malignancy	**9**[27,43,74,120,121,141,193,194,203]				
Total Overall	**30**[9,13,15,17,27,35,43,47,51,62,74,82,92,96,97,106,120,121,124,137,141,156,163,169,185,193,194,201-203]				

NRSI = Non-Randomized Study of Interventions; NR = Not Reported; OVCF = Osteoporotic Vertebral Compression Fracture; VP = Vertebroplasty; UC = Usual Care; KP = Kyphoplasty; SP = Sacroplasty

Number of systematic reviews included for each comparison of efficacy of included interventions for treatment of vertebral fracture

Comparisons	NRSIs (publications)	Funding : No. NRSIs (Publications)			
		Industry	Other*	None	NR
SIF					
SP	1[31]	---	---	1[31]	---
Total Overall	**1**[31]				

NRSI = Non-Randomized Study of Interventions; NR = Not Reported; SP = Sacroplasty; SIF = Sacral Insufficiency Syndrome

4.2 Osteoporotic Vertebral Compression Fractures

4.2.1 KQ1 Effectiveness

4.2.1.1 Vertebroplasty

4.2.1.1.1 Vertebroplasty versus Sham

Six RCTs (across 11 publications, N=641)[24,29,41,42,50,66,67,80,89,99,157] compared vertebroplasty with a sham procedure (Table 6). Three studies were conducted in the Netherlands,[29,66,80] and two were conducted in Australia. [41,89] Two were industry funded.[41,67] The majority of patients were female (mean 75%) with a mean age of 75 years old. Mean pain duration was 9 weeks or less in four RCTs.[24,41,67,80] One of these trials enrolled patients with a mean fracture duration of 2.6 weeks[41] and reported sub-analysis of patients with fracture duration of ≤ 3 weeks in a subsequent population.[50] Mean pain durations in the other two RCTs were 18 weeks[89] and 26 weeks[29] respectively corresponding to chronic pain. Evidence of bone marrow edema (BME) was required for study inclusion in three trials[29,67,80] as a measures of fracture acuity, but was not reported in the other three trials.[24,41,89] Single level interventions were most common (60% to 87%) across the four trials reporting numbers of levels treated.[29,41,67,89] PMMA volumes ranged from 1.4 to 7 ml.

In all RCTs, patients randomized to sham procedures received generally similar pre-procedure preparations as those randomized to VP. Methods to simulate the VP procedure included verbal and physical cues consistent with PMMA injection, such as needle insertion or pressure on the back to simulate needle insertion and tapping to simulate entry of the needle into bone. Patients randomized to the sham procedure received the same periosteal infiltration of local anesthetic as patients randomized to vertebroplasty in four RCTs,[24,29,66,89] one of which termed this an active control did local anesthetic injection into the pedicular periosteum.[29] Local anesthetic was injected into the vertebral body in a fifth trial[80] and local anesthesia was confined to subcutaneous infiltration without periosteal numbing and a 4 mm skin incision was made in the sixth trial.[41] For sham procedures, PMMA was prepared[24,29,67,80] or the methacrylate monomer was opened in the procedure room[89] to create the odor of mixing the cement. One trial did not report attempts to simulate the cement odor reporting that there was conversation about PMMA mixing and injection suggesting that VP was being done.[41] Four RCTs did not allow patients to cross over from one group to the other,[24,29,41,67], one trial did not report on cross-over.[80] One allowed cross-over after 1 month, with substantially more patients crossing over to the VP group compared with sham by 3 months (51% vs. 13%). [89] Our report focuses on results prior to cross over for this trial. Four RCTs were considered good quality [24,41,67,89] and two were fair.[29,80] Methodologic limitations for fair studies included unclear concealment of treatment allocation and some baseline differences between treatment groups.

Table 6. Study Characteristics of Trials Comparing Vertebroplasty versus Sham in Patients with Fractures due to Osteoporosis

Study, year Country Quality	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)[*] Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME MRI	Duration follow-up (months)	Industry funding	In MA
Carli, 2023 The Netherlands Fair	71	68%	80	7.5 (1.7) Inclusion: VAS ≥5	26[†]	NR[‡] (Treated: 1: 60% 2: 19% 3: 16% 4: 4% 5: 3%)[§]	1.4	Yes	12	NR	Yes
Clark, 2016; Diamond, 2020 Australia Good	80	73%[**]	120	8.2 (1.7) Inclusion: NRS ≥7	2.6	NR (Treated: 1: 87% 2: 13%)	7.5	No	6	Yes	Yes
Firanescu, 2018; Firanescu, 2019 The Netherlands Good	76	76%	180	7.8 (2.5) Inclusion: VAS ≥5	≤9	1.26 (0.55)[††] (Treated: 1: 78% 2: 17% 3: 6%)	5.1	Yes	12	Yes	Yes
Hansen, 2019 The Netherlands Fair	70	87%	52	7.5 (2.1) Inclusion: VAS >7	≤8	NR[‡‡]	2 to 4	Yes	12	No	Yes
Kallmes, 2009; Comstock, 2013 UK, Australia Good	74	76%	131	7.0 (1.9) Inclusion: VAS ≥3	18	NR (Treated: 1: 68% 2: 20% 3: 11%)	2.6	No	12	No	Yes

Study, year Country Quality	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)* Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME MRI	Duration follow-up (months)	Industry funding	In MA
Buchbinder, 2009; Kroon, 2014; Staples, 2015 Australia Good	74	76%	78	7.3 (2.2) Inclusion: NR	9.3	NR (Treated: 1: 82% 2: 18%)	3	No	24	No	Yes

BME = bone marrow edema; MA = meta-analysis; ml = milliliter; MRI = magnetic resonance imaging; NR = not reported; NRS = numerical rating scale; PMMA = polymethylmethacrylate; RCT = randomized control trial; SD = standard deviation; VAS = visual analogue scale; VCF = vertebral compression fracture; VP = vertebroplasty.

* For most studies, weighted mean and SD was calculated using mean estimates, SDs or 95% confidence intervals, and n's for each group at baseline.

† Median.

‡ Mean number not reported.

Carli: Number of VCFs at baseline were 72 vs. 63.

§ Some baseline differences existed between groups for number of levels treated:

Carli, 2023: 1: 70% vs. 50%, 2: 10% vs. 28%, 3: 13% vs. 18%, 4: 5% vs. 3%, 5: 3% vs. 3%.

** Some baseline differences existed between groups for % female:

Clark, 2016: 79% vs. 68%.

†† Calculated using the proportion of patients in each group with 1, 2, and 3 fractures at baseline.

‡‡ Authors do not report mean number or % of levels treated:

Hansen, 2019: 27 vs. 28 levels were treated in total.

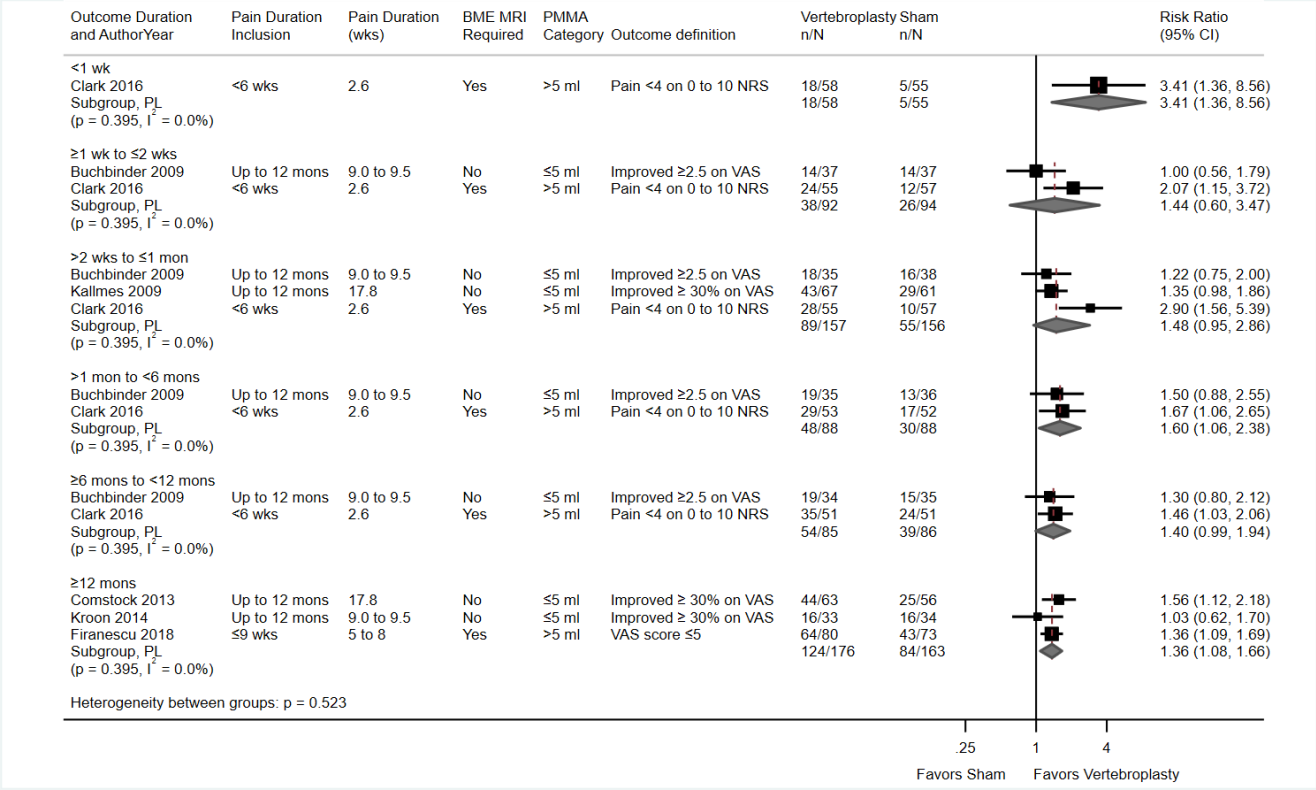
4.2.1.1.1.1 Primary Outcomes

Pain

Four trials (across 6 publications) reported the proportion of patients who were considered pain responders, defined as those experiencing a reduction in pain relative to baseline of ≥30% on a 0-10 VAS or NRS scale.[24,41,67,89] VP was associated with a much greater likelihood of meeting that threshold compared with sham treatment within 1 week of treatment in one RCT[41] (1 RCT N=113, 31% vs. 8.5%, RR 3.41, 95%CI 1.36 to 8.56). The higher likelihood of response with VP persisted to time frames up to 12 months in this trial. This trial differed from the other sham trials in several ways: mean fracture duration was ≤3 weeks in 79% of patients, local anesthesia was confined to subcutaneous infiltration without periosteal numbing and used a higher PMMA volume compared with the other trials reporting this outcome. Table 6. At baseline, more patients in the VP group had previous osteoporotic fractures (62% vs. 54%) and more severe fractures (Genant Grade 3, 74% vs.66%) compared with sham in this trial but baseline pain scores were similar between groups.

In contrast, another trial [24] in patients with mean pain duration of 9 weeks found similar likelihood of response at all time frames ≥1 week and at follow up to ≥12 months.[99] (Figure 2) Pooled analyses across time frames, show a statistically similar likelihood of pain response at ≤1 week to ≥2 weeks (2 RCTs, N=186, 41% vs. 27.7%, RR 1.44 95% CI 0.60 to 3.47, I^2=0%)[24,41], moderate likelihood of improvement with VP at >2 weeks to ≤1 month (3 RCTs, N= 313, 57.7% vs. 35.2%, RR 1.48, 95% CI 0.95 to 2.86, I^2=0%)[24,41,89], and at >1 month to <6 months (2 RCTS, N=176, 54.5% vs. 34%, RR 1.60 95% CI 1.06 to 2.38 I^2=0%)[24,41], with a slightly higher likelihood at with VP vs. sham at to ≥6 months to 12 months (2 RCTs, N = 171, 63.5% vs. 45.3% RR 1.40. 95% CI 0.99 to 1.94, I^2=0%)[24,41], and at ≥12 months (3 RCTs, N=339, 70.5% vs. 51.5%, RR 1.36, 95% CI 1.08 to 1.66, I^2=0%).[42,67,99]

Figure 2. Vertebroplasty versus sham procedures: Pain Response (≥30% VAS pain reduction from baseline, 0-10 scale)

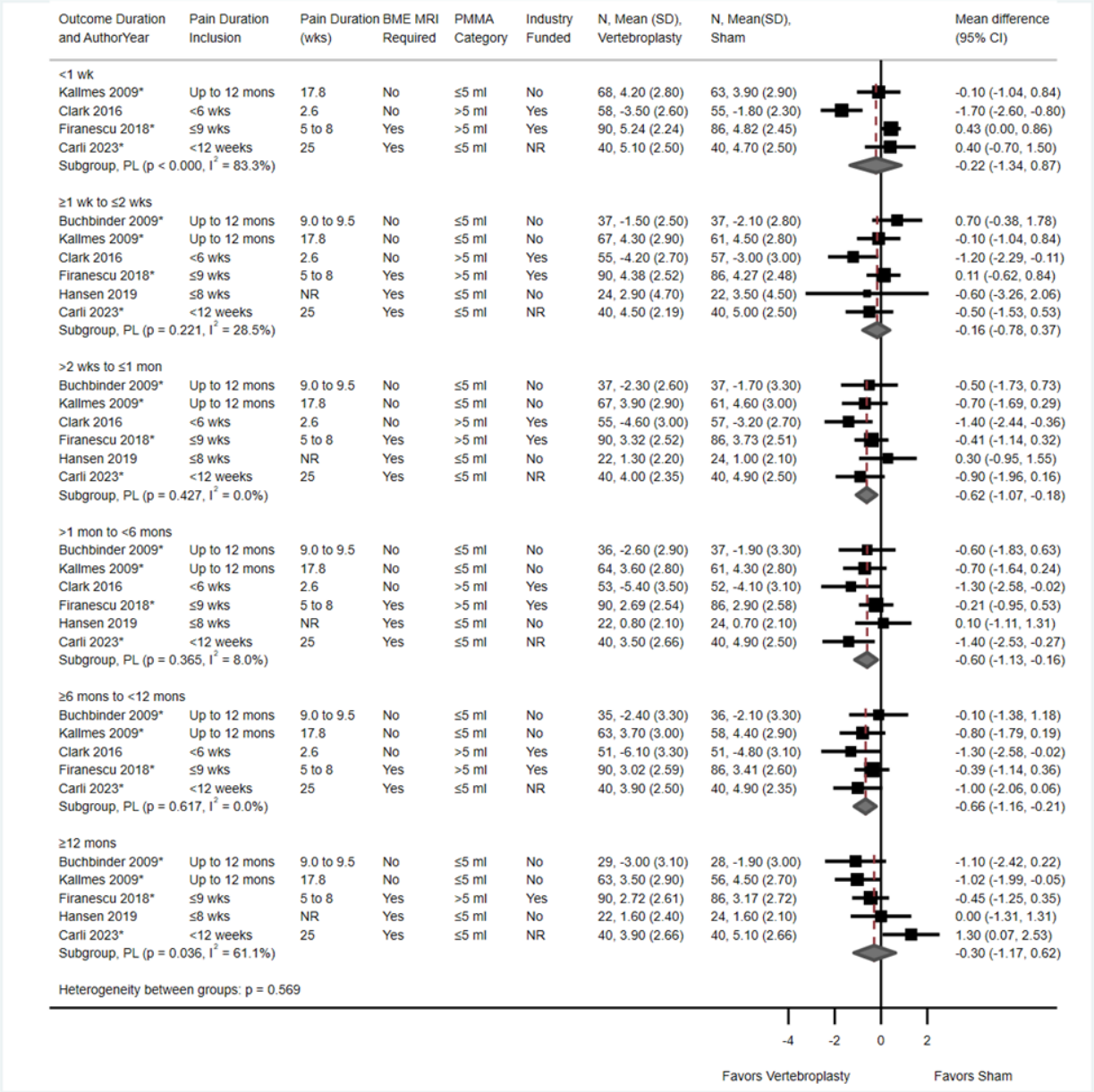


Comstock is follow-up to Kallmes; Kroon is follow-up to Buchbinder 2009

BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; RR = risk ratio, wks = weeks

All six RCTs reported VAS or NRS pain scores at one or more time frames.[24,29,41,67,80,89] Pooled estimates across trials show that pain improvement was similar between VP and sham at the two earliest time frames, namely <1 week (4 RCTs, N= 500) and ≥1week to ≤2 weeks (6 RCTs, N= 616) or at long term (≥ 12 months, 5 RCTs, N= 478). (Figure 3) At three intermediate time frames, vertebroplasty was associated with small improvement in pain intensity versus sham procedures at >2 weeks to ≤1 month (6 RCTs N = 616, MD -0.62 95% CI-1.07, -0.18, I^2 = 0%)[24,29,41,67,80,89] at >1 month to <6 months (6 RCTs, N = 605, MD: -0.60 95% CI -1.13, -0.16, I^2 = 8.0%)[24,29,41,67,80,89], at ≥6 months to <12 months (5 RCTs, N = 550 MD -0.66 95% CI -1.16, -0.21, I^2 = 0%)[24,29,41,67,89] Substantial heterogeneity (83%) is noted at the earliest follow-up (<1 week) with one sham trial[41] which had shorter mean fracture duration (≤3 weeks in 79% of patients), used a higher PMMA volume (7.5 ml vs. range of 1.4 ml to 5 ml) compared with the other trials reporting this outcome and confined local anesthesia to subcutaneous infiltration without periosteal numbing. As noted above, baseline differences in prior osteoporotic fractures (62% vs. 54%) and fracture severity (Genant Grade 3, 74% vs.66%) were noted in VP versus sham recipients; baseline pain scores were similar between groups. There was also substantial heterogeneity (61%) at ≥12 months with one trial[29] tending to favor sham treatment. Pain duration at baseline was longest in this trial (median 26 weeks), reported the smallest PMMA volume (1.4ml) and reported treatment of more vertebral levels than other trials. TABLE 6. At baseline, difference between groups on fracture severity were reported for VP versus sham recipients with fewer VP fractures rated as mild (Genant Grade 1, 35% versus 46%) and more VP fractures rated as moderate (37% vs. 25%) in this trial; baseline pain scores were similar.

Figure 3. Vertebroplasty versus sham procedures: Pain scores (VAS or NRS pain, 0-10 scale)



BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; RDQ = Roland-Morris Disability Questionnaire, MRDQ = Modified Roland-Morris Disability Questionnaire SD = standard deviation, wks = weeks

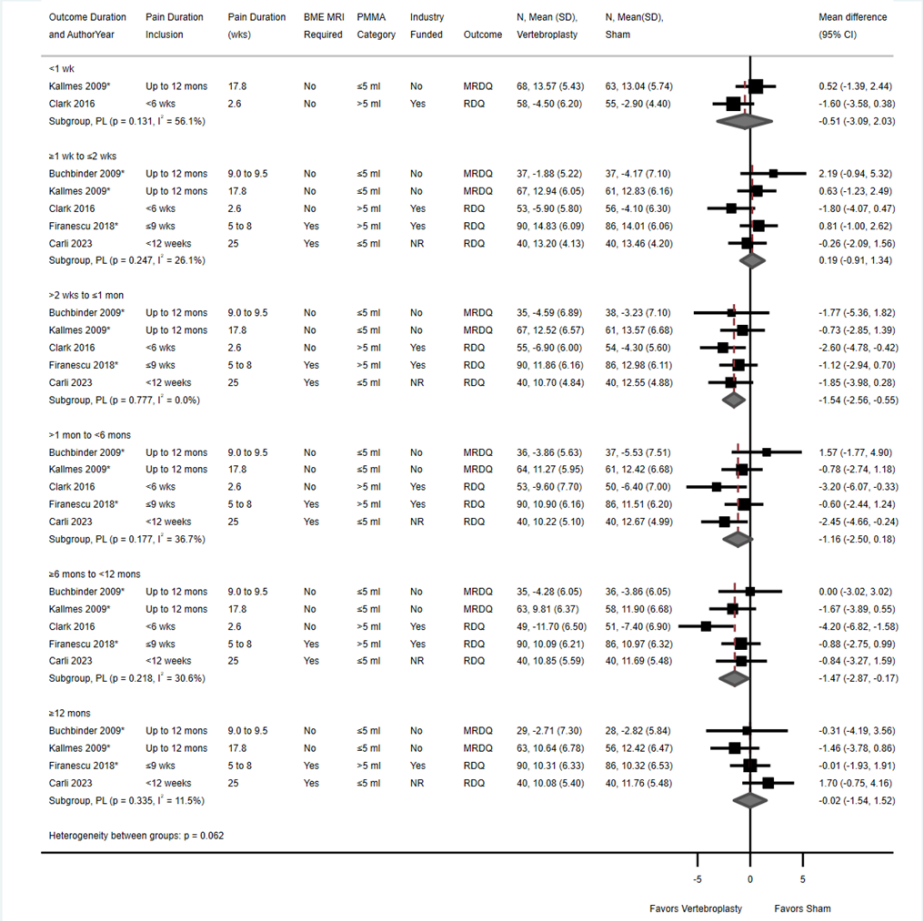
Function

No RCTs reported on the proportion of patients who were considered responders on the RDQ or other measure of function.

Five RCTs used the original RDQ or modified RDQ to report function; all RDQ versions were converted to a 0-24 scale for analyses. VP was associated with small functional improvement at two time frames across all five RCTs, [24,29,41,67,89] namely at >2 weeks to ≤1 month (5 RCTs, N = 566, MD (1.54, 95% CI -2.56 to -0.55, I^2 = 0%) and at ≥6 months to <12 months (5 RCTs, N= 548 MD -1.47, 95%CI -2.87 to -0.17, I^2 =

30 6%). Some heterogeneity at these time frames is seen. As noted above for the pain outcomes, one sham trial[41] which had shorter mean fracture duration (≤3 weeks in 79% of patients), used a higher PMMA volume (7.5 ml vs. range of 1.4 ml to 5 ml) compared with the other trials reporting this outcome and confined local anesthesia to subcutaneous infiltration without periosteal numbing. Baseline differences in prior osteoporotic fractures (62% vs. 54%) and fracture severity (Genant Grade 3, 74% vs.66%) were noted in VP versus sham recipients; baseline pain scores were similar between groups. Functional improvement was similar for VP and sham at all other time frames (Figure 4)

Figure 4. VP vs. sham procedures: Function based on Roland-Morris Disability Questionnaires, 0-24 scale



BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; RDQ = Roland-Morris Disability Questionnaire, MRDQ = Modified Roland-Morris Disability Questionnaire SD = standard deviation, wks = weeks

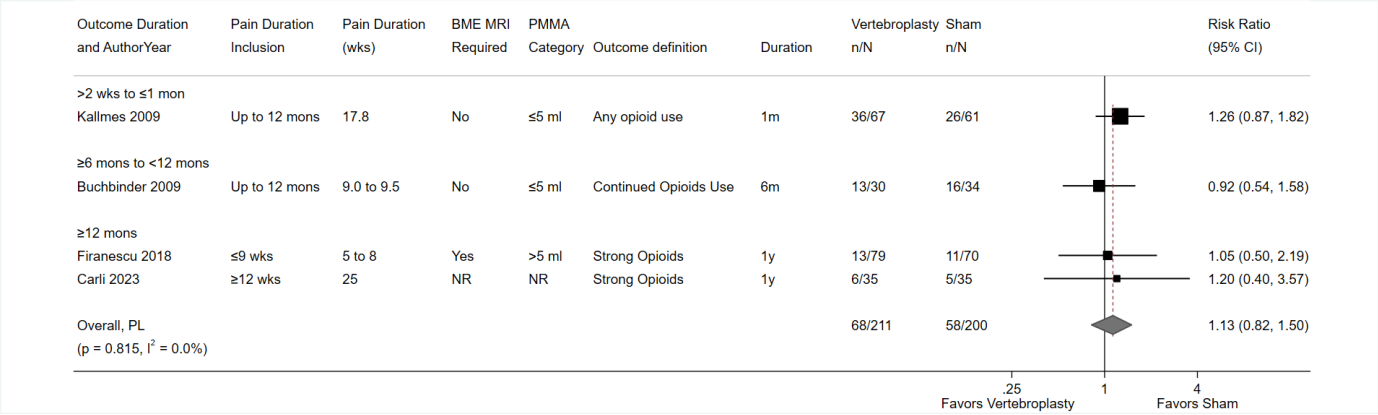
4.2.1.1.1.2 Secondary Outcomes

Opioid use

Four RCTs reported on opioid use. [24,29,67,89] The proportion of patients using strong opioids (e.g., morphine, fentanyl) and weaker opioids (e.g., codeine, tramadol) was similar between patients

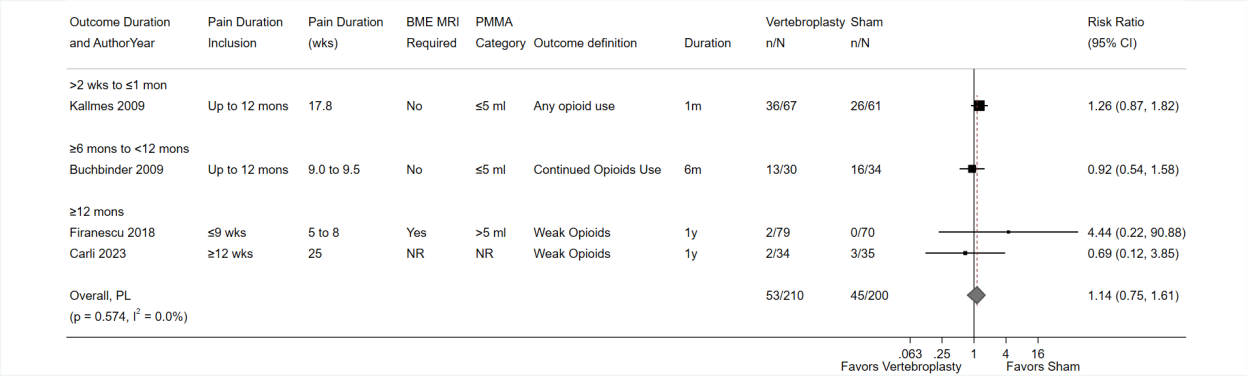
receiving VP and those receiving sham treatment at final follow-up across all trials. (Figures 5 and 6). Opioid use was also similar between VP and sham in analyses done by individual time frames (Appendix P, Figures P1 and P2)

Figure 5. VP vs. sham procedures: Strong opioid use by latest follow-up



BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; RR = risk ratio, wks = weeks

Figure 6. VP vs. sham procedures: Weak opioid use by latest follow-up

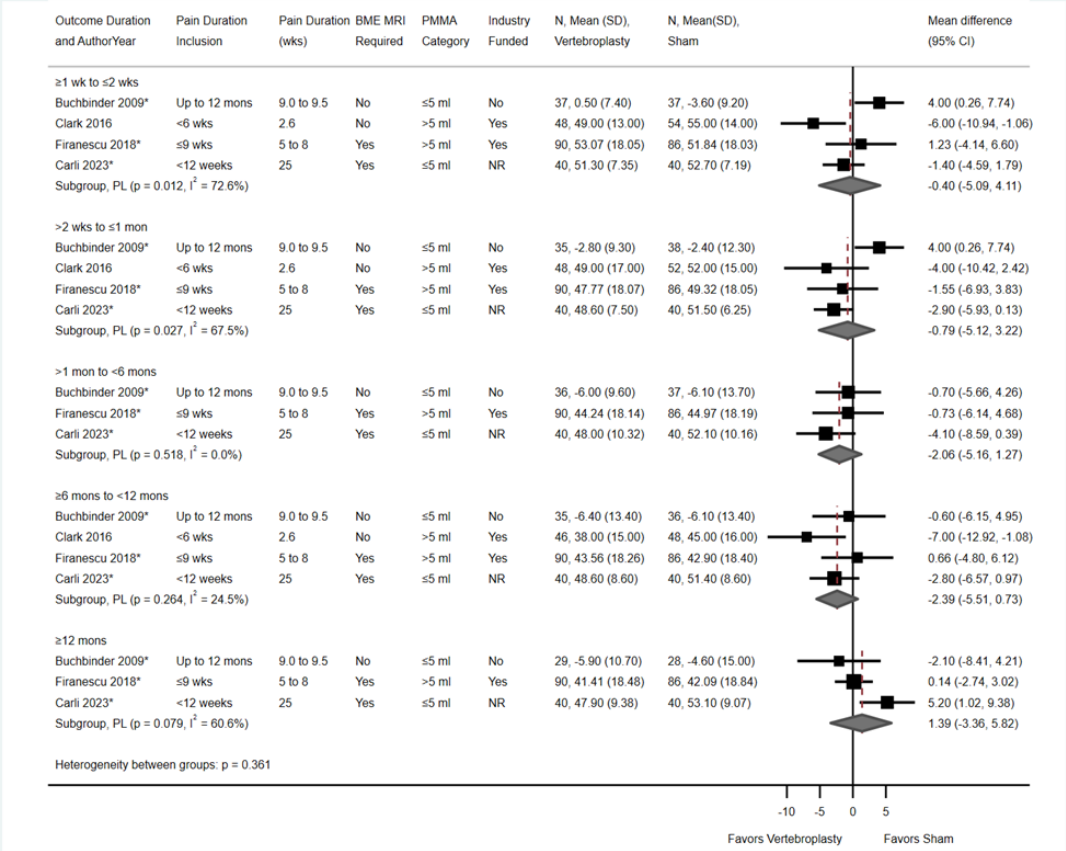


Duration indicates time at last follow-up.
BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; RR = risk ratio, wks = weeks

Quality of Life

Quality of life measures reported across trials included the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO, 0-100 scale), European Quality of Life–5 Dimensions (EQ–5D 0 to 1 scale) and the Short-Form-36 Mental Component Score (MCS) and Physical Component Score PCS (0-100 scales). The QUALEFFO (0-100 scale) was most frequently used with four RCTs reporting it.[24,29,41,67]. Across time frames, quality of life was similar for VP and sham procedures based on the QAULEFFO. (Figure 7) Similarly, quality of life was similar for VP and sham procedures measured using EQ-5D and SF-36 MCS and PCS. (Table 7) For EQ-5D, one trial used 0.074 as the minimal clinically important difference.[24] Based on this, the effect sizes at <1 week, >2 weeks to ≤1 month were below the threshold for meaningful effect.

Figure 7. VP vs. sham procedures: Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO, 0-100 scale)



BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; RDQ = Roland-Morris Disability Questionnaire, MRDQ = Modified Roland-Morris Disability Questionnaire SD = standard deviation, wks = weeks

Table 7. Vertebroplasty versus sham procedures: Additional Quality of Life Measures

Outcome	Time	Studies (N)	VP vs. Sham MD (95% CI)
European Quality of Life–5 Dimensions (EQ–5D 0 to 1 scale)	<1 week	1 RCT (N = 110) Clark 2016	MD 0.04 (0.00, 0.08)
	≥1 week to ≤2 weeks	2 RCTs (N = 179) Clark 2016, Buchbinder 2009	0.01 (-0.05, 0.07)
	>2 weeks to ≤1 month	3 RCTs (N = 299) Clark 2016, Buchbinder 2009, Kallmes 2009	0.04 (0.00. 0.08)
	>1 month to <6 months	3 RCTs (N = 219) Clark 2016, Buchbinder 2009, Hansen 2019	0.02 (-0.04, 0.07)
	≥6 months to <12 months	2 RCTs (N = 168) Clark 2016, Buchbinder 2009	0.06 (-0.03, 0.11)
	≥12 months	2 RCTs (N = 113) Buchbinder 2009, Hansen 2019	0.08 (-0.18, 0.10)
	>2 weeks to ≤1 month	1 RCT (N = 128) Kallmes 2009	1.00 (-3.11, 5.11)

Outcome	Time	Studies (N)	VP vs. Sham MD (95% CI)
SF-36 MCS (0-100 scale)	>1 month to <6 months	1 RCT (N = 46) Hansen 2019	1.70 (-8.37, 4.97)
	≥12 months	1 RCT (N = 46) Hansen 2019	5.00 (-11.11, 1.11)
SF-36 PCS (0-100 scale)	>2 weeks to ≤1 month	1 RCT (N = 113) Kallmes 2009	1.00 (-1.67, 3.67)
	>1 month to <6 months	1 RCT (N 46 =) Hansen 2019	-2.50 (-8.45, 3.45)
	≥12 months	1 RCT(N = 46) Hansen 2019	-3.30 (-9.42, 2.82)

CI = confidence interval; MD = mean difference; RCT = randomized controlled trial; SF-36 = Short-Form 36, MCS = Mental Component Score, PCS = Physical Component Score

4.2.1.1.2 Vertebroplasty versus Usual Care

Nine trials (reported in 12 publications) compared VP versus UC for the treatment of osteoporotic compression fractures (patients with fracture due to cancer were excluded) (Table 8 and Appendix F Table F1).[18,33,64,93-95,101,144,145,180,195,197] Five trials were conducted in Europe[18,93,101,144,180] three trials in China,[33,195,197] and one trial in Iran.[64] Only one trial received partial funding from industry but stated that the sponsors were not involved in any aspect of the trial.[93] Sample sizes ranged from 34 to 400 (N=1,334). The majority of patients were female (65% to 100%) and the mean patient age ranged from 66 to 80 years; in one trial[101] that did not report the mean age the range was 56 to 82 years. Mean pain duration was 4 week or less in four RCTs[93,101,144,195] and 12 weeks or greater in four RCTs [18,33,64,180]; one RCT did not indicate pain duration prior to study entry.[197] Evidence of bone marrow edema (BME) was required for study inclusion in seven RCTs.[18,33,64,93-95,101,180,195] Patients presented with both single and multiple level fractures (range, mean 1 to 2.5 vertebra treated, across trials reporting this data)[18,33,64,93,144,180,195,197] in all but one RCT that included only single fractures.[101] In the trials that reported mean volume of PMMA injected during VP, all used less than 5 ml (range, 3.2 to 4.5 ml).[33,64,93,101,180,195,197] Usual care consisted of various conservative therapies (e.g., analgesics, physical therapy, graded activity, and braces or walking aids), but only one trial described specific medications and doses.[64] The duration of follow up ranged from 6 to a mean of 49 months.

Five RCTs were considered fair quality[18,33,93,144,180] and four were poor quality[64,101,195,197] Methodologic limitations for fair-quality trials included unclear concealment of treatment allocation and some baseline differences between treatment groups. Additional limitations for poor-quality trials included high or unknown attrition and/or lack of intent-to-treat analysis.

Two poor quality trials were excluded from efficacy meta-analyses but were included for data on harms. One trial[101] only reported efficacy outcomes for the VP arm and the other[64] had serious data discrepancies—implausible values for standard deviations or results (mean differences, 95% confidence intervals [CIs], and p values) inconsistent with reported data. One trial included both VP and KP procedures versus UC but provided limited data for each treatment separately (most analyses combined the VP and KP arms); this trial is included in the VP vs. KP and KP vs. UC sections as well.[197]

Table 8. Study Characteristics of RCTs comparing VP versus UC for the treatment of vertebral compression fractures due to osteoporosis.

Study, year Country Quality	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)* Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME required on MRI	Duration F/U (months)	Industry funding	In MA
Blasco, 2012 Spain Fair	73	78%	125	6.8 (0.4)† VAS ≥4	20.3	3.3 (2.5)‡ (treated: 2.5 (1.6))‡	NR	Yes	12	No	Yes
Yang, 2016 China Poor	77	65%	135	7.6 (1.1) VAS ≥5	0.8	1: 85% 2: 15% (all treated)	4.5	Yes	12	No	Yes
Leali, 2016 Italy, France, Switzerland Poor	NR (range 56-82)	100%	400	NR§	NR (“acute”)	1: 100% (all treated)	4	Yes	6	NR	No (efficacy) Yes (safety)
Chen, 2014 China Fair	66	70%	96	6.5 (0.9) (inclusion NR)	30.2	2.1 (0.72) (treated NR)	3.6	Yes	12	No	Yes
Farrokhi, 2011 Iran Poor	73	73%	82	7.8 (1.7)† (inclusion NR)	28.5	NR**	3.5	Yes	36	No	No (efficacy) Yes (safety)
Klazen, 2010, 2010, 2010 The Netherlands, Belgium Fair	75	69%	202	7.5 (NR) VAS ≥5	4	2.3 (1.7) (treated NR)	4.1	Yes	12	Yes (partial)††	Yes
Rousing, 2009, 2010 Denmark Poor	80	82%	49	8.1 (1.5)† (inclusion NR)	1.1	1: 76% 2: 20% 3: 4%‡‡	NR	No	12	No	Yes

Study, year Country Quality	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)* Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME required on MRI	Duration F/U (months)	Industry funding	In MA
						(treated NR)					
Voormolen, 2007 The Netherlands Fair	73	82%§§	34	7.3 (NR) (inclusion NR)	11.6	3.2 (NR) (treated: 1.4 (NR))	3.2	Yes	12	NR	Yes
Yi, 2014 China Poor	NR (71, combined VP/KP)	NR (68%, combined VP/KP)	211	NR	NR	NR (1.4 (0.75), combined VP/KP; treated NR)	4.0	No	Mean 49.4	NR	Yes

BME = bone marrow edema; MA = meta-analysis; ml = milliliter; MRI = magnetic resonance imaging; NR = not reported; PMMA = polymethylmethacrylate; RCT = randomized control trial; SD = standard deviation; VAS = visual analogue scale; VCF = vertebral compression fracture; VP = vertebroplasty.

* For most studies, weighted mean and SD was calculated using mean estimates, SDs or 95% confidence intervals, and n's for each group at baseline.

† Some baseline differences existed between groups for VAS pain scores:

Blasco, 2012, mean (SD): 7.2 (0.3) vs. 6.3 (0.4);

Farrokhi, 2011, mean (SD): 8.4 (1.6) vs. 7.2 (1.7);

Rousing, 2009, mean (95% CI): 7.5 (95% CI 6.6 to 8.4) vs. 8.8 (95% CI 8.2 to 9.3).

‡ In the VP group there were a mean 3.55 (2.82) VCFs at baseline; the mean number of vertebral bodies treated was 2.46 (1.56).

4.2.1.1.2.1 Primary Outcomes

Pain

Two trials reported the proportion of patients who were considered pain responders at 12 months but used different criteria. In one trial, a similar proportion of patients in the VP and UC groups achieved a score of less than 4 on a 0-10 VAS (N=95, 44.7% vs. 47.9%, RR 0.93, 95% CI 0.60 to 1.44).[18] In the second trial, VP was associated with a large increase in the likelihood of achieving "complete pain relief" (N=89, 84.8% vs. 34.9%, RR 2.43, 95% CI 1.59 to 3.72).[33]

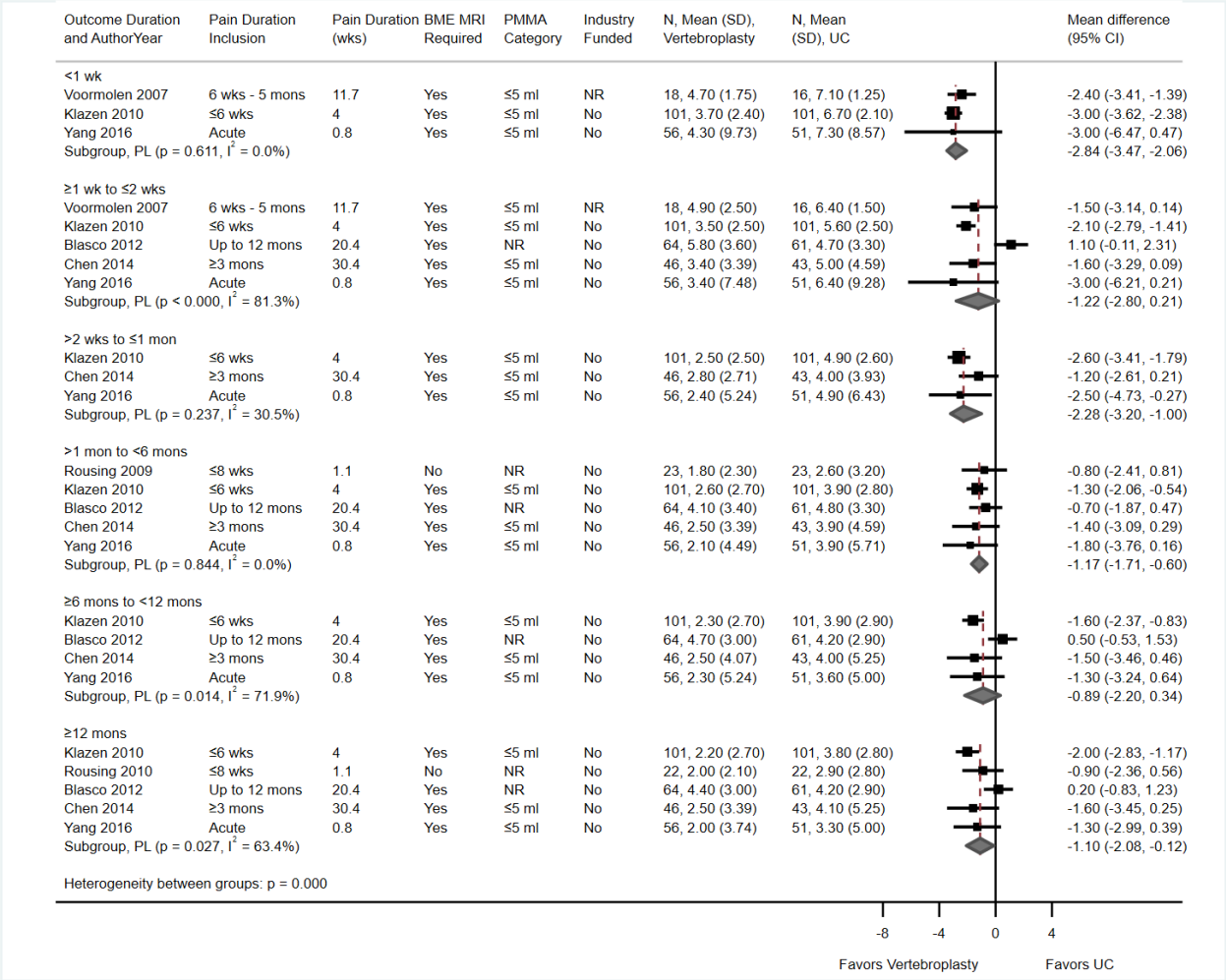
Six RCTs reported VAS or NRS pain scores (scale 0-10).[18,33,93,144,180,195] In general, vertebroplasty was associated with improvement in pain intensity versus usual care with the largest improvements seen at earlier timepoints (Figure 8): a large improvement at less than 1 week (3 RCTs, N=343) and at greater than 2 weeks to 1 month (3 RCTs, N=398); and a moderate improvement at greater than 1 month to less than 6 months (5 RCTs, N=569) and at 12 months or longer (5 RCTs, N=567). At 1 to 2 weeks and 6 months to less than 12 months, there was no difference in pain improvement between groups in the pooled estimates; heterogeneity was substantial, and the estimates were imprecise. Removal of one outlier trial[18] resulted in a moderate improvement in pain favoring vertebroplasty at these latter timepoints (the estimate at ≥12 months remained moderate) and eliminated heterogeneity (Appendix P, Figure P5). It is unclear why this trial showed different results than the other trials (tended favor UC); however, the vertebroplasty group had slightly more vertebral fractures (52% vs. 45% with >2 fractures) and greater pain at baseline (7.2 vs. 6.3) which may have impacted outcomes. Sensitivity analyses excluding the poor-quality trial[195] and all timepoints showed similar results to the original analyses, however, there was more heterogeneity and estimates were more imprecise; at 12 months or longer (4 RCTs, N=460) there was no longer a difference between the groups (Appendix P, Figure P6).

Function

No trial reported the proportion of patients who were considered function responders.

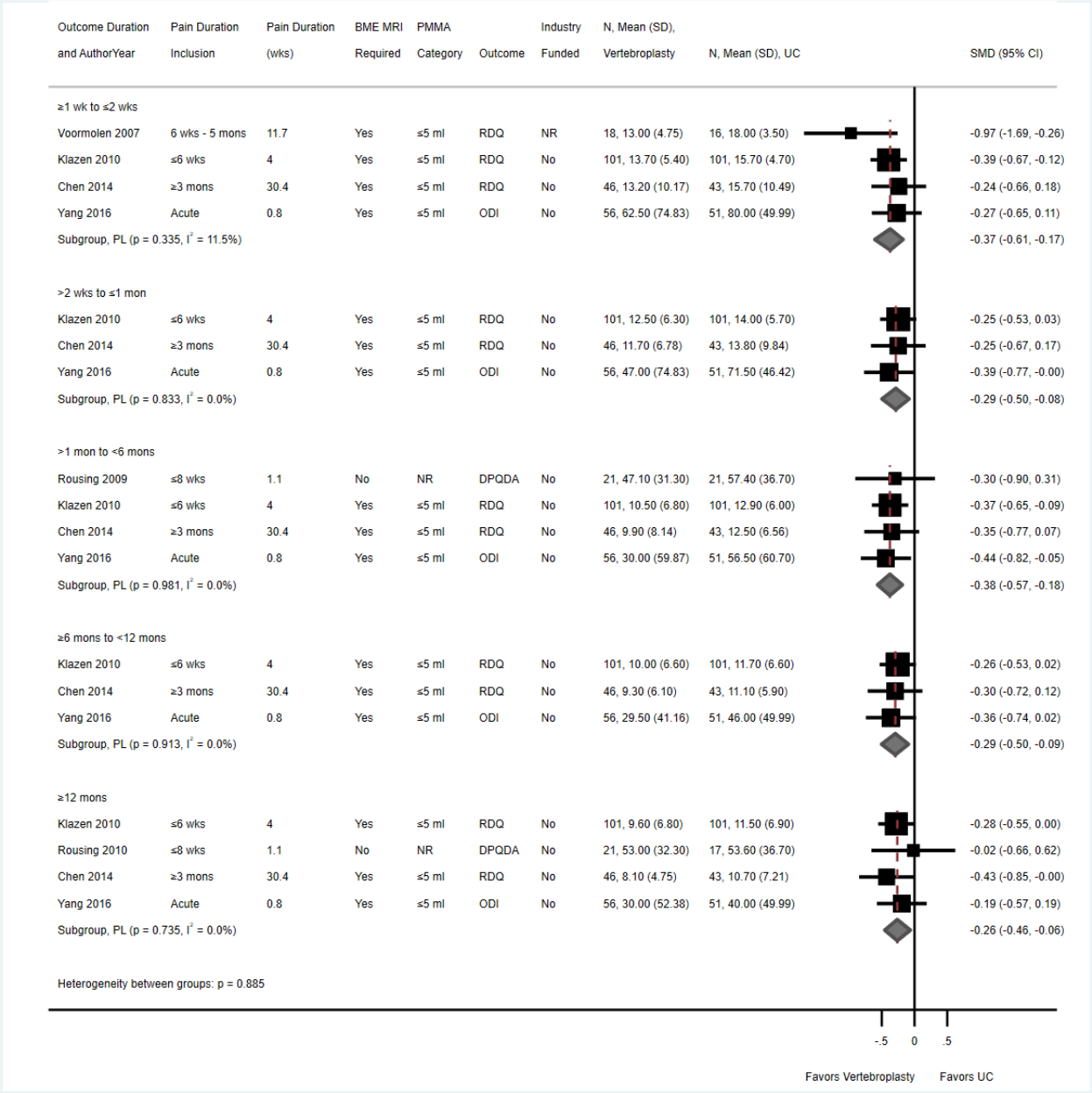
Five RCTs reported function scores using different outcomes measures (RDQ, ODI, DPQ).[33,93,144,180,195] Vertebroplasty was associated with a small improvement in function versus usual care across all timepoints measured (Figure 9); standardized mean differences (SMDs) ranged from -0.26 to -0.38 across 3 to 4 RCTs (N range, 398 to 440). Exclusion of the poor-quality trial[195] did not change conclusions and analyses confined only to RDQ scores were generally similar (Appendix P, Figures P7 and P8).

Figure 8. VP vs. UC: Pain scores (VAS or NRS, 0-10 scale) from RCTs



BME = bone marrow edema; CI = Confidence Interval; mons = months; MD = mean difference; MRI = magnetic resonance imaging; NRS = numerical rating scale; PMMA = polymethylmethacrylate; SD = standard deviation; UC = usual care; VAS = visual analogue scale; wks = weeks.

Figure 9. VP vs. UC: Function scores from RCTs



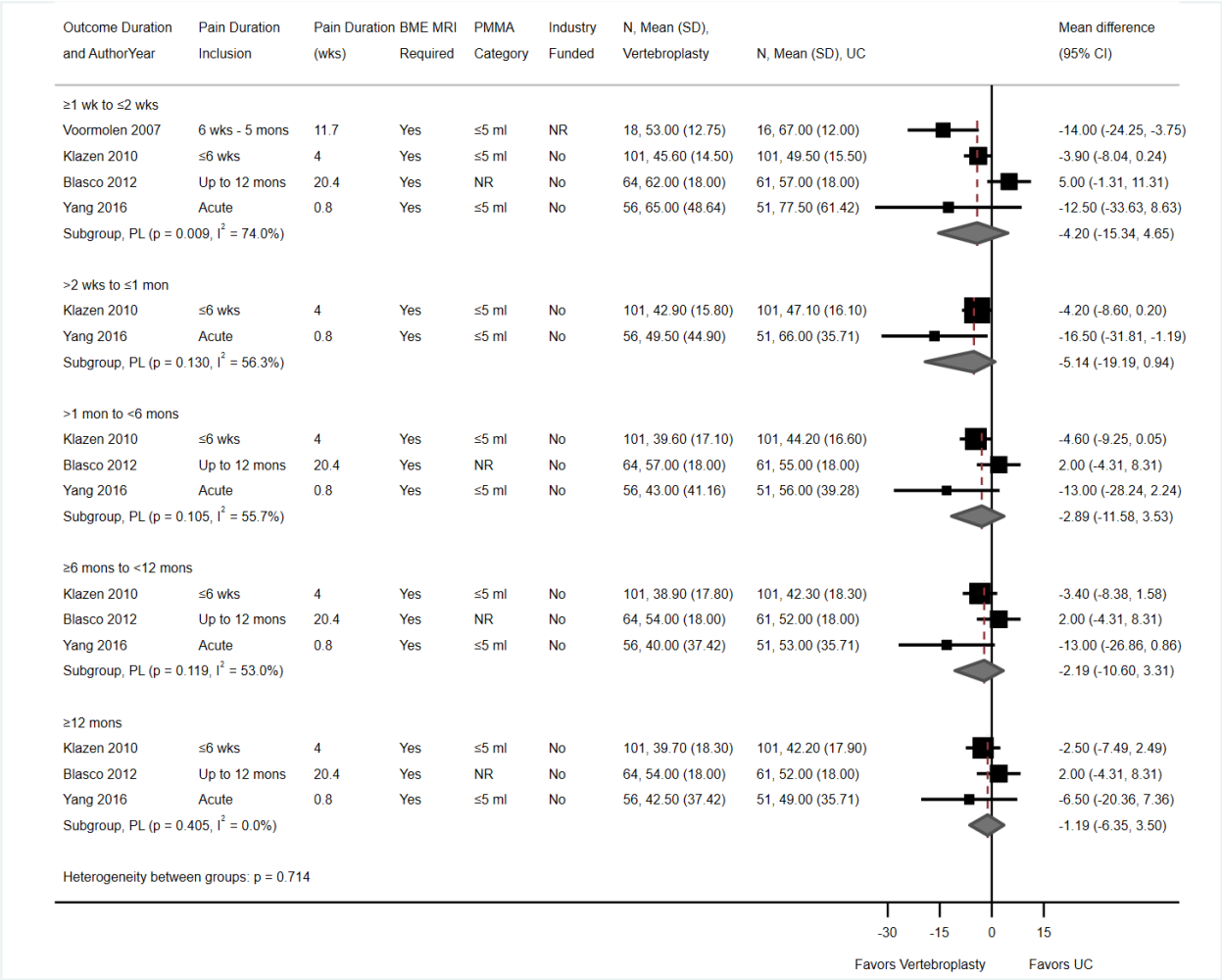
BME = bone marrow edema; CI = Confidence Interval; DPQDA = Dallas Pain Questionnaire Disability Assessment; mons = months; MRI = magnetic resonance imaging; ODI = Oswestry Disability Index; PMMA = polymethylmethacrylate; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SMD = standardized mean difference; UC = usual care; wks = weeks.

4.2.1.1.2.2 Secondary outcomes

Quality of Life

Four RCTs (N=468)[18,93,180,195] reported quality of life using the Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO, 0-100 scale), a vertebral fracture-specific measure of quality of life, with no differences between groups in pooled analyses at any timepoint (Figure 10); heterogeneity was high and estimates were imprecise. Exclusion of the poor-quality trial[195] resulted in similar, though slightly attenuated, estimates (Appendix P, Figure P9). A sensitivity analysis excluding the outlier trial[18] (Appendix P, Figure P10) showed a small improvement in QOL favoring VP versus UC at the earliest timepoint (1 to 2 weeks, 3 RCTs, N=343, MD -5.55, 95% CI -18.02 to -0.24) but there remained no difference between groups at all other times.

Figure 10. VP vs. UC: QUALEFFO scores from RCTs



BME = bone marrow edema; CI = Confidence Interval; mons = months; MD = mean difference; MRI = magnetic resonance imaging; NRS = numerical rating scale; PMMA = polymethylmethacrylate; SD = standard deviation; UC = usual care; VAS = visual analogue scale; wks = weeks.

Other QOL outcomes reported by two fair-quality RCTs are summarized in Table 9. Both trials[93,144,145] reported the EQ5D, a general quality of life measure. At all timepoints except for 3 months, VP was associated with a small improvement EQ5D scores versus UC. One trial[144,145] reported SF-36 PCS and MCS scores and found no difference between groups at 3 and 12 months.

Table 9. VP vs. UC: Summary of other quality of life outcomes:

Outcome	Time	Studies	VP vs. UC MD (95% CI)*
EQ5D (0-1, higher = better)	≥1 week to ≤2 weeks (1 week)	1 RCT (N=202) Klazen 2010	0.10 (0.02 to 0.18)
	>2 weeks to ≤1 month (1 month)	1 RCT (N=202) Klazen 2010	0.10 (0.03 to 0.17)
	>1 month to <6 months (3 months)	2 RCTs (N=234) Klazen 2010 Rousing 2009	Pooled MD 0.07 (-0.13 to 0.32), I^2=76.8%
	≥6 months to <12 months (6 months)	1 RCT (N=202) Klazen 2010	0.10 (0.02 to 0.18)
	≥12 months (12 months)	2 RCTs (N=234) Klazen 2010 Rousing 2009	Pooled MD 0.10 (0.02 to 0.19), I^2=0%
SF-36 PCS (0-100, higher= better)	>1 month to <6 months (3 months)	1 RCT (N=43) Rousing 2009	4.70 (-1.12 to 10.52)
	≥12 months (12 months)	1 RCT (N=41) Rousing 2010	1.60 (-4.73 to 7.93)
SF-36 MCS (0-100, higher= better)	>1 month to <6 months (3 months)	1 RCT (N=43) Rousing 2009	2.70 (-5.45 to 10.85)
	≥12 months (12 months)	1 RCT (N=41) Rousing 2010	0.30 (-7.64 to 7.04)

CI: confidence interval; EQ5D: EuroQoL 5 dimensions questionnaire; f/u: follow-up; MD: mean difference; mos: months; RCT: randomized controlled trial; SF-36 MCS: Short-Form 26 Mental Component Score; UC: usual care; VP: vertebroplasty; wks: weeks; SF-36 PCS: Short-Form 26 Physical Component Score;
*calculated

Opioid use

VP was associated with a large increase in the likelihood of using major opioids at 12 months compared with usual care (N=83, 36.6% vs. 16.7%, RR 2.20, 95% CI 1.00 to 4.82) but there were no differences between groups at earlier timepoints (range, 2 weeks to 6 months) in one RCT[18]; minor opioid use was similar between groups (Table 10).

Three trials found VP associated with a decrease in pain medication use (not restricted to opioids) compared with UC; differences were significant at all timepoints up to 12 months in one trial[33] only up to 1 month in the other trial (with 12-month follow up),[93] and at 2 weeks in the third trial[180] (Appendix G, Table G1).

Table 10. VP vs. UC: Summary of opioid use from Blasco 2012

		VP	UC	
Outcome	**F/U**	**% (n/N)**	**% (n/N)**	**RR (95% CI)**
Major opioids	≥1 week to ≤2 weeks (2 weeks)	35.7% (20/56)	29.3% (17/58)	1.22 (0.72 to 2.08)
	>1 month to <6 months (2 months)	30.8% (16/52)	30.4% (17/56)	1.01 (0.57 to 1.79)
	≥6 months to <12 months (6 months)	36.7% (18/49)	32.7% (17/52)	1.12 (0.66 to 1.92)
	≥12 months (12 months)	36.6% (15/41)	16.7% (7/42)	2.20 (1.00 to 4.82)
Minor opioids	≥1 week to ≤2 weeks (2 weeks)	23.2% (13/56)	32.8% (19/58)	0.71 (0.39 to 1.29)
	>1 month to <6 months (2 months)	26.9% (14/52)	28.6% (16/56)	0.94 (0.51 to 1.73)
	≥6 months to <12 months (6 months)	16.3% (8/49)	26.9% (14/52)	0.61 (0.28 to 1.32)
	≥12 months (12 months).	17.1% (7/41)	23.8% (10/42)	0.71 (0.30 to 1.70)

UC = usual care; F/U = follow-up; RR = risk ratio; VP = vertebroplasty.

4.2.1.1.3 Vertebroplasty versus Minimally Invasive Procedures (Nerve Blocks)

Two RCTs[168,181] and one NRSI[8] were identified that met inclusion criteria and compared VP to other minimally invasive surgeries (i.e., nerve or facet blocks).

Two trials compared vertebroplasty to other minimally invasive surgeries[168,181]. In one study[168], vertebroplasty, using 2 to 5 ml of PMMA, was compared to a medial branch spinal nerve block. The medial branch spinal nerve block targeted the facet joints above and below the fracture. This procedure involved a mixture of 0.5% bupivacaine with 40 mg depomedrone, with each medial branch blocked using 1 to 1.5 ml of the solution. The other study[181] compared vertebroplasty to a facet block. Facet block targeted the facet joint capsule of the fractured vertebral body. All patients in both trials were required to wear a brace to aid ambulation for three months following their procedures.

One trial was conducted in the UK[168] and the other in China[181]. The mean age of participants ranged from 63 to 82 years; notably, the UK trial[168] focused on frail, older, hospitalized patients and excluded those younger than 70 years of age. Baseline pain levels differed significantly between trials: in the Chinese trial[181], the mean baseline pain was 7.7±1.1 on the VAS, while in the UK trial[168] involving older patients, the mean baseline pain was reported as 9.0 (SD not reported) on the NRS, with exclusion criteria set for patients with pain <7 on the NRS. The UK trial[168] included only 30 patients, whereas the Chinese trial[181] included 217 patients at baseline. See Table 11.

Between 70% and 83% of patients in these trials were female, although the UK trial[168] showed a significant difference in the proportion of female patients between groups (57% in the vertebroplasty group and 85% in the nerve block group). Both trials included patients with either acute or subacute fractures (<6 to 8 weeks), though in the Chinese trial[181], 86% of patients had acute fractures (≤2 weeks). Both trials required participants to have MRI findings consistent with bone marrow edema at the fracture site. The duration of follow-up varied, with up to 8 weeks in the UK trial[168] and 12 months in the Chinese trial[181].

Both trials were rated as fair (Appendix Table E4). Limitations included the inability to clearly mask participants, care providers, or outcome assessors in either trial[168,181], as well as a lack of clarity in allocation concealment in the Chinese trial[181]. Additionally, the UK trial[168] faced challenges in recruiting similar numbers of female participants across groups, though this trial had a small sample size of 30. The UK trial[168] reported funding from the government, while the Chinese trial[181] reported no funding.

Table 11. Study Characteristics of Trials Comparing Vertebroplasty versus Nerve Block in Patients with Fractures due to Osteoporosis

Study, year Country Quality (intervention)	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)* Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME MRI	Duration follow-up (months)	Industry funding	In MA
Tan, 2023 UK Fair (VP vs. Medial branch spinal nerve block)	82	70%†	30	9.0 (NR) Inclusion: NRS ≥7	≤6	NR‡	2 to 5	Yes	1.8	No	Yes
Wang, 2016 China Fair (VP vs. Facet block)	63	83%	217	7.7 (1.1) Inclusion: NR	≤6§	NR	3 to 9	Yes	12	No	Yes

BME = bone marrow edema; MA = meta-analysis; ml = milliliter; MRI = magnetic resonance imaging; NR = not reported; NRS = numerical rating scale; PMMA = polymethylmethacrylate; RCT = randomized control trial; SD = standard deviation; VCF = vertebral compression fracture; VP = vertebroplasty.

* For most studies, weighted mean and SD was calculated using mean estimates, SDs or 95% confidence intervals, and n's for each group at baseline.

† Some baseline differences existed between groups for % female:

Tan, 2023: 57% vs. 85%

‡ Inclusion criteria allowed for up to three fractures.

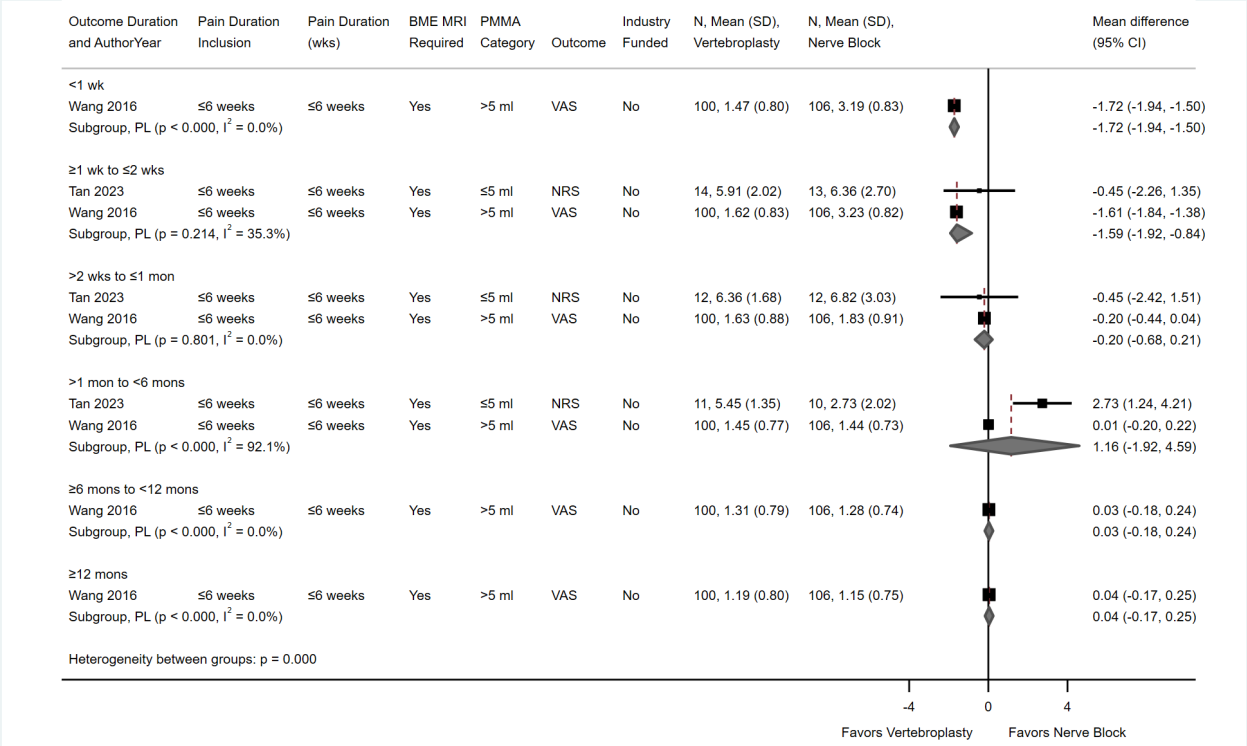
§ 85.9% of patients had acute fractures ≤2 week

4.2.1.1.3.1 Primary Outcomes from RCTs

Pain

VP was associated with moderate improvement in pain compared with nerve or facet block at <1 week (1 RCT, 1 day post-treatment)[181] and at ≥1 to ≤2 weeks (2 RCTs, 2 week follow-up)[168,181], but the effect did not persist at later timepoints (Figure 11). At >1 to <6 months (2 to 3 month follow-up), there was notable inconsistency (I^2=92.1%) between the trials and the estimate was imprecise; the smaller trial[168] that included older patients reported a large effect favoring spinal nerve block while the larger study reported similar improvement between VP and facet block.[181] All pooled results were driven by the larger trial.[181]

Figure 11. VP versus Nerve Block: Pain score (VAS or NRS, 0-10 scale) from RCTs



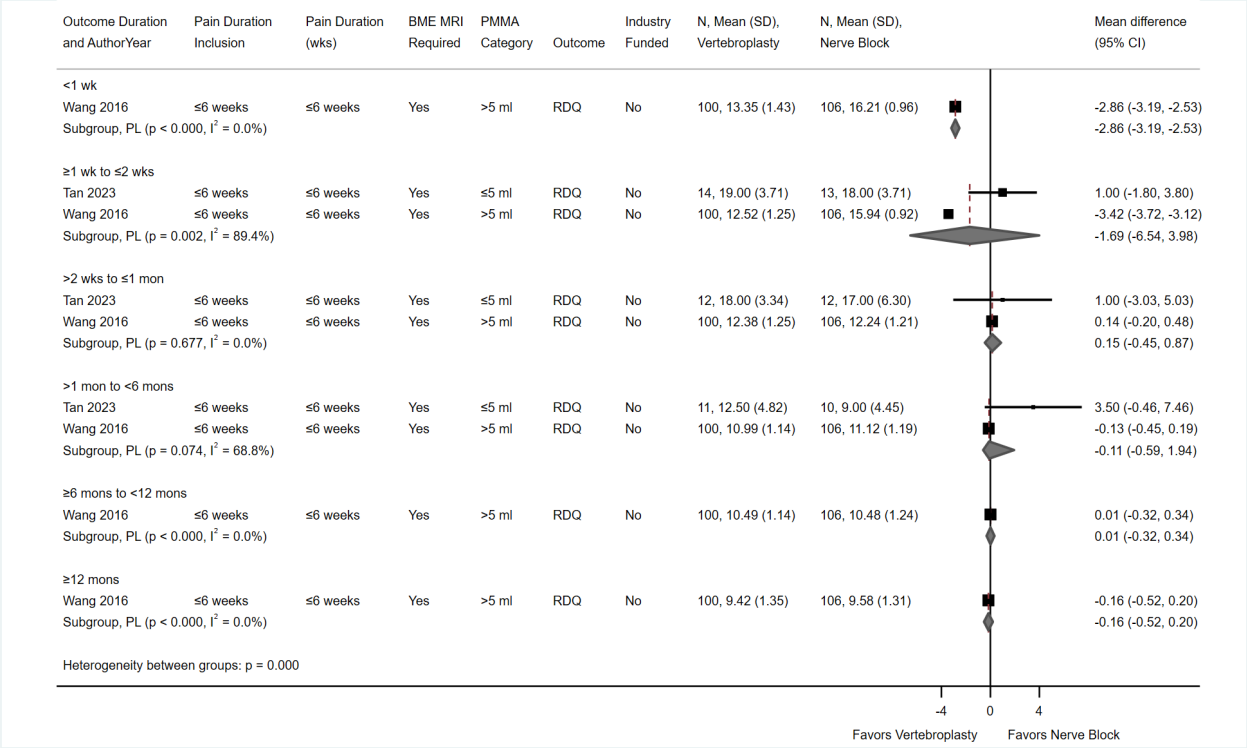
BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; NRS = numerical rating scale; PL = profile likelihood; PMMA = polymethylmethacrylate; RCT = randomized control trial; SD = standard deviation; VAS = visual analogue scale; wks = weeks.

Function

VP was associated with a moderate improvement in RDQ scores compared with facet block at <1 week (1 day) and ≥1 to ≤2 weeks (1 week) in one large trial,[181] (Figure 12). The smaller trial found no difference between VP and spinal nerve block at ≥1 to ≤2 weeks (1 month) and pooled analysis across both trials[168,181] at this timepoint resulted in no difference between groups but the point estimate was very imprecise and there was substantial heterogeneity (89.4%). There were no differences between groups in RDQ scores at other timepoints (Figure 12).

Both trials reported function using additional measures and reported similar results (Appendix G, Table G1). VP was associated with a small improvement in ODI scores at <1 week (1 day) and ≥1 to ≤2 weeks (1 week) compared with facet joint block in the large trial,[181], but this effect did not persist at later timepoints. There was no difference between VP and medial branch nerve root block in Extended Activities of Daily Living (NEADL) scores at any time point in the other trial.[168]

Figure 12. VP vs. Nerve Block. Function Scores (RDQ, 0-24).



BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; PL = profile likelihood; PMMA = polymethylmethacrylate; RCT = randomized control trial; RDQ = Roland Morris Disability Index; SD = standard deviation; wks = weeks.

4.2.1.1.3.2 Secondary Outcomes from RCTs

Quality of Life

VP was associated with similar improvement in quality of life compared with spinal nerve/facet blocks at all follow-up times in both trials; one trial reported EQ-5D scores[168] and the other trial reported SF-36 PCS and MCS scores,[181] (see Appendix Table G1).

Opioid use

In one trial, all but one patient was taking strong opioids at baseline (VP, 93%; nerve block, 100%), which continued through 1 week post-procedure.[168] Similar proportions of patients who received VP versus medial branch nerve block were still using opioids (and/or other analgesics) regularly at 4 weeks (66.6% [8/12] vs. 75% [9/12]; RR 0.89, 95% CI 0.53 to 1.49) and 8 weeks (50% [4/11] vs. 66.6% [6/10]; RR 0.61, 95% CI 0.24 to 1.54).

NRSI

Only one retrospective NRSI (N=164) reported on vertebroplasty versus medial branch block in patients with single-level osteoporotic fractures in the thoracolumbar vertebrae[8]. Medial branch block (n=72) targeted the facet joint capsule of the fractured vertebral body (See appendix X for details). All patients were prescribed bed rest prior to the procedure. The mean age of participants was 75 years, with 26% being female. Notably, 15.2% (14/92) of vertebroplasty patients were admitted due to a new (not defined) fracture, compared to 4% (3/72) of medial branch block patients. No other demographic data were reported. The study was rated as poor. The primary limitations were related to the retrospective design, including the inability to ascertain sampling methods, attrition, and blinding.

4.2.1.1.3.3 Primary Outcomes from NRSI

Function

While vertebroplasty was associated with improvements in ODI scores compared with medial branch block at all time points up to 24 months, the mean differences between the groups were consistently below the threshold for a small effect[8] (see Appendix Table I1).

Pain

At 1 week, there was a small improvement in VAS pain scores (0-10) favoring medial branch block versus VP (N=164, MD 0.8, 95% CI 0.4 to 1.2).[8] However, at 1 month (N=164, MD -0.5, 95% CI -0.9 to -0.1) and 3 months (N=164, MD -1.0, 95% CI -1.3 to -0.7), VP was associated with a small improvement in pain compared with medial branch blocks. By 12 and 24 months, there was no difference between groups (Appendix Table I1).

4.2.1.1.4 Vertebroplasty versus Kyphoplasty

Ten trials (reported in 11 publications) compared VP versus KP for the treatment of osteoporotic compression fractures (patients with fracture due to cancer were excluded) (Table 12 and Appendix F, Table F1, Appendix G, Tables G1 and G4).[52,59,61,77,108,109,179,182-184,197] Two trials were conducted in the U.S.,[52,61] three in Europe[59,77,179] and five in Asia, primarily China.[109,182-184,197] Both U.S. trials and one trial from Germany were funded by industry. [52,61,179] Sample sizes ranged from 66 to 404 (N=1,337). The majority of patients were female (65% to 100%) in all but one trial (44%),[184] and the mean patient age ranged from 42 to 82 years. The trial that enrolled only women included patients who had undergone bilateral resection of ovarian cancer. Mean pain duration was 6 weeks or less in six RCTs[52,59,61,109,179,184] (range, mean 2.5 to 3.6 weeks across the 3 trials that reported this), 4 weeks or greater in two RCTs[77,182] and not reported in two RCTs.[183,197] Evidence of bone marrow edema (BME) was required for study inclusion in only two RCTs.[52,77] The mean number of vertebral fractures at baseline or treated was not always clear but most trials appeared to enrolled patients with 1-3 fractures; two trials treated only single fractures.[59,109] In the trials that reported the volume of PMMA injected during the procedures, means ranged from 3.1 to 4.9 ml for VP and 3.8 to 5.6 ml for KP.[52,59,108,109,179,182,183,197] The VP and KP procedures were conducted according to the standard practice of the institution and the treating physician. In all trials, the KP procedures used a standard balloon to create space/regain vertebral height; one trial had three arms and evaluated a second type of kyphoplasty, shield kyphoplasty, that uses a permanent implant rather than a balloon.[59] The duration of follow-up ranged from 3 months to a mean of 49.4 months.

Three RCTs were considered fair quality[61,77,182] and the other seven were poor quality.[52,59,109,179,183,184,197] Methodological limitations for fair-quality trials included unclear concealment of treatment and some baseline differences between treatment groups. Additional limitations for poor-quality trials included unclear randomization method, poor reporting of baseline demographics, high or unknown attrition and lack of intent-to-treat analysis. Only three trials stated that patients were masked to the treatment received.[59,179,182] One poor-quality RCT was a quasi-randomized trial and there were large differences between treatment groups in baseline pain and function.[59]

Two trials[108,109,184] did not report the sample size at follow-up; we assumed that the follow-up sample sizes were the same as at baseline for the primary analysis and conducted sensitivity analyses around this assumption.

Table 12. Study Characteristics of Trials Comparing Vertebroplasty versus Kyphoplasty in Patients with Fractures due to Osteoporosis

Study, year Country Quality	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)* Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME required on MRI	Duration F/U (months)	Industry funding	In MA
Wang, 2015 China Fair	69	76%	107	8.1 (1.2) VAS ≥5†	≥4	NR‡ (≥1 VCF)	VP: 3.3 KP: 4.2	No	12	No	Yes
Liu, 2010, 2015 Taiwan Poor	73	77%	100	8.0 (0.8) (Inclusion NR)	2.6	1: 100%	VP: 4.9 KP: 5.6	No	6	No	Yes
Griffoni, 2020 Italy Fair	73	82%	113	7.9 (6.9) (Inclusion NR)	≥4	1: 65% 2-3: 35% (treated NR‡)	NR	Yes	12	NR	Yes
Evans, 2016 USA Fair	76	71%§	115	7.6 (2.0) NRS ≥5	2.5	NR (≥1 VCF)	NR	NR	12	Yes	Yes
Endres, 2012 Germany Poor	68	68%	66	8.5 (1.1) (Inclusion NR)	≤6	1: 100%	VP: 3.1 KP: 3.9 SKP: 4.6	NR	6	NR	Yes
Dohm, 2014** USA Poor	66	77%	404	7.7 (NR) (Inclusion NR)	3.6	1: 57% 2: 24% 3: 19% Treated: 1: 79% 2: 17% 3: 5%	VP: 4.0†† KP: 4.6††	Yes	24	Yes	Yes

Study, year Country Quality	Mean age (years)	% Female	N randomized	Mean baseline pain (SD)* Inclusion criteria	Mean Pain duration (weeks)	No. VCFs at baseline (no. treated) mean (SD) or No. (%)	Mean PMMA volume (ml)	BME required on MRI	Duration F/U (months)	Industry funding	In MA
Vogl, 2013 Germany Poor	73	71%	77	8.4 (1.1) (Inclusion NR)	≤6	NR‡ (≤3 VCFs)	VP: 4.0 KP: 3.8	No	12	Yes	Yes
Yi, 2014 China Poor	NR (71, combined VP/KP)	NR (68%, combined VP/KP)	169	NR (Inclusion NR)	NR	NR (1.4 (0.75), combined VP/KP; treated NR)	NR (4.0, combined VP/KP)	No	Mean 49.4	NR	No (Efficacy) Yes (Safety)
Wang, 2018 China Poor	42	100%‡‡	86	SF-MPQ 57.1 (3.2) (Inclusion NR)	NR	1: 100%§§	VP: 3.9 KP: 3.8	NR	12	No	Yes
Wang, 2023 China Poor	82	44%	100	7.4 (1.2) (Inclusion NR)	≤3	NR	NR	NR	3	No	Yes

BME = bone marrow edema; KP = kyphoplasty; MA = meta-analysis; ml = milliliter; MRI = magnetic resonance imaging; NR = not reported; NRS = numerical rating scale; PMMA = polymethylmethacrylate; RCT = randomized control trial; SD = standard deviation; SF-MPQ = Short Form McGill Pain Questionnaire; SKP = Shield Kyphoplasty; VAS = visual analogue scale; VCF = vertebral compression fracture; VP = vertebroplasty.

* For most studies, weighted mean and SD was calculated using mean estimates, SDs or 95% confidence intervals, and n's for each group at baseline.

† Patients need to be experiencing ≥5 on VAS pain scale after at least 4-weeks conventional therapy.

‡ Authors do not report mean number or % of levels treated:

- Wang, 2015: 68 vs. 72 levels were treated in total;
- Griffoni, 2020: 87 vs. 59 levels were treated in total;
- Dohm, 2014: 244 vs. 235 levels were treated in total. Additionally, patients receiving unilateral procedures were 0.8% vs. 34.9%, and bilateral 99.2% vs. 65.1%;
- Vogl, 2013: 65 vs. 38 levels were treated in total.

§ Some baseline differences existed between groups for % female: Evans, 2016: 78% vs. 65%.

** Study is reported as being in a mixed population. 45% of VP patients and 37.2% of KP had osteoporosis. 0% had malignancies, and there is no reporting of trauma, so it is unclear if this population is actually mixed, or if the rest had osteopenia.

†† Median.

‡‡ All patients post bilateral resection of ovarian cancer.

§ Inferred by adding together number of fracture sites in the cervical, thoracic, and lumbar vertebra which equaled the same total number of patients.

4.2.1.1.4.1 Primary Outcomes

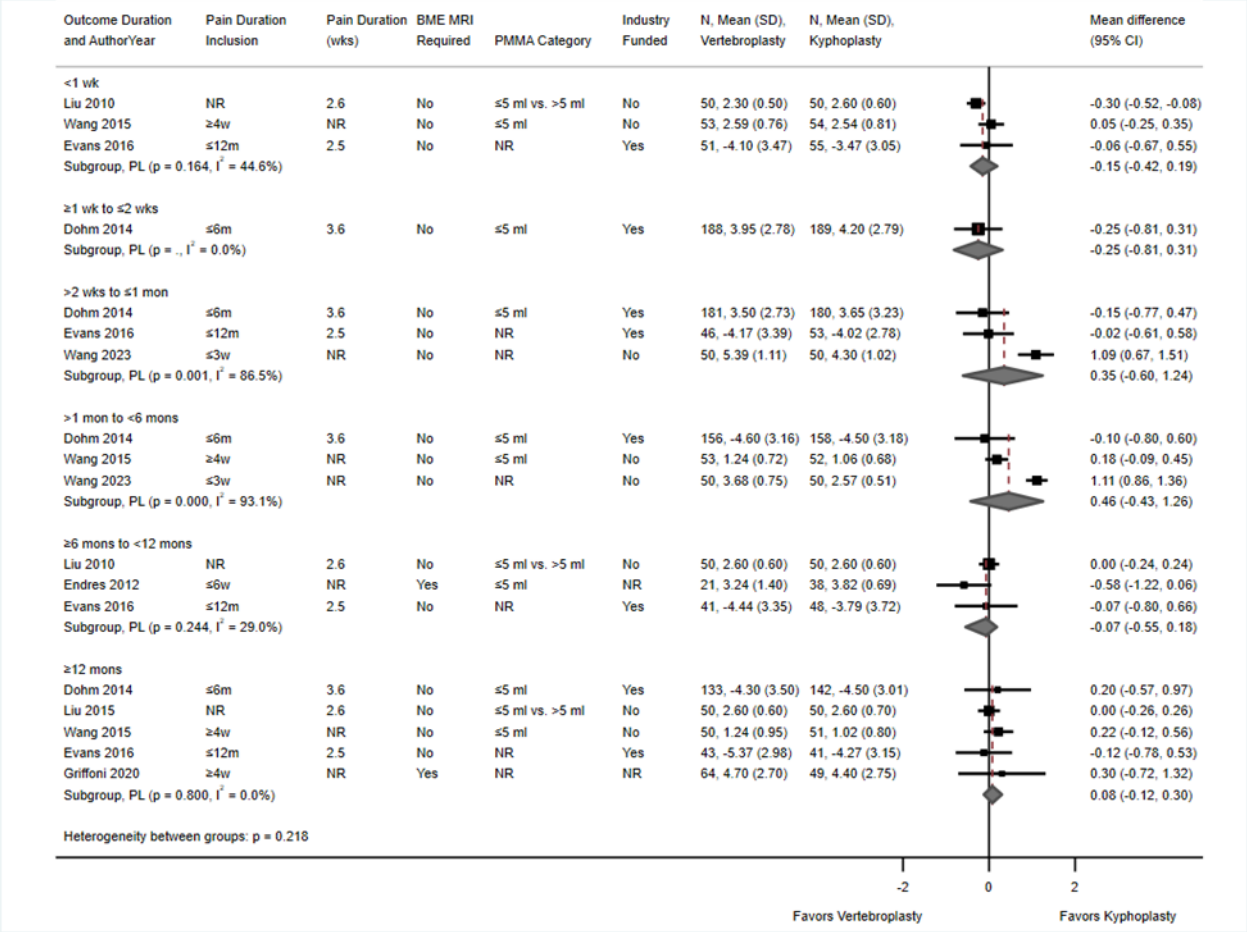
Pain

VP was associated with a small decrease in the likelihood of achieving treatment response, defined as complete ("cure"), excellent or effective (not defined) improvement in clinical symptoms, compared with KP at 3 months in one poor-quality trial (N=100, 74% vs. 94%, RR 0.79, 95% CI 0.66 to 0.94)[184]; when only complete or excellent improvement was considered, the difference was still clinically relevant (56% vs. 74%; RR 0.76, 95% CI 0.56 to 1.02).

Seven trials (in 8 publications) reported VAS or NRS pain scores (scale 0-10).[52,59,61,77,108,109,182,184] VP and KP were associated with similar improvement in pain at all time points evaluated (Figure 13). At 1 and 3 months, heterogeneity was substantial (>86%) due to an outlier trial[184]; removal of this poor-quality trial eliminated heterogeneity and decreased imprecision, resulting in estimates closer to zero (1 month: 2 RCTs, N=460, MD -0.08, 95% CI -0.58 to 0.41, I^2=0%; 3 months: 2 RCTs, N=419, MD 0.14, 95% CI -0.29 to 0.45, I^2=0%) and more consistent with other timepoints (Appendix P, Figure P11). Sensitivity analysis excluding poor-quality trials showed similar results (Appendix P, Figure 12). Two trials[52,108] reported pain outcomes past 12 months (for the meta-analysis all follow-up was at 12 months for the ≥12 month category); pain improvement remained similar between VP and KP at 24 months (2 RCTs, N=320, MD -0.16, 95% CI -0.67 to 0.42, I^2=0%)[52,108] but VP was associated with small improvement in pain compared with KP at 60 months in one poor-quality trial (N=100, MD -0.60, 95% CI -1.13 to -0.07).[108]

One poor-quality trial[183] reported pain using the Short-form McGill Pain Questionnaire-2 (SF-MPQ-2) and could not be pooled with the other trials. KP was associated with greater pain improvement on the SF-MPQ-2 at 1, 6 and 12 months compared with VP (Appendix G, Table G1).

Figure 13. VP vs. KP: Pain scores (VAS or NRS, 0-10 scale) from RCTs



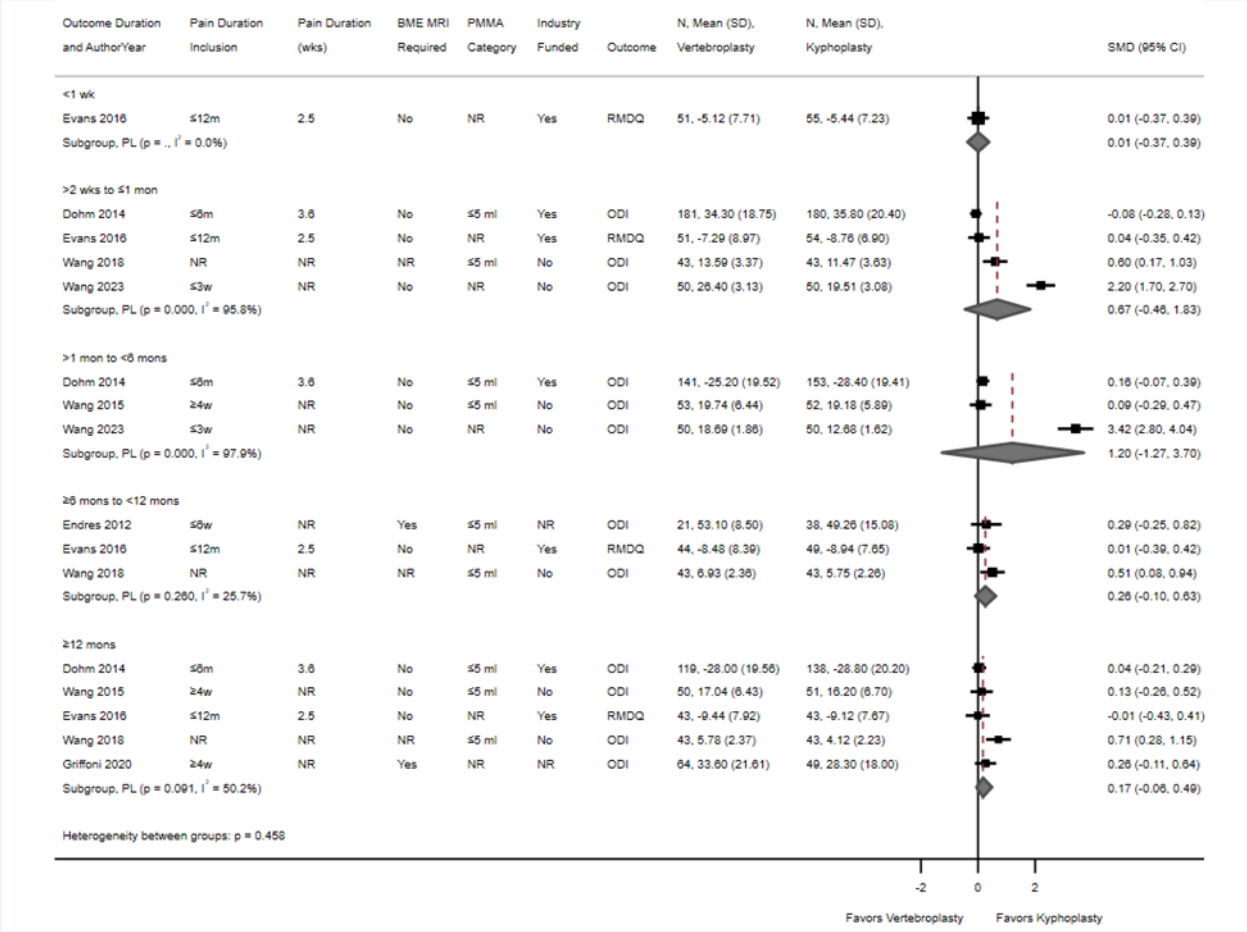
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; NR = not reported; NRS = numerical rating scale; PMMA = polymethylmethacrylate; SD = standard deviation; VAS = visual analogue scale.

Function

No trial reported the proportion of patients who were considered function responders.

Seven RCTs reported function scores using the ODI or the RDQ.[52,59,61,77,182-184] VP and KP resulted in similar improvement in pain at all time points evaluated, however estimates were imprecise (Figure 14). At 1 and 3 months, heterogeneity was substantial (>95%) and removal of a poor-quality outlier trial[184] reduced heterogeneity somewhat at 1 month (3 RCTs, N=552, SMD 0.13, 95% CI -0.28 to 0.63, I^2=73.9%) and eliminated it at 3 months (2 RCTs, N=399, SMD 0.14, 95% CI -0.11 to 0.38, I^2=73.9%) but estimates remained imprecise and not statistically significant (Appendix P, Figure P14). Sensitivity analysis excluding poor-quality trials showed similar, though attenuated, results and there was no heterogeneity; however, only one trial contributed data at each timepoint prior to 12 months (Appendix P, Figure P13)

Figure 14. VP vs. KP: Function scores from RCTs



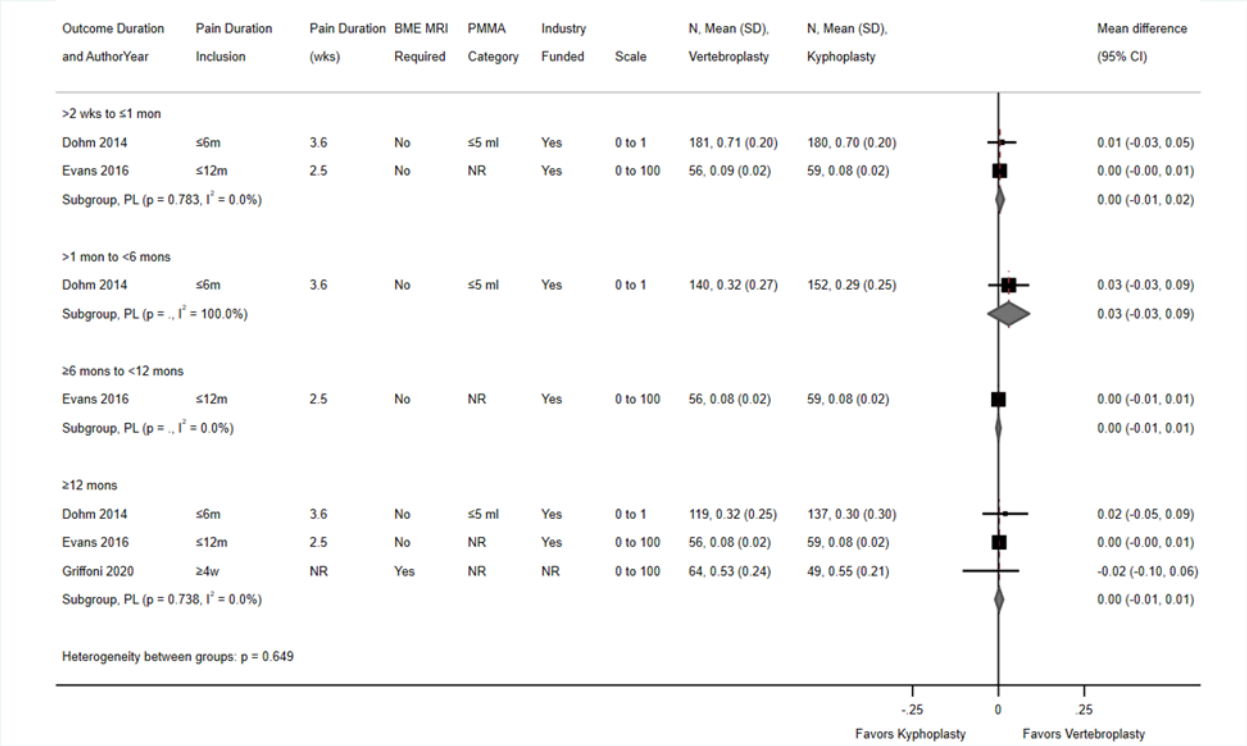
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; NR = not reported; ODI = Oswestry Disability Index; PMMA = polymethylmethacrylate; RMDQ = Roland Morris Disability Index; SD = standard deviation; SMD = standardized mean difference.

4.2.1.1.4.2 Secondary outcomes

Quality of Life

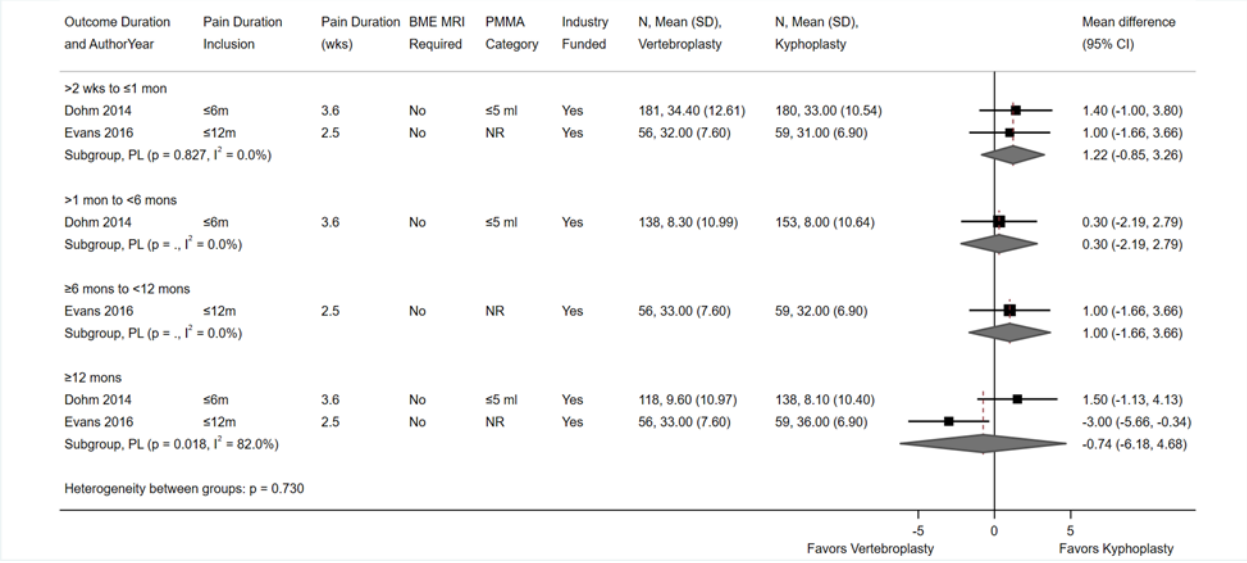
VP was associated with similar improvement in EQ-5D QoL scores (scale 0-1) across three RCTs (Figure 15)[52,61,77] and SF-36 PCS scores (scale 0-100) across two RCTs (Figure 16)[52,61] at all timepoints compared with KP. Sensitivity analyses excluding the poor-quality trial[52] showed similar results (Appendix P, Figures P15 and P16)(while the SF-36 PCS scores at 12 months favored vertebroplasty in one RCT [N=115, MD -3.00, 95% CI -5.66 to -0.34][61] the estimate was below the threshold for a small effect). There was no difference in SF-36 MCS scores (scale 0-100) at 1, 6 and 12 months in one RCT (MD range, -1.00 to 3.00); estimates were imprecise.[61]

Figure 15. VP vs. KP: EQ-5D quality of life scores from RCTs



BME = bone marrow edema; CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; MRI = magnetic resonance imaging; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation.

Figure 16. VP vs. KP: SF-36 PCS quality of life scores from RCTs



BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; NR = not reported; PCS = physical component score; PMMA = polymethylmethacrylate; SD = standard deviation; SF-36 = 36-Item Short Form Questionnaire.

Opioid use

VP was associated with a similar risk of opioid use at 6 months (N=284, 23.9% vs. 17.6%, RR 1.36, 95% CI 0.86 to 2.16) and 24 months (N=173, 25.6% vs. 17.6%, RR 1.46, 95% CI 0.82 to 2.60) compared with KP in one poor-quality trial.[52] No other trial reported opioid or analgesic use.

4.2.1.2 Kyphoplasty

4.2.1.2.1 Kyphoplasty versus Usual Care

Four RCTs (in 6 publications)[19,105,110,174,187,197] compared kyphoplasty to usual care in patients experiencing fractures due to osteoporosis. Kyphoplasty was performed via a variety of approaches and most frequently with PMMA cement, though one study[19,174,187] did not report cement type. No study reported cement volume. All trials reported providing analgesics, bed rest, bracing or orthoses, and physical therapy or related treatment to participants in usual care groups; One trial[19,174,187] also reported supplying calcium, vitamin D, antiresorptive, and anabolic agents as needed and indicated that usual care treatments were provided to both groups. Three trials took place in China[105,110,197] and one trial[19,174,187] in a number of countries (Austria, Belgium, France, Germany, Italy, UK, USA, Sweden).

The four trials included a total of 696 participants (range, 80 to 300). Mean age ranged from 66 to 74 and percentage of participants who were female ranged from 30 to 70 percent. One trial[110] did not report patient totals clearly enough to determine sex distribution. Two trials reported number of vertebral levels treated, with one trial[197] reporting an average of 1.4 levels per participant and the other[19,174,187] reporting distribution (0 levels: 7%, 1 level: 67%, 2 levels: 19%, 3 levels: 7%). Baseline pain ranged from 6.9 to 8.6 (scale 0 to 10) in three trials[19,105,110,174,187] and was not reported in one.[197] One trial[19,174,187] additionally reported a mean duration of pain at enrollment of less than three weeks.

Three trials[105,110,197] were rated poor and one trial[19,174,187] was rated fair (Appendix Table E5). Limitations included unclear or absent blinding of participants and researchers in all studies[19,105,110,174,187,197] as well as unclear randomization and considerable between-group heterogeneity in the three poor trials[105,110,197]. Additionally, two of these trials[105,110] did not clearly describe attrition or whether they incorporated intention-to-treat analyses.

No NRSIs that compared KP versus UC were identified that met inclusion criteria.

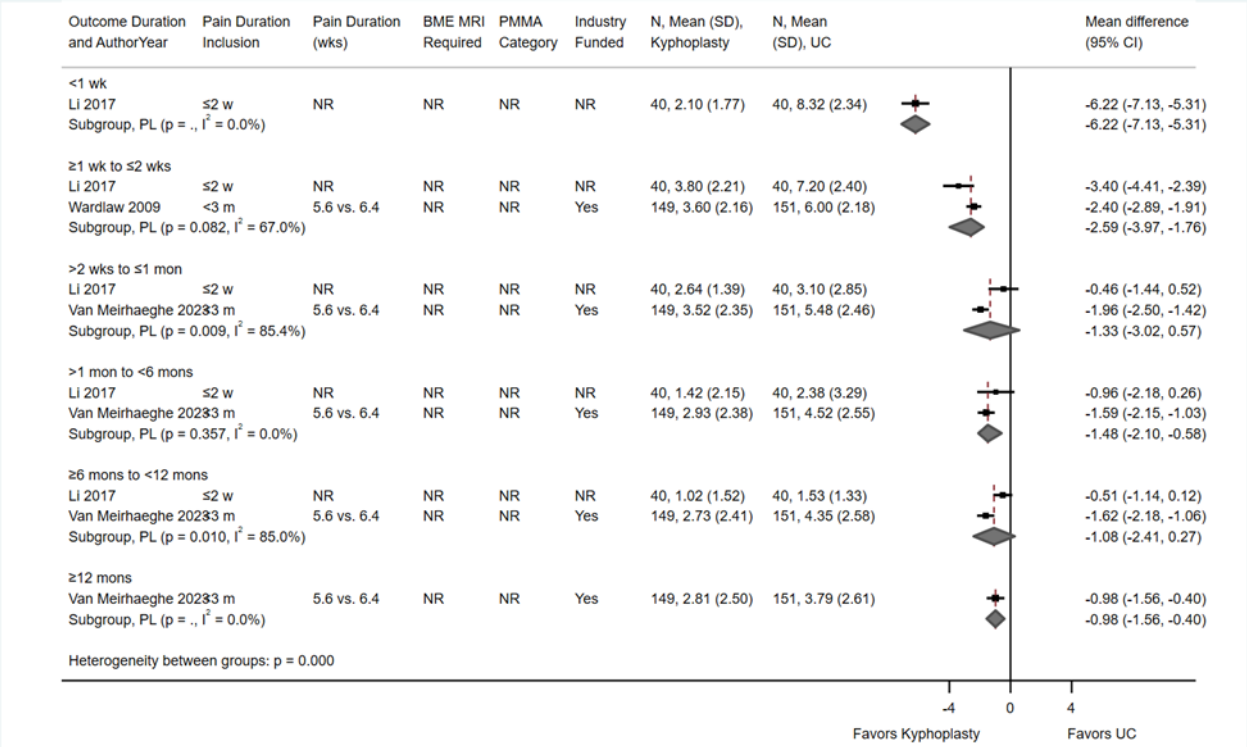
4.2.1.2.1.1 Primary Outcomes

Pain

Two RCTs reported VAS or NRS pain scores (scale, 0-10).[105,174,187] In general, KP was associated with improvement in pain versus UC with the largest improvements seen at earlier timepoints, though not all differences were statistically significant (Figure 17). KP was associated with a large improvement in pain at <1 week (3 days, 1 RCT, N=80)[105] and at 1 to 2 weeks (1 week, 2 RCTs, N=380),[105,187] a moderate improvement at >1 month to <6 months (3 months, 2 RCTs, N=380),[105,174] and a small improvement at ≥12 months (1 RCT, N=300; 12 months, see Figure 17; 24 months: MD -0.83, 95% CI -1.341 to -0.25).[174] There was no difference in pain improvement between groups in the pooled estimates at >2 weeks to 1 month (1 month) and ≥6 months to <12 months (6 months); heterogeneity was substantial (85%), and the estimates were imprecise.[105,174] In both instances, the larger, fair-quality trial[174] reported a moderate improvement in pain with KP while the smaller, poor-quality trial[105] tended to favor KP but the

difference between groups was not statistically significant. When just the fair-quality trial was considered, KP was associated with significant improvement compared with UC at all timepoints, again with the largest effects seen at earlier timepoints (Appendix G2).

Figure 17. KP versus Usual Care: Pain Scores

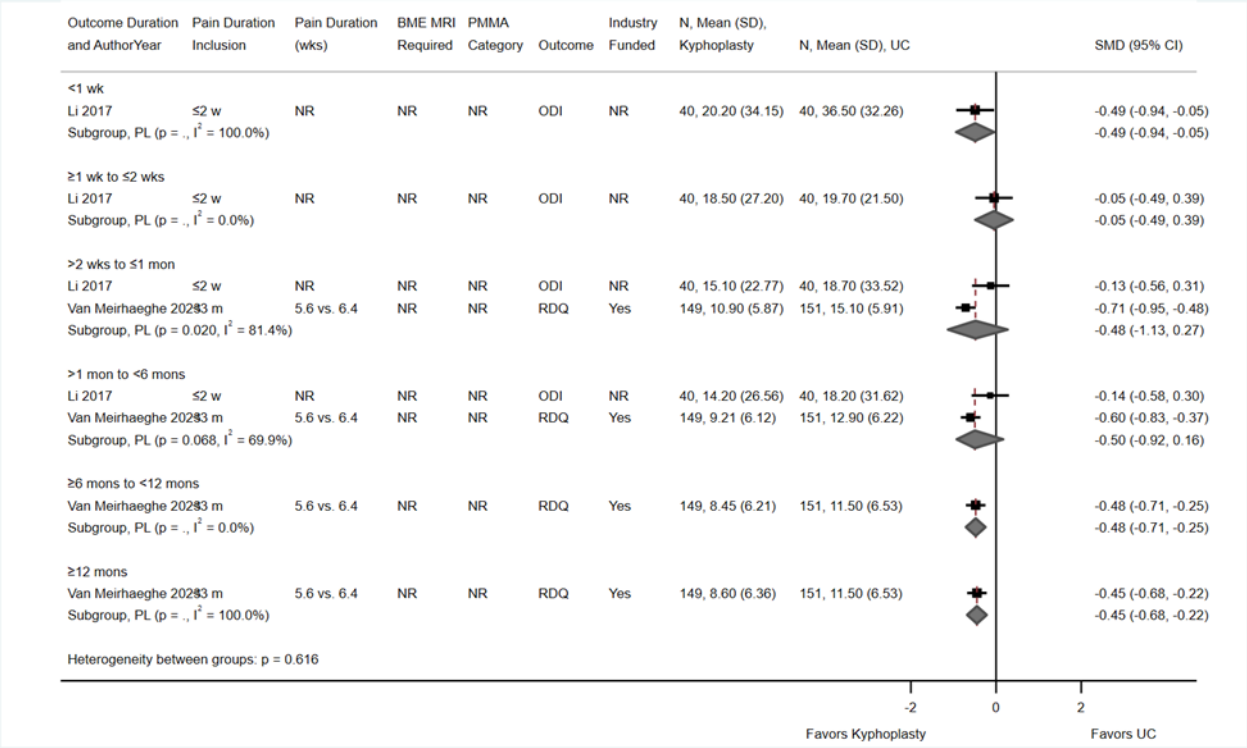


BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; NR = not reported; NRS = numerical rating scale; PMMA = polymethylmethacrylate; SD = standard deviation; VAS = visual analogue scale.

Function

Two RCTs[105,174] reported function using RDQ or ODI scores (Figure 18); results tended to favor KP at most timepoints, though not all differences were statistically significant. There was no difference in function improvement between groups in the pooled estimates at >2 weeks to 1 month (1 month) and >1 month to <6 months (3 months); heterogeneity was substantial (≥70%), and the estimates were imprecise.[105,174] In both instances, the larger, fair-quality trial (n=300)[174] reported a moderate improvement in function with KP versus UC while the smaller, poor-quality trial[105] showed similar improvement between groups. Likewise, when only the fair-quality trial was considered at later timepoints (6 months to <12 months [6 months] and ≥12 months [12 months]), KP was associated with a small improvement in function (no difference at 24 months) while the poor-quality trial showed similar improvement in function between groups at all timepoints, with the exception of <1 week (3 days) which showed a small effect favoring KP.

Figure 18. KP versus Usual Care: Function Scores



BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; NR = not reported; ODI = Oswestry Disability Index; PMMA = polymethylmethacrylate; RDDQ = Roland Morris Disability Index; SD = standard deviation; SMD = standardized mean difference.

4.2.1.2.1.2 Secondary Outcomes

Quality of Life

One trial[19,174,187] reported SF-36 PCS and EQ-5D scores up to 24 months. Both kyphoplasty and usual care showed slight improvement in SF-36 PCS scores at all time points except for a moderate improvement at 6 months in the kyphoplasty group. Authors report a small association between kyphoplasty and improvement in SF-36 PCS at 1 month, and small differences at 3 and 6 months, but no difference beyond that. Additionally, both groups reported large improvements in EQ-5D scores at all time points, with a significant difference between group scores favoring kyphoplasty at all time points. See Table 13 and Appendix G2.

Opioid Use

At baseline, just over half of the patients (57%) were using a combination of opioid and nonopioids and 14% were using strong opioids in one RCT.[187] Analgesic use decreased over time and was similar between groups, except for any opioid use at 6 months[19] and combined opioid and nonopioid use at 1 month, which were less common (small effect) in the KP versus the UC group (Table 13).

Table 13. KP vs. UC: Summary of quality of life outcomes and opioid use from the FREE trial.

Outcome Author, year	Time	KP vs. UC MD (95% CI) or % (n/N) and RR (95% CI)*
EQ5D (0-1, higher = better) Van Meirhaeghe, 2013	>2 weeks to ≤1 month (1 month)	MD 0.17 (95% CI 0.09 to 0.25)
	>1 month to <6 months (3 months)	MD 0.10 (95% CI 0.02 to 0.18)
	≥6 months to <12 months (6 months)	MD 0.13 (95% CI 0.05 to 0.21)
	≥12 months (12 months)	MD 0.10 (95% CI 0.02 to 0.18)
	≥12 months (24 months)	MD 0.08 (95% CI 0.00 to 0.16)
SF-36 PCS (0-100, higher= better) Van Meirhaeghe, 2013	>2 weeks to ≤1 month (1 month)	MD 5.90 (95% CI 3.64 to 8.16)
	>1 month to <6 months (3 months)	MD 4.50 (95% CI 2.16 to 6.84)
	≥6 months to <12 months (6 months)	MD 3.80 (95% CI 1.50 to 6.10)
	≥12 months (12 months)	MD 2.10 (95% CI –0.24 to 4.44)
	≥12 months (24 months)	MD 2.00 (95% CI –0.34 to 4.34)
Any Opioids Boonen 2011	≥6 months to <12 months (6 months)	29.8% (37/124) vs. 42.9% (48/112) RR 0.69 (95% CI 0.49 to 0.98)
	≥12 months (12 months)	28.0% (33/118) vs. 33.7% (34/101) 0.83 (95% CI 0.56 to 1.24)
	≥12 months (24 months)	8.8% (10/114) vs. 9.5% (10/105) 0.92 (95% CI 0.40 to 2.12)
Strong Opioid Use Wardlaw, 2009	>2 weeks to ≤1 month (1 month)	5% (6/114) vs. 8% (9/115) RR 0.67 (95% CI 0.25 to 1.83)
	≥12 months (12 months)	4% (5/117) vs. 5% (5/101) RR 0.86 (95% CI 0.26 to 2.90)
Combined Opioid and Non-opioid Use Wardlaw, 2009	>2 weeks to ≤1 month (1 month)	41% (47/114) vs. 57% (65/115) RR 0.73 (95% CI 0.56 to 0.96)
	≥12 months (12 months)	24% (28/117) vs. 29% (29/101) RR 0.83 (95% CI 0.53 to 1.30)

CI: confidence interval; EQ5D: EuroQoL 5 dimensions questionnaire; f/u: follow-up; MD: mean difference; mos: months; RCT: randomized controlled trial; SF-36 MCS: Short-Form 26 Mental Component Score; UC: usual care; VP: vertebroplasty; wks: weeks; SF-36 PCS: Short-Form 26 Physical Component Score;

* Calculated.

† Authors report mean change from baseline. Presented are MDs for follow-up scores, calculated by converting 95% CIs to SDs and using means and SDs to calculate MD from follow up scores.

4.2.2 KQ2 Harms and Safety

Adverse events were variably and sparsely reported across the six RCTs (in 11 publications) of VP versus sham,[24,29,41,42,50,66,67,80,89,99,157] nine RCTs (reported in 10 publications) of VP vs. UC,[18,33,64,93,101,144,145,180,195,197] four RCTs (in 6 publications) of KP versus UC,[19,105,110,174,187,197] and nine RCTs (in 10 publications) of VP versus KP.[52,59,61,77,108,109,179,182-184,197]. Descriptions of these studies are provided in KQ 1 in patients with osteoporotic VCF.

Fourteen retrospective administrative database studies provided safety data related to mortality, SAEs and reoperation comparing vertebral augmentation (VP or KP) with nonoperative care or comparing VP and KP.[32,35,55,56,87,92,100,103,106,118,129,136,192,198] Five database studies used U.S. Medicare Claims (CMS) data.[32,55,56,118,129] Three of these were industry-funded, had substantial potential overlap in data based on years sampled (2005 to 2014) and were performed by the same primary author group from a corporate engineering and consulting firm and used similar methodology.[55,56,118,129] One study did not receive outside funding [32] and another reported government and professional society funding.[118] One sampled Medicare data from 2006 [32] and the other study used a 20% random sample of Medicare data from 2002 to 2006, representing the least potential for overlap.[118] Given the overlap in sampling data from two studies [32,55], they will not be reported. Two other administrative data studies sampled the American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) database to compare 30 day adverse event for VP versus KP.[35,92] Overlap in sampling periods is noted: one study sampled from 2012 to 2014[35] the other sampled from 2011 to 2013.[92] Two additional US-based studies were identified. Two used National Inpatient Sample (NIS) data with funding not reported in one[198] and no funding received in the other[136] and the other used administrative data from Emory University Hospital (partial NIH funding).[103] Four non-US studies were also identified. Three studies sampled the Taiwan National Health Insurance Research Database (NHIRD), with no funding received in one,[192] funding not reported in one,[106] and government funding reported for the other[87]; these studies also had overlap in populations. The third study from Germany sampled a major private health insurance database (industry funded).[100] Some studies did not separate outcomes by type of vertebral augmentation. Studies varied in approaches to adjustment for confounding and details of analyses, particularly related to propensity matching, were not consistently provided by studies. Although most studies report a focus on osteoporotic fractures, it is possible that databases included a combined osteoporotic and malignant vertebral compression fracture. Residual confounding and selection bias may still be present in all these studies. Causality cannot be inferred from these studies.

Five retrospective comparative NRSIs[8,34,63,163,189] that met inclusion criteria were also included that reported on safety: VP vs. KP (2 studies),[34,163] VP vs. nerve block[8], KP vs. UC (1 study)[63] and KP versus screw fixation.[189] Sample sizes ranged from 164 to 497.

Data from 18 case series[9,15,17,47,51,62,96,106,124,137,156,169,177,185,201,202] that met inclusion criteria and evaluated safety outcomes following VP or KP for osteoporotic (primarily) vertebral fractures are also summarized. Six studies evaluated VP only,[9,51,62,96,169,177] nine studies evaluated KP only[15,17,47,106,124,137,156,202] and five studies evaluated VP or KP together as vertebral augmentation,[35,92,163,185,201] Case series data are presented in their own sections at the end of the VP and KP sections. Sample sizes ranged from 1,932 to 2,433 in the database studies and from 263 to 1,512 in the case series. Follow-up periods ranged broadly from peri-operative to 43 months. All case-series were considered to be at high risk of bias.

4.2.2.1 Vertebroplasty

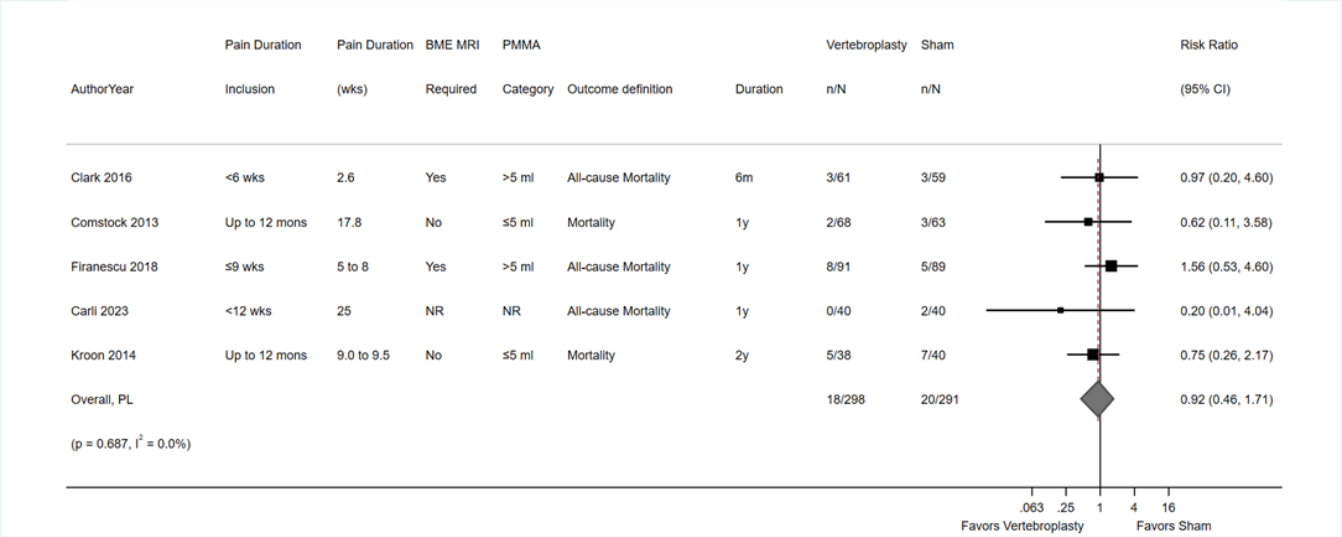
4.2.2.1.1 Vertebroplasty versus Sham

All six RCTs of VP versus sham reported harms.[24,29,41,67,80,89]

Mortality

Five RCTs of VP versus sham reported on mortality (N= 589)[29,41,42,67,99] Mortality risk at last follow-up was similar between patients receiving VP and those receiving sham (RR 0.92, 95%CI 0.46 to 1.17, I^2= 0%). Figure 19 All authors report that deaths were not procedure or device related. Studies may have been underpowered, or follow-up time may not have been long enough to a detect difference. Mortality was also similar between VP and sham in analyses done by individual time frames. (Appendix Figure P3) In one additional RCT (not in the figure) one participant developed respiratory insufficiency the day after the procedure related to underlying chronic obstructive pulmonary disease.[67] Sample sizes in some trials may have been too small to detect rare events and estimates are imprecise.

Figure 19. Vertebroplasty versus sham procedures: Cumulative mortality by last follow-up



Comstock is follow-up to Kallmes; Kroon is follow-up to Buchbinder 2009; Duration indicates time at last follow-up.
BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; SD = standard deviation; wks = weeks

Serious Adverse Events

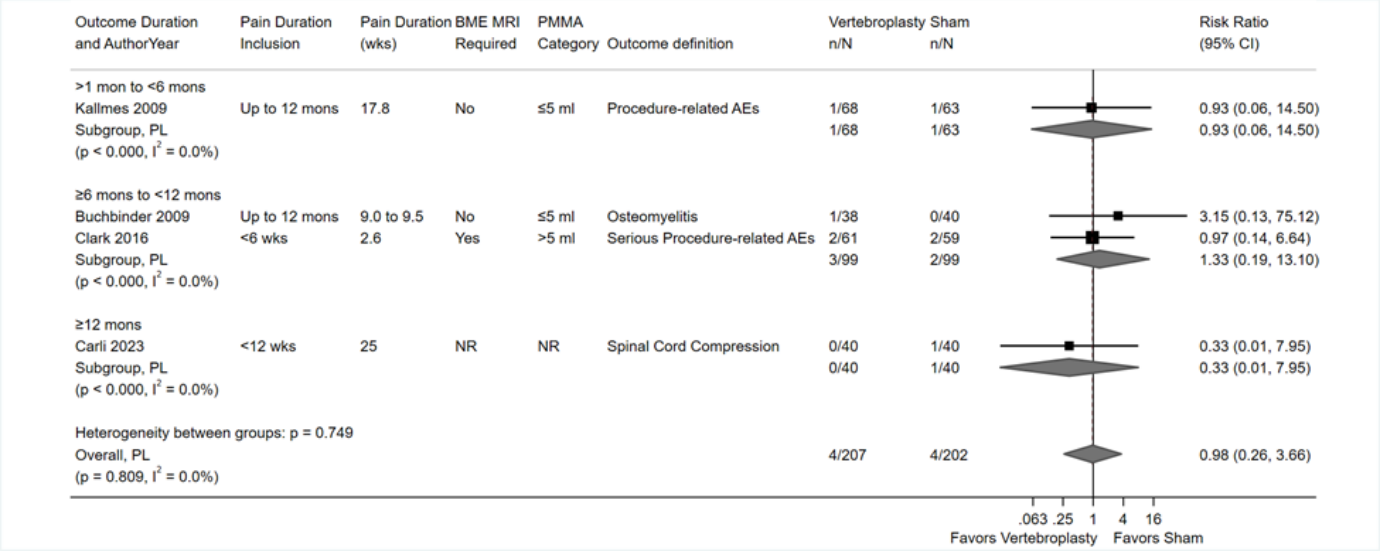
Four RCTs reported occurrence of any serious AE which may be procedure related. Risk of any SAE at any time was similar for VP and sham interventions across time frames (4 RCTs, N=409, RR 0.96, 95% CI 0.26 to 3.66, I^2=0%). [24,29,89] Studies may have been underpowered to detect uncommon or rare SAEs. Figure 20 Reported SAEs were as follows:

Table 14. Serious Adverse Events

Study	VP (n/N), SAE	Sham (n/N) SAE
Kallmes 2009	1/68, Thecal sac injury requiring hospitalization	1/63: Tachycardia, rigors requiring hospitalization
Clark 2016	2/61; respiratory arrest after sedation (n=1); supracondylar humerus fracture from transfer to radiology table (N=1)	2/59: 2 spinal cord compression due to fracture collapse and retropulsion several weeks after enrollment; (one became paraplegic, the other had decompressive surgery to resolve the neurological deficit
Buchbinder 2009	1/38, Osteomyelitis (requiring surgical drainage and antibiotic)	0/40
Carli 2023	0/40	1/40 Spinal Cord Compression

KP = kyphoplasty; SAE = serious adverse events; VA = vertebral augmentation VP = vertebroplasty.

Figure 20. Vertebroplasty versus sham procedures: Cumulative risk serious adverse events by last follow-up



BME = bone marrow edema was an inclusion criterion; CI = confidence interval, ml = milliliter, mos = months, NR= not reported, PMMA = polymethylmethacrylate; SD = standard deviation, wks = weeks

New Vertebral Fracture

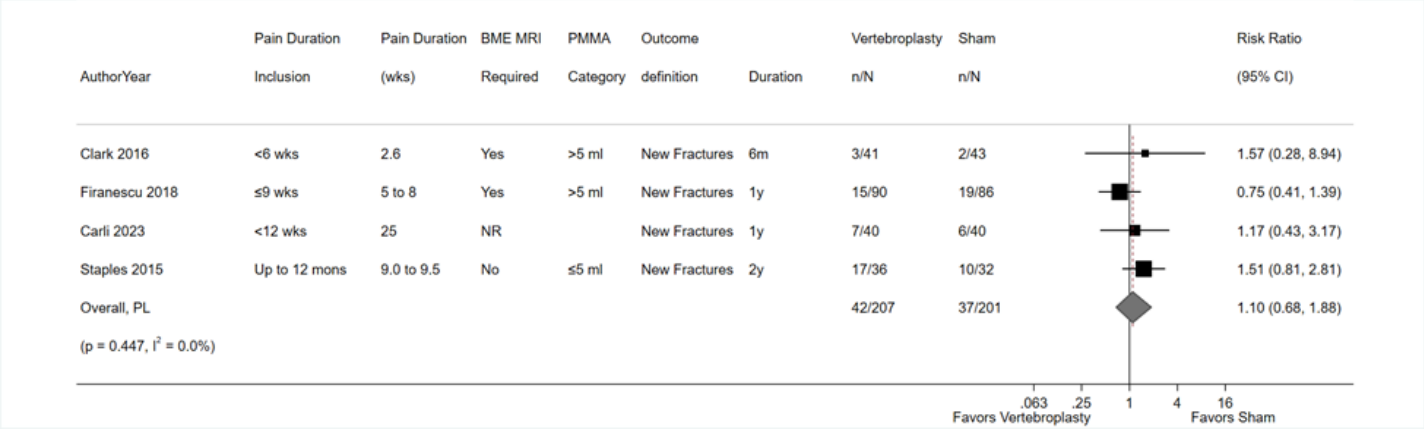
Four RCTs reported on risk of new vertebral fracture (N=408) [29,41,67,157] The risk of any new fracture by time of last follow-up was similar between patients receiving VP and those receiving sham (Figure 21). Fracture risk was also similar between VP and sham in analyses done by individual time frames. Sample sizes in some trials may have been inadequate to determine new fractures. One trial reported that six participants in each group underwent re-intervention during follow-up for one or more new symptomatic fractures and were treated based on original group assignment.[67] Two trials provided information on new fracture locations.[67,157] Risks for new fracture by location were similar between VP and sham groups and there were no difference between groups in the proportions of new fractures that occurred in treated levels or untreated levels, Table 15.

Table 15. VP versus sham: New vertebral fracture

Fractures	Study	VP % (n/N) patients	Sham % (n/N) patients	Effect Size (95% CI)
Any level	Firanescu, 2018	16.6% (15/90)	22.1% (19/86)	RR 0.75 (95% CI 0.41 to 1.39)
New Fracture	**Study**	**VP % (n/N) fractures**	**Sham % (n/N) fractures**	**Effect Size (95% CI)**
Any, Symptomatic with Bone Edema	Firanescu, 2018	40% (6/15)	31% (6/19)	RR 0.84 (0.38 to 1.84)
Any level -total	Staples, 2015	N = 17	N = 10	HR 1.80 (0.83 to 3.94)
Treated level		6% (1/17)	0	RR/HR not calculable
Untreated level		94% (16/17)	100% 10/10	HR 1.69 (0.77 to 3.74)
Adjacent level	Firanescu, 2018	52% (16/31)	46% (13/28)	RR 1.03 (95% CI 0.62 to 1.71)
	Staples, 2015	35.3% (6/17)	30% (3/10)	HR 2.30 (0.57 to 9.29)
Non-adjacent	Staples, 2015	58.8% (10/17)	70% (7/100)	HR 1.45 (0.55 to 3.81)
Distant	Firanescu, 2018	45% (14/31)	46% (13/28)	RR 0.97 (95% CI 0.56 to 1.70)
Between (sandwich)	Firanescu, 2018	3% (1/31)	4% (1/28)	RR 0.90 (95% CI 0.06 to 13.77)
New or Progressed fracture	**Study**	**VP % (n/N) fractures**	**Sham % (n/N) fractures**	**Effect Size (95% CI)**
Any level -total	Staples, 2015	N = 40	N = 33	HR 1.29 (0.80to 2.08)
Treated level		35% (14/40)	36.4% (12/33)	HR 1.05 (0.47 to 2.34)
Untreated level		65% (26/40)	63.6% (21/33)	HR 1.44 (0.80 to 2.61)
Adjacent level		25% (10/40)	15.2% (5/33)	HR 2.18 (0.74 to 6.42)
Non-adjacent		40% (16/40)	48.5% (16/33)	HR 1.18 (0.58 to 3.94)

CI = confidence interval; HR = hazard ratio; RR = risk ratio; VP = vertebroplasty.

Figure 21. Vertebroplasty versus sham procedures: Cumulative risk of vertebral fracture by latest follow-up



Staples 2015 Is follow-up to Buchbinder 2009; Duration = duration of follow up
BME = bone marrow edema; CI = confidence interval; mons = months; NR = not reported; PL = profile likelihood; PMMA = polymethylmethacrylate; wks = weeks.

Cement Leakage

Cement leakage ranged from 0% to 91% of treated levels across four RCTs.[24,29,67,80](Table 16) One study reported that all leaks were asymptomatic and provided information on location of the leaks.[67] None of the other RCTs provided information on symptomatology or location. Two RCTs did not report on cement leakage [41,89]

Table 16. Cement leakage following vertebroplasty reported in RCTs of VP versus Sham

Author, year Quality	Mean PMMA volume (ml)	Cement leakage % (n/N)	Comments
Carli, 2023 Fair	1.4 ml	70% of treated levels (n=72 treated)	Detected on CT; Symptomatology NR
Firanescu, 2018 Good	5.1 ml	91.3% (105/115) of treated levels; Leakage type/location, % treated levels • Type 1 = disc above treated level (20%) • Type 2 = disc under treated level (15%) • Type 3 = perivertebral tissue (10%) • Type 4 = perivertebral veins (39%) • Type 5 = pulmonary (7%) • Type 6 = spinal canal (8%)	Any perceptible on post-procedural CT; recorded even very small cement traces outside the target vertebra. All leaks reported as asymptomatic
Hansen, 2019 Fair	2 to 4 ml	None observed (number of treated levels NR)	No further detail; Symptomatology NR
Buchbinder, 2009 and Staples 2015 Good	3 ml	36.8% (14/38) of patients 40.0% (18/45) of treated levels	Authors report as minimal; based on postprocedural images; Symptomatology NR

NR = not reported; PMMA = polymethylmethacrylate; RCT = randomized controlled trial.

4.2.2.1.2 Vertebroplasty versus Usual Care

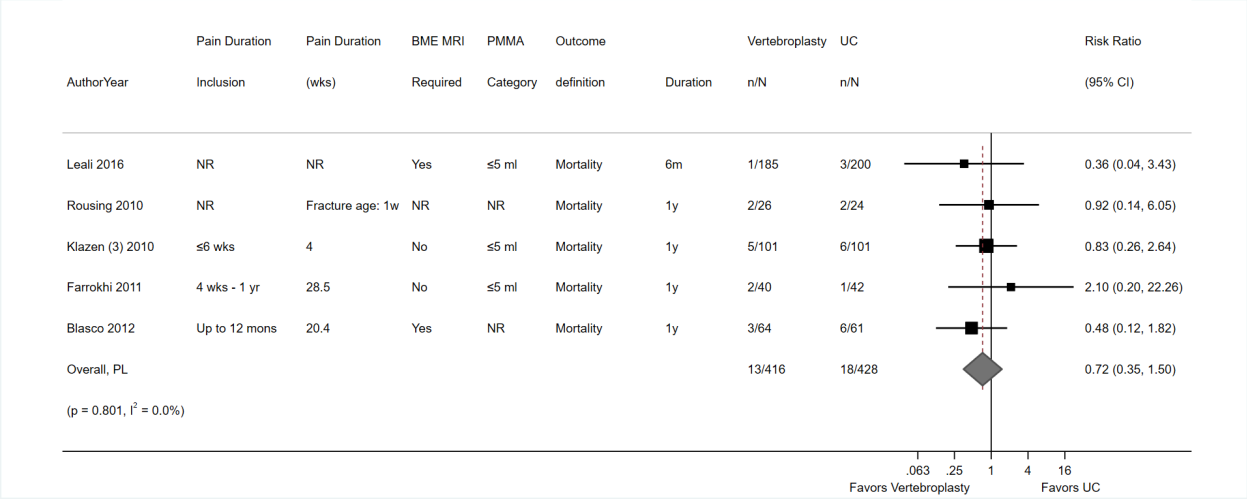
All nine trials (reported in 12 publications) that compared VP versus UC for the treatment of osteoporotic compression fractures included for efficacy reported safety outcomes (Table 8 and Appendix Table F1).[18,33,64,93-95,101,144,145,180,195,197]

In addition, four retrospective administrative database studies (using overlapping populations from the U.S. Medicare and Taiwan NHRID databases) evaluated adverse events for VP compared with UC.[56,87,106,129] Further details of these studies are provided above at the beginning of the Safety section and in Appendix Table H2.

Mortality

VP was associated with a similar risk of mortality at latest follow-up (primarily 12 months) compared with UC across five RCTs (N=844),[18,64,93,101,145] (Figure 22). Results were similar when mortality was analyzed by timepoint (Appendix P, Figure 12) and when poor quality trials[64,101] were excluded (Appendix P, Figure 13). In general, deaths were contributed to causes unrelated to the procedures or the trial. Sample sizes in some trials may have been too small to detect rare events and estimates are imprecise.

Figure 22. VP vs. UC: Cumulative risk of mortality at latest follow-up



BME = bone marrow edema; CI = confidence interval; mons = months; NR = not reported; PL = profile likelihood; UC = usual care.

In two studies[56,129] using Medicare data with overlapping sample frames reported that VP was associated with slightly lower mortality risk compared with UC. The estimate for the based on the larger of these studies was: HR 0.87 95% CI 0.87, 0.88.[129] Similarly, The two studies with overlapping samples from the Taiwan National Health Insurance Research Database reported slightly lower mortality with VP versus nonoperative care[87,106] Table 17

Table 17. Summary of mortality findings across administrative data studies

Database	Study Database search dates	N	Author Findings
Mortality Within 30 Days			
Medicare	McCullough 2013 (2002-2006)20% random sample	Propensity-score matched cohort VP or KP: 9017 Non-operative: 9017	Propensity matched 0.3% (31/9017) vs 0.6% 51/9017), Adj OR 0.61 (95% CI 0.39-0.95)
ACS-NSQIP (Possible overlap in data)	Choo, 2018 (2012-2014)	VP: 242 (10%) KP: 2,191 (90%)	30-day mortality: 2% (n=49): analyses indicate that augmentation was not an independent risk factor for mortality
	Kim, 2022 (2011-2011)	N = 1932 VP: 197 (10%) KP: 1769 (90%)	KP vs. VP: Adj. OR 0.94 (0.27-3.24); Procedure type was not a risk factor for mortality
Nationwide Inpatient Sample	Zampini, 2010 (2005)	N = 5766 KP: 15% Nonoperative: 84.7%	KP vs. Nonoperative: 0.3% vs. 1.6% Adj OR: 0.52, p= 0.003 (CI NR)
Mortality at Longer Follow -up (>30 days)			
Medicare	Ong, 2018 (2005-2014) [Data overlaps with Edidin, 2018]	VP: 117,232 KP: 261,756 Non-operated: 1,698,956) (authors do not clearly provide n's or data for propensity matched cohort)	Mortality risk overall at 10 years: 85.1% (95% CI, 84.7–85.5%) Propensity-adjusted results comparing groups: 19% (95% CI, 19–19%; p < 0.001) and 7% (95% CI, 7–8%; p < 0.001) lower 10-year mortality risk for KP and VP respectively versus the non-operated group. KP cohort: 13% (95% CI, 12–13%; p < 0.001) lower 10-year mortality risk than the VP cohort; Authors state that results were statistically significant at other times (data not provided) HRs (95% CI) reported in Hinde (any time) Any VA vs Nonop HR: 0.83 (0.82, 0.83) VP vs. Nonop: HR 0.926 (0.926, 0.917) KP vs. Nonop: HR 0.81 (0.813, 0.806) KP vs. VP: HR 0.87 (0.87, 0.88)
	Edidin 2015 (2005-2009) [Data overlaps with Ong, 2018]	Propensity-score matched (osteoporotic fractures) VP: 37,252 KP: 36,286 Non-Operated matches: VP: Nonop n =107,930 KP: Nonop n = 163791	Adjusted HR at 4 years: Nonop vs. VP: HR 1.30 (1.28–1.33) Nonop vs. KP: HR 1.62 (1.60–1.64) KP vs. VP: HR 0.83 (0.81–0.85) Higher risk of mortality reported in non-operated group versus VP or KP; KP associated with lower mortality vs. VP
	McCullough 2013 (2002-2006, 20% random sample) [Some overlap with Edidin and Ong]	Propensity-score matched cohort VP or KP: 9017 Non-operative: 9017	Mortality at 1 year: Adjusted HR 5.2% 469/9017 vs 5.6% 505/9017), HR 0.92 (95% CI, 0.81-1.04); not statistically significant
Emory University Hospital	Levy, 2012 (1998 to 2007)	N=250 VA (VP or KP): Non-operative (medical): No treatment:	Multivariate analyses (no treatment group reference group) VA: Adj HR: 0.81 (0.42, 1.59) p 0.55 Non-op: Adj HR 0.83 (0.36, 1.89)

Private health insurance (Germany)	Lange, 2014 (2006-2010)	**N =298** matched patients Characteristics across full cohort of 3607: VA (KP or VP): 598 non-operative:3009	Kaplan-Meier plot shows similar survival between VA and nonoperative management up to 36 months since diagnosis (data NR). Any VA vs. non-op by 60 months Survival rates: VA vs. Non-op: 69.9% vs. 53.8% VA vs. non-operated: HR (adj) 0.58 (0.48, 0.70)
Taiwan National Health Insurance Research Database (NHIRD)	Lin 2017 (2002 to 2013) Overlap with Huang	Matched cohort: Early VP (≤3 months: 1773 Non-VP*:5324	Mortality incidence at 1 year: 0.46 (0.38–0.56) vs.0.63 (0.57–0.70) per 100 person-months Non-VP vs. VP: HR1.39 (1.09–1.78)
	Huang 2020 (2003 to 2013) Overlap with Lin	VP:1389 Open surgery: 1219 or Conservative: 6017	Follow-up times: Conservative vs. VP vs. Surgery (years) 4.8 vs. 3.2 vs. 4.7 VP vs. conservative: 19.2% (267/1389) vs. 26.2% (1576/6017), Adj HR 0.87 (0.77–0.99) Open surgery vs. conservative care Adj HR 0.80 (0.70–0.93)

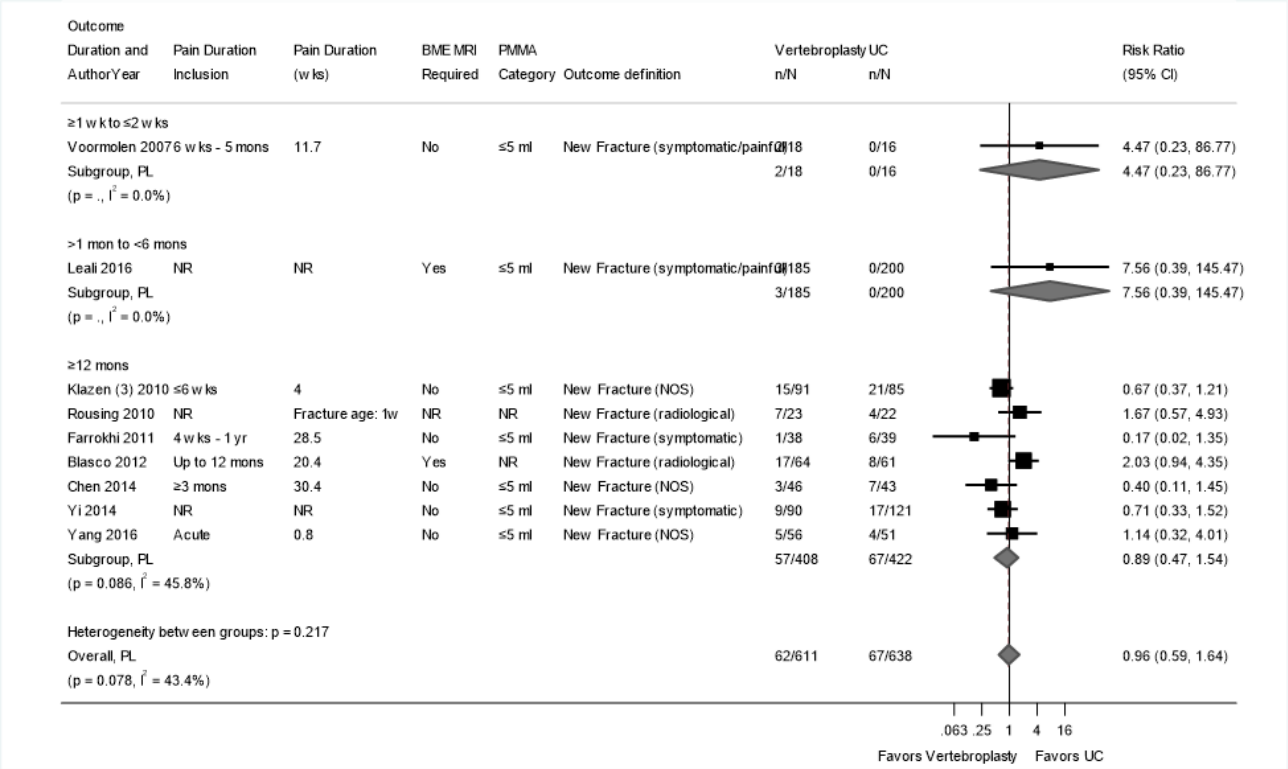
Adj = adjusted; CI = confidence interval; HR = hazard ratio; KP = kyphoplasty; NR= not reported; OR = odds ratio; VA = vertebral augmentation VP = vertebroplasty.
* Defined as those that did not receive VP within 3 months of VCF. Assumed to be non-VP patients, but it is not clearly defined.

A recent AHRQ-funded comparative effectiveness review[38] that included the majority of RCTs of vertebroplasty included in this updated HTA found moderate strength of evidence of no increased mortality risk for vertebroplasty versus sham or usual care (as a combined comparison group) across 7 trials, (N= 1159), RR 0.88, 95% CI, 0.50 to 1.53, I^2=0%). Mortality risk was also similar between groups 6 to 12 months (3 trials, N=598, RR 0.76, 95% CI, 0.23 to 2.65, I^2=0%) and 12 months and longer (5 trials, N=639, RR 0.98, 95% CI, 0.51 to 1.87, I^2=0%). Confidence in the results of this review is high based on modified AMSTAR-2 criteria.[48,49]

New Vertebral Fracture

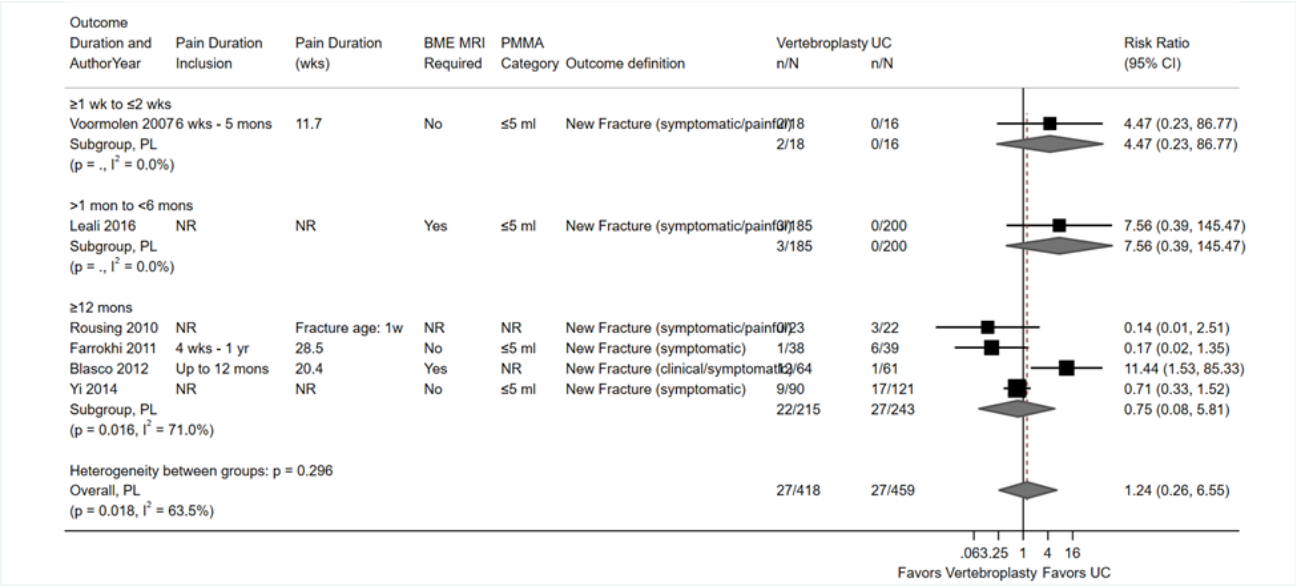
VP and UC were associated with a similar risk of ***any new vertebral fractures*** (9 RCTs, N=1,249, 10.1% vs. 10.5%)[18,33,64,94,101,145,180,195,197] and any new ***symptomatic/clinical vertebral fractures*** (6 RCTs, N=877; 6.5% vs. 5.9%),[18,64,101,145,180,197] at latest follow-up though some imprecision was present (Figures 23 and 24). Results were similar for both outcomes when the analyses excluded poor-quality trials[64,101,195,197] and were restricted to trials with ≥12 months followup,[18,33,64,94,145,195,197] though in some instances the estimates were more imprecise. Removal of one outlier trial[18] from the analysis of symptomatic fractures resulted in pooled estimates that tended to favor vertebroplasty at latest follow-up and when restricted to trials with ≥12 months follow up and reduced imprecision and heterogeneity, however the differences remained not statistically significant. Estimates at the earlier timepoints were confined to one trial and very imprecise (Figure E and F). See Appendix P for additional analyses.

Figure 23. VP vs. UC: Cumulative risk of new vertebral fractures at latest follow-up



BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; NOS = not otherwise stated; NR = not reported; PL = profile likelihood; PMMA = polymethylmethacrylate; UC = usual care.

Figure 24. VP vs. UC: Cumulative risk of new symptomatic vertebral fractures at latest follow-up



BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; NOS = not otherwise stated; NR = not reported; PL = profile likelihood; PMMA = polymethylmethacrylate; UC = usual care.

Five RCTs[64,94,101,145,180] reported the incidence of new *adjacent vertebral fracture**s***, with no statistical difference between VP and UC (Table 18). New adjacent vertebral fractures tended to occur more often following VP in three of the trials.[101,145,180] One poor quality trial found that VP was associated with a lower risk of adjacent level fracture but the difference was not statistically significant.[64] The fifth trial only reported adjacent fractures out of the total number of fractures that occurred, and patients had multiple fractures; the incidence was similar between VP and UC according to the authors.

Table 18. VP vs. UC: Summary of new adjacent level fractures from RCTs

Outcome	Author, year Quality	F/U	VP % (n/N)	UC % (n/N)
Symptomatic, new adjacent level fracture	Voormolen 2007 Fair	2 weeks	11.1% (2/18)	0% (0/16)
	Leali 2016 Poor	6 weeks	1.6% (3/185)	0% (0/200)
	Farrokhi 2011 Poor	24 months	2.6% (1/38)	15.4% (6/39)
Radiologic, new adjacent level fracture	Rousing, 2010 Fair	3 months	4.2% (1/24)	0% (0/23)
		12 months	8.7% (2/23)	0% (0/22)
	Klazen 2010 Fair	12 months	7/18 fractures* (n=91)	11/30 fractures* (n=85)

F/U = follow-up; VP = vertebroplasty; UC = usual care.

*Patients had multiple fractures, not reported out of patients. In total, there were 18 fractures in 15 patients in the VP group (n=91) and 30 fractures in 21 patients in the UC group (n=85). Adjacent level fractures incidence was similar between VP and UC according to authors.

One large database study from Taiwan[87] found no difference between VP and UC in the risk of any new fracture (N=7,406; <0.3% vs. <0.1%) over a mean follow-up of 4.5 years (Table 21); the UC group was followed a mean of 1 year longer than the VP group.

Serious adverse events (SAEs)

SAEs were not well reported by RCTs (Table 19). Sample sizes were likely too small to detect rare events.

Any SAE

Two RCTs reported that no SAEs/AEs occurred during follow-up.[144,145,197]

DVT

The risk of deep vein thrombosis (DVT) did not differ between VP and UC in one poor-quality RCT (3.6% vs. 7.8%, respectively), however, all patients in the UC care group required treatment to resolve the DVT; authors did not indicate that treatment was required in the VP patients.[195]

Cement related complications and cement emboli

Three trials reported AEs related to cement; only one was symptomatic. In one case (2.5%), cement leakage into the epidural space caused lower extremity weakness that required immediate decompression through a bilateral laminectomy and evacuation of the bone cement; the patient could walk unassisted with no radicular pain after 2 months.[64] This same trial mentioned that no cases of cement embolism occurred. Two trials reported cement migration (4.3% [2/46])[33] or deposition (1.0%

[1/101])[93] involving the pulmonary system, though in all three cases the patients remained asymptomatic through follow-up and no complications were noted.

An observational follow-up study[93] to one of the above mentioned RCTs[93] evaluated the longer term risk of pulmonary cement embolism (PCE) in a subset of VP patients. This study included only half of the patients (53.5%, 54/101) originally randomized to PV. After a mean follow-up of 22 months, PCE was detected in 26% of patients (14/54); all were asymptomatic and distributed in the periphery of the lungs (none in the heart and central pulmonary vessels). Results from this subgroup analysis should be interpreted with caution.

Any AEs

One poor-quality trial found VP associated with decreased risk of any adverse event versus UC (N=107, 16.1% vs. 35.3%, RR 0.46, 95% CI 0.23 to 0.92).[195]

Table 19. VP vs. UC: Summary of serious adverse events and reoperation from RCTs

Outcome	Author, year Quality	F/U	VP % (n/N)	UC % (n/N)	RR (95% CI)
SAEs					
SAEs, not specified	Yi, 2014 Poor	Mean 49.4 mos.	0% (0/90)	0% (0/121)	N/A
	Rousing, 2009/2010 Fair	12 mos.	0% (0/26)	0% (0/24)	N/A
DVT/thrombophlebitis*	Yang, 2016 Poor	12 mos.	3.6% (2/56)	7.8% (4/51)	0.46 (0.09 to 2.38)
Epidural cement leakage causing LE pain and weakness requiring surgical intervention†	Farrokhi, 2011 Poor	1 wk.	2.5% (1/40)	N/A	N/A
Reoperation					
Epidural cement leakage causing lower extremity pain and weakness requiring surgical intervention†	Farrokhi, 2011 Poor	1 wk.	2.5% (1/40)	N/A	N/A
Painful new fracture (adjacent) requiring reoperation after failed conservative care	Voormolen, 2007 Fair	2 wks.	11.1% (2/18)	NR	N/A
Symptomatic new fracture requiring VP (repeat or new)	Yi, 2014 Poor	Mean 49.4 mos.	10.0% (9/90)	9.1% (11/121)	1.10 (0.48 to 2.54)

CI = confidence interval; DVT = deep vein thrombosis; F/U = follow-up; LE = lower extremity; mos. = months; N/A = not applicable; NR = not reported; RR = risk ratio; SAEs = serious adverse events; VP = vertebroplasty; wks. = weeks; UC = usual care.
*The 4 UC patients required treatment to resolve DVTs; authors do not say anything about treatment being required in the VP patients.
†Included under both SAEs and Reoperation. Reoperation included immediate decompression through a bilateral laminectomy and evacuation of bone cement; the patient could walk unassisted with no radicular pain after 2 months.

Two large administrative database studies that used CMS data and had overlapping populations reported the incidence of SAEs using propensity score adjusted analyses. In the larger dataset,[129] VP was associated with lower adjusted relative risks of cardiac complications (e.g., MI; 7% to 20% lower), pulmonary complications (1% lower) and infection (1% to 6% lower) but a higher adjusted risk of thromboembolic events (PE and DVT, 3% to 7% higher) compared with UC across timepoints; not all

differences may be clinically meaningful. The second, smaller dataset[56] that provided subgroups analyses specifically in osteoporotic vertebral compression fractures found that VP was associated with a higher adjusted risk of pulmonary complications (adjust HR 1.07, 95% CI 1.05 to 1.10) compared with UC but a similar risk of all other SAEs. Two additional studies used data from the NHIRD from Taiwan and had overlapping populations but reported different outcomes. In one study[87] there was no difference between VP and UC in the risk of PE (0.4% vs. 0.7%) or vertebral osteomyelitis or infection (1.0% vs. 0.9%) over a mean follow-up of 4.5 years; the UC care group was followed a mean of 1 year longer than the VP group. In the second study, VP was associated with a lower risk of respiratory failure versus UC.[106] See Table 21 for details.

Reoperation

In addition to the cement leakage requiring reoperation mentioned above, two RCTs noted reoperation for symptomatic new fractures (Table 19). One small, fair-quality RCT reported two cases (11.1%) of painful adjacent level fracture after VP that required reoperation after failed conservative care.[180] A second, poor-quality RCT reported a similar rate of (re)operation for new symptomatic fractures in patients randomized to VP versus UC (10% vs. 9.1%, respectively).[197]

One large Medicare database study that reported propensity score adjusted analyses found that VP was associated with a substantially greater risk of subsequent augmentation procedures over 4 years compared with UC, HR 11.1 (95% CI 11.1 to 12.5).[56]

Cement Leakage

In all but one case (1%, 1/100),[64] the cement leakage was asymptomatic and there were no cases of leakage into the spinal canal. The incidence of cement leakage per vertebra/levels treated ranged widely from 14.0% to 72.4% across five RCTs (508 total vertebra/levels) that provided clear data (Table 20).[18,33,64,93,144,145,195] When only the three fair-quality trials were considered (343 total vertebra/levels),[18,33,93] the incidence ranged from 49.3% to 72.4%. Two additional trials (one fair and one poor quality) noted that cement leakage occurred but did not provide details or sufficient data to calculate an incidence (Table 20).[144,145,197] The volume of cement injected during VP was less than 5 ml in all trials that provided this information.

An observational follow-up study[93] to one of the above mentioned RCTs[93] evaluated the longer term risk of cement leakage in a subset of VP patients. This study included only half of the patients (53.5%, 54/101) originally randomized to PV; perivertebral cement leakage occurred in 64 of 80 treated vertebra (80%). After a mean follow-up of 22 months, there was no change on CT scan, the patients remained asymptomatic and there were no cases of late cement migration. Results from this subgroup analysis should be interpreted with caution.

Table 20. VP vs. UC: Cement Leakage from RCTs

Author, year Quality	F/U	Mean PMMA volume	% with cement leakage after VP (n/N vertebra)	Comments
Chen, 2014 Fair	12 mos.	3.6 ml	52.2% (36/69)	All asymptomatic; most discal or paravertebral, or into veins or the puncture path, none into spinal canal
Blasco, 2012 Fair	NR	NR	49.3% (69/140)	All asymptomatic; most discal or into veins
Klazen, 2010 Fair	12 mos.	4.1 ml	72.4% (97/134)	All asymptomatic; most discal or into segmental veins, none into spinal canal
Rousing, 2009/2010 Fair	3 and 12 mos.	NR	Data NR (63 levels)	No adverse events except for extra-vertebral cemental leaks (no data provided); all asymptomatic
Farrokhi, 2011 Poor	1 wk.	3.5 ml	14.0% (14/100)	13 asymptomatic, 1 symptomatic*; paravertebral, discal or epidural, none into spinal canal
Yang, 2016 Poor	12 mos.	4.5 ml	33.8% (22/65)	All asymptomatic; most discal or paravertebral, none into spinal canal
Yi, 2014 Poor	Mean 49.4 mos.	4.0 ml	4 cases† N unclear (90 patients, levels NR)	All asymptomatic; most discal or paravertebral, none into spinal canal

F/U = follow-up; mos. = months; NR = not reported; PMMA = polymethylmethacrylate; VP = vertebroplasty; wk. = week.
*Reoperation included immediate decompression through a bilateral laminectomy and evacuation of bone cement; the patient could walk unassisted with no radicular pain after 2 months.
†This trial also included a kyphoplasty (KP) arm and for some outcomes the authors did not report data for vertebroplasty (VP) and KP separately (i.e., combined them into one interventional arm). It is unclear whether these 4 cases were all in VP arm.

Table 21. Adverse Events Other than Mortality from Comparative Database Studies Comparing Vertebroplasty versus Usual Care for Osteoporotic Vertebral Compression Fractures

Adverse Event	Database	Study Database search dates	N	Finding and conclusion
SAE				
PE	NHIRD	Huang, 2019 (2003-2013)	VP: 1,389 UC: 6,017	Mean 3.2 (2.5) vs. 4.8 (3.2) years: 0.4% (6/1389) vs. 0.7% (42/6017), p=NS
	Medicare	Ong, 2018 (2005-2014)	VP: 117,323 UC: 1,698,956 All patients*	Propensity-adjusted risk: 1 year: 3% higher with VP vs. UC, p<0.05 2 years: 7% higher with VP vs. UC, p<0.001 5, 8, 10 years: 6% higher with VP vs. UC, p<0.001
		Edidin 2015 (2005-2009)†	VP: 36,657 UC: 107,930 Propensity-score matched	4 years: Adj. HR 1.07 (95% CI 0.98 to 1.18)
DVT	Medicare	Ong, 2018 (2005-2014)	VP: 117,323 UC: 1,698,956 All patients*	Propensity-adjusted risk: 1 year: 5% higher with VP vs. UC, p<0.001 2 years: 3% higher with VP vs. UC, p<0.001 5, 8, 10 years: 0% difference

		Edidin 2015 (2005-2009)†	VP: 36,657 UC: 107,930 Propensity-score matched	4 years: Adj. HR: 1.03 (0.97 to 1.08)
Cardiac complications	Medicare	Ong, 2018 (2005-2014)	VP: 117,323 UC: 1,698,956 All patients*	Propensity-adjusted risk, p<0.001: 1 year: 20% lower with VP vs. UC 2 years: 13% lower with VP vs. UC 5 years: 9% lower with VP vs. UC 8 and 10 years: 7% lower VP vs. UC
		Edidin 2015 (2005-2009)†	VP: 36,657 UC: 107,930 Propensity-score matched	4 years: Adj. HR 0.96 (95% CI 0.90 to 1.03)
Pulmonary/ respiratory complications	Medicare	Ong, 2018 (2005-2014)	VP: 117,323 UC: 1,698,956 All patients*	Propensity-adjusted risk: 1, 2 years: 1% lower with VP vs. UC, p<0.01 5, 8, 10 years: 0% difference
		Edidin 2015 (2005-2009)†	VP: 36,657 UC: 107,930 Propensity-score matched	4 years: Adj. HR 1.07 (95% CI 1.05 to 1.10)
Respiratory failure	NHIRD	Lin 2017 (2000-2013)	VP: 1,773 UC: 5,324 Propensity-score matched	1 year, VP vs. UC: 0.26 (95% CI 0.20-0.34) vs. 0.36 (95% CI 0.31-0.41) per 100-person months Adj. HR 0.68 (0.50 to 0.96)
Infection	Medicare	Ong, 2018 (2005-2014)	VP: 117,323 UC: 1,698,956 All patients*	Propensity-adjusted risk: 1 year: 6% lower with VP vs. UC, p<0.001 2, 5, 8, 10 years: 1% lower with VP vs. UC, p=NS
		Edidin 2015 (2005-2009)†	VP: 36,657 UC: 107,930 Propensity-score matched	4 years: Adj. HR 1.00 (95% CI 0.88 to 1.14)
Vertebral osteomyelitis or infection	NHIRD	Huang, 2019 (2003-2013)	VP: 1,389 UC: 6,017	Mean 3.2 (2.5) vs. 4.8 (3.2) years: 1.0% (14/1389) vs. 0.9% (54/6017), p=NS
New Fracture				
Any	NHIRD	Huang, 2019 (2003-2013)	VP: 1,389 UC: 6,017	Mean 3.2 (2.5) vs. 4.8 (3.2) years <0.3% (NR/1389) vs. <0.1% (NR/6017), p=NS
Reoperation				
Subsequent augmentation	Medicare	Edidin 2015 (2005-2009)†	VP: 36,657 UC: 107,930 Propensity-score matched	4 years Any subsequent augmentation: Adj. HR 11.1 (95% CI 11.1 to 12.5) Subsequent augmentation or fusion: Adj. HR 11.1 (95% CI 11.1 to 12.5)

Adj. HR = adjusted hazard ratio; CI = confidence interval; DVT = deep vein thrombosis; N/A = not applicable; NHIRD = National Health Insurance Research Database of Taiwan; NR = not reported; NS = not significant; PE = pulmonary embolism; SAE = serious adverse event; UC = usual care; VP = vertebroplasty.

* Authors do a propensity score matched analysis but only provide data for the larger population; n's unclear for adjusted analyses.

† Data are for the OVCF cohort (osteoporotic and pathologic) which excludes the traumatic VCF patients.

4.2.2.1.3 Vertebroplasty versus Minimally Invasive Procedures

Two RCTs[168,181] and one retrospective comparative NRSI[8] that compared VP to other minimally invasive surgeries (i.e., nerve or facet blocks) and were included for efficacy also reported safety. Adverse events were not well reported. None of studies reported mortality or serious adverse events.

New vertebral fractures

PV was associated with a similar risk of new vertebral fractures at 12 months compared with facet block in one RCT (13% [13/100] vs. 10.4% [11/106], RR 1.25, 95% CI 0.59 to 2.67).[181]

In the NRSI, VP was associated with a large increase in the risk of any new vertebral fractures compared with medial branch block at 24 months (15.2% [14/92] vs. 4.2% [3/72], RR 3.65, 95% CI 1.09 to 12.23); the association remained significant after adjustment for confounding factors (e.g., age, sex, bone mineral density, use of drugs to treat osteoporosis).[8]

Cement leakage

One case (1%, 1/100) of asymptomatic cement leakage following VP was reported through 12 months in one RCT[181]; symptomatic cement leakage was not reported (or did not occur)(See Appendix Table G4).

In the NRSI, five patients (5.3%) in the VP group experienced cement leakage; only one case (1%) was symptomatic (subjective leg weakness) and resolved after 1 week.[8] The remaining asymptomatic leakages did not require additional treatment.

Other adverse events

One small RCT (N=27) reported readmission at 8 weeks due to continued back pain in two patients (15.4%) who received medial branch nerve root block; none of the patients who received VP required readmission.[168] Overall, 48.1% (13/27; not reported by group) of patients experienced any adverse event throughout the 8 week follow-up in this same trial.[168] (See Appendix Table G4).

4.2.2.1.4 Vertebroplasty versus Surgical Procedures

Two large administrative database studies that had overlapping populations from the Taiwan NHIRD database but reported different outcomes compared VP with conventional open surgery[87] and VP with other non-VP management[192] to include surgery or conservative medical care. Both studies found a similar risk of SAEs, to include PE, stroke, infection and new fracture following both procedures. One study found no difference in the risk of mortality over a mean of 4.5 years between VP and open surgery (Table 22).

Table 22. Adverse Events Other than Mortality from Comparative Database Studies Comparing Vertebroplasty versus Other Surgery for Osteoporotic Vertebral Compression

Adverse Event	Database	Study Database search dates	N	Follow-up Findings
Various SAEs	NHIRD	Huang, 2019 (2003-2013)	VP: 1,389 Open Surgery: 1,219	Mean 3.2 (2.5) vs. 4.7 (3.1) years, p=NS for all Mortality: 19.2% (267/1389) vs. 17.4% (212/1219) PE: 0.4% (6/1389) vs. ≤0.3% (NR/1219) Vertebral osteomyelitis or infection:1.0% (14/1389) vs. 1.0% (12/1219) New Fracture: <0.3% (NR/1389) vs. <0.3% (NR/1219)
Any Stroke		Wu, 2012 (2000-2008)	VP: 334 Other surgery or medical treatment: 1,655 Propensity- score matched	≤5 years: Any stroke: Adj. HR 1.22 (95% CI 0.67 to 2.24) Hemorrhagic Stroke: Adj. HR 3.17 (95% CI 0.97 to 10.3) Ischemic stroke: Adj. HR 0.96 (95% CI 0.49 to 1.91)

Adj. HR = adjusted hazard ratio; NHIRD: National Health Insurance Research Database of Taiwan; NS = not significant; PE = pulmonary embolism; VP = vertebroplasty.

4.2.2.1.5 Case Series of Vertebroplasty

Six case series[9,51,62,96,169,177] evaluated adverse events following VP. Sample sizes ranged from 292 to 1,512. Follow-up periods ranged broadly from peri-operative to 36 months. See Table 23 for a summary of AEs and Appendix Table K2 for further details.

Mortality

One case series of VP reported 1 year mortality of 1.2% (6/485)[96] and another reported no death due to embolism at a mean of 8 months (0/1512).[62]

SAEs

Across two case series,[62,177] the frequency of symptomatic SAEs (cardiopulmonary arrest, cement embolism) was very low (≤0.3%). Procedure-related AEs were unclear or not well reported; only one study stated that no procedure related AEs occurred.[96]

Cement Embolism

Three studies[62,163,177] looked specifically at the risk of cement embolism. In one study[163] (N=373) the incidence of pulmonary cement embolism (PCE) on post-procedural CT was 17.2%; authors state that most cases were asymptomatic, and the incidence was similar for VP and KP. In another study,[177] 3.7% of patients (N=299) had an asymptomatic PCE during VP; follow-up after 12 months showed no further sequelae or symptoms. In the third study[62] (N=1512), which included a mixed population of osteoporotic and malignant fractures, the incidence of any intercardiac cement embolism was 4.8% following VP but symptomatic embolisms were rare (0.3%).

New fractures

The incidence of any new vertebral fracture following VP ranged from 11.6% to 22.1% across three studies[9,96,169] and of any new adjacent fracture ranged from 6.6% to 7.8% across two of these studies.[9,96] No trial reported on symptomatic new fractures.

Cement leakage

No cases of symptomatic cement leakage were reported in one study.[96] Any cement leakage (primarily asymptomatic) is common following VP and the frequency varies widely (16.0% to 77.7%).[9,51,96,169]

Reoperation

One large case series[169] (N=1090) reported that 22.1% of patients required reoperation for new vertebral fractures but did not indicate if the fractures were symptomatic.

Table 23. Adverse Events in Single Arm Studies of VP for Osteoporotic fractures

Adverse Event	Follow Up	Study	% (n/N), VP
Mortality			
Any	1 year	Kobayashi, 2021*	1.2% (6/485)
Mortality Due to Embolism	Mean 8.1 months	Fadili Hassani, 2019	0% (0/1512)
SAE			
Cardiorespiratory arrest	Mean 8.1 months	Fadili Hassani, 2019†	<0.1% (1/1512)
Any intercardiac cement embolism	Mean 8.1 months	Fadili Hassani, 2019†	4.8% (72/1512)
Intercardiac cement embolism with PCE	Mean 8.1 months	Fadili Hassani, 2019†	4.1% (62/1512)
Symptomatic intercardiac cement embolism	Mean 8.1 months	Fadili Hassani, 2019†‡	0.3% (6/1512)
Asymptomatic PCE	Perioperative	Venmans, 2008	3.7% (11/299)
New Fracture			
Any	1 year	Kobayashi, 2021	18.6% (67/361)
	36 months	Bae, 2017	11.6% (34/293)
	NR	Tang, 2021	22.1% (241/1090)
Adjacent Fracture	1 year	Kobayashi, 2021	6.6% (24/361)
	36 months	Bae, 2017	7.8% (23/293)
Distant Fracture	1 year	Kobayashi, 2021	12.7% (46/361)
	36 months	Bae, 2017	3.8% (11/293)
Symptomatic Cement Leakage			
Any	1 year	Kobayashi, 2021	0% (0/485)
Any Leakage	≥2 years	Ding, 2016	77.7% (227/292)
Asymptomatic Leakage	1 year	Kobayashi, 2021	35.7% (173/485)
Spinal canal leakage	1 year	Kobayashi, 2021	5.3% (26/485)
Cortical Leakage	NR	Tang, 2021	20.3% (295/1456 levels)
Venous Leakage	NR	Tang, 2021	56.2% (819/1456 levels)
Adjacent disc space leakage	36 months	Bae, 2017	16.0% (41/256) [1 level]
Reoperation			
Reoperation for new fracture	NR	Tang, 2021	22.1% (241/1090)
Procedure-Related AE			
Any	1 year	Kobayashi, 2021*	0% (0/485)

AE = adverse event; PCE = pulmonary cement embolism; NR = not reported; SAE = serious adverse event; VP = vertebroplasty.
* Not related to procedure
† Mixed pop of osteo (34%) and malignant (40%)
‡ None of the embolisms resulted in death.

4.2.2.1.6 Vertebroplasty versus Kyphoplasty

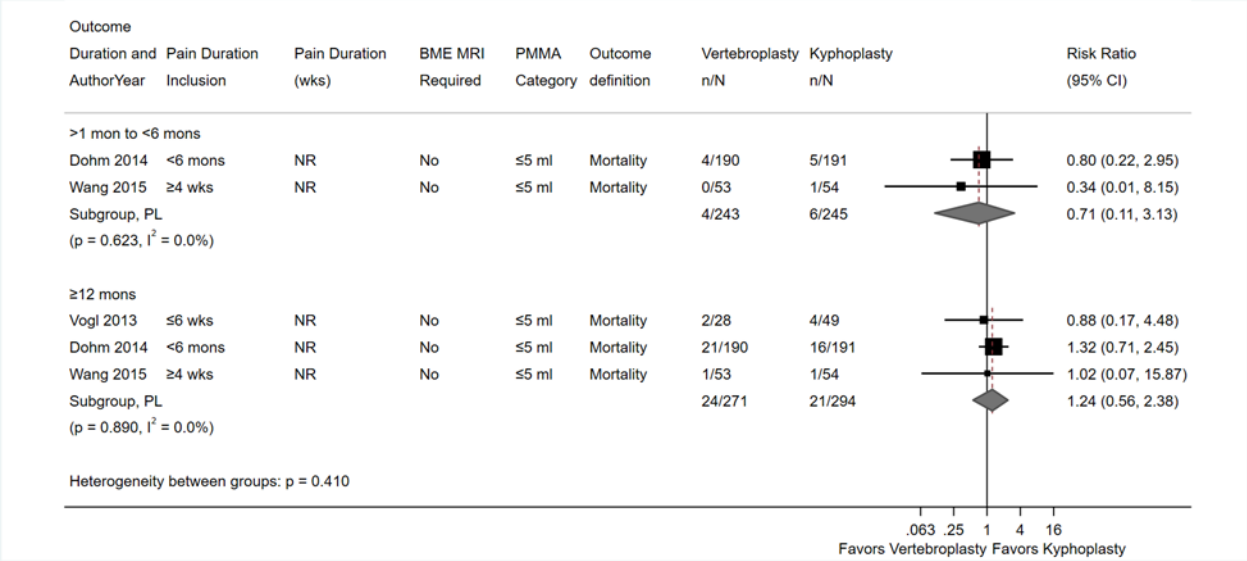
Ten trials (reported in 11 publications) compared VP versus KP for the treatment of osteoporotic compression fractures (patients with fracture due to cancer were excluded) (Table 12 and Appendix G-H).[52,59,61,77,108,109,179,182-184,197]

In addition, four retrospective administrative database studies (using overlapping populations from the U.S. Medicare and ACS-NSQIP databases)[56,87,106,129] and two retrospective comparative NRSIs[34,163] evaluated adverse events for VP compared with KP. Further details of these database studies are provided above at the beginning of the Safety section. Appendix Table K2 provide details of all nonrandomized studies.

Mortality

In RCTs, VP and KP were associated with a similar cumulative risk of mortality at 3 months (2 RCTs, N=488, 1.6% vs. 2.4%)[52,182] and at latest follow-up (12 to 24 months; 3 RCTs, N=565, 8.9% vs. 7.1%),[52,179,182] Figure 25. One trial (N=66)[59] reported two deaths (3%) by 6 months but did not indicate to which treatment group the patients were randomized. In general, deaths were contributed to causes unrelated to the procedures or the trial.

Figure 25. VP vs. KP: Cumulative risk of mortality



BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PL = profile likelihood; PMMA = polymethylmethacrylate; wks = weeks.

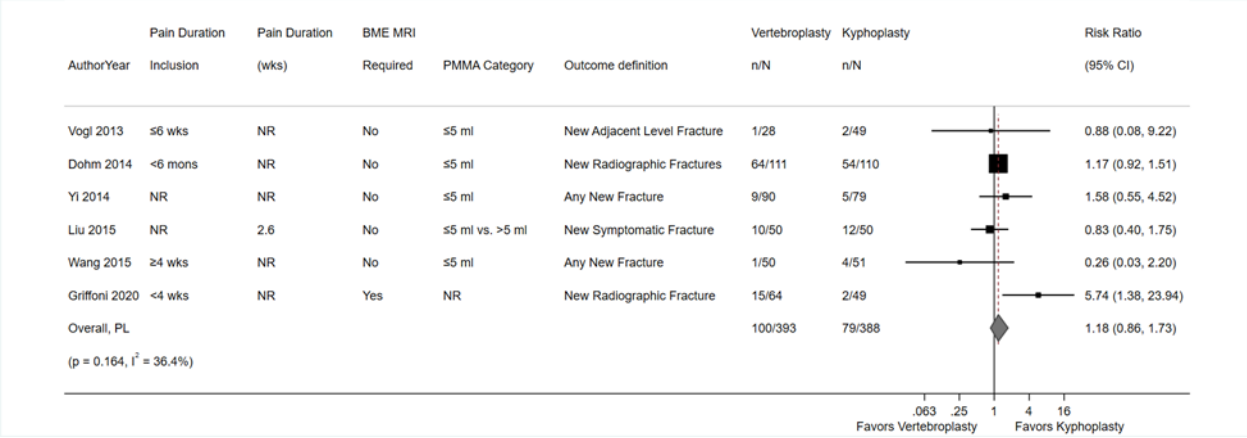
Two studies using the same administrative database[35,92] report that procedure type (KP vs. VP) was not an independent risk factor for 30-day mortality. At longer time fames two studies[56,129] using Medicare data with overlapping sample frames reported that KP was associated with slightly lower mortality risk compared with VP based on estimates for the larger of these studies (adjusted HR 0.87 95% CI 0.87, 0.88),[129] Table 17.

Subsequent fractures

Any new fracture and new symptomatic fracture

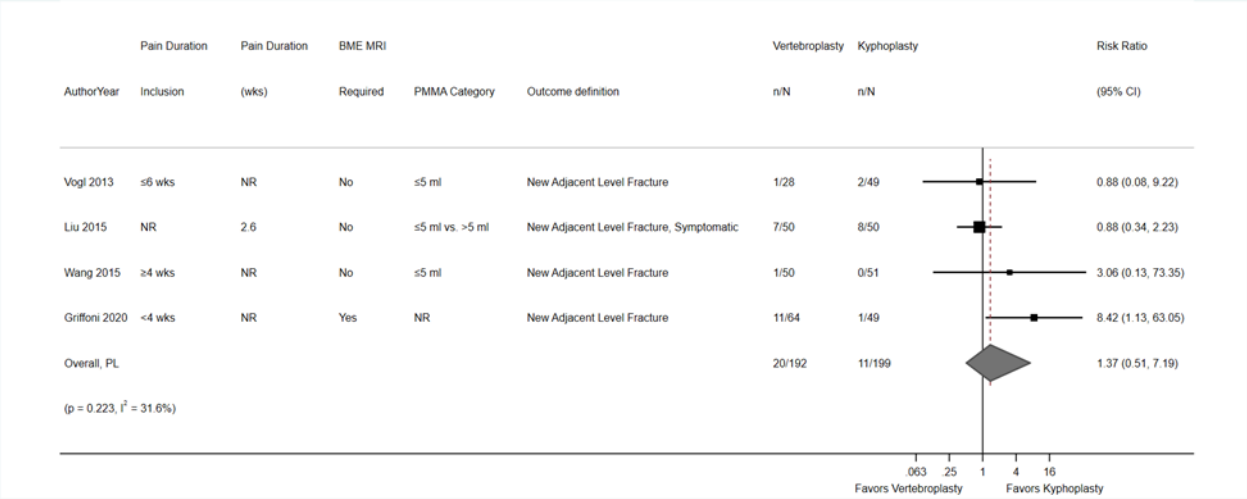
VP and KP were associated with a similar risk of any new fracture (6 RCTs, N=781, 25.4% vs. 20.4%; follow-up range, 12 to a mean of 49 months),[52,77,108,179,182,197] (Figure 26) and new adjacent level fracture (4 RCTs, 10.4% vs. 5.6%, 12 to 60 months)[77,108,179,182] (Figure 27) at latest follow-up, though adjacent level fractures tended to occur more frequently following VP. Results were similar across sensitivity analyses that stratified by timing (RR 1.21, 95% CI 0.85 to 1.83, I^2=36.4%), that excluded an outlier trial[77] (RR 1.15, 95% CI 0.74 to 1.52, I^2=0%) and that limited analysis to new radiographic only fractures (RR 1.22, 95% CI 0.82 to 2.57, I^2=56.3) (Appendix P, Figures 17 to 24). Similarly, there was no difference in the risk of new symptomatic fractures following VP versus KP; however, only two poor-quality trials specified symptomatic fractures, and the estimate was imprecise (Figure 28).

Figure 26. VP vs. KP: Cumulative risk of any new vertebral fractures at latest follow-up



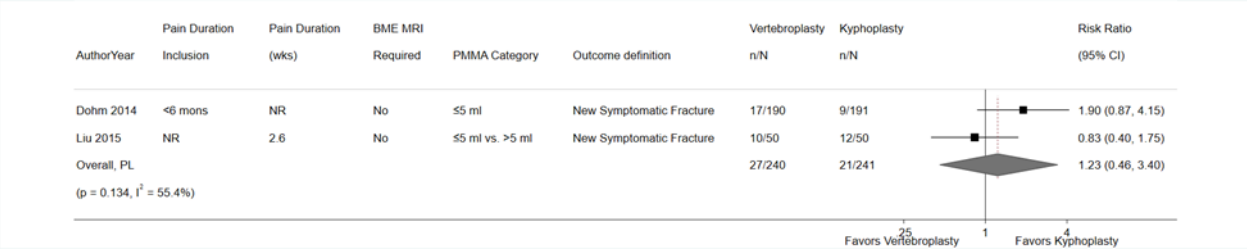
BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; PL = profile likelihood; PMMA = polymethylmethacrylate; wks = weeks.

Figure 27. VP vs. KP: Cumulative risk of any new adjacent level vertebral fractures



BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; PL = profile likelihood; PMMA = polymethylmethacrylate; wks = weeks.

Figure 28. VP vs. KP: Cumulative risk of any new symptomatic vertebral fractures



BME = bone marrow edema; CI = confidence interval; mons = months; MRI = magnetic resonance imaging; PL = profile likelihood; PMMA = polymethylmethacrylate; wks = weeks.

One comparative NRSI (N=338)[34] found that VP was associated with a large decrease in the risk of adjacent level fractures by 1 year (3.3% vs. 9.8%, RR 0.33, 95% CI 0.13 to 0.83) (See Appendix Table I3)

Refracture or Worsening Index Level Fracture

There was no difference between VP and KP in the risk of refracture or worsening fracture at the index level across two poor-quality trials (N=348, 6.3% [10/159] vs. 2.6% [5/189], RR 2.24, 95% CI 0.29 to 8.49, I^2=0%)[52,179]; however, the estimate was imprecise (Appendix Figure X). One trial[179] reported a case of a symptomatic refracture at the index level in a patient who was randomized to KP; it was re-treated using VP. In the second trial,[52] more patients who received VP versus KP experienced worsening index fractures at 12 months, but the difference was not statistically significant; there were no additional worsening index fractures at 24 months. These fractures are likely included in the count above for any new vertebral fractures.

Serious Adverse Events

Any SAE

One poor-quality RCT (N=381)[52] reported a similar frequency of any SAE (27.4% vs. 26.2%) and of procedure- or device-related or possibly related SAE (4.2% vs. 4.2%) within 30 days after VP versus KP; by 24 months results remained similar between groups, respectively (65.8% vs. 65.5% and 5.8% vs. 6.3%), Table 24. For this study, serious adverse events (SAEs) included death, serious deterioration in health, life threatening injury/illness, hospitalization or prolonged hospitalization, or resulted in medical or surgical intervention. A second poor quality trial[197] reported no SAE occurred in any patient over a mean 49.4 months.

Two large administrative database studies (Medicare) with overlapping populations found that VP was associated with a greater risk of most SAEs compared with KP over 10 years of follow-up in propensity score matched analyses. In the larger dataset,[129] VP was associated with higher adjusted relative risks of thromboembolic events (i.e., PE and DVT), cardiac complications (e.g., MI), pulmonary complications, and infection with the biggest difference in relative risk seen for PE (9% to 12% higher risk) at all timepoints, and DVT at 1 year (7% higher risk). The other differences in risks ranged from 1% to 3% across timepoints and some may not be clinically relevant. The second, smaller dataset[56] that provided subgroup analyses specifically in osteoporotic fractures found that VP was associated with higher risks of PE (adjusted HR 1.16, 95% CI 1.01 to 1.35) and pulmonary complications (adjusted HR 1.05, 95% CI 1.01 to 1.09) over 4 years compared with KP, but not DVT, cardiac complications (e.g., MI), or infection which were similar between both cohorts. Two additional database studies using overlapping

populations from the ACS-NSQIP database found that VP was associated with a 3-fold increase in risk of pulmonary/respiratory complications (1 study)[35] but a similar risk of any SAE (1 study)[92] by 30 days in adjusted analyses compared with KP. See Table 26 for details.

Reoperation

Reoperation for any new fracture or refracture occurred with similar frequency following VP and KP across the RCTs except for one trial[77] which found that VP was associated with a large increase in the risk of reoperation compared with KP (23.4% vs. 4.1%; RR 5.74, 95% CI 1.38 to 23.94), Table 24.

Table 24. VP vs. KP: Summary of serious adverse events and reoperation from RCTs

Outcome	Author, year Quality	F/U	VP % (n/N)	KP % (n/N)	RR (95% CI)
SAEs					
Any SAE	Dohm, 2014 Poor	30 days	27.4% (52/190)	26.2% (50/191)	1.04 (0.75 to 1.46)
		24 mos.	65.8% (125/190)	65.5% (125/191)	1.00 (0.87 to 1.16)
	Yi, 2014 Poor	mean 49.4 mos.	0% (0/90)	0% (0/79)	NA
Procedure or device related SAEs†	Dohm, 2014 Poor	30 days	4.2% (8/190)	4.2% (8/191)	1.01 (0.39 to 2.62)
		24 mos.	5.8% (11/190)	6.3% (12/191)	0.92 (0.42 to 2.04)
Reoperation					
Reoperation for any new or refracture	Vogl, 2013‡ Poor	12 mos.	3.6% (1/28)	2.0% (1/49)	1.75 (0.11 to 26.90)
	Wang, 2015§ Fair	12 mos.	2.0% (1/50)	7.8% (4/51)	0.26 (0.03 to 2.20)
	Griffoni, 2020**	12 mos.	23.4% (15/64)	4.1% (2/49)	5.74 (1.38 to 23.94)
	Yi, 2014††	Mean 49.4	10.0% (9/90)	6.3% (5/79)	1.58 (0.55 to 4.52)
Discectomy with posterior spinal fusion for severe discogenic back pain related to cement leak	Wang, 2015 Fair	12 mos.	0% (0/50)	2.0% (1/51)	NA
Other AEs (not including serious)					
Any nonserious AE	Dohm	24 mos.	36.3% (69/190)	28.8% (55/191)	1.26 (0.94 to 1.69)

AE = adverse event; CI = confidence interval; F/U = follow-up; KP = kyphoplasty; mos. = months; NA = not applicable; RR = risk ratio; SAEs = serious adverse events; VP = vertebroplasty.

* For this study, serious adverse events (SAEs) included death, serious deterioration in health, life threatening injury/illness, hospitalization or prolonged hospitalization, or resulted in medical or surgical intervention.

† Not further defined.

‡ VP: reoperation with VP for 3 adjacent level fractures.

KP: sx refracture at the index level, treated with VP.

§ VP: adjacent level fracture, surgically treated.

KP: nonadjacent level fractures, surgically treated.

** Any new radiographic fracture, all required reoperation.

†† Reoperation with VP or KP for new fracture.

One large Medicare database study that reported propensity score adjusted analyses found that VP was associated with a similar risk of subsequent augmentation procedures over 4 years compared with KP, to include retreatment for subsequent VCFs, HR 1.03 (95% CI 0.97 to 1.09), Table 26.[56]}

Cement Leakage

Symptomatic cement leakage was rare and occurred with similar frequency following treatment with VP (0% to 1.1%) versus KP (0% to 1.9%) across five RCTs (Table 25).[52,59,179,182,197] Three RCTs reported no cases of symptomatic leakage in any patient; across the other two trials, there were two cases in both groups. One fair-quality trial[182] reported one case of discogenic back pain related to cement leakage that required discectomy with posterior lateral fusion in a patient who underwent KP. A second poor-quality trial[52] reported one case of inferior cement leakage that possibly contributed to a new symptomatic fracture at 2 days postoperatively in one KP patient.

In general, VP was associated with a higher risk of any (mostly asymptomatic) cement leakage compared with KP across six trials [52,59,77,179,183,184] but the difference was not always statistically significant (Table 25). Two poor-quality trials[52,179] reported the rate of leakage out of the number of levels treated (range, 74% to 81.6% with VP vs. 48% to 73.4% with KP) and four trials (1 fair and 3 poor quality) [8,59,77,183,184] reported it out of the number of patients (range, 4.7% to 30.2% with VP vs. 4.1% to 11.4% with KP). One fair-quality trial[182] found that VP was associated with a significantly lower risk of cement leakage versus KP based on the number of levels treated; it is unclear why this trial's findings differed from the others. Another poor-quality trial[197] reported only four cases of cement leakage (out of 169 patients, 217 levels) but did not report to which groups the patients were randomized.

Similarly, one comparative NRSI (N=338)[34] reported a higher frequency of asymptomatic cement leakage following VP versus KP (7.0% vs. 0%, p=0.003). See Appendix Table I3.

Cement Embolism

Symptomatic cement embolism was rare (< 1%) and occurred with similar frequency following VP versus KP in one poor-quality RCT (N=381, 0.5% [1/190] vs. 0.5% [1/191], RR 1.01, 95% CI 0.06 to 15.96)[52]; the estimate was imprecise, and the trial was likely unpowered to detect this rare event. A second RCT (N=101)[182] reported one case (2.0%) of asymptomatic cement embolism in the right lung in a patient who received KP; there were no cases of embolism in the VP group.

One comparative NRSI (N=373)[163] specifically designed to look for pulmonary cement embolism using postprocedural CT scans reported an overall incidence of 17.2% with similar incidences a median of 412 days after VP versus KP (18.2% vs. 14.8%, RR 1.23, 95% CI 0.74 to 2.05); authors infer that most were asymptomatic.

Table 25. VP vs. KP: Cement Leakage

		PMMA volume (ml)	Any cement leakage, % (n/N)			Symptomatic cement leakage, % (n/N)		
Author, year Quality	F/U	VP vs. KP	VP	KP	Conclusion	VP	KP	Conclusion
Wang 2015[*] Fair	12 mos.	3.3 vs. 4.2	13.2% (9/68 levels)	30.6% (22/72 levels)	Any leakage **Lower with VP** (p=0.013) - primarily into disc space	0% (0/53)	1.9% (1/54)[†]	**Similar** between groups, (p=0.68)
Vogl, 2013[‡] Poor	Post-tx	4.0 vs. 3.8	74% (29/39 levels[§] Spinal canal, leaks per level: 37% [**]	48% (31/65 levels[††] Spinal canal, leaks per level: 3.5%	Any leakage **higher with VP** (p=0.013) - primarily lateral cortical leaks Spinal canal leakage **higher with VP** (p=0.026)	0% (0/28)	0% (0/49)	**Similar** between groups
Dohm, 2014[‡] Poor	NR	4.0 vs. 4.6	81.6% (164/201 levels)	73.4% (157/214 levels)	Any leakage **higher with VP** (p=0.047) - primarily discal and intravascular leaks Spinal canal leakage **similar** between groups, (p=0.12, data NR)	1.1% (2/190[‡‡]	0.5% (1/191[§§]	**Similar** between groups
Griffoni, 2020[***] Fair	12 mos.	NR	4.7% (3/64)	4.1% (2/49)	Any leakage **similar** between groups (RR 1.2, 95% CI 0.2 to 6.6); No other information provided	NR	NR	NR
Wang, 2023 Poor	3 mos.	NR	24% (12/50)	8% (4/50)	Any leakage **higher with VP** (RR[†††] 3.00, 95% CI 1.04 to 8.67); No other information provided	NR	NR	NR
Endres, 2012[‡‡‡] Poor	6 mos.	3.1 vs. 4.3	27.3% (6/22)	11.4% (5/44)[§§§]	Any leakage **similar** between groups (RR 2.40, 95% CI 0.82 to 7.00); - all discal or lateral leaks	0% (0/22)	0% (0/44)	**Similar** between groups
Wang, 2018[***] Poor	Peri-op	3.9 vs. 4.0	30.2% (13/43)	9.3% (4/43)	Any leakage **higher with VP** (RR 3.25, 95% CI 1.15 to 9.18); No other information provided	NR	NR	NR

		PMMA volume (ml)	Any cement leakage, % (n/N)			Symptomatic cement leakage, % (n/N)		
Author, year Quality	F/U	VP vs. KP	VP	KP	Conclusion	VP	KP	Conclusion
Yi, 2014**** Poor	Mean 49.4 mos.	4.0 ml	4 cases††††, N unclear (169 patients, 217 levels) Most discal or paravertebral, none into spinal canal			0% (0/90)	0% (0/79)	**Similar** between groups

CI = confidence interval; F/U = follow-up; KP = kyphoplasty; mos. = months; NR = not reported; peri-op = perioperative; PMMA = polymethylmethacrylate; RR = risk ratio; tx = treatment; VP = vertebroplasty.

* Radiographs supplemented by CT scans.

† Discogenic back pain related to leakage, required discectomy with posterior lateral fusion.

‡ CT scan.

§ 12 levels had multiple leaks for a total of 54 leaks in 39 levels.

** Multiple leaks possible per level

†† 6 levels had multiple leaks for a total of 42 leaks in 65 levels.

‡‡ 1 cement embolism and 1 new symptomatic fracture occur within 2 days postoperatively (inferior to the index level), with inferior cement leakage that was considered possibly bone cement–relate

§§ Cement embolism

*** NR (CT or radiograph)

††† Calculated.

‡‡‡ Radiographs supplemented by CT scans

§§§Balloon: 4/22, Shield: 1/22.

**** Radiograph and MRI.

†††† Authors did not report data for VP and KP separately (i.e., combined them into one interventional arm).

Table 26. Adverse Events Other than Mortality from Comparative Database Studies Comparing Vertebroplasty versus Kyphoplasty for Osteoporotic Vertebral Compression

Adverse Event	Database	Study Database search dates	N	Finding and conclusion
SAE				
PE	Medicare	Ong, 2018 (2005-2014)	VP: 117,232 KP: 261,756 All patients*	Propensity-adjusted risk, p<0.001: 1, 5, 8, 10 years: 9% higher with VP vs. KP 2 years: 12% higher with VP vs. KP
		Edidin 2015 (2005-2009)†	VP: 37,252 KP: 36,286 Propensity-score matched	4 years: Adj. HR: 1.16 (95% CI 1.01 to 1.35)
DVT	Medicare	Ong, 2018 (2005-2014)	VP: 117,232 KP: 261,756 All patients*	Propensity-adjusted risk, p<0.001: 1 year: 7% higher with VP vs. KP 2 years: 4% higher with VP vs. KP 5, 8, 10 years: 2% higher VP vs. KP
		Edidin 2015 (2005-2009)†	VP: 37,252 KP: 36,286 Propensity-score matched	4 years: Adj. HR 1.05 (95% CI 0.96 to 1.15)
Cardiac complications (to include MI)	Medicare	Ong, 2018 (2005-2014)	VP: 117,232 KP: 261,756 All patients*	Propensity-adjusted risk: 1 year: 1% lower with VP vs. KP, p=NS 2 years: 1% higher with VP vs. KP, p=NS 5 years: 2% higher with VP vs. KP, p<0.01 8, 10 years: 3% higher with VP vs. KP, p<0.001
		Edidin 2015 (2005-2009)†	VP: 37,252 KP: 36,286 Propensity-score matched	4 years: Adj. HR 1.05 (95% CI 0.94 to 1.16)
Pulmonary/ respiratory complications	Medicare	Ong, 2018 (2005-2014)	VP: 117,232 KP: 261,756 All patients*	Propensity-adjusted risk: 1 year: 1% higher with VP vs. KP, p<0.05 2 years: 2% higher with VP vs. KP, p<0.001 5, 8, 10 years: 3% higher with VP vs. KP, p<0.001
		Edidin 2015 (2005-2009)†	VP: 37,252 KP: 36,286 Propensity-score matched	4 years: Adj. HR 1.05 (95% CI 1.01 to 1.09)
	ACS-NSQIP	Choo, 2018 (2012-2014)	VP: 242 KP: 2,191	30 days: Adj. OR 3.28 (95% CI 1.56-6.88)
Infection	Medicare	Ong, 2018 (2005-2014)	VP: 117,232 KP: 261,756 All patients*	Propensity-adjusted risk: 1 year: 3% lower with VP vs. KP, p<0.05 2 years: 1% higher with VP vs. KP, p<0.001 5, 8, 10 years: 0% difference
		Edidin 2015 (2005-2009)†	VP: 37,252 KP: 36,286 Propensity-score matched	4 years: Adj. HR 1.05 (95% CI 0.87 to 1.27)
Any SAE	ACS-NSQIP	Kim 2022 (2011-2013)	VP: 191 KP: 1741	Adj. OR 1.93 (95% CI 0.58 to 6.41)

Reoperation				
Subsequent augmentation	Medicare	Edidin 2015 (2005-2009)†	VP: 37,252 KP: 36,286 Propensity-score matched	4 years: Any subsequent augmentation: Adj. HR 1.03 (95% CI 0.97 to 1.09) Subsequent augmentation or fusion: Adj. HR 1.03 (95% CI 0.97 to 1.09) Subsequent VCF with repair: Adj. HR 1.03 (95% CI 0.97 to 1.09)

Adj. HR = adjusted hazard ratio; Adj. OR = adjusted odds ratio; CI = confidence interval; DVT = deep vein thrombosis; KP = kyphoplasty; MI = myocardial infarction; NS = not significant; PE = pulmonary embolism; SAE = serious adverse event; VCF = vertebral compression fracture; VP = vertebroplasty.

* Authors do a propensity score matched analysis but only provide data for the larger population; n's unclear for adjusted analyses.

† Data are for the OVCF cohort (osteoporotic and pathologic) which excludes the traumatic VCF patients.

4.2.2.2 Kyphoplasty

4.2.2.2.1 Kyphoplasty versus Usual Care

All four RCTs (in 6 publications)[19,105,110,174,187,197] that compared KP to UC for osteoporotic compression fractures included for efficacy also reported safety outcomes. In addition, three large administrative database studies (two with overlapping populations from the Medicare database and one using NIS data)[56,129,198] and one retrospective comparative NRSI[63] evaluated harms following KP and UC.

Mortality

KP was associated with a similar risk of mortality compared to UC through 24 months (8.1% vs. 7.2%) in one fair-quality RCT.[19,187]

One study using the Nationwide Inpatient Sample reported slightly lower mortality with KP versus non-operative care but estimate precision could not be determined (0.3% vs. 1.6%, Adj OR: 0.52, p= 0.003)[198]. At longer time frames, two studies[56,129] using Medicare data with overlapping sample frames reported that KP was associated with slightly lower mortality risk compared with nonoperative care based on estimates for the larger of these studies (adjusted HR 0.81, 95% CI 0.813, 0.806)[129], Table 17

Serious Adverse Events and Withdrawals due to Adverse Events

KP was associated with a similar risk of any SAE compared to UC through 24 months (49.7% vs. 48.3%) in one fair-quality RCT,[19] defined as death, life threatening injury, or permanent impairment, or an event that required extended hospital stay or intervention to prevent impairment (Table 27). Most SAEs were not attributed to the procedure. Over 24 months, three patients who received kyphoplasty experienced four treatment-related SAEs: a soft tissue hematoma at the surgical site, anterior cement migration that likely caused a new fracture, and a postoperative UTI requiring intervention; the same patient who had a UTI developed spondylitis near the cement that required treatment by 24 months. In the UC group, there was one case of serious back pain at 30 days attributed to treatment. The risk of treatment-related SAEs was similar between groups (2.0% vs. 0.7%) as was the risk of withdrawal due to AEs (0.6% in both groups). A poor-quality trial reported that no SAEs occurred through a mean of 49 months.[197] The trials were likely underpowered to detect rare events.

Three large database studies reported adverse events; two used the Medicare database and had overlapping populations[56,129] and the other used the NIS database.[198] In the largest Medicare database study,[129] the propensity-adjusted risk of cardiac (e.g. MI), pulmonary and thromboembolic complications

(PE and DVT) and infection were significantly lower in the KP group versus the UC group at all time points up to 10 years (Table 29); the risk of cardiac complications was lowest (range, 11% to 19% lower). The second, smaller Medicare dataset[56] that provided subgroup analyses specifically in osteoporotic fractures only found that KP was associated with a lower risk of cardiac complications (adjusted HR 0.88, 95% CI 0.83 to 0.93) and DVT (adjusted HR 0.92, 95% CI 0.87 to 0.96) compared with UC but the risk of PE, pulmonary complications, and infection was similar between groups at 4 years. The third study (NIS database)[198] found no difference between KP and UC in the 30-day risk of DVT (0.2% in both groups) or infection (0.1% in both groups) calculated out of the number of total fractures.

New vertebral fractures

Two RCTs[19,174,187,197] reported no difference in the incidence of new symptomatic vertebral fractures following treatment with KP (range, 6.3% to 17.4%) versus UC (range, 11.3% vs. 14.0%) at latest follow-up. (Table 28). In the fair-quality RCT, new symptomatic vertebral fractures that were possibly or probably related to cement occurred in 3.4% of KP patients through 30 days and 7.4% through 24 months. In this same trial, KP and UC were associated with a similar risk of new radiographic vertebral fractures, to include index level and adjacent level vertebral fractures through 24 months (Table 28).

Conversely, one retrospective comparative NRSI[63] reported significantly fewer incidences of new vertebral fracture over 2 years after treatment with KP versus UC (26.3% vs. 47.1%, OR 0.44, 95% CI NR, p<0.05).

Cement Leakage

Cement leakage that possibly or probably caused symptomatic vertebral compression fractures following KP occurred in 11 patients (7.4%) by 24 months in the fair-quality trial; five cases (34%) occurred by 30 days.[19,174] A poor-quality trial reported no cases of symptomatic cement leakage.[197] Asymptomatic cement leakage was reported by three trials and the frequency varied.[110,174,187,197] See Table 27 for data.

Reoperation

The fair-quality RCT found no difference in the risk of reoperation to treat new symptomatic vertebral compression fractures following KP vs. UC at 24 months (8.1% vs. 4.0%), Table x2.[19,187]

One large Medicare database study that reported propensity score adjusted analyses found that KP was associated with a substantially greater risk of subsequent augmentation procedures over 4 years compared with UC, HR 12.5 (95% CI 12.5 to 14.3), Table 29.[56]

Any adverse event

KP was associated with a moderate increase in the risk of any adverse event compared with UC at 30 days but there was no difference between groups at longer term follow-up in one fair-quality trial (Table 27).[19,174,187] Conversely, in the poor-quality trial, KP was associated with large decrease in the risk of any adverse event compared UC, but the estimate was imprecise and follow-up was unclear.[110]

One large database study (NIS database)[198] found no difference between KP and UC in the 30-day risk of any SAE (1.7% vs. 1.0%) calculated out of the number of total fractures (N=5,766), though the difference approached statistical significance favoring UC (p=0.061).

Table 27. KP vs. UC: Summary of adverse events (other than vertebral fracture) from RCTs

Outcome (cumulative)	Author, year*	F/U	KP % (n/N)	UC % (n/N)	RR (95% CI)
SAEs					
Mortality†	Wardlaw 2009, Boonen 2011	12 mos.	6.0% (9/149)	4.6% (7/151)	1.30 (0.50 to 3.41)
		24 mos.	8.1% (12/149)	7.2% (11/151)	1.11 (0.50 to 2.43)
SAEs‡	Yi, 2014	Mean 49.4 mos.	0% (0/79)	0% (0/121)	-----
	Wardlaw 2009, Boonen 2011, Van Meirhaeghe, 2013	30 days	16.1% (24/149)	11.2% (17/151)	1.43 (0.80 to 2.55)
		12 mos.	38.9% (58/149)	35.7% (54/151)	1.09 (0.81 to 1.46)
		24 mos.	49.7% (74/149)	48.3% (73/151)	1.02 (0.82 to 1.29)
Treatment-related SAEs§	Boonen 2011, Van Meirhaeghe 2013	30 days	1.3% (2/149)	0.7% (1/151)	2.03 (0.19 to 22.12)
		12 mos.	1.3% (2/149)	NR	-----
		24 mos.	2.0% (3/149)	NR	-----
Withdrawal due to AEs	Wardlaw 2009, Boonen 2011	12 mos.	0.6% (1/149)	0.6% (1/151)	1.01 (0.06 to 16.05)
		24 mos.	0.6% (1/149)	0.6% (1/151)	1.01 (0.06 to 16.05)
Cement leakage, symptomatic**	Boonen 2011, Van Meirhaeghe 2013	30 days	3.4% (5/149)	N/A	N/A
		24 mos.	7.4% (11/149)	N/A	N/A
	Yi 2014	Mean 49.4 mos.	0% (0/79)	N/A	N/A
Cement leakage, asymptomatic††	Wardlaw 2009, Van Meirhaeghe 2013	12 mos.	32.2% (48/149)	N/A	N/A
	Yi 2014	Mean 49.4 mos.	1.8% (4/217 vertebra)	N/A	N/A
	Liu, 2019	NR	1.7% (1/58)	N/A	N/A
Any adverse event	Wardlaw 2009, Boonen 2011, Van Meirhaeghe 2013	30 days	63.1% (94/149)	36.4% (55/151)	1.73 (1.36 to 2.21)
		12 mos.	87.2% (130/149)	80.8% (122/151)	1.07 (0.98 to 1.19)
		24 mos.	89.9% (134/149)	88.7% (134/151)	1.01 (0.94 to 1.10)
	Liu 2019	NR	1.7% (1/58)	15.5% (9/58)	0.11 (0.01 to 0.85)

*For the FREE trial, Van Meirhaeghe 2013 provided 30-day data; Wardlaw 2009, 12-month data; and Boonen 2011, 24-month data. The FREE trial is considered fair quality and remaining trials are poor quality.

†None considered device or procedure related and included the following events that resulted in death through 24 months: cardiovascular event: 3.3% (5/149) vs. 3.3% (5/151); respiratory event: 0.6% (1/149) vs. 1.3% (2/151); cancer: 2% (3/149) vs. 1.3% (2/151); other (NR) event: 2% (3/149) vs. 1.3% (2/151).

‡Yi 2014 did not specify serious adverse events; the FREE trial (Wardlaw, et al) defined SAEs as death, life threatening injury, or permanent impairment, or required extended hospital stay or intervention to prevent impairment. Most SAEs were not attributed to the procedure.

§At 30 days (and 12 months): 2 patients experienced serious adverse events that were attributed to kyphoplasty – a soft tissue hematoma at the surgical site and a postoperative urinary tract infection that needed intervention; 1 serious event was attributed to UC – back pain. At 24 months, the same patient who had UTI by 12 months developed spondylitis near the cement that required treatment; in another patient there was anterior cement migration that likely caused a recurrent fracture.

**Possibly/probably causing symptomatic vertebral fracture; included under new vertebral fractures also.

††Two trials[187,197] reported that there were no cases of cement leakage into the spinal canal and one reported no cases of cement embolism.[187]

Table 28. KP vs. UC: Incidence of new vertebral fractures from RCTs

Author, year Quality	Outcome	F/U	KP % (n/N)	UC % (n/N)	RR (95% CI)
Boonen 2011, Van Meirhaeghe 2013 FREE trial Fair	**New radiographic vertebral fracture**	30 days	7.4% (11/149)	4.6% (7/151)	1.59 (0.63 to 4.00)
		3 mos.	22.0% (27/118)	27.0% (27/102)	0.86 (0.54 to 1.37)
		12 mos.	38.1% (45/118)	38.4% (38/102)	1.02 (0.73 to 1.44)
		24 mos.	47.5% (56/118)	44.1% (45/102)	1.08 (0.81 to 1.44)
	New radiographic index level vertebral fracture	24 mos.	4.2% (5/118)	10.8% (11/102)	0.39 (0.14 to 1.09)
	New radiographic adjacent vertebral fracture	24 mos.	23.7% (28/118)	16.7% (17/102)	1.42 (0.83 to 2.45)
	New clinical/symptomatic vertebral fracture	24 mos.	17.4% (26/149)	11.3% (17/151)	1.55 (0.88 to 2.74)
	New clinical/symptomatic vertebral fractures, possibly or probably related to cement	30 days	3.4% (5/149)	N/A	N/A
		24 mos.	7.4% (11/149)	N/A	N/A
Yi, 2014 Poor	**New clinical/symptomatic fractures**	Mean 49 mos.	6.3% (5/79)	14.0% (17/121)	0.45 (0.17 to 1.17)
Wardlaw, 2009, Boonen 2011 FREE trial Fair	**Reoperation, new clinical/symptomatic fractures**	3 mos.	4.0% (6/149)	NR	N/A
		6 mos.	6.0% (9/149)	NR	N/A
		24 mos.	8.1% (12/149)	4.0% (6/151)	2.03 (0.78 to 5.26)

CI = confidence interval; F/U = follow-up; KP = kyphoplasty; mos. = months; N/A = not applicable; NR = not reported; RR = risk ratio; UC = usual care.

Table 29. Adverse Events Other than Mortality from Comparative Database Studies Comparing Kyphoplasty versus Usual Care for Osteoporotic Vertebral Compression

Adverse Event	Database	Study Database search dates	N	Finding and conclusion
PE	Medicare	Ong, 2018 (2005-2014)	KP: 261,756 UC: 1,698,956 All patients*	Propensity-adjusted risk, p<0.001: 1 year: 7% lower with KP vs. UC 2 years: 5% lower with KP vs. UC 5, 8, 10 years: 3% lower with KP vs. UC
		Edidin 2015 (2005-2009)†	KP: 55,770 UC: 163,791 Propensity-score matched	4 years: Adj. HR 0.99 (95% CI 0.92 to 1.08)
DVT	Medicare	Ong, 2018 (2005-2014)	KP: 261,756 UC: 1,698,956 All patients*	Propensity-adjusted risk: 1, 2 years: 2% lower with KP vs. UC, p<0.01 5, 8, 10 years: 3% lower with KP vs. UC, p<0.001
		Edidin 2015 (2005-2009)†	KP: 55,770 UC: 163,791 Propensity-score matched	4 years: Adj. HR 0.92 (95% CI 0.0.87 to 0.96)
	NIS	Zampini, 2010 (years NR)	n's NR	Inpatient 0.2% (n=882 fractures) vs. 0.2% (n=4884 fractures), p=0.899
Cardiac complications	Medicare	Ong, 2018 (2005-2014)	KP: 261,756 UC: 1,698,956 All patients*	Propensity-adjusted risk, p<0.001: 1 year: 19% lower with KP vs. UC 2 years: 15% lower with KP vs. UC 5, 8, 10 years: 11% lower with KP vs. UC
		Edidin 2015 (2005-2009)†	KP: 55,770 UC: 163,791 Propensity-score matched	4 years: Adj. HR 0.88 (95% CI 0.83 to 0.93)
Pulmonary/ respiratory complications	Medicare	Ong, 2018 (2005-2014)	KP: 261,756 UC: 1,698,956 All patients*	Propensity-adjusted risk, p<0.001: 1 year: 2% lower with KP vs. UC 2, 5 years: 3% lower with KP vs. UC 8, 10 years: 4% lower with KP vs. UC
		Edidin 2015 (2005-2009)†	KP: 55,770 UC: 163,791 Propensity-score matched	4 years: Adj. HR 1.00 (95% CI 0.98 to 1.02)
Infection	Medicare	Ong, 2018 (2005-2014)	KP: 261,756 UC: 1,698,956 All patients*	Propensity-adjusted risk 1, 2 years: 2% lower with KP vs. UC, p<0.05 5 years: 1% lower with KP vs. UC, p<0.05 8,10 years: 0% difference
		Edidin 2015 (2005-2009)†	KP: 55,770 UC: 163,791 Propensity-score matched	4 years: Adj. HR 1.00 (95% CI 0.90 to 1.10)
	NIS	Zampini, 2010 (years NR)	n's NR	Inpatient 0.1% (n=882 fractures) vs. 0.1% (n=4884 fractures), p=0.929
Any SAE	NIS	Zampini, 2010 (years NR)	n's NR	Inpatient 1.7% (n=882 fractures) vs. 1.0% (n=4884 fractures), p=0.061

Subsequent augmentation	Medicare	Edidin 2015 (2005-2009)†	KP: 55,770 UC: 163,791 Propensity-score matched	4 years Subsequent augmentation: Adj. HR 12.5 (95% CI 12.5 to 14.3) Subsequent augmentation or Fusion: Adj. HR 12.5 (95% CI 12.5 to 14.3)

Adj. HR = adjusted hazard ratio; CI = confidence interval; DVT = deep vein thrombosis; KP = kyphoplasty; NIS = National Inpatient Sample; NR = not reported; PE = pulmonary embolism; SAE = serious adverse event; UC = usual care.
* Authors do a propensity score matched analysis but only provide data for the larger population; n's unclear for adjusted analyses.
† Data are for the OVCF cohort (osteoporotic and pathologic) which excludes the traumatic VCF patients.

4.2.2.2.2 Kyphoplasty versus Minimally Invasive Procedures

One NRSI was identified that met inclusion criteria and compared KP with pedicle screw fixation for the treatment of severe osteoporotic compression fractures and reported harms.[189] There were no deaths in either group and no difference between groups in the incidence of any adverse event to include DVT, new vertebral fracture and reoperation (Table 30). Cement leakage was reported in 30.1% of patients who received KP, but all were asymptomatic and there were no cases of symptomatic pulmonary embolism.

Table 30. Adverse events from one NRSI comparing KP with pedicle screw fixation.

Study	F/U	Adverse Event	KP % (n/N)	PSF % (n/N)	Adj HR/OR/RR (95% CI) Analysis
Wen, 2021	3 years	Morality	0% (0/376)	2.5% (0/121)	
		DVT	0% (0/376)	2.5% (3/121)	-
		Adjacent or distant new vertebral fracture	7.7% (29/376)	5.8% (7/121)	Unadjusted RR 1.33 (0.60 to 2.96)
		Cement leakage, asymptomatic	30.1% (113/376)	-	-
		Reoperation, any	7.7% (29/376)*	5.8% (7/121)	Unadjusted RR 1.33 (0.60 to 2.96)
		Reoperation, removal of device due to back pain	-	0.8% (1/121)	-

PSF = Adj HR = adjusted hazard ratio; CI = confidence interval; DVT = deep vein thrombosis; F/U = follow-up; KP = kyphoplasty; OR = odds ratio; PSF = pedicle screw fixation; RR = risk ratio.
* All for new fracture

Case series

Nine case series[15,17,47,106,124,137,156,202] evaluated adverse events following KP for the treatment of osteoporotic vertebral fractures. Sample sizes ranged from 263 to 1,752. Follow-up periods ranged broadly from peri-operative to 43 months. See Table 31 for AE details.

Mortality

Only one death (0.3%) was reported in one case series (N=297); it is unclear if it was procedure related or not.[15]

SAEs

Perioperative SAEs for KP were rare (≤1.0%) as reported by one case series (N=297)[15] and included allergic reaction to the balloon (severe hypotension and tachycardia) and subcutaneous hematoma requiring release.

New fractures

The incidence of new vertebral fracture after KP was as follows: any new vertebral fracture, range 12.1% to 22.2% across three studies;[17,106,124] any new symptomatic fracture, range from 8.1% to 10.6% across two studies;[15,47] any new adjacent fracture, range 4.6% to 10.5% across four studies;[106,124,156,202] and symptomatic adjacent fracture from 0.3% to 6.6% across two studies.[15,47] Only one study[124] reported refracture at the index level (0.7%, n=921). Longest follow-up was 43 months.

Cement leakage

The incidence of symptomatic cement leakage appears to be rare but was not commonly reported and ranged from 0% to 2.3% across two studies.[15,137] Any cement leakage (primarily asymptomatic) is common following KP.

Reoperation

Repeat KP for symptomatic fractures occurred in 8% and 10.6% of patents in two studies.[15,47]

Table 31. Adverse Events in Single Arm Studies of KP for Osteoporotic fractures

Adverse Event	Follow Up	Study	% (n/N)
KP			
Mortality			
Any	NR	Bergmann, 2012	0.3% (1/297)
SAE			
Allergic reaction to balloon	Intraoperative	Bergmann, 2012	0.3% (1/297)
Subcutaneous hematoma requiring release	Postoperative	Bergmann, 2012	1.0% (3/297)
New Fracture			
Any	1 year	Lin, 2017	22.2% (110/495)
	Mean 43 months	Ning, 2021	12.1% (111/921)
	≥1 year	Bian, 2022	21.7% (57/263)
Any symptomatic fracture	Mean 350 days	Deibert, 2016	10.6% (77/726)
	NR	Bergmann, 2012	8.1% (23/293)
Adjacent fracture	3.6 months	Spross, 2014	9.9% (37/375)
	1 year	Zhao, 2022	4.6% (80/1752)
	Mean 43 months	Ning, 2021	5.5% (51/921)
	1 year	Lin, 2017	10.5% (52/495)
Adjacent symptomatic fracture	NR	Bergmann, 2012	0.3% (1/293)
	Mean 350 days	Deibert, 2016	6.6% (48/726)
Refracture of index level	Mean 43 months	Ning, 2021	0.7% (6/921)
Symptomatic Cement Leakage			
Any	6 months	Qi, 2022	2.3% (21/896)[*]
	Intraoperative	Bergmann, 2012	0% (0/297)
Cement Leakage			
Any	1 year	Lin, 2017	20% (99/495)[†]

	1 year	Zhao, 2022	11.5% (202/1752)‡
	≥1 year	Bian, 2022	28.3% (105/371)
	6 months	Qi, 2022	6.3% (56/896)
	Intraoperative	Bergmann, 2012	40.1% (129/297)
Reoperation			
Repeat KP for symptomatic fracture	≥1 year	Deibert, 2016	10.6% (77/726)
	NR	Bergmann, 2012	8% (23/293)

F/U = follow-up; KP = kyphoplasty; NR = not reported; SAE = serious adverse event.

* An additional 35 had asymptomatic bone cement displacement

† Calculated using the totals for cement leakage in the different fracture groups.

‡ Calculated by combining adjacent and non-adjacent fracture groups.

4.2.3 KQ3 Differential Effectiveness

4.2.3.1 Vertebroplasty

Evidence on differential effectiveness or harms of VP reported in included RCTs for subpopulations defined by gender, age, psychological or psychosocial co-morbidities, provider characteristics, or payer type or by fracture age, pain duration or intervention characteristics is sparse. In addition to data from three trials of VP, [24,50,89] results discussed below include those from an AHRQ review[38] that reported stratified analyses for VP. These analyses included all but one of the RCTs of VP versus sham or usual care that are included in this HTA update. No RCTs of KP reported stratified analyses for subpopulations. One RCT comparing VP with KP briefly described such analyses. [61] Analyses in all trials were likely to have low power for detecting effect modification by factors that were evaluated. Confidence in findings from stratified analyses from included studies is very low.

Key Points

- ***VP versus Sham or Usual Care***
    - Fracture age/pain duration: There does not appear to be modification of the treatment effect for vertebroplasty (versus sham) in patients with acute osteoporotic fractures, compared with those with more chronic fractures, based on reported subgroup analyses from included RCTs or from reported stratified analyses of RCTs comparing VP to a combined UC and sham across RCTs for the outcomes of pain or function.
    - Other factors:
        - No modification of treatment effect based on sex, presence or absence of pervious fractures or treating center were reported by one RCT of VP versus sham
        - For the outcomes of pain and function, there appears to be no modification of treatment effect based for the following subgroups based on stratified analysis of RCTs comparing VP with sham or usual care: PMMA volume, study enrollment requirement of MRI findings of bone marrow edema
        - In stratified analysis of RCTs of VP, control type appeared to modify treatment at 2-4 weeks with a smaller difference in effect size for pain observed in trials with sham control versus usual care as a control. Interaction between control types was not statistically significant for function.
- ***VP versus KP***

- One RCT comparing VP versus KP reported that no appreciable differences in the magnitude of pain reduction were seen for subgroup analysis on sex, age, preoperative pain scores or preoperative RDQ scores. Authors do not provide data or p-values for interaction.

4.2.3.1.1 Vertebroplasty versus Sham or Placebo

Detailed analysis

Fracture age or pain duration

Three RCTs comparing VP with sham reported post-hoc stratified analyses based on pain or fracture duration.[24,50,89] Across two of the RCTs of VP versus sham, baseline pain duration did not modify treatment effect for pain.[24,89] One RCT of VP[24] reported no modification of treatment effect based on duration of symptoms when stratified by ≤6 weeks vs. ≥ 6 weeks or as a continuous variable. Data were not presented. Another trial of VP versus sham[89] reported similar treatment effects for pain (0 to 10 scale) at one month across three pain duration categories (p = 0.58 across the three groups). Results by pain duration categories were imprecise: <13 weeks duration, MD -0.8 scale, 95% CI, -2.5 to 0.8), 14 to 26 weeks duration, MD -1.3, 95% CI, -3.4 to 0.8) and 7 to 52 weeks duration (MD 0.0, 95% CI, -1.6 to 1.7).

One RCT of VP versus sham (N= 120) [41] reported no modification of treatment effect (p for interaction =0.12) based on risk differences (RD) for patients achieving an NRS (0-10 scale) of <4 was based on fracture age of ≤3 weeks (RD 31, 95%CI 12 to 50) >3 weeks (RD -4, 95%CI -39 to 31). Exploratory subanalysis from this trial of patients (N=85) with a fracture age of ≤3 weeks [50] suggests clinically important improvements in pain and function to patients with osteoporosis when vertebroplasty is performed within 3 weeks of facture to 6 months, (APPENDIX Table P2) however differential efficacy (effect modification) by fracture age cannot be assessed as the analysis does not compare with data in patients with older fractures. Estimates across follow-up times are imprecise.

Stratified analyses from a recent AHRQ review[37] across 10 RCTs (N=1093), comparing VP to sham or usual care control (most of which are included in this HTA update) found no statistically significant interaction at for subgroups of baseline pain duration by study inclusion criteria based on pain for the outcome of pain. Similarly, for the outcome of function, no statistically significant interaction for either of these pain duration subgroups was observed. **Table X** At 2 to 4 weeks, a significant interaction based on control type (sham, usual care) was observed for pain. Effect sizes for VAS pain (0-10 scale) were substantially smaller for VP versus sham (5 RCTs, N= 536, MD -0.57, 95% CI -1.09 to -0.05, I^2=0%), compared with VP versus usual care (3 RCTs, N=382, MD -2.27 95% CI -3.20 to -0.94, I^2=0%) with an interaction p-value or 0.01. In general, estimates were imprecise across factors. Additional data are found in the AHRQ Report.

Other factors

One RCT of VP [24] reported no modification of treatment effect based on sex, presence or absence of pervious fractures or treating center (p> 0.10 for all tests of interaction). Data were not presented.

Stratified analyses updates from the recent AHRQ review[37] comparing VP to sham or usual care control found no statistically significant interaction at any time frame for subgroups-based inclusion criteria

requiring MRI evidence of bone marrow edema, PMMA volume or study quality on pain. Similarly, no statistically significant interaction was seen at any time frame for subgroups based on study quality or use of the original RDQ (-0-24 scale) vs. modified RDQ (0-23 scale) for function.

Control type (sham or usual care)

In general across pooled analyses of RCTs included in this HTA update, effect estimates for pain and function were smaller for VP versus sham than for VP versus usual care at all follow-up times. In stratified analyses from the AHRQ report,[38] a significant interaction based on control type (sham, usual care) was observed for pain at 2 to 4 weeks but not at other time frames. Effect sizes for VAS pain (0-10 scale) were substantially smaller for VP versus sham (5 RCTs, N= 536, MD -0.57, 95% CI -1.09 to -0.05, I^2=0%), compared with VP versus usual care (3 RCTs, N=382, MD -2.27 95% CI -3.20 to -0.94, I^2=0%) with an interaction p-value or 0.01. In general, estimates were imprecise across factors. Interaction between control types at other times was not statistically significant for function. Additional data are found in the AHRQ Report.

4.2.3.1.2 Vertebroplasty versus Kyphoplasty

One RCT comparing VP versus KP [61] provided limited information on subgroup analysis, reporting only that no appreciable difference in the magnitude of pain reduction were seen for subgroup analysis on sex, age (<75 years vs. ≥75 years), preoperative pain scores (<7 vs. ≥7 on 0 to 10 scale) or preoperative RDQ scores (<17 vs. ≥17, 0-24 scale). Authors do not provide data or p-values for interaction.

4.3 Vertebral Fractures Due to Malignancies or Tumors

4.3.1 KQ1 Effectiveness

4.3.1.1 Vertebroplasty

A total of 16 studies[10,14,21,27,43,69,74,98,102,119-121,141,154,193,194] were identified that met inclusion criteria and reported on VP or KP for patients with malignant fractures. One RCT[14] compared KP with UC. The remaining studies – three comparative NRSIs (2 from the prior report,[69,98] 1 newly identified[10]), one recent systematic review,[154] and four case series[74,121,141,194] not included the SR – that evaluated the effectiveness of PV vs. KP or of PV or KP for treatment of VCFs due to cancer were identified. In addition to the above studies for efficacy/effectiveness, three systematic reviews[21,102,119] from the prior report and four case series[27,43,120,193] provided information on safety specifically in this population

No RCTs were identified comparing VP with KP for the treatment of vertebral fractures due to malignancy or tumor.

Three comparative NRSIs (2 from the prior report,[69,98] 1 newly identified[10]), one recent systematic review,[154] and four case series[74,121,141,194] not included in the SR that evaluated the effectiveness of PV and KP for malignant vertebral fractures were included. Many of the studies included in the SR did not meet our inclusion criteria (e.g., case reports, not specifically treating vertebral fractures, ineligible comparators) but given the lack of high-quality evidence for this population we included it for completeness.

Comparative NRSIs

Three retrospective NRSIs[10,69,98] (N range, 34 to 342) compared VP with KP for the treatment of vertebral compression fractures due to malignancy. One study was conducted in the U.S,[69] one in the Republic of Korea[10] and one in Turkey;[98] funding was not reported. Median patient age ranged from 61 to 64 years and 45% to 53% were female. The most common primary cancers across the studies were multiple myeloma (100% in one study),[98] lung and breast. Only one study reported the duration of spinal pain (or fracture age) which was 3.2 months.[69] All populations had a mix of single and multiple level fractures, with as many as six fractures undergoing treatment. Further details of the study populations, treatments, and inclusion and exclusion criteria can be found in Appendix Tables H3, I4, and I5.

4.3.1.1.1.1 Primary Outcomes

Pain

VP and KP were associated with a similar likelihood of achieving pain response and similar improvement in VAS pain scores at all timepoints across two NRSIs (Table 32).[10,69] The third NRSI[98] reported an aggregate of mean VAS pain scores during five activities of daily living (pain at rest, walking, sitting-standing, taking a shower and wearing clothes) and found that VP was associated with less improvement compared with KP at 6 and 12 months (there was no difference between groups at 6 weeks).

Table 32. Retrospective Comparative NRSI

Outcome*	Author, year	F/U	VP Mean (SD) or % (n/N)	KP Mean (SD) or % (n/N)	Effect Size (95% CI)
Pain Response					
Responders VAS (score ≥3)	Bae 2016 (N=342)	NR*	62% (148/238)	57% (59/104)	RR 1.10 (0.90 to 1.33)
Complete or improved pain relief†	Fourney 2003 (N=49)	24 hours	86% (30/35 sessions)	80% (12/15 sessions)	RR 1.07 (0.80 to 1.43)
Complete pain relief			23% (8/35 sessions)	7.0% (1/15 sessions)	RR 3.43 (0.47 to 25.06)
Improved pain relief			63% (22/35 sessions)	73% (11/15 sessions)	RR 0.86 (0.58 to 1.28)
Pain Improvement					
VAS pain (0-10 scale)	Bae 2016 (N=342)	NR*	2.5 (1.8) (n=238)	2.8 (2.1) (n=104)	MD -0.30 (-0.74 to 0.14)
	Fourney 2003 (N=49)	1 month	median 2 (NR) (n=34)	median 2.5 (NR) (n=15)	p=NS for all
		3 months	median 2 (NR) (n=34)	median 2.5 (NR) (n=15)	
		6 months	median 2 (NR) (n=34)	median 4 (NR) (n=15)	
		12 months	median 1 (NR) (n=34)	median 2 (NR) (n=15)	
VAS pain (0-50 scale)‡	Köse 2006 (N=34)	6 weeks	15.3 (4.1) (n=16)	12.1 (3.6) (n=18)	MD 3.2 (0.51 to 5.89)
		6 months	12.2 (3.0) (n=16)	8.6 (2.3) (n=18)	MD 3.6 (1.74 to 5.46)
		12 months	13.5 (2.9) (n=16)	9.7 (2.4) (n=18)	MD 3.8 (1.95 to 5.65)

CI = confidence interval; F/U = follow-up; KP = kyphoplasty; MD = mean difference; NR = not reported; NS = not significant; RR = risk ratio; SD = standard deviation; VAS = visual analogue scale; VP = vertebroplasty.

* Time of discharge or first follow-up visit.

† Refers to an analysis of documented VAS pain scores within first 24 hours.

‡ Average of pain during 5 activities of daily living: pain at rest, walking, sitting-standing, taking a shower, and wearing clothes.

Systematic review and case series

One poor-quality SR published in 2019[154] included two RCTs, 60 cohorts (16 prospective, 44 retrospective) and 25 case series/case reports (N=3,426) in patients with vertebral compression fractures due to malignancy and summarized information on pain and function. The authors divided each article into two groups and pooled results for VP and KP separately. A total of 2,091 patients were treated with VP and 1,335 with KP. The weighted mean age was 63 years, 51% of patients were female and the mean number of treated levels per patient was 2.2. The most common primary malignancies were multiple myeloma (36%), lung (19%), and breast (19%).

Four case series not included in the recent SR evaluated VP (2 studies)[121,141] and KP (2 studies)[74,194] for the treatment of vertebral fractures due to cancer. Sample sizes ranged from 44 to 92, mean patient age

from 57 to 68 years and proportion female from 47% to 67%. The most common primary cancers were multiple myeloma, breast, lung, and prostate. Many of the patients had multiple fractures; one case series[121] enrolled only patients undergoing treatment for 6 or more fractures (up to 13).

Results (Function, Pain, Quality of Life, Opioid Use)

The SR reported that both VP and KP resulted in similar, clinically relevant improvements in pain, ODI, and KPS across all follow-up intervals (Table X4). Measures of variance and p-values were not reported.

Consistent with the findings from the SR, across all four case series,[74,121,141,194] patients who received VP or KP experienced significant improvement in VAS/NRS pain scores compared with baseline over various follow-up times up to 24 months (Table X4). Results for function, opioid use and quality of life also showed significant improvement following both VP and KP across follow-up times, with the exception of KPS scores (all time points) and opioid use (MEDD) after 1 month in one study,[141] which showed some improvement following VP but the difference was not statistically significant compared with baseline. This study enrolled only patients with multiple myeloma

Table 33. Effectiveness outcomes from one SR and four case-series evaluating VP or KP for vertebral fracture due to malignancy

Author, year	Outcome*	F/U	VP Mean (SD) or % (n/N)	KP Mean (SD) or % (n/N)	Effect Size (95% CI)
Pain					
Sorensen 2019 SR (N=3,426; 87 studies)	VAS (0-10 scale) VP: 35 studies KP: 21 studies	Baseline	7.48 (NR) (n=1,445)	7.05 (NR) (n=1,103)	p=NR
		<4 weeks	3.00 (NR) (n=1,147)	2.96 (NR) (n=814)	
		≤6 weeks	2.90 (NR) (n=606)	2.99 (NR) (n=222)	
		<6 months	2.50 (NR) (n=370)	3.12 (NR) (n=318)	
		< 12 months	2.85 (NR) (n=784)	3.55 (NR) (n=204)	
		≥12 months	2.98 (NR) (n=260)	3.09 (NR) (n=375)	
Moulin 2020 (N=50)†	NRS pain (0-10 scale)	Baseline	5.0 (NR) (1.8)	NA	p<0.001 compared with baseline
		1 month	1.7 (NR) (1.4)	NA	
Rocha Romero 2020 (N=44)	NRS pain (0-10 scale)	Baseline	5.16 (NR)	NA	p<0.001 for all compared with baseline
		1 month	1.07 (95% CI 1.00-1.14)	NA	
		3 months	1.48 (95% CI 1.40-1.56)	NA	
		12 months	1.77 (95% CI 1.70-1.84)	NA	
		24 months	1.68 (95% CI 1.59-1.77)	NA	
Wu 2023 (N=92)	VAS pain (0-10 scale)	Baseline	NA	6.3 (2.0)	p<0.001 for all compared with baseline
		3 days	NA	3.3 (1.5)	
		1 month	NA	2.3 (1.1)	
		3 months	NA	2.8 (1.2)	
		12 months	NA	3.4 (1.1)	

Garcia-Maroto 2015 (N=75)	VAS pain (0-10 scale)	Baseline	NA	7.49 (1.19)	p<0.05 compared with baseline
		9-12 months	NA	3.21 (0.95)	
Function					
Sorensen 2019 SR (N=3,426; 87 studies)	ODI (0-100 scale) VP: 5 studies KP: 13 studies	Baseline	74.68 (NR) (n=226)	66.02 (NR) (n=592)	p=NR
		<4 weeks	17.73 (NR) (n=190)	34.73 (NR) (n=275)	
		≤6 weeks	32.25 (NR) (n=67)	38.54 (NR) (n=156)	
		<6 months	31.68 (NR) (n=67)	37.35 (NR) (n=381)	
		< 12 months	29.88 (NR) (n=81)	30.16 (NR) (n=162)	
		≥12 months	28.93 (NR) (n=103)	32.45 (NR) (n=301)	
	KPS (0-100 scale) VP and KP: 8 studies	Baseline	66.99 (NR) (n=611)		p=NR
		<4 weeks	80.28 (NR) (n=609)		
		≤6 weeks	83.11 (NR) (n=263)		
		<6 months	83.92 (NR) (n=263)		
		< 12 months	82.02 (NR) (n=265)		
		≥12 months	79.08 (NR) (n=110)		
Rocha Romero 2020 (N=44)	KPS (0-100)	Baseline	78.6 (NR)	NA	p=NS for all compared with baseline
		1 month	78.0 (95% CI 77.8-78.2)	NA	
		3 months	76.7 (95% CI 76.4-77.0)	NA	
		12 months	77.2 (95% CI 76.9-77.5)	NA	
		24 months	77.9 (95% CI 77.6-78.2)	NA	
Wu 2023 (N=92)	ODI (0-100 scale)	Baseline	NA	70.9 (7.1)	p<0.001 for all compared with baseline
		3 days	NA	31.4 (4.7)	
		1 month	NA	31.2 (3.5)	
		3 months	NA	31.2 (3.5)	
		12 months	NA	30.4 (3.2)	
Garcia-Maroto 2015 (N=75)	KPS (0-100)	Baseline	NA	60.2 (10)	p=0.03 compared with baseline
		9-12 months	NA	80.7 (12.1)	
Opioids					
Moulin 2020 (N=50)[†]	Opioid consumption (mean, mg/d)	Baseline	76 (42)	NA	p=0.0003 compared with baseline
		1 month	45 (38)	NA	
Rocha Romero 2020 (N=44)	MEDD (mean, mg)	Baseline	33.4 (NR)	NA	p<0.001 for 1 month; p=NS for 3-24 months compared with baseline
		1 month	24.0 (95% CI 23.1-24.9)	NA	
		3 months	29.4 (95% CI 28.2-30.6)	NA	
		12 months	28.2 (95% CI 26.8-29.6)	NA	

		24 months	21.0 (95% CI 19.7-22.3)	NA	
Garcia-Maroto 2015 (N=75)	Major opioid use	Baseline	NA	53% (40/75)	p<0.001 compared with baseline
		12 months	NA	12% (9/75)	
Quality of Life					
Wu 2023 (N=92)	SF-36 (scale NR)	Baseline	NA	89.7 (16.1)	p<0.001 compared with baseline
		12 months	NA	99.5 (19.7)	

CI = confidence interval; F/U = follow-up; KP = kyphoplasty; KPS = Karnofsky Performance Status index; MEDD = morphine equivalent daily dose; NA = not applicable; NR = not reported; NRS = numerical rating scale; NS = not significant; ODI = Oswestry Disability Index; SD = standard deviation; SF-36 = 36-item Short Form Questionnaire; SR = systematic review; VAS = visual analogue scale; VP = vertebroplasty

* Lower score is better for all outcomes except KPS and SF-36 for which a higher score is better.

† Enrolled patients with ≥6 fractures treated simultaneously.

4.3.1.2 Kyphoplasty

4.3.1.2.1 Kyphoplasty versus Usual Care

One RCT (CAFE trial) (N=134)[14] compared KP to UC for the treatment of symptomatic vertebral compression fractures due to malignancy (see Appendix Table F3). The trial was conducted across 22 sites in the U.S., Canada, Europe, and Australia and was funded by industry. Mean patient age was 64 years, 58% were female, and the majority were White (88%). Most patients had multiple fractures (3: 31%; 2: 29%; 1: 39%) and more patients randomized to KP (38%) versus UC (23%) had three fractures. Median fracture age was 3.4 months. Primary cancer types included multiple myeloma (38%), breast (22%) and other (26%; colon, ovarian, esophageal, and bladder cancer) and cancer was considered stable in 38% of patients, progressive in 36% and in remission in 8%. Previous treatments before enrollment included spine radiation (21%), bone radiation (16%), surgeries (51%), chemotherapy/hormonal therapy (67%), and steroids (35%), with some differences between groups observed at baseline. Balloon KP was performed using standard techniques and PMMA cement (volume not reported). Patients in both treatment groups received UC, which could include analgesics, bed rest, bracing, physiotherapy, rehabilitation programs, walking aids, radiation treatment, and other antitumor therapies at the discretion of the treating physician. Patients also received treatment for concurrent osteoporosis or bone metastases as needed.

After 1 month, patients in the UC group were offered KP. Given the high rate of cross-over (59%), we focused our efficacy analyses on outcomes up to 1 month (total follow-up period was 12 months). For harms (see Key Question 2), all patients were analyzed, both as randomized and as treated.

The trial was rated as fair (up to one month). Major limitations included differences between groups at baseline, and the lack of blinding.

No comparative observational studies meeting our inclusion criteria were identified.

4.3.1.2.1.1 Primary Outcomes

Pain

Kyphoplasty was associated with a large improvement in NRS pain scores (scale 0-10) at 1 week (N=117, MD -3.50, 95% CI -4.27 to -2.73) and 1 month (N=114, MD -3.50, 95% CI -4.37 to -2.63) compared with usual care (see Table 34).[14] The authors did not report the proportion of patients considered pain responders, defined as a decrease of 1 to 2.5 points on the NRS, but stated that KP patients improved by a mean of 3.8 points (p<0.05) at 1 week, while usual care patients showed no significant improvement

Function

At 1 month, KP was associated with a large increase in the likelihood of achieving function response on both the RDQ, defined as a ≥2-point improvement (N=113, 80.9% vs. 28%, RR 2.89, 95% CI 1.82 to 4.58), and the Karnofsky Performance Status (KPS) scale, defined as a ≥10-point improvement (N=112, 65.1% vs. 26.5%, RR 2.45, 95% CI 1.49 to 4.04), compared with usual care (Table 34).[14] Significantly more patients who received KP had a KPS score of ≥70 (clinically meaningful ability to care for oneself) at 1 month compared with usual care. KP was also associated with a large improvement in RDQ scores (0-24 scale, N=113, MD -8.90, 95% CI -9.49 to -8.31) and KPS scale scores (0-100 scale, N=112, MD 14.5, 95% CI 12.83 to 16.17) compared with usual care.

4.3.1.2.1.2 Secondary Outcomes

Quality of life

KP was associated with a large improvement in quality-of-life scores as measured by the SF-36 PCS (0-100 scale, N=105, MD 8.0, 95% CI 7.18 to 8.82) and MCS (0-100 scale, N=105, MD 10.0, 95% CI 8.74 to 11.26), Table 34.[14] The authors did not report the proportion of patients considered responders, defined as an increase of 3.5 to 4.3 points on the SF-36 PCS, but stated that KP patients improved by a mean of 9.4 points (p<0.05) at 1 month, while usual care patients showed no change.

Table 34. Summary of efficacy results: Kyphoplasty versus Usual Care in Patients with Fractures due to Tumors and Malignancies from the CAFE Trial (Berenson, 2011)

Outcome*	F/U	KP Mean (SD) or % (n/N)	Usual Care Mean (SD) or % (n/N)	Effect Size (95% CI)
Pain				
NRS Pain (0-10; worse)	1 week	3.5 (2.4)† (n=63)	7.0 (1.7)† (n=54)	MD -3.50 (-4.27 to -2.73)‡
	1 month	3.3 (2.9)† (n=64)	6.8 (1.4)† (n=50)	MD -3.50 (-4.37 to -2.63)‡
Function				
Responders RDQ (≥2 points)	1 month	80.9% (51/63)	28% (14/50)	RR 2.89 (1.82 to 4.58)‡
Responders KPS (≥10 points)	1 month	65.1% (41/63)	26.5% (13/49)	RR 2.45 (1.49 to 4.04)‡
Proportion with KPS score ≥70 (ability to care for oneself)	1 month	74.6% (47/63)	38.8% (19/49)	RR 1.92 (1.32 to 2.81)‡
RDQ (0-24; worse)	1 month	9.1 (1.9)§ (n=63)	18.0 (1.0)§ (n=50)	MD -8.9 (-9.49 to -8.31)‡
KPS (0-100; better)	1 month	73.0 (4.5)§ (n=63)	58.5 (4.5)§ (n=49)	MD 14.5 (12.83 to 16.17)‡

Quality of Life				
SF-36 PCS (0-100; better)	1 month	35 (2.5)† (n=58)	27 (1.5)† (n=47)	MD 8.0 (7.18 to 8.82)‡
SF-36 MCS (0-100; better)	1 month	46.5 (3.0)† (n=58)	36.5 (3.5)† (n=47)	MD 10.0 (8.74 to 11.26)‡

AE = Adverse event; CI = Confidence interval; KP = Kyphoplasty; KPS = Karnofsky Performance Status; MCID = Minimally clinically important difference; MCS = Mental component scale; MD = Mean difference; NA = Not applicable; NC = Not calculable; NR = Not reported; NRS = Numerical Rating Scale; NS = Not significant; PCS = Physical component scale; RDQ = Roland Morris Disability Questionnaire; RR = Risk ratio; SF-36 = 36 Item Short-Form Survey.
* All outcomes are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (KP); with the exception of the KPS.
† Mean estimated from graphs, SD calculated from 95% CIs.
‡ Calculated from Means and SDs (MDs, 95% CI), or proportions (RR, 95% CI).
§ Estimated from graphs.

4.3.2 KQ2 Harms and Safety

4.3.2.1 Vertebroplasty

4.3.2.1.1 Vertebroplasty versus Kyphoplasty

A total of two comparative NRSIs,[69,98] one recent systematic review[154] and three reviews[21,102,119] from the prior report, and eight case series[27,43,74,120,121,141,193,194] provided information on safety specifically in this population following VP and KP.

Comparative studies

Two of the three retrospective NRSIs included for effectiveness reported safety outcomes; samples sizes were very small (N range, 34 to 49).[69,98] Overall, the incidence of adverse events was low and occurred with similar frequency between treatments (VP vs. KP), except for asymptomatic cement leakage and reoperation which occurred more frequently following VP vs. KP in one study (Table 35).

Table 35. Adverse Events in Comparative NRSIs evaluating VP vs. KP for treatment of Malignant vertebral Fractures

Adverse Event	Study	Follow Up	VP % (n/N)	KP % (n/N)
Mortality				
	Fourney 2003	30 days	0% (0/34)	0% (0/15)
		2.5 months	2.9% (1/34)	0% (0/15)
SAE				
Neurological complications	Kose 2006	12 mos.	0% (0/16)	0% (0/18)
Pulmonary complications		12 mos.	0% (0/16)	0% (0/18)
Paraplegia due to metastasis	Fourney 2003	30 days	2.9% (1/34)	0% (0/15)
Readmission for CHF		30 days	0% (0/34)	6.7% (1/15)
Device-or procedure related complication				
Balloon rupture (asx)	Kose 2006	12 mos.	n/a	5.6% (1/18)
Any	Fourney 2003	30 days	0% (0/34)	0% (0/15)
New Fractures				
Adjacent level fracture	Kose 2006	12 mos.	0% (0/16)	0% (0/18)
Symptomatic fracture requiring reoperation				
Cement leakage				

Asymptomatic leakage	Fourney 2003	30 days	9.2% (6/65 levels)	0% (0/32 levels)
Reoperation				
Repeat VP or KP	Fourney 2003	4.5 months	2.9% (1/34)	6.7% (1/15)
Subsequent spinal surgery			5.9% (2/34)	0% (0/15)

asx = asymptomatic; CHF = congestive heart failure; F/U = follow-up; KP = kyphoplasty; mos. = months; n/a = not applicable; SAE = serious adverse event; VP = vertebroplasty.

Systematic review and case series

One poor-quality SR published in 2019 included for effectiveness and three SRs from the prior report summarize data on any complication, new vertebral fractures and cement leakage. Data from these SRs of non-randomized studies were from a combination of case series and cohort studies which were pooled to provide summary estimates. Data from included reviews is summarized in Table 37.

Eight case series not included in the SRs were identified that met inclusion criteria and evaluated safety outcomes following VP (3 studies),[43,121,141] KP (4 studies)[74,120,193,194] or either VP or KP (1 study)[27] for vertebral fractures due to malignancy (Table 36).

New fracture

The incidence of any new vertebral fracture was similar after VP and KP in one SR (18% vs. 17% of patients); when considered out of the number of levels treated, the VP arms had a slightly higher incidence of new fracture (21% vs. 13%). A second SR that included only studies evaluating KP, reported an overall fracture rate of 10.2%. Across the case series not included in the SRs, the incidence of any new vertebral fracture ranged from 10.0% to 29.5% across four studies; incidence ranged from 10% to 29.5% by 12 months in two studies of VP only[121,141] and was 24.6% in one study of VP or KP[27] and 14.7% in one study of KP only.[74] New adjacent vertebral fractures occurred in 15.9% (1 study of VP) to 17.6% (1 study of VP/KP) over 2 to 12 months.[27,141]

Cement leakage

Cement leakage was more common with VP (37.9%-79.0%) than KP (6.0%-13.6%). Differences across studies in the extent of surveillance and reporting of leakage may influence the range of rates. Symptomatic leakage was uncommon for both procedures, ranging from 0.26% to 3.1% for VP and 0% for KP. Across the case series not included in the SRs, one study[27] reported four cases (1.0%) of symptomatic epidural extravasation at 2 months following VP or KP and another study[193] reported no cases of symptomatic cement leakage.

Reoperation

One case series[27] reported any reoperation in 19.2% of patients by 2 months following VP or KP but the reason is unclear. This same study reported two cases (0.5%) reoperation for symptomatic cement leakage causing lower extremity weakness and another study reported that 6.0% of patients had repeat procedures to treat new fractures at 1 month.[121]

Mortality

Across four case series,[74,120,121,193] the incidence of mortality ranged from 0% (by 3 months) to 34.0% (at a mean 401 days); excluding the study that reported no deaths, the risk ranged from 9.3% (by 9 months) to 34.0%.

Any SAE

The incidence of any symptomatic complication was rare as reported by the most recent SR and occurred with similar frequency between VP (1.7%) and KP (0.8%).[154] A second SR reported four cases of neurological complications (not further specified) in the VP groups compared with none in the KP groups (4.1% vs. 0%).[119] Across the case series not included in the SRs, the incidence of individual SAEs was low and ranged from 0% to 2.0% across four studies.[27,121,193,194] In one study[121] there was one case (2.0%) each of symptomatic cement PE, lumbar artery puncture and retroperitoneal hematoma requiring blood transfusion during or up to 1 month after VP.

Table 36. Adverse Events in Single Arm Studies for Malignant Fracture

Adverse Event	Follow Up	Study	% (n/N)
Mixed VP/KP			
SAE			
Increase in pain and radicular symptoms, treated conservatively	2 months	Burton, 2011*	0.5% (2/407)
Vertebral body infection	2 months	Burton, 2011	0.2% (/407)
New Fracture			
Any	2 months	Burton, 2011	24.6% (100/407)
Adjacent Fracture	2 months	Burton, 2011	17.6% (72/407)
Cement Leakage			
Any	2 months	Burton, 2011	33.7% (134/407)
Symptomatic epidural extravasation	2 months	Burton, 2011	1.0% (4/407)
Reoperation			
Any	2 months	Burton, 2011	19.2% (78/407)
Weakness due to compression from symptomatic epidural extravasation	2 months	Burton, 2011	0.5% (2/407)
VP			
Mortality			
Any	100 days	Moulin, 2020	10.0% (5/50)†
	Mean 401 days	Moulin, 2020	34.0% (17/50)†
SAE			
Any grade 4/5	1 month	Moulin, 2020	0% (0/50)
Symptomatic Cement PE	1 month	Moulin, 2020	2.0% (1/50)
Lumbar artery puncture	Perioperative	Moulin, 2020	2.0% (1/50)
Retroperitoneal hematoma requiring blood transfusion	Perioperative	Moulin, 2020	2.0% (1/50)
New Fracture			
Any	1 month	Moulin, 2020	10.0% (5/50)
	1 year	Rocha Romero, 2020	29.5% (13/44)
Adjacent level	1 year	Rocha Romero, 2020	15.9% (7/44)
Cement Leakage			
Any	NR	Cui, 2022	34.9% (185/530)
Reoperation			
For new fracture	1 month	Moulin, 2020	6.0% (3/50)
KP			
Mortality			
Any	3 months	Garcia-Maroto, 2015	1.3% (1/75)
	9 months	Garcia-Maroto, 2015	9.3% (7/75)
	3 months	Molloy, 2016	0% (0/158)

	1 year	Wu, 2022	18.8% (22/117)
SAE			
Any	1 year	Wu, 2022	0% (0/117)
	1 year	Wu, 2023	0% (0/92)
New Fracture			
Any	12 months	Garcia-Maroto, 2015	14.7% (11/75)
Cement Leakage			
Symptomatic	1 year	Wu, 2022	0% (0/215)
Any	Mean 11 months	Garcia-Maroto, 2015	5.7% (7/122)
	1 year	Wu, 2022	13.0% (28/215 procedures)
	1 year	Wu, 2023	13.0% (12/92)
Other AE			
Any	1 year	Wu, 2022	74.3% (87/117)

AE = adverse event; KP = kyphoplasty; NR = not reported; SAE = serious adverse event; VP = vertebroplasty.

* Burton 2011 includes a mixed population, but primarily malignancy 65% versus 35% OVCF.

† No deaths in Moulin 2020 were related to treatment.

Table 37. Summary of pooled estimates of cement leakage, new fractures, and any complications reported in systematic reviews of comparative studies and case series evaluating VP and KP for treatment of pathologic vertebral fractures

Author (year)	Number of studies	Any Cement leak*		Symptomatic Cement Leak*		Any new Fracture*		Any complications*	
		VP	KP	VP	KP	VP	KP	VP	KP
Pathologic fractures†									
Sorensen (2019)	•VP N=62 (11 prospective) •KP N = 27 (7 prospective)	•37.9% (439/1157)‡	•13.6% (28/206)‡	NA	NA	NA	NA	•1.7% (35/2024)‡§	•0.8% (8/909)‡§
Lee (2009)**	•PV N = 13 (1 prospective) •KP N = 7 (2 prospective)	All studies (per level) •79.07% (601/760 levels)	All studies (per level) •6.07% (13/214 levels)	All studies (per level) •0.26% (21/760 levels)	All studies (per level) •0.0% (0/214 levels)	All studies •18.0% (490/2781 pts) •21% (830/3912 levels) Prospective studies •18.1% (122/672 pts) •16.3% (154/941 levels)	All studies •17.0% (123/727 pts) •13.0% (158/1192 levels) Prospective studies •16.1% (11/68 pts) •11.2% (12/107 levels)	NA	NA
Bouza (2009)**	•KP Only •N = 7 studies (4 prospective)	NA	All studies •5.8% (1.96, 9.64%) (41 leaks, presume levels reported), Prospective studies	NA	•0.0%	•NA	•10.23% (95% CI, 2.8%, 17.7%) (21/172 patients)	NA	NA

			•11.2% • Retrospective studies •0.51%						
Mendel (2009)**	•PV N = 5 prospective •KP N = 6 prospective	Prospective studies •58.4% (59/101 levels)	Prospective studies •12.1% 12/2391 levels)	Prospective studies •3.1% (3/98 patients)	Prospective studies •0%	Prospective studies NR	Prospective studies NR	4.1% (4/98 patients)††	0%

KP = kyphoplasty; NA = not applicable; VP = vertebroplasty.

* Authors may report rate per number of patients or number of levels treated (level) or number of vertebrae as noted in the table.

† Pathologic fractures may include multiple myeloma, hemangioma or metastases.

‡ We assume that the denominator is the total n they report in table 1.

§ Included radiating pain, transient chest pain, radiculopathy without palsy, hemothorax, hematoma, radicular neuritis, asymptomatic and symptomatic pulmonary embolisms, bilateral leg motor deficits, cauda equina, and complete paraplegia.

** From prior report.

†† All neurological, not specified.

4.3.2.2 Kyphoplasty

4.3.2.2.1 Kyphoplasty versus Usual Care

One RCT (CAFE trial) (N=134)[14] that compared KP to UC for the treatment of symptomatic vertebral compression fractures due to malignancy reported adverse events. Details of this trial have been reported in Key Question 1 (also see Appendix Table F3). Given the high rate of cross-over (59%) from UC to KP after 1 month, we reported harms for patients both as randomized and as treated.

New vertebral fracture

KP was associated with a similar risk of new symptomatic vertebral fractures compared with UC at 1 month (N=134, 2.8% vs. 4.7%, RR 0.61, 95% CI 0.11 to 3.53).[14] After one month up to 12 months, a total of 18 patients (16.7%), nine (12.8%) originally randomized to KP and nine (23.7%, 9/38) who crossed over from UC to KP, experienced a symptomatic fracture compared to no patient who received UC only (Table 38). In the original randomized cohort, KP was associated with a similar risk of non-index radiographic vertebral fractures at 1 month compared with UC in patients who had radiographic data: KP (19.4%, 12/62) and UC (17.0%, 8/47); RR 1.14, 95% CI 0.51 to 2.56. An additional five patients who crossed over from UC to KP at 1 month experience a new radiographic fracture by 12 months; it is unclear how many of these patients had radiographs available (i.e., denominator is unclear).

Cement Leakage

Only one KP patient (1.4%) experienced a symptomatic cement leakage by the 1-month follow-up; this patient also suffered an adjacent fracture the day after the procedure, which was classified as a serious device-related event.[14] (see Table 38 and Appendix Table G6 for details).

Table 38. Summary of safety results: Kyphoplasty versus Usual Care in Patients with Fractures due to Tumors and Malignancies from the CAFE Trial (Berenson, 2011)

Outcome*	Follow up	Analysis group	KP % (n/N)	Usual Care % (n/N)	RR (95% CI)
Mortality*	1 month	As randomized	2.8% (2/70)	1.5% (1/64)	1.82 (0.17 to 19.69)
	≥1 month and ≤12 months	As randomized	30.0% (21/70)	19.2% (5/26)	1.56 (0.66 to 3.71)
		As treated (after crossover)	25.0% (27/108)	19.2% (5/26)	1.30 (0.55 to 3.05)
Any Serious AEs†	1 month	As randomized	NR	NR	NR
	≥1 month and ≤12 months	As randomized	52.8% (37/70)	30.7% (8/26)	1.72 (0.93 to 3.19)
		As treated (after crossover)	50.9% (55/108)	30.7% (8/26)	1.66 (0.90 to 3.03)
Symptomatic Fracture	1 month	As randomized	2.8% (2/70)	4.7% (3/64)	0.61 (0.11 to 3.53)
	≥1 month and ≤12 months	As randomized	12.8% (9/70)	0% (0/26)‡	NC, p=0.056
		As treated (after crossover)	16.7% (18/108)	0% (0/26)	NC, p=0.026
Cement Leakage	1 month	As randomized	1.4% (1/70)	NA	NC

Any AEs[§]	1 month	As randomized	37.1% (26/70)	29.7% (19/64)	1.25 (0.77 to 2.03)
	≥1 month and ≤12 months	As treated (after crossover)	NR	NR	NR

AE = Adverse event; CI = Confidence interval; KP = Kyphoplasty; NA = Not applicable; NC = Not calculable; NR = Not reported; RR = Risk ratio.

* No deaths were determined to be related to the device or procedure.

† Serious AEs defined as any event that resulted in death, life-threatening injury or permanent impairment, needed intervention to prevent impairment, or resulted in prolonged hospitalization. Some patients had multiple serious AEs.

‡ 9 patients that crossed over to KP from usual care experienced symptomatic fractures between 1 and 12 months; one patients had vertebral fracture before the crossover procedure, but was counted in the crossover group, and another had a new adjacent fracture 13 days after the crossover procedure which was possibly device related.

§ Included Blood and lymphatic disorders, cardiac disorders, eye disorders, gastrointestinal disorders, general disorders, infections, balloon rupture (asymptomatic), myocardial infarction, procedure-related pain, postoperative urine retention, metabolic/nutritional disorder, musculoskeletal disorders, neoplasms, nervous system disorders, psychiatric disorders, respiratory disorders, vascular disorders, myocardial infarction resulting in death, cardiac failure resulting in death, and general disorders resulting in death.

Mortality

KP and usual care were associated with a similar risk of mortality by 1 month, though results were imprecise (N=131, 2.8% vs. 1.5%, RR 1.82, 95% CI 0.17 to 19.69).[14] In the KP group, one death was attributed to a myocardial infarction (MI) that occurred prior to treatment and the other to general disorders with no further information. The one death in the usual care group was a result of cardiac failure. The risk of mortality after 1 month and up to 12 months remained similar between treatment groups in both the "as randomized" and the "as treated" analyses (Table 38). Most deaths were the result of neoplasms and general disorders and none were judged to be related to treatment.

Serious Adverse events

The total number of patients with a SAE by 1 month was not reported (patients could have more than one AE). By 1 month, there were three events that occurred in the KP group described by the authors as serious: two cases of myocardial infarction (2.7%, 2/70) (one occurred before treatment and resulted in death and the other occurred intraoperatively with intermittent atrial fibrillation and was attributed to anesthesia and resolved with medical therapy) and one case (1.4%, 1/70) of cement leakage into the adjacent disc resulting in an adjacent fracture 1 day after the procedure (Table 38).

The risk of SAEs (defined as any event that resulted in death, life-threatening injury or permanent impairment, required intervention to prevent impairment, or resulted in prolonged hospitalization) after 1 month up to 12 months was higher in the KP group compared to the UC group, but the difference was not statistically significant in both the "as randomized" (52.8% vs. 30.7%) and the "as treated" analyses (50.9% vs. 30.7%), Table 38.[14]. None of the SAEs in the patients initially randomized to KP were considered device related. In the crossover group (n=38), two patients (5.3%) had serious device-related events: an airway complication caused by anesthesia (resolved by mask ventilation) and a new vertebral compression fracture VCF 13 days after KP that the local investigator reported as possibly device related. Overall,

2.8% (2/108) of patients who received KP at any point during the 12-month follow-up had a serious device-related adverse event.

Other Adverse events

KP was associated with a similar risk of any adverse event compared to UC by 1 month (N=134, 37.1% vs. 29.7%, N=RR 1.25, 95% CI 0.77 to 2.03), Table 38.[14] Four cases of device-related AEs were reported but were not serious: one asymptomatic balloon rupture, two cases of extravasation to the disc (one asymptomatic), and one superficial wound infection. The incidence of any AE after 1 month was not reported.

4.3.2.3 Mixed Vertebroplasty and Kyphoplasty Populations

Studies that did not report VP and KP separately but analyzed data across both augmentation procedure types are summarized here. Five large administrative database studies were included. Three provided comparative data; one used a 20% random sampling of the Medicare database,[118] one used the NIS database[136] and one used private health insurance data from Germany [100] to compare adverse events following vertebral augmentation with UC or operative treatment. The other two database studies provided primarily single arm data from the ACS-NSQIP database, have overlapping populations and are summarized with the case series below.[35,185] In addition, one retrospective comparative NRSI[103] evaluated safety following KP and UC.

Mortality

One study of Medicare claims data[118] reported slightly lower 30-day mortality with vertebral augmentation versus conservatively treated propensity matched controls, but estimates were imprecise (adjusted OR 0.61, 95 CI 0.39-0.95) At 12 months mortality risk was similar between groups (HR of 0.92 (95% CI, 0.81-1.04). This study used a 20% random sample of Medicare data. Similarly, data from a small hospital-based study in the US reported no difference in mortality between vertebral augmentation and no treatment (Adj HR 0.81 95% CI 0.42, 1.59).[103] In contrast, another study using data from private health insurance in Germany reported slightly lower mortality with vertebral augmentation versus nonoperative care by 60 months (HR (adj) 0.58, 95% CI 0.48, 0.70), however author's Kaplan-Meier plot shows similar survival between VA and nonoperative management up to 36 months since diagnosis(data NR).[100] Table 39

Other SAEs

Across the three studies comparing **KP and UC**, there were no difference in the risk of any or specific SAEs across the database studies[118,136] or in the risk of recurrent fracture in the comparative NRSI,[103] except for 30-day outcomes in the Medicare database study[118] which showed that KP was associated with fewer SAEs (any) in propensity score adjusted analysis (adj. OR 0.90, 95% CI 0.81-0.99).

KP was associated with significantly fewer SAEs, both any SAE (adj. OR 0.48, 95% CI 0.41 to 0.56) and individual SAEs, i.e., stroke, MI, PE, shock, after adjusted analyses compared with open surgery in the large NIS database study.[136]

Table 39. Adverse Events Other than Mortality from Comparative Database and Comparative Nonrandomized Studies Evaluating Any Vertebral Augmentation (i.e., Vertebroplasty/Kyphoplasty) for Osteoporotic Vertebral Compression

Adverse Event	Database	Study Database search dates	N	Finding and conclusion
SAE				
Specific SAEs	NIS	Purvis, 2018 (2002-2011)	VP/KP: 11,116 UC: 46,962	Post-op, all p=NR Stroke: 0.1% (11/11116) vs. 0% (0/46962) MI: 0.6% (67/11116) vs. 0.8% (376/46962) PE: 0.2% (22/11116) vs. 0.3% (141/46962) Shock: 0.2% (22/11116) vs. 0.2% (94/46962)
Any SAE	NIS	Purvis, 2018 (2002-2011)	VP/KP: 11,116 UC: 46,962	Post-op 8.1% (900/11116) vs. 8.7% (4086/46962), Adj. OR 0.95 (95% CI 0.87–1.03)
	Medicare	McCullough, 2013* (2002-2006, 20% random sample)	VP/KP: 9,017 UC: 9,017 propensity-score matched	30 days 9.5% (860/9017) vs. 10.5% (947/9017), Adj. OR 0.90 (95% CI 0.81-0.99) 1 year 29.8% (2691/9017) vs. 30.0% (2709/9017), Adj. HR 1.00 (95% CI 0.94-1.06)
Recurrent Fracture	NA	Levy 2012† (NA)	VP/KP: 57 UC: 27	17.5% (10/57) vs. 25.9% (7/27), unadjusted RR 0.68 (0.29 to 1.58); p=NS in adjusted analyses
Mixed VP/KP vs. Operative Treatment				
SAEs	NIS	Purvis, 2018 (2002-2011)	VP/KP: 11,116 Open Surgery: 1,487	Post-op Stroke: 0.1% (11/11116) vs. 0.3% (4/1487), p<0.001 MI: 0.6% (67/11116) vs. 2.2% (33/1487), p<0.001 PE: 0.2% (22/11116) vs. 1.2% (18/1487), p<0.001 Shock: 0.2% (22/11116) vs. 1.0% (15/1487), p<0.001 Any SAE: 8.1% (900/11116) vs. 16.3% (242/1487); Adj. OR 0.48 (95% CI 0.41-0.56)

Adj. HR = adjusted hazard ratio; Adj. OR = adjusted odds ratio; CI = confidence interval; KP = kyphoplasty; MI = myocardial infarction; NA = not available; NIS = National Inpatient Sample; NR = not reported; PE = pulmonary embolism; RR = risk ratio; SAE = serious adverse event; UC = usual care; VP = vertebroplasty.

* Major medical complications included diagnosis codes for cardiorespiratory arrest, acute myocardial infarction, respiratory failure, pulmonary embolism, pneumonia, and stroke as well as relevant procedural codes.

† This trial had a 3rd arm of VP/KP + medical that included 49 patients but was excluded b/c it does not meet inclusion criteria.

Case series

Two large single arm database studies with overlapping populations[35,92] and four case series[163,177,185,201] evaluated safety for VP or KP together as vertebral augmentation. Sample sizes ranged from 1,932 to 2,433 in the database studies and from 299 to 358 in the case series. Follow-up periods ranged broadly from 1 to 31 months across the studies that reported follow-up. See Table 40 for AE details.

Mortality

The incidence of mortality following VP or KP was low, 2.0% across two large databases with overlapping populations.[35,185]

SAEs

The frequency of any SAE following VP or KP across the two database studies ranged from 4.9% to 5.8% and the incidence of individual SAEs (e.g., thromboembolic events, cardiac events, cerebrovascular events, etc.) was very low (≤1.0%); it is unclear if any SAEs were related to the treatment.[35,92] One additional case series reported that no SAE occurred in any patient.[185]

Cement Embolism

Two case series studies looked specifically at the risk of pulmonary cement embolism (PCE); there were no symptomatic cases in either study. In one study (N=373) the incidence of PCE on post-procedural CT was 17.2% and author state the incidence was similar for VP and KP.[163] In another study,[177] 3.7% of patients (N=299) had a PCE during VP; follow-up after 12 months showed no further sequelae or symptoms.

New fractures

Only one study reported the incidence of new vertebral fracture which was 12.6% for any new fracture and 7.3% for any new symptomatic adjacent level fracture by a mean of 31 months.[185]

Cement leakage

The incidence of symptomatic cement leakage or leakage into the spinal canal was rare (≤2%) across two studies but asymptomatic or any cement leakage was common following VP (32.5% to 40.8%).[185,201]

Reoperation

The rate of any reoperation across the two database studies ranged from 3.2% to 3.6%.[35,92] Repeat VP or KP for symptomatic adjacent level fractures was 7.3% in one case series.[185]

Table 40. Adverse Events in Single Arm Studies for mixed VP/KP for Osteoporotic fractures

Adverse Event	Follow Up	Study	% (n/N)
Mortality			
Any	1 month	Choo, 2018	2.0% (49/2433)
	1 month	Kim, 2022*	2.1% (40/1932)
SAE			
Any	1 month	Choo, 2018	5.8% (140/2433)
	1 month	Kim, 2022	4.9% (95/1932)
	NR	Wang, 2014	0% (0/358)
Thromboembolic events			
Any thromboembolic event	1 month	Choo, 2018	1.0% (24/2433)
PE	1 month	Kim, 2022	0.7% (13/1932)
DVT	1 month	Kim, 2022	0.7% (14/1932)
Cardiac events			
Cardiac arrest	1 month	Kim, 2022	0.2% (4/1932)
MI	1 month	Kim, 2022	0.1% (1/1932)
CVA events			
Stroke	1 month	Choo, 2018	0.1% (3/2433)
CVA with neurologic deficit	1 month	Kim, 2022	0.1% (1/1932)
Infection			
Deep infection	1 month	Kim, 2022	0% (0/1932)
Septic complication	1 month	Choo, 2018	0.8% (20/2433)
Sepsis	1 month	Kim, 2022	0.5% (9/1932)
Septic shock	1 month	Kim, 2022	0.2% (4/1932)
Bleeding			
Bleeding requiring transfusion	1 month	Choo, 2018	0.7% (16/2433)
Pulmonary Cement Embolism			
Asymptomatic PCE	Median 412 days	Sun, 2023	17.2% (64/373)
	Perioperative	Venmans, 2008	3.7% (11/299)
New Fracture			
Any fracture	Mean 31 months	Wang, 2014	12.6% (45/358)
Adjacent level, symptomatic fracture	6 months	Wang, 2014	3.1% (11/358)
	Mean 31 months	Wang, 2014	7.3% (26/358)
Cement Leakage			
Any symptomatic leakage requiring intervention	Mean 31 months	Wang, 2014	0% (0/358)
Any leakage	Mean 31 months	Wang, 2014	40.8% (146/358)
	NR	Zhang, 2020	32.5% (96/295 levels)
Spinal canal leakage	NR	Zhang, 2020	2.7% (8/295 levels)
Reoperation			
Any	1 month	Choo, 2018	3.6% (88/2433)
	1 month	Kim, 2022*	3.2% (61/1932)
Repeat VP/KP for symptomatic adjacent level fracture	Mean 31 months	Wang, 2014	7.3% (26/358)
Any AE			
	1 month	Kim, 2022	8.6% (166/1932)

AE = adverse event; CI = confidence interval; CVA = cerebrovascular accident; DVT = deep vein thrombosis; F/U = follow-up; KP = kyphoplasty; MI = myocardial infarction; NR = not reported; PCE = pulmonary cement embolism; PE = pulmonary embolism; SAE = serious adverse event; VP = vertebroplasty.

* Some overlap with Choo 2018 these are databases - also, these are included in the comparative database table very minimal outcomes.

4.4 Sacroplasty

A total of six studies were identified for this update review that met inclusion criteria and evaluated the effectiveness and safety of sacroplasty for treatment of sacral insufficiency fractures: four comparative NRSIs, one prospective[71] and three retrospective,[4,147,196] one SR of case series,[31] and one single arm registry study.[13] The SR of case series includes the sacroplasty arm from one of our included comparative NRSIs.[71] See Appendix H for study details.

In addition, results from the prior report from one SR[12] and nine case series[5,11,28,70,75,90,91,162,191] (4 that were included in the SR)[28,70,162,191] are summarized for completeness.

4.4.1 KQ1 Effectiveness

Description of Included comparative studies

One prospective[71] and three retrospective,[4,147,196] NRSIs were included that compared sacroplasty with nonsurgical management consisting of usual care, i.e., analgesics (to include muscle relaxants and opioids), bracing, walking aids and/or bed rest (3 studies)[4,71,147]

and percutaneous teriparatide injections (20 µg once a day for 26 weeks) (1 study).[196]

Sacroplasty was performed at the discretion of the treating physician and included cement (PMMA) sacroplasty in three studies[71,147,196] in the fourth study a variety of methods were used and included balloon sacroplasty, radiofrequency sacroplasty, vertebrosacroplasty and cement sacroplasty.[4] Sample sizes ranged from 27 to 244, the mean age of patients ranged from 70 to 81 years and the majority were female (range, 81% to 95% across 3 studies and 100% in one study). Duration of pain ranged from mean of 6.8 to 11.2 weeks in two studies[147,196] and was ≥3 weeks in another[71] based on inclusion criteria; the fourth study did not indicate pain duration. [4] Only one was conducted in the U.S.[71] the others took place in Austria,[4] Turkey,[147] and Taiwan.[196] Two studies received no funding for their work[4,71] and the other two did not report funding.[147,196]

One of the above studies[4] included a third treatment arm and also compared sacroplasty with screw fixation in 178 patients (83% female) with a mean age of 70 years. Sacroplasty was performed using a variety of methods and included balloon sacroplasty, radiofrequency sacroplasty, vertebrosacroplasty and cement sacroplasty. The alternative surgical method consisted primarily of iliosacral screw fixation (64%), most with cement augmentation. Mean duration of pain was 11.2 weeks.

These studies were all consider poor-quality (high risk of bias) due serious confounding by indication and lack of controlling for this and for baseline differences. See Appendix H for study details.

4.4.1.1 Sacroplasty versus Non-surgical Management

4.4.1.1.1 Primary Outcomes

Pain

Across the three studies comparing sacroplasty with usual care,[4,71,147] sacroplasty resulted in significantly greater improvement in VAS pain scores (0-10 scale) across most timepoints and was sustained longer term (follow-up range, 6 months to 10 years), (Table 41). One of the studies[4] presented data stratified by a number subgroups but did state that patients receiving sacroplasty experienced pain reductions rapidly and significantly ($p<0.001$), while patients receiving conservative therapy either did not or

experienced a delayed in response. Of note, one of these studies had a large imbalance in pain scores at baseline (8.82 for sacroplasty vs. 4.18 for UC) and based its results on change scores.[147]

The study that compared sacroplasty with daily percutaneous injections of teriparatide over 26 weeks reported significantly less improvement in VAS pain scores after sacroplasty at 12 and 26 weeks of follow-up but scores were not statistically different between treatment groups at earlier timepoints (Table 41).[196]

Function

Two of the studies reported function using the ODI[147] and the Hamburg Barthel Index (HBI)[4] and found that sacroplasty resulted in significantly greater improvement in function scores compared with UC at all timepoints (Table 41). There were large imbalances between groups in baseline scores in both trials, such that the sacroplasty group had greater disability; one study based its results on change scores.[147]

The study that compared sacroplasty with daily percutaneous injections of teriparatide over 26 weeks reported significantly less improvement in ODI scores after sacroplasty at 4, 12 and 26 weeks of follow-up; scores were not statistically different between treatment groups after 2 weeks (Table 41).[196]

Opioid use

One study[71] reported opioid use only in those who received sacroplasty. At study entry, 77.1% (162/210) of patients were using opioids which decreased to 32.9% (69/210) during the postoperative period; at the 10-year follow-up, none of the patients (0/117) reported using opioids for sacral pain (Appendix Table I1).

Table 41. Sacroplasty vs. Usual Care/Non-surgical Management

Outcome*	Author, year Type of NSM	F/U	SP Mean (SD)	UC/NSM Mean (SD)	Effect Size (95% CI)
Pain					
VAS pain (0-10, lower better)	**Frey, 2017** UC: analgesics including opioids, corsets, and/or bed rest	Baseline	8.29 (0.13) (n=210)	7.47 (0.38) (n=34)	p>0.05
		2 weeks	2.82 (SE 0.17) (n=NR)	5.44 (SE 0.44) (n=34)	NC
		4 weeks	2.39 (SE 0.15) (n=NR)	4.24 (SE 0.42) (n=34)	NC
		12 weeks	1.93 (SE 0.14) (n=NR)	3.47 (SE 0.46) (n=34)	NC
		24 weeks	1.45 (SE 0.13) (n=NR)	2.47 (SE 0.42) (n=34)	NC
		1 year	0.89 (SE 0.10) (n=NR)	1.44 (SE 0.28) (n=34)	NC
		2 years	0.66 (SE 0.08) (n=82)	1.12 (SE 0.25) (n=34)	MD -0.46 (-0.86 to -0.06)
		10 years	0.50 (SE 0.08) (n=117)	NA	NA
	Sarigul 2023 UC: analgesics, muscle relaxants, and bed rest	Baseline	8.82 (NR) (n=83)	4.18 (NR) (n=102)	p<0.05
		10 days	5.91 (NR) (n=83)	1.48 (NR) (n=102)	p=0.77 for change scores
		12 weeks	4.22 (NR) (n=83)	1.36 (NR) (n=102)	p=0.02 for change scores
		1 year	1.15 (NR) (n=83)	2.82 (NR) (n=102)	p<0.001 for change scores
	Yang 2023 Percutaneous teriparatide injection 20 µg 1x/day for 26 weeks	Baseline	7.7 (0.8) (n=13)	8.0 (1.0) (n=14)	p>0.05
		2 weeks	4.7 (1.3) (n=13)	5.0 (0.8) (n=14)	p>0.05
		4 weeks	4.6 (1.2) (n=13)	3.8 (1.1) (n=14)	P=NR
		12 weeks	3.8 (1.5) (n=13)	1.8 (0.6) (n=14)	p<0.001
		26 weeks	2.7 (1.4) (n=13)	0.6 (0.8) (n=14)	p<0.001
Function					
ODI (0-100, lower better)	**Sarigul 2023** UC: analgesics, muscle relaxants, and bed rest	Baseline	78.64 (NR) (n=83)	51.79 (NR) (n=102)	p<0.05
		10 days	24.31 (NR) (n=83)	48.76 (NR) (n=102)	p=0.04 for change scores
		12 weeks	14.28 (NR) (n=83)	42.94 (NR) (n=102)	p=0.03 for change scores
		1 year	8.44 (NR) (n=83)	21.16 (NR) (n=102)	p<0.001 for change scores
	Yang 2023	Baseline	82.6 (9.1) (n=13)	82.7 (9.7) (n=14)	p>0.05
		2 weeks	68.3 (3.5) (n=13)	64.6 (8.2) (n=14)	p>0.05
		4 weeks	56.9 (4.1) (n=13)	48.8 (8.0) (n=14)	P=0.010

Outcome*	Author, year Type of NSM	F/U	SP Mean (SD)	UC/NSM Mean (SD)	Effect Size (95% CI)
	Percutaneous teriparatide injection 20 µg 1x/day for 26 weeks	12 weeks	32.4 (4.8) (n=13)	22.6 (9.4) (n=14)	p=0.005
		26 weeks	20.7 (4.9) (n=13)	11.2 (3.5) (n=14)	p<0.001
HBI (0-100, higher = better)	**Andresen 2022** UC: bed rest, analgesic therapy and mobilization using a walker or crutches	Baseline	48 (14) (n=109)	65 (10) (n=88)	p<0.001
		2 years*	83 (6) (n=109)	76 (13) (n=88)	P<0.05

CI = confidence interval; F/U = follow-up; HBI = Hamburg Barthel Index; MD = mean difference; NA = not applicable; NC = not calculable; NR = not reported; NSM = non-surgical management; ODI = Oswestry Disability Index; SD = standard deviation; SE =standard error of the mean; SP = sacroplasty; UC = usual care; VAS = visual analogue scale.
*Data not provided for other timepoints but authors indicate significant differences between treatment groups favoring sacroplasty from 2 days to 18 months

4.4.1.2 Sacroplasty versus Surgery

One study compared sacroplasty with screw fixation (primarily iliosacral screw fixation with cement augmentation)[4] and found that patients in both groups experienced significant improvement in pain (VAS scores) and function (HBI scores) but it was more rapid following sacroplasty; authors state that patients who received screw fixation benefited after 6 months with sustained benefits. Data was not well reported, especially for pain, and comparisons between the two groups were not distinctly made. HBI function scores at 2 years (N=154) were similar, mean 83 (SD 6) vs. 84 (SD 6), respectively. Those in the surgical group had more complex and severe fractures which likely impacted the recovery time. See Appendix Table I1 for study details and results.

Systematic review and case series

One SR published in 2019 included one comparative NRSI and 18 cases series that evaluated sacroplasty for treatment of sacral insufficient fracture and neoplastic lesions.[31] The sole NRSI[71] included in the SR is a study we've included above for comparative effectiveness; only the sacroplasty arm of this study was used in the SR's analyses. A total of 861 patients were included (range of N's, 6 to 243). The majority of patients were female (79%) with a weighted mean the age of 74 years (range, 58 to 83 years). Most of studies evaluated patients with osteoporotic fracture (63%, 12/19); the remainder included patients with malignant lesions only (21%, 4/19) or both osteoporotic and malignant fracture (16%, 3/19). Six patients underwent two procedures for a total of 867 sacroplasties performed. Funding was not reported.

One prospective study[13] conducted an interim analysis of the first 102 patients included in the Vertebral Augmentation Sacroplasty Fracture Registry, a U.S. national ongoing registry involving 10 sites designed to assess the effectiveness and safety of sacroplasty as an as-treated, on-label procedure. The mean patient age was 74 years, 69% were female and 98% received sacroplasty for osteoporotic sacral insufficient fractures. Most patients had failed nonoperative treatment of up to 4 months duration. Sacroplasty and other procedural techniques used were at the discretion of the investigational site and the treating physician. The study was funded by an academic foundation.

See Appendix Table J1, K1, and K2 for study and results details.

Studies from the 2010 report

A systematic review[12] of sacroplasty for the treatment of sacral insufficiency fractures caused by osteoporosis included 15 papers; seven of these were case series, three were case reports, and five were technical reports. A total of 108 patients (range 1-52 per study) with a mean age of 75.5 years were included across all the studies in this review, with follow-up ranging from 24 hours to 42 months.

In addition, nine case series[5,11,28,70,75,90,91,162,191](several of which were also included in the systematic review)[12] were also included. As with this update report, only studies with five or more patients were considered for inclusion. Two studies[70,162] were of patients with osteoporosis (N = 65 total patients), three[5,11,75] were of patients primarily with multiple myeloma or other tumors (N = 28 total patients), and four[28,90,91,191] were of patients with sacral insufficient fractures of undefined or mixed causes (N = 48 total patients). Summarizing the results of these studies was made difficult by the lack of consistency in the outcomes reported or in length of follow-up.

4.4.1.2.1 Primary Outcomes

Pain

Pooled analyses of VAS pain scores in the SR[31] demonstrated statistically significant improvement in pain levels from baseline to 24 to 48 hours and 6 and 12 months post-procedure; cumulative pain scores are in provided in Table 42. Results remained robust after adjusting for potential publication bias. The authors also performed a meta-regression which showed that none of the covariates studied (sacroplasty indication [osteoporosis vs. malignancy], study design [prospective vs. retrospective], and technical modifications [none, radiofrequency augmentation, balloon dilation) were associated with VAS-study effect size. Clinical success was achieved in 95.7% of patients but the timing is unclear.

Consistent with the findings from the SR, the registry study[13] showed statistically significant improvement in VAS pain scores at all time points (1, 3 and 6 months) following sacroplasty with 91.8% of patients achieving a clinically meaning improvement (≥2 points) on the VAS pain scale by 6 months (Table 42).

Studies from the 2010 report

Pain results from the prior report are consistent with those of this update report, showing significant improvement in pain following sacroplasty. In the SR[12] included in the prior report VAS pain scores were significantly improved in the 62 patients for whom it was measured, improving from 8.9 pre-operatively to 2.6 post-operatively (across a range of follow-up times). Six of the nine case series reported VAS pain scores which improved following sacroplasty for both osteoporotic and malignant fractures, from a mean of 8.1 to 9.1 pre-operatively to 0.8 to 3.8 at varying follow-up periods. [5,11,70,75,90,191] (Across two studies,[28,162] 11 of 19 patients (58%) reported complete or significant pain relief at follow-up of approximately two weeks.

Function

In the registry study,[13] RDQ scores improved significantly at all time points (1, 3 and 6 months) post-procedure with 83.7% of patients achieving a clinically meaning improvement (≥5 points) on the RDQ pain scale by 6 months (Table 42). The SR did not report function outcomes.

Table 42. Effectiveness outcomes from one SR and one registry study of SP for sacral insufficient fractures or malignant fractures

Author, year	Outcome	F/U	SP, mean (SD) or % (n/N)
Chandra 2019 SR (N=861; 19 single arm studies)	VAS pain (0-10 scale)	Baseline	19 studies (N=861): pooled mean 8.35 (95% CI 8.08 to 8.63); range of mean scores: 5.3 to 9.3[*]
		24-48 hours	15 studies (N=749): pooled mean 2.70 (95% CI 2.19 to 3.20); range of mean scores: 0.67 to 4.1
		6 months	8 studies (N=352): pooled mean 2.26 (4.5)[†]; range of mean scores NR
		12 months	9 studies (N=357): pooled mean 2.01 (95% CI 1.35 to 2.67); range of mean scores: 0.89 to 2.4
	Clinical success[‡]	NR	18 studies: 95.7% (623/651)
Beall 2023[§] (N=102)	NRS pain (0-10 scale)	Baseline	N=102: mean 7.8 (2.4)
		1 month	N=51: mean 2.4 (3.3), p<0.001 vs. baseline

The Sacroplasty Registry		3 months	N=52: mean 1.2 (2.5), p<0.001 vs. baseline
		6 months	N=49: mean 0.9 (2.2); p<0.001 vs. baseline
	≥2-point improvement on NRS	1 month	72.6% (37/51)
		3 months	90.4% (47/52)
		6 months	91.8% (45/49)
	RDQ (0-24 scale)	Baseline	N=102: mean 17.7 (6.4)
		1 month	N=51: mean 8.4 (4.9), p<0.001 vs. baseline
		3 months	N=52: mean 6.9 (4.9), p<0.001 vs. baseline
		6 months	N=49: mean 5.2 (5.2); p<0.001 vs. baseline
	≥5-point improvement on RDQ	1 month	76.5% (39/51)
		3 months	78.8% (41/52)
		6 months	83.7% (41/49)

CI = confidence interval; F/U = follow-up; NRS = numerical pain rating scale; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SP = sacroplasty; UC = usual care; VAS = visual analogue scale.
*Excluding the 5.3 score, the range is 7.5 to 9.3.
†Standard deviation calculated from standard error given.
‡Defined as: the patient's pain improved, stayed the same, or if remobilization was achieved after the sacroplasty procedure.
§N indicated the number of patients with completed follow-ups at time of data collection.

Studies from the 2010 report

Function results from the prior report are consistent with those of this update report, showing significant improvement in function after sacroplasty. One case series[90] measured function via a 5-point mobility scale (1 = normal, 5 = bedridden; mean 4.3 [SD 1] at baseline vs. 2.3 [SD 1.2] posttreatment) and another study[191] reported improvement in all activities of daily living (data not reported).

4.4.1.2.2 Secondary Outcomes

Opioid Use

Neither the SR of case series or the registry study[13] included in this update report reported opioid use.

Studies from the 2010 report

Three studies reported decreases in the use of opioid pain medication following sacroplasty, from a range of 71%-58% at baseline to 10%-21% at follow up in two studies of patients with osteoporotic fractures[70,90] and from 100% to 0% (most were using only nonsteroidal anti-inflammatory medications) in a series of eight patients with fractures due to malignancy.[5]

4.4.2 KQ2 Harms and Safety

Safety was not well reported across the included studies. See Table 43 for a summary of reported AEs.

Mortality

The only comparative data provided was for mortality in one retrospective NRSI[4] which showed that sacroplasty was associated with significantly fewer deaths compared with usual care over 12 months (8.4% vs. 21.7%); while there were fewer deaths in patients who received sacroplasty versus pedicle screw fixation surgery (8.4% vs. 13.6%) in this same study, the difference was not statistically significant. The single arm registry study reported no deaths through 6 months.[13]

SAE

Only one retrospective NRSI called out a specific adverse event and noted that there were no cases of pulmonary embolism in any patient following sacroplasty.[147] The study that compared sacroplasty with daily injections of teriparatide only stated that no specific complication related symptoms were noticed in either group.

Symptomatic Fracture

There were three cases (3.0%) of new symptomatic fracture reported over 6 months in the registry study; two patients suffered new sacral fractures and had revision sacroplasty and one patient presented with severe back pain and underwent vertebral augmentation to treat a new VCF.[13] These patients are included under reoperation below. None of the studies reported the incidence of new fracture (symptomatic or asymptomatic).

Reoperation

In the SR of case series, there were three patients (0.9%) with radicular pain due to cement leakage who required decompression to relieve the symptoms.[31] All three patients with symptomatic fractures in the registry study (see above) required either repeat sacroplasty or vertebral augmentation.[13]

Cement Leakage

Symptomatic cement leakage was rare as reported by the SR, the registry study and two NRSIs (0% to 1.0%).[4,13,31,147] Across just the SR and the registry study the incidence ranged from 0.6% to 1.0%.[13,31]

Table 43. Summary of safety results for NRSIs of sacral insufficiency fractures

Outcome	Author, year Study Design	Follow up	Sacroplasty % (n/N)
Comparative			
Mortality	**Andresen 2022** Comparative NRSI	12 months	8.4% (10/119) vs. *UC*: 21.7% (25/114); RR 0.38 (95% CI 0.19 to 0.76)
			8.4% (10/119) vs. *Surgery*: 13.6% (8/59); RR 0.62 (95% CI 0.23 to 1.49)
SP arm only			
Mortality			
Any	**Beall 2023** Single arm Registry	6 months	0% (1/102)
SAE			

PE	**Sarigul 2023** Comparative NRSI	Perioperative	0% (0/83)
Symptomatic Fracture			
Any	**Beall 2023** Single arm Registry	6 months	3% (3/102), required surgery
New sacral fracture			2% (2/102), required surgery
New VCF			1% (1/102), required surgery
Reoperation*			
Radicular pain due to cement leakage (SAE)	**Chandra 2019** SR of single arm studies	3-18 months	0.3% (3/861)
Any new sacral or VCF	**Beall 2023** Single arm Registry	6 months	3% (3/102)
Cement Leakage			
Symptomatic	**Chandra 2019** SR of single arm studies	3-18 months	0.6% (5/861), radicular pain
	Beall 2023 Single arm Registry	6 months	1.0% (1/102), new neurologic deficit
	Sarigul 2023 Comparative NRSI	Perioperative	0% (0/83)
	Andresen 2022 Comparative NRSI	Perioperative	0% (0/119)
Asymptomatic	**Chandra 2019** SR of single arm studies	3-18 months	2.2% (19/861)
	Beall 2023 Single arm Registry	6 months	17.7% (18/102)
	Sarigul 2023 Comparative NRSI	Perioperative	2.4% (2/83)
	Andresen 2022 Comparative NRSI	Perioperative	8.4% (10/119)

AE = Adverse event; CI = Confidence interval; NA = Not applicable; NC = Not calculable; NR = Not reported; NRSI = Nonrandomized study of interventions; RR = Risk ratio; SAE = serious adverse event; SR = systematic review; UC = usual care; VCF = vertebral compression fracture.
*Also included under cement leak complications and symptomatic fracture.

Safety results reiterated from the 2010 report

Very few adverse outcomes were reported in the SR[12] included in the prior report (clinically insignificant cement leakage and S1 radiculopathy). No major complications were reported in any of the case series of sacroplasty.[5,11,28,70,75,90,91,162,191] Asymptomatic cement leakage was reported in 7 of 34 patients across four series.[11,28,75,91] One patient developed radicular pain during cement injection, which was relieved 7 days later with an epidural steroid injection.[70] Two patients had radicular pain during the procedure from tumor extension into neural foramen, which was treated with selective nerve root block.[75]

4.5 KQ4 Cost-Effectiveness

4.5.1 Evidence of Cost Implications and Cost-Effectiveness of Vertebroplasty, Kyphoplasty, and Sacroplasty

Summary of studies and key points:

Three full economic studies comparing either vertebroplasty or kyphoplasty with conventional treatment were included in the 2010 HTA report: One evaluated kyphoplasty and was of reasonable quality[161]. One moderate quality[93], and one was poor quality study evaluated vertebroplasty[114]. All evaluated populations with osteoporotic vertebral compression fractures. All suggest that in the short term, vertebroplasty (two studies[93,114]) and kyphoplasty (one study[161]) may be of at least comparable cost and may provide earlier pain relief compared with conventional treatment, however confidence in the evidence was very low. None examined the cost effectiveness of either balloon kyphoplasty or vertebroplasty in a U.S. setting, thus generalizability to the U.S populations was unknown. Studies provided little information on the impact of various factors on overall cost effectiveness, and all were limited by lack of long-term data on effectiveness and safety. Two studies were industry funded[93,161] and funding was not stated in the third[114]. No full economic studies in patients with tumor-related fractures or sacroplasty were identified for the 2010 HTA.

For this update, three reviews[20,134,135] and six full economic studies relevant to populations with osteoporotic vertebral compression fractures[57,72,86,160,165,167] and one relevant to cancer-related VCF[131]were identified. Summaries of the three reviews are in Appendix N. Individual studies described in these reviews that met our inclusion criteria are summarized individually below or were included in the prior report. Two studies were U.S. based[57,86]. Both were industry funded (Medtronic). Of the non-US based studies two were reported by government entities[131,160] in the UK and Canada. One[167] was performed in Japan and received no funding. The other two were performed in Sweden[72] and the UK[165] and were industry funded (Medtronic). No economic studies on sacroplasty were identified. Critical appraisal of the individual studies is found in Appendix Tables E12 to E14

4.5.1.1 Key Findings Across New Economic Studies

In general, most economic studies suggest that vertebral augmentation may be cost effective versus nonoperative conventional management. Mortality was modeled in many of these studies. Several economic studies evaluating associations between VP, KP and usual care and mortality used data from or analyses of Medicare/CMS data for mortality, which is an important limitation of these studies. Sensitivity analyses in most studies suggest that assumptions regarding mortality had important impacts on cost-effectiveness. Well known limitations of such administrative database studies include selection bias, inability to control confounding including confounding by indication, missing data and misclassified data. Although authors report various methods of adjusting for bias, such as propensity score matching, residual confounding and selection bias may persist. Causal inference for mortality benefit is not possible. Some studies modeled a life-time horizon or longer-term horizons (5 years) however long-term RCT data from are sparse. Patient populations modeled were generally >65 and changes in health status and co-morbidities may impact life years and quality of life. The impact of adverse events and potential for subsequent fractures were infrequently modeled or considered in sensitivity analyses.

Given the differences in healthcare systems and reimbursement between the U.S. and other countries, the generalizability of findings from studies from outside of the U.S. is unclear.

4.5.1.1.1 U.S. Based Economic Studies

Two US based economic studies were identified

- One poor-quality, industry-funded study [57] evaluated cost effectiveness in terms cost per life-years gained based on analysis of Medicare data.
  - Cost per life-year-gained for VP ranged from $2452 to $13,543 and from $1863 to $6687 for VP versus nonoperative care. Based on Medicare enrollment information, survivorship was modeled from the time of VCF diagnosis until death, being censored or the end of the study period.
  - A primary limitation of this study is that causal inference that augmentation reduces mortality is not possible given the limitations of administrative data and lack of detailed information on causes of death and the possibility of residual confounding and selection bias even after adjusted for these. Sensitivity analyses were limited to the impact of discount rate.
- One good-quality, US-based study[86] also used Medicare data (paid for by industry) used a model based on that of the UK study described below[165] to compare VP and KP with conventional medical management (CMM) using Medicare data. The authors' results suggest that both KP and VP are more cost-effective in the outpatient setting than in the inpatient setting. Four groups based on treatment and inpatient versus outpatient setting were constructed (KP inpatients vs CMM, KP outpatients vs CMM, VP inpatients vs CMM and VP outpatients vs CMM)
  - In all four of the treatment scenarios tested, surgical intervention was predicted to be cost-effective compared to CMM with incremental cost-effectiveness ratios (ICERs) ranging from USD $11,000 to $43,000 per quality adjusted life-year (QALY) gained across groups with the highest ICER for inpatient KP versus nonoperative management and lowest for outpatient.
  - As was the case with several of other new economic studies, mortality "benefit" was a key driver of cost effectiveness; authors report ICER ranges of $55,485 per QALY for outpatient KP versus nonoperative care to $314,958 per QALY for inpatient VP versus nonoperative care which exceed the willingness-to-pay (WTP) threshold.
  - A primary limitation of this study relates to the limitations of administrative data including potential selection bias and residual confounding despite propensity scoring and the inability to make causal inference regarding mortality benefit associated with augmentation.

4.5.1.1.2 Non-U.S. Based Economic Studies

Two government-funded economic studies outside of the US were identified

- The most comprehensive economic analysis was performed by the UK National Institute for Health Research (NIHR)[160] which included data from both sham controlled trials and trials comparing VP with usual care in patients with osteoporotic VCF. It was rated as good quality.
  - Based on extensive sensitivity analyses, including consideration of whether sham involving local anesthetic might be considered a more "active" control, authors conclude that ICERs are driven by the clinical scenarios chosen:
    - KP was consistently cost-effective (at WTP below £20,000) if modeling included differential mortality benefit versus UC. When no mortality benefit was assumed, the method for utility determination influenced cost/QALY

        - ICERs for VP and KP were often greater than £20,000 when blinded trials were used
        - PV was constantly cost effective at ICER below £20,000 when a pooled beneficial effect was used.
        - Authors note that while vertebral augmentation may lead to decreased mortality, the data for this is from administrative data (registry) and that causal inference is not possible given lack of detailed information on causes of death.
    - The only cost-utility analysis (CUA) in patients with malignant VCF was performed by Health Quality Canada[131] concluded that KP and VP may be cost-effective with ICERs of Canadian Dollars (CAD) $33,471/QALY gained for KP and CAD $17,870/QALY gained for VP, both in comparison to nonsurgical management. It was rated as good quality.

Three cost utility studies from outside of the U.S. were identified.

- Two industry-funded CUAs, one performed in the UK[165] which used data from FREE and VERTOS II RCTs and the other performed in Sweden[72] using data from the FREE trial.
    - One fair-quality study was conducted in Sweden[72], had industry funding, and evaluated the cost-effectiveness of KP compared to UC based on data of 63 patients from the FREE trial.
        - The base case ICER was Swedish Krona (SEK) 884,682 (USD $134,043) per QALY. Sensitivity analysis showed that adjusting QALY benefits could make KP cost-effective, with costs ranging from SEK 359,146 to SEK 745,812 per QALY.
        - Limitations include potential selection bias, reliance on cost diaries, and a small sample size, making the findings specific to the Swedish healthcare system and limiting their generalizability.
    - A good-quality UK based study, Svedbom 2013[165] used a Markov model with data from the FREE and VERTOS II RCTs to evaluate the cost-effectiveness of KP versus VP and non-surgical management (NSM) in the UK. Funding was from Industry
        - KP showed higher costs (£9,313) and more quality-adjusted life years (5.473) than NSM and VP, with ICERs of £2,706 and £15,982 per QALY, respectively.
        - Sensitivity analysis confirmed KP's cost-effectiveness was robust but was sensitive to changes in mortality benefit assumptions. A 75% reduction in mortality benefit increased the ICER to £3,104 per QALY for NSM and £32,419 per QALY for VP. Without mortality benefit, KP was less cost-effective, particularly compared to VP.
        - The study's limitations include reliance on retrospective administrative mortality data, potential placebo effects in health-related quality of life (HRQoL), and differences between patient populations in the trials, limiting generalizability outside the UK.
    - One good-quality study from Japan[167] (which received no funding) was a propensity score matching study evaluating the cost-effectiveness of KP compared to NSM. The study used a Markov model with a lifetime horizon to assess costs and QALYs for 71 matched patients.
        - The base case analysis indicated that KP was associated with higher costs (402,988 Japanese Yen [JPY]) compared to NSM and a gain in QALYs of 0.033 at 6 months and 0.089 over 3 years. The ICER for KP was 4,404,158 JPY per QALY at 3 years and 2,416,406 JPY per QALY at 20 years.
        - Sensitivity analyses showed that the ICER ranged from 652,181 JPY to 4,896,645 JPY depending on variations in HRQoL benefits and mortality reductions. With a

50% reduction in HRQoL benefit, the ICER increased to 4,896,645 JPY. When full mortality benefit was assumed, the ICER decreased to 871,450 JPY. The study highlighted that KP remained cost-effective under most scenarios, but its cost-effectiveness was notably lower in patients aged over 80 years.

- The study was not based on RCT data with some data coming the author's institution, some from historical controls and mortality data was sourced from the Japanese government.

4.5.1.2 Detailed Analysis of New Economic Studies

Data for included studies and identified reviews are in Appendix O.

4.5.1.2.1 U.S. Studies

Edidin 2012:[57]

Study Overview: This industry-funded, poor-quality study (QHES 53/100) evaluated the cost-effectiveness of KP or VP following a diagnosis of OVCF (costing year: 2010) versus NSM by comparing the cost per life-year gained from the US Medicare payer perspective. Authors used cumulative costs from the 2005-2008 Centers for Medicare & Medicaid Services (CMS) database (i.e., inpatient and outpatient claims data) with three years of data of patients aged 65 years and older diagnosed with newly diagnosed OVCFs. The OVCF patients were stratified in two groups: 21.3% (n=182,946) in the "operated" treatment group (i.e., n=119,253 for KP and n=63,693 for VP) and the remaining 78.7% (n=676,032 patients) in the "non-operated" control group.

The Weibell survival model was used to estimate survival and life-years gained. Cost life-year gained was determined as the ratio between the discounted incremental cost and the discounted years of life gained. A discount rate of 3% was used for the base case in the cost-effectiveness analysis, with 0% and 5% discount rates used in sensitivity analyses. Authors analyzed the cost-versus-mortality benefit alone (not incorporating quality-adjusted life-years), finding that the cost per life-year gained ranged from USD $1,863 to $6,687 for KP and from USD $2,452 to $13,543 for VP compared with NSM. VP treatments were found to be cost effective in the Medicare population when compared with NSM. Among patients for whom surgical treatment was indicated, KP was found to be cost effective and cost saving compared with VP.

Base case and sensitivity analyses: Although the 3-year cumulative costs were higher for KP and VP, this study shows that the OVCF treatments are cost effective in the Medicare population when compared with NSM. Their results indicate that the survival rates are higher for the "operated patients" than the "non-operated ones (i.e., 50.3% for NSM vs 54.8% and 59.1% for VP and KP, respectively). This study also shows higher median life expectancies among the "operated" patients versus the "non-operated" patients. Patients in the kyphoplasty group had the longest median life expectancies, followed by vertebroplasty patients, and the non-operated patients. Among the oldest patients (85 years old and older), both KP and VP treatments were found to be cost effective (in terms of cost per life-year gained).

After accounting for the differences in median costs and using a discount rate of 3%, the cost per life-year gained for kyphoplasty and vertebroplasty patients ranged from USD $1,863 to $6,687 and from USD $2,452 to $13,543, respectively, compared with "non-operated" patients. The cost per life-year gained for kyphoplasty compared with vertebroplasty ranged from -$4,878 (cost saving) to $2,763.

A sensitivity analysis was performed to assess if differences in life expectancy between the study cohorts could be attributed to selection bias but provide limited detail of their modeling. Authors also evaluated the effects of the exclusion criteria on the life expectancy and did not find any differences in the life expectancy between the patient groups.

Limitations: Authors present only limited sensitivity analyses and do not describe drivers of cost or cost-effectiveness. This study was based on the Medicare population aged 65 years and older. A primary limitation of this study is that causal inference that augmentation reduces mortality or confers a benefit is not possible given the limitations of administrative data and lack of detailed information on causes of death and the possibility of residual confounding and selection bias even after adjusted for these. Confounding by indication could at least partially explain the findings. The impact of complications and new fractures or new comorbidities was addressed in this analysis. Authors acknowledge the potential risk of selection bias as the control population in the claims may receive different types of conservative care, but do not describe the potential impact of biases in detail. This analysis is limited to the first three years after an OVCF diagnosis. Costs associated with that diagnosis remain unchanged from the fourth year onward.

Hopkins 2020[86]

Study overview: This industry funded, good-quality study (QHES 82/100) evaluated the cost-effectiveness VP and KP following a diagnosis of OVCF (costing year: 2016) compared with CMM from the US Medicare payer perspective. Medicare data were paid for by industry and three co-authors were full-time employees of that company. Modeling was based on that of the UK study described below. [165] This analysis evaluated patient subgroups by treatment setting (i.e., inpatient or outpatient).

Using a propensity-score model, the authors stratified CMS patients into four groups based on treatment and inpatient versus outpatient setting (i.e., KP inpatients vs CMM with n= 2,071 x 2; KP outpatients vs CMM with n= 3,708 x 2; VP inpatients vs CMM with n= 720 x 2; VP outpatients vs CMM with n= 1,042 x 2). The authors evaluated patient subgroups by treatment setting using 2014 through 2016 CMS claims data. Patients were selected based on the date of their first VCF diagnosis, with a 2-year follow-up.

The average age in the KP inpatients group is 81.6 years old (vs 82.2 in the matching CMM group), 78.9 years old among the KP outpatients (vs. 79.3 in the matching CMM group), 81.4 years old VP inpatients (vs. 81.8 in the matching CMM group), 79.5 years old among the VP outpatients (vs. 80.4 in the matching CMM group). Overall and when looking at the four groups (based on treatment and inpatient versus outpatient setting), the proportion of female patients is higher (e.g., 82.3% in the KP inpatients group vs 83.23% in the matching CMM group). The Charlson Score ranges from 0 to 2+. 56.6% of the KP outpatients have a Charlson Score of 0 (vs 56.3% in the matching CMM group). Compared to the KP outpatients and VP outpatients, both the KP inpatients and the VP inpatients had a higher proportion of a Charlson Score of 2+ (i.e., 42.2% for the KP inpatients and 46.8% for the VP inpatients). The percentage of diagnosis is the highest among the KP inpatients (i.e., 70.8% vs 71.0% in the matching CMM group) and is the lowest among the VP outpatients (i.e., 56.0% vs 56.1% in the matching CMM group). The time to visit (measured in number of days) from first visit with diagnosis of VCF to surgery is the highest among the VP inpatients (i.e., 401.7 days) and is the lowest among the KP inpatients with 13.3 days. All four groups (based on treatment and inpatient versus outpatient setting) had a higher proportion of 0 inpatient visits (i.e., 68.2% for the KP inpatients, 78.6% for the KP outpatients, 64.8% for the VP inpatients, 78.9% for the VP outpatients).

Although each of the four groups seem to look similar in size and in demographic characteristics against the number of patients in the control group (i.e., CMM), there is a statistically significant difference for the following pairwise treatment comparison:

- Mean age for the KP inpatients versus CMM patients (*p<0.05*)
- Mean age for the KP outpatients versus CMM patients (*p<0.05*)
- Mean age for the KP outpatients versus CMM patients (*p<0.001*)
- Proportion of female patients in the KP outpatients vs CMM patients (*p<0.05*).

Base case and sensitivity analyses: The authors developed a 6-month Markov simulation model based on an existing UK-based model using demographic, clinical, and cost inputs to reflect a US Medicare population. Demographic data were sourced from the CMS database. US-specific VCF risks such as age and gender, and health state utilities were derived from other studies and published literature to calculate and reflect US utility values. Based on published literature, the authors also calculated the relative risk of subsequent fracture(s) in their 6-month Markov microsimulation.

By stratifying patients into four groups based on treatment and inpatient versus outpatient setting, the results from this study show that KP and VP are cost-effective at a willingness-to-pay threshold (WTP) of < $50,000 per QALY. For each pairwise comparison, the authors computed the mean cost per patient, the mean 2-year costs of post-fracture outpatient care, and the mean inpatient cost. Costs and QALY were discounted at 3% per year.

The authors performed some sensitivity analyses by changing and varying the model inputs and assessing the impact on the outcome of interest, the cost-effectiveness of the treatment. The base-case model assumed that the utility weights for patients undergoing VP were the same as those receiving KP. The authors changed both the VP and KP utilities within their confidence limits to assess the impact and did the same for the age-specific relative risks of subsequent fractures. ICERS did not exceed the WTP threshold of $50,000/QALY for most sensitivity analyses. As was the case with several of these new economic studies, mortality "benefit" was a key driver of cost effectiveness; authors report ICER ranges of $55,485 per QALY for outpatient KP versus nonoperative care to $314,958 per QALY for inpatient VP versus nonoperative care which exceed the WTP threshold.

They also ran a probabilistic sensitivity analysis to evaluate the cumulative impact of all the model inputs on the cost-effectiveness results. For each treatment comparison, the model results show that the probabilistic results cluster around the deterministic results.

The authors' results suggest that both KP and VP are more cost-effective in the outpatient setting than in the inpatient setting. In all four of the treatment scenarios tested, surgical intervention was predicted to be cost-effective compared to CMM with ICERs ranging from USD $11,000 to $43,000 per QALY gained. For KP in the outpatient setting at the willingness-to-pay threshold of USD $50,000, outpatient KP had an ~100% probability of being cost-effective compared with CMM

Limitations: A limitation of this study is that causal inference that augmentation reduces mortality or confers a benefit is not possible given the limitations of administrative data, including selection bias and confounding. Parameters in the model, such as health state utilities, were derived from other studies and published literature to calculate and reflect US utility values. However, those studies only included a KP treatment arm. While those for the VP treatment were not observed in the previously published literature, this study used the same utility weights for both KP and VP patients. The authors observed similar cost-effectiveness profiles for both KP and VP compared with CMM which might have not been otherwise the case; had the authors not used the same utility weights for both KP and VP patients. Additionally, some assumptions were made regarding health utilities. For the KP and VP groups, the

authors assumed that the patients' utility between 24 months and 36 months would decrease linearly and that there would not be any additional utility benefit between the treatment and control group after three years (i.e., for the KP and VP groups, the patients' utility would equal the CMM group's utility at 0.668).The authors made some additional adjustments for the ongoing risks of mortality and the subsequent risk(s) of fracture while using the CMS data in their model. The base case in this study used Medicare data to develop a microsimulation model whose population is older than the one from the previously published literature (i.e., mean age 79–82 years vs mean 72–74 years). There was no explicit discussion on the impact of specific biases. In particular, selection bias may be present, and propensity score matching may not fully account for confounders.

4.5.1.2.2 Non-U.S. Studies

Two analyses from government reports were identified. One analysis from Health Quality Ontario [131] in populations with VCF due to cancer and the other in patients with osteoporotic fractures from the UK National Institute for Health Research (NIHR).[160] Two industry-funded studies [72,165] and one that receive no industry funding [167]were also identified .

Cameron 2016 (Health Quality Ontario): [131]

Study overview: This good-quality cost-utility study (QHES 80/100) evaluated the cost-effectiveness of KP or VP for VCF due to cancer in a Canadian healthcare setting (costing year: 2015), compared with NSM. The authors conducted a systematic review (SR) of health economic studies assessing the cost-effectiveness of KP or VP versus NSM for the treatment of VCF in patients with cancer, from the perspective of the Ontario Ministry of Health. The study utilized a 1-year Markov simulation model with 1-month health state transitions, categorizing patients into the following states: 1) alive without subsequent vertebral fractures, 2) alive with subsequent vertebral fractures, and 3) death. The model estimated the cost-utility of KP and VP compared to NSM, focusing on a population of cancer patients (e.g., lung cancer, breast cancer, prostate cancer, multiple myeloma) using clinical outcomes including QALYs, mortality reduction, subsequent fractures, and costs.

The population was comprised of 90% outpatients with a mean age of 65. The survival rates by cancer type were derived from published literature, with model data indicating that over a 60-month period, survival rates for all cancer types decreased by nearly 10%, with breast cancer having the highest survival rate. The average cancer care cost was sourced from published literature, and utility estimates were derived from an industry-sponsored abstract of the CAFE trial, [14] where SF-36 scores were mapped to utilities. All direct costs and outcomes were discounted at 5%.

Base case and sensitivity analyses: At baseline, patients undergoing NSM had a utility of 0.27, which increased to 0.30 after one month. Similarly, patients receiving KP had a nearly identical baseline utility to those in the NSM group (0.30 vs. 0.27, respectively). However, after one month, the utility for patients treated with KP was significantly higher, more than double that of the NSM group (0.63 vs. 0.30). Although NSM resulted in lower 1-year costs compared to both KP and VP, it also led to fewer QALYs. The base case analysis demonstrated ICERs of CAD $33,471 per QALY gained for KP and CAD $17,870 per QALY gained for VP, both in comparison to NSM.

The authors performed univariate sensitivity analysis using the model inputs (i.e., HRQOL benefit, cancer type, time horizon, costs, and mortality reduction). HRQOL range had the biggest impact on the ICER with the upper range reaching CAD $75,000/QALY based on estimate from author's graph. The ICER range by cancer type ranged from about CAD $25,000/QALY to around CAD $50,000/QALY, again based on author's figure. Mortality benefit had the least impact on the based case ICER and enhanced cost effectiveness.

Limitations: This study evaluates the impact of KP or VP interventions compared to NSM for treating OVCF in cancer patients, but it only includes one year of data. The authors made assumptions about the proportion of cancer patients in their analysis. Despite examining different cancer types, they did not consider any adverse events, and it remains unclear whether the utility gains extend beyond one year. Although they mentioned varying the discount rate in their sensitivity analysis, the specific results were not provided. Similarly, despite describing their findings as robust, they did not present results for VP versus NSM. The study relied on procedure cost data from a single Ontario hospital, which may limit the generalizability of the findings to the broader Canadian or U.S. healthcare systems. Additionally, the study estimated the total costs of kyphoplasty and vertebroplasty, without adjusting for the number of levels treated per procedure, at CAD $7,240 and $3,870, respectively. These estimates may vary depending on factors such as hospital experience, case complexity, referral patterns, or operating costs.

Stevenson 2014: UK National Institute for Health Research (NIHR)[160]

Study overview: This was the most comprehensive economic analysis identified. This very good-quality economic analysis (HTAS) (QHES 99/100) evaluated the cost-effectiveness of KP or VP as active surgical interventions following a diagnosis of OVCF compared with NSM in England and Wales. The related systematic review is registered as PROSPERO number CRD42011001822 and was funded by the National Institute for Health Research Health Technology Assessment Programme. This systematic review and cost-effectiveness analysis included data from the following RCTs with patients with painful OVCF: Blasco 2012[18], Buchbinder et al. 2009[24], Farrokhi 2011[64], FREE[187], INVEST[89] Liu 2010[109], Rousing 2009[144], VERTOS[180], and VERTOS II[93]. The proportion of female patients in those RCTs was ~70% with an average age of 70 years old. A mathematical model was used to assess the cost-effectiveness of KP and VP (using low-viscosity cement in 85% of patients and high-viscosity cement in 15% of patients) versus sham, which they term operative placebo with local anesthesia (OPLA) and optimal pain management (OPM).

Base case and sensitivity analyses:

The authors evaluated six scenario analyses including mortality reduction and utility benefits to evaluate the cost-effectiveness of KP and VP treatments versus optimal medical management (OPM. They included extensive one-way sensitivity analyses around numerous input parameters assumptions including age, mortality risk, utilities, OPM, pan and function and use of bisphosphonates. (See Appendix O) One-way, univariate analyses did not alter the authors' overall conclusions for most parameters. However, they found that the assumption of a mortality benefit significantly influenced the relative cost-effectiveness of the treatment. Additionally, the source of utility values for EQ-5D—whether mapped from VAS or obtained directly from the trials also impacted the results.

Authors evaluate six scenarios that included data from FREE trial[19,187] and two sham-controlled trials, Buchbinder et al [24] and INVEST[89]. Details of sensitivity analyses for each scenario and reported the corresponding ICERs are in Appendix O.

Overall, this study did not reach a definitive conclusion regarding whether or not KP or VP is/are cost-effective as such a conclusion is tied to assumptions chosen in the analyses. The authors reported that:

- If differential mortality effects with KP being more effective than VP were assumed, then KP is a more cost-effective treatment at the GBP £20,000 WTP per QALY gained (scenarios 1 & 2).
- If differential mortality effects of KP and VP were identical, with OPM providing half the benefit, VP was dominating KP at the GBP £10,000 WTP per QALY gained (scenarios 3 & 4).
- If OPM, KP, and VP have identical mortality benefits, then OPM dominates VP when VP costs are higher than OPM and when QALY losses due to AEs for VP and the EQ-5D data from the RCTs were used.

- Wherever OPM was not deemed to be a comparator, VP is more cost-effective than OPM at the GBP £10,000 WTP per QALY gained.
- In scenarios 5 & 6, assuming equal hospitalization costs for all interventions, VP has a cost per QALY gained higher than the standard GBP £20,000 WTP.

Limitations: Authors mentioned that the RCTs and systematic reviews used for data were specific to patients with painful OVCF and, therefore, those results cannot be generalized to other VCFs such as myeloma, traumatic, and/or metastatic deposits. Authors note that while vertebral augmentation may lead to decreased mortality, the data for this is from administrative data (registry) and that causal inference is not possible given lack of detailed information on causes of death.

Fritzell 2011 Sweden[72]

Study overview: This industry-funded, Fair-quality study (QHES 79/100) evaluated the cost-effectiveness of active surgical intervention (with KP) following a diagnosis of OVCF (costing year: 2008) compared with UC in a Sweden setting from a societal perspective and health care perspective. Given a SEK 600,000 WTP threshold, the authors concluded that KP was not cost-effective compared to UC in patients with OVCF.

Base case and sensitivity analyses: This study included Swedish patients (n=63) from the FREE trial and utilized patient-reported EQ-5D values after 24 months. The difference in QALYs gained over this period between the treatment and control groups was 0.085 (95% CI -0.132 to 0.306) in favor of KP. There was no difference in indirect costs between the groups, as all patients were on pension. The authors also considered adverse events, such as infection in the index-cemented vertebra, and adjusted costs accordingly in their analyses.

The base case ICER in this study was at SEK 884,682 (€92,154 and US $134,043) per QALY gained for undergoing KP vs UC. The uncertainty in the ICER estimate was assessed through bootstrapping and presented in a CEAC. At an ICER of SEK 884,682, the CEAC indicates a 50% probability that KP is cost-effective compared to UC. However, with a willingness-to-pay (WTP) threshold set at SEK 600,000 (€62,500 and US $90,910) per QALY gained, the probability of KP being cost-effective drops to less than 40%.

Sensitivity analysis, after varying costs, resulted in a cost per QALY gained ranging from SEK 622,800 (€64,875 and US $94,364) to SEK 745,812 (€77,689 and US $113,002). Adjusting the QALY benefit from 0.085 to 0.21 (based on the FREE trial) led to a cost per QALY gained of SEK 359,146 (€37,411 and US $54,416), potentially making KP cost-effective compared to UC.

Limitations: The authors acknowledged the potential for selection bias, as the treatment could not be masked, possibly influencing patients' responses. The analysis covered only a short time period and relied primarily on 'cost diaries' to measure and report costs, a method susceptible to manual and human errors. Additionally, the authors noted that the inclusion and exclusion criteria might have been too restrictive, potentially affecting the analysis results. A further issue is that the model, along with its assumptions and limitations, was not clearly described or justified, nor was there sufficient explanation of how the data were parameterized within the model. The authors emphasized that these findings should be interpreted within the context of the Swedish healthcare system and cannot be generalized to the US healthcare perspective. One final issue may result from a small sample size: with only 63 patients, outliers (particularly those identified as having high costs due to complications and additional procedures) may be disproportionately influencing the overall cost-effectiveness results.

Takahashi et al. (2019) Japan[167]

Study Overview: This fair-quality study (QHES 77/100) evaluated the cost-utility of KP following a diagnosis of OVCF compared with NSM in an elderly population in Japan from a healthcare perspective over a lifetime time horizon. In this cost utility analysis (costing year: 2018), the authors used a Markov simulation model and a propensity-score matched analysis (i.e., 71 matched patients with a propensity score match tolerance of 0.015). They compared outcomes between the KP and NSM groups to estimate the probability of undergoing KP using patient characteristics such as age, gender, number of baseline old fractures, fracture level, and SF-6D score for QALY. Inclusion criteria included presence, severity, and duration of pain (i.e., VAS pain score greater or equal to 4 and T scores less or equal to -1). Cost data in this analysis came from published literature and prior studies from the authors. All direct costs and outcomes have been discounted at 3.5%. Mortality data was sourced from the Japanese government (i.e., Statistics and Information Department of the Minister's Secretariat, Ministry of Japanese Health, Labor, and Welfare). Health utilities came from a SF-6D survey. The study was not based on RCT data with some data coming the author's institution and some from historical controls; they cite two [115,173] cohorts used for propensity score matching for NSM data.

Base case and sensitivity analyses:

100% of the patients in the treatment group required hospitalization while 66% of the patients in the control group were hospitalized (*p<0.001*). The difference in the duration of hospitalization in the KP vs NSM groups was also statistically significant (*p<0.001*). There is no statistically significant difference in the 71 matched patients' age (*p=0.456*), gender (*p=1*), level of fracture (*p=0.068*), baseline prevalent fractures (*p=0.978*), duration of back pain (*p=0.320*), bone density measured in T-score (*p=0.665*), osteoporosis medication before injury (*p=0.603*).

The difference in costs, in favor of NSM, was 402,988 JPY (*p=0.001*) and the gains in QALY at 6-month follow-up, in favor of KP, was 0.033 (i.e., 0.153 for KP vs 0.120 for NSM). The incremental cost over the incremental QALY benefit represents the ICER. In their analysis, the authors divided the patients into three subgroups based on the patients' age to test differences in ICER per QALY gained. Prior to any sensitivity analysis, a smaller ICER was achieved over a longer time horizon (i.e., 4,404,158 JPY for a 3-year time period vs 2,416,406 JPY for a 20-year time period). Both the 3-year and 20-year ICERs are the base case scenario in this study. The authors reported a cost-effectiveness of KP vs NSM in patients with OVCF in Japan aged between 60 and 79 years old, but not in patients > 80 years old. ICERs for three and 20 years were 4,404,158 JPY and 2,416,406 JPY, respectively.

The authors considered 5,000,000 JPY (i.e., £33,404) to be an acceptable WTP. They evaluated the magnitude of ICERs when varying model inputs for mortality reduction from 50% to 100% (i.e., full mortality reduction), QALY benefits, use of fracture prevention medication (i.e., bisphosphonate), and 0% to 7% discount rates. The sensitivity analyses show an overall range of ICERs from 652,181 JPY to 4,896,645 JPY (£4,418 – £33,168):

- Sensitivity analyses with a 100% and a 50% mortality reduction reported lower ICERs (i.e., 871,450 and 1,202,067 JPY, respectively).
- A 50% reduction in bisphosphonate use combined with mortality reduction achieved a slightly higher ICER (i.e., 897,668 JPY) than an ICER with a full mortality reduction alone (i.e., 871,450 JPY).
- With a 50% QALY benefit, the ICER increased to 4,896,645 JPY (£33,168).

- Absent any mortality reduction, the difference in ICER when the discount rate changed from 0% and 7% was relatively small (i.e., 2,349,185 JPY and 2,529,388 JPY, respectively).
- On the other hand, with a mortality reduction, the ICER increased nearly twofold when the discount rate varied from 0% to 7%.
- According to the authors, the best scenario achieved is an ICER of 652,181 JPY (£4,418) with a 44% mortality reduction in KP.

The cost-effectiveness acceptability curve (CEAC) indicates that a 50% probability for KP to be cost-effective is associated with a cost of 1,121,453 JPY (£7,596). From the CEAC, KP is a cost-effective treatment with a probability higher than 80% for that same level of WTP.

Limitations The authors acknowledged the lack of data supporting the cost-effectiveness of KP (vs. NSM) in patients aged 80 years or older with OVCF, noting that their sample consisted exclusively of female patients. They also highlighted that the observed mortality reduction benefit could be influenced by confounders or unobserved treatment biases. While the authors recognized the risk of adverse events, such as cement leakage, these were not included in their analysis. Additionally, the study raises concerns about relying solely on the ICER as a unique and reliable measure of cost-effectiveness in an elderly population. Further limitations include a lack of justification in model parameters and data sources, and no discussion on biases or their effect on analyses. In particular, a 'super-aging' society is unlikely to be generalizable to other populations.

Svedbom et al. (2013), UK[165]

Study Overview: This industry-funded, good-quality study (QHES 84/100) evaluated the cost-effectiveness of active surgical intervention with balloon kyphoplasty (KP) following a diagnosis of osteoporotic vertebral compression fracture (OVCF) (costing year: 2009) compared with NSM and VP in a UK setting.

In this cost utility analysis, the authors built a Markov model to estimate the cost utility of KP compared to NSM and VP (VP compared to NSM is not explicitly modeled). Markov models allow for transitions between health states over time. This is an extension and revision of the 2010 Strom study (i.e., including VP); the latter had been included in our prior HTA. This study was updated to include VP as an intervention and includes the potential beneficial effect of KP and VP on mortality.

This cost-effectiveness study relied on functional outcomes data from previous randomized trials (i.e., FREE[19,187] and VERTOS II[93] with duration of 2 and 1 year, respectively) and compared KP with NSM and VP. Cost data in this analysis came from published literature (e.g., general practitioner, referral costs, analgesics, cost per bed day) and from the National Health Service (NHS). All costs have been discounted at 3.5%. Health utilities were derived from these previous randomized trials. The NSM health utilities during each cycle (from cycle 1 to cycle 7) were lower than those for KP and VP. The health utilities for KP were similar to those for VP for each cycle. Average age (70) of the all-female population in the base case was also derived from the mean age in these previously mentioned RCTs. Mortality hazard ratios for KP, VP and NSM modeled in the sensitivity analysis were from published literature.

The authors estimated that KP was cost effective with cost per QALY gained of GBP £2,706 and £15,982 compared to NSM and VP, respectively. The authors noted that this measure fell within the threshold range of GBP £20,000 to £30,000 in the UK for the willingness to pay for a QALY.

Base case and sensitivity analyses: The base-case population was an all-female population of 70-year-old UK patients with at least one vertebral fracture (requiring hospitalization) and a T-score of -3 (a

clinical measure of bone density where -2.5 or lower indicates osteoporosis); health states of additional fracture or death were possible every six months until age 100 or death. The primary impact of kyphoplasty was assumed to be through improvements in quality of life, so the first 12 months of available data from the FREE trial were used.

The base case analysis indicated increased cost (GBP £9,313) and increased quality-adjusted life years (QALY) (5.473) for KP compared with NSM and VP, with an incremental cost-effectiveness ratio (ICER) of GBP £2,706 and £15,982 per QALY gained, respectively.

The authors did a one-way sensitivity analysis to evaluate KP vs VP and KP vs VP and NSM using modeling inputs (i.e., mortality rates (hazard ratios), QALY levels, risk of fracture, medication, costs, age, and discount rate). Sensitivity analysis showed evidence of the cost-effectiveness of KP vs NSM when mortality and QALY inputs were varied. Sensitivity analysis indicated that the ICER for KP vs NSM continued to be less than GBP £30,000/QALY under variations of patient age (60-80 years), QALY benefit (25%, 50%, 75% QALY benefit), mortality benefit (25%, 50%, 75% mortality benefit), discount rates (0% to 7%), relative risk of fracture, bisphosphonate treatment, and hospital length of stay benefit (0-9 days difference from NSM).

The authors ran a probabilistic sensitivity analysis (1,000 simulations) using the model inputs (i.e., QALY, mortality reduction, length of stay in hospital) under six scenarios, including the base case scenario. The cost effectiveness acceptability curves show that, given the base case inputs, KP had a probability of ~60% and ~75% of being the optimal intervention at a WTP threshold of GBP £20,000 and £30,000 per QALY, respectively. When considering six different scenarios and when comparing KP vs NSM and VP, the PSA results indicate that:

- In the base scenario, KP had a probability of ~60% of being the optimal intervention at the WTP threshold of GBP £20,000.
- Excluding mortality benefit, NSM had a probability of ~80% of being the optimal intervention at the WTP threshold of less than GBP £20,000.
- When comparing KP to NSM alone and when excluding the mortality benefit, KP had a probability greater than 80% of being the optimal intervention at the WTP threshold of GBP £20,000.
- When comparing KP vs NSM, excluding mortality benefit and after removing ¾ of QALY gain, there is uncertainty, KP is cost-effective at the GBP £30,000 WTP threshold with a 60% probability.
- Excluding reduction in length of stay in hospital and QALY benefit, in both scenarios, KP had a probability of ~60% of being the optimal intervention at the WTP threshold of GBP £20,000.
- In the scenario where mortality reduction was removed, VP had the highest probability of being the optimal intervention given any level of cost-effectiveness threshold (i.e., from GBP £0 to £30,000).

Between the one-way sensitivity analysis and the PSA, the authors found that, across the six scenarios, the maximum difference between the deterministic and mean probabilistic ICERs was 6%.

Limitations of this study include lack of blinding for outcomes assessment and a potential that results may at least in part be due to placebo effect. The base case assumed a 12-month benefit in QALY and a 6-day improvement in hospital length of stay for kyphoplasty. Fracture incidence and mortality were

modeled from UK and Swedish registry data; cost inputs were from published literature and U.K. NHS data. The main limitations are the lack of available long-term data on QALY, effectiveness, or complications and mortality associated with KP. The use of a 100-year time horizon may not be realistic. The study sample consists only of female patients. The authors' unclear presentation of utility at less than 12 months and lack of presentation of sensitivity analysis of costs are limitations, especially for assessing generalizability to a US health care setting. From the one-way analysis and PSA, the authors concluded that kyphoplasty was a cost-effective intervention, but this should be revisited as additional evidence becomes available. The mortality reduction with KP and VP were not obtained from a RCT but from a US retrospective study conducted in an inpatient and outpatient settings and then used in an outpatient setting in the UK. The health systems and treatment pathways might be different in the US vs in the UK. Some results were discussed in the study but not shown (e.g., impact of a longer offset period of KP and VP over NSM and KP over VP). Adverse events were not evaluated by the authors.

5 Strength of Evidence (SOE)

5.1 Strength of Evidence Summary: Osteoporotic Vertebral Compression Fractures

5.1.1 Strength of Evidence Summary: Effectiveness from Vertebroplasty versus Sham in Patients with Osteoporotic Vertebral Compression Fractures

Outcome*	Time	Studies (N)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. Sham Effect estimate (95% CI) Conclusion	Quality (SoE)
Pain Response ≥30% reduction in pain from baseline	<1 week	1 RCT (N=113) Clark 2016	No	Unknown	No	Yes (-1)	31% (18/58) vs. 8.5% (5/55); RR 3.41 (1.36 to 8.56) Conclusion: Large likelihood of pain improvement with VP	⊕⊕◯◯ LOW
	≥1 week to ≤2 weeks	2 RCTs (N=186) Buchbinder 2009, Clark 2016	No	No	No	Yes (-1)	41% (38/92) vs. 27.7% (26/94) RR 1.44 (0.60 to 3.47), I^2=0% Conclusion: Similar likelihood of response	⊕⊕⊕◯ MODERATE
	>2 weeks to ≤1 month	3 RCTs (N= 313) Buchbinder 2009, Kallmes 2009, Clark 2016	No	No	No	Yes (-1)	57.7% (89/157) vs. 35.2% (55/156) RR 1.48 (0.95 to 2.86), I^2=0% Conclusion: Moderate likelihood of pain improvement with VP vs. sham	⊕⊕⊕◯ MODERATE
	>1 month to <6 months	2 RCTs (N= 176) Buchbinder 2009, Clark 2016	No	No	No	Yes (-1)	54.5% (48/88) vs. 34% (30/88) RR 1.60 (1.06 to 2.38), I^2=0% Conclusion: Small likelihood of pain improvement with VP vs. sham	⊕⊕⊕◯ MODERATE
	≥6 months to <12 months	2 RCTs (N=171) Buchbinder 2009, Clark 2016	No	No	No	Yes (-1)	63.5% (54/85) vs. 45.3% (39/86) RR 1.40 (0.99 to 1.94), I^2=0%	⊕⊕⊕◯ MODERATE

							Conclusion: Small increase in likelihood of pain improvement with VP vs. sham	
	≥12 months	3 RCTs (N=339) Comstock 2013, Firanescu 2018 Kroon 2014	No	No	No	Yes (-1)	70.5% (124/176) vs. 51.5% (84/163) RR 1.36 (1.08 to 1.66), I^2=0% Conclusion: Small increase in likelihood of pain improvement with VP vs. sham	⊕⊕⊕◯ MODERATE
Pain (0-10 scale)	<1 week	4 RCTs (N = 500) Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023	No	Yes (-1)	No	Yes (-1)	**Pooled MD:** -0.22 (-1.34 to 0.87), I^2=83.3% Conclusion: Similar pain improvement vs. sham	⊕⊕◯◯ LOW
	≥1 week to ≤2 weeks	6 RCTs (N=616) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Hansen 2019 Carli 2023	No	No	No	Yes (-1)	**Pooled MD:** -0.16 (-0.78 to 0.37), I^2=28.5% Conclusion: Similar pain improvement vs. sham	⊕⊕⊕◯ MODERATE
	>2 weeks to ≤1 month	6 RCTs (N = 616) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Hansen 2019	No	No	No	No	**Pooled MD:** -0.62 (-1.07 to -0.18), I^2=0% Conclusion: Small improvement with VP vs. sham	⊕⊕⊕⊕ HIGH

		Carli 2023						
	>1 month to <6 months	6 RCTs (N = 605) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Hansen 2019 Carli 2023	No	No	No	No	**Pooled MD: -0.60 (-1.13, -0.16)** I^2=8.0% Conclusion: Small improvement with VP vs. sham	⊕⊕⊕⊕ HIGH
	≥6 months to <12 months	5 RCTs (N = 550) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023	No	No	No	No	**Pooled MD:** -0.66 (-1.16 to -0.21), I^2=0% Conclusion: Small improvement with VP vs. sham	⊕⊕⊕⊕ HIGH
	≥12 months	5 RCTs (N = 478) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023	No	Yes (-1)	No	Yes (-1)	**Pooled MD**: -0.30 (-1.17 to 0.62), I^2=61.1% Conclusion: Similar pain improvement with VP vs. sham	⊕⊕◯◯ LOW
Function **RDQ scores (0-24 scale)†**	<1 week	2 RCTs (N = 244) Kallmes 2009 Clark 3018	No	Yes (-1)	No	Yes (-1)	**Pooled MD:** -0.51 (-3.09 to 2.03), I^2=56.1% Conclusion: Similar improvement in function with VP vs. sham	⊕⊕◯◯ LOW
	≥1 week to ≤2 weeks	5 RCTs (N = 531)	No	Yes (-1)	No	Yes (-1)	**Pooled MD**: 0.19 (-0.91 to 1.34), I^2=26.1%	⊕⊕◯◯ LOW

		Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023					Conclusion: Similar improvement in function with VP vs. sham	
	>2 weeks to ≤1 month	5 RCTs (N = 566) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023	No	No	No	Yes (-1)	**Pooled MD**: -1.54 (-2.56 to -0.55), I^2=0% Conclusion: Small improvement in function with VP vs. sham	⊕⊕⊕◯ MODERATE
	>1 month to <6 months	5 RCTs (N = 557) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023	No	Yes (-1)	No	Yes (-1)	**Pooled MD**: -1.16 (-2.50 to 0.18), I^2=36.7% Conclusion: Similar improvement in function with VP vs. sham	⊕⊕◯◯ LOW
	≥6 months to <12 months	5 RCTs (N = 548) Buchbinder 2009 Kallmes 2009 Clark 3018 Firanescu 2018 Carli 2023	No	Yes (-1)	No	Yes (-1)	**Pooled MD**: -1.47 (-2.87 to -0.17), I^2 = 30.6% Conclusion: Small improvement in function with VP vs. sham	⊕⊕◯◯ LOW
	≥12 months	4 RCTs (N = 432) Buchbinder 2009	No	Yes (-1)	No	Yes (-1)	**Pooled MD** -0.02 (-1.54 to 1.52), I^2=11.5%	⊕⊕◯◯ LOW

		Kallmes 2009 Firanescu 2018 Carli 2023					Conclusion: Similar improvement in function	

CI = confidence interval; MD = mean difference; RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; RR = risk ratio; SD = standard deviation; SoE = strength of evidence; VP = vertebroplasty.
*VAS and RDQ are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (VP).
†Modified RDQ scaled/converted to 0-24 scale.

5.1.2 Strength of Evidence Summary: Safety from RCTs of Vertebroplasty versus Sham in Patients with Osteoporotic Vertebral Compression Fractures

Outcome	Time	Studies (N)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. Sham Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	At last follow-up	5 RCTs (N= 289) Clark 2016 Comstock 2013 Firanescu 2018 Carli 2023 Kroon 2014	No	No	No	Yes (-1)	6.0% (18/298) vs. 6.9% (20/291) RR 0.92 (0.46 to 1.71), I^2 = 0% Conclusion: Similar mortality in each group; Results were similar when mortality was analyzed by timepoint.	⊕⊕⊕◯ MODERATE
Any New Vertebral Fracture	By last follow-up	4 RCTs (N=408) Clark 2016 Firanescu 2018 Carli 2023 Staples 2015	No	No	No	Yes (-1)	20.3% (4/207) vs. 18.4% (37/201) RR 1.10 (0.68 to 1.88), I^2 = 0% Conclusion: Similar risk of new fracture	⊕⊕⊕◯ MODERATE
Any, Symptomatic with Bone Edema	By last follow-up	1 RCT (N=34) Firanescu 2018	No	Unknown	No	Yes (-1)	40% (6/15) vs. 31% (6/19) RR 0.84 (0.38 to 1.84) Conclusion: Similar frequency between groups	⊕⊕◯◯ LOW

Outcome	Time	Studies (N)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. Sham Effect estimate (95% CI) Conclusion	Quality (SoE)
Any Serious Adverse Events	Across time frames, by last follow-up	4 RCTs (N=409) Kallmes 2009 Buchbinder 2009 Clark 2018 Carli 2023	No	No	No	Yes (-1)	1.9% (4/207) vs. 2.0% (4/202) RR 0.98 (0.26 to 3.66) I^2 = 0% <u>Conclusion</u>: Similar frequency of serious adverse event in each group	⊕⊕◯◯ LOW
Cement Leakage (symptomatology not reported)	Any	3 RCTs (N=232 treated levels) Carli 2023 Firanescu 2018 Buchbinder, 2009	No	Yes -1	no	Yes (-1)	Range across studies was 40% (to 91% of treated levels. 1 RCT (Firenescu) reported: location/type: Type 1 = disc above treated level (20%) Type 2 = disc under treated level (15%) Type 3 = perivertebral tissue (10%) Type 4 = perivertebral veins (39%) Type 5 = pulmonary (7%) Type 6 = spinal canal (8%) A fourth RCT reported than none were observed (Hansen 2019) <u>Conclusion</u>: Cement leakage is common	⊕⊕⊕◯ MODERATE

CI = confidence interval; RCT = randomized controlled trial; RR = risk ratio; SoE = strength of evidence; VP = vertebroplasty.

5.1.3 Strength of Evidence Summary: Effectiveness for Vertebroplasty versus Usual Care in Patients with Osteoporotic Vertebral Compression Fractures

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
Pain Responders (<4 on 0-10 VAS)	12 months	1 RCT (N=95) Blasco 2012	Yes (-1)	Unknown	No	Yes (-1)	44.7% (21/47) vs. 47.9% (23/48); RR 0.93 (0.60 to 1.44) Conclusion: Similar likelihood of achieving pain response pain for VP vs. UC.	⊕◯◯◯ INSUFFICIENT
Pain Responders (Complete Pain Relief)	12 months	1 RCT (N=89) Chen 2014	Yes (-1)	Unknown	No	Yes (-1)	84.8% (39/46) vs. 34.9% (15/43); RR 2.43 (1.59 to 3.72) Conclusion: Evidence is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Pain Scores (VAS/ NPRS (0-10))	<1 weeks	3 RCTs (N=343) Klazen 2010 Voormolen 2007 Yang 2016	Yes (-1)	No	No	No	**Pooled MD:** -2.84 (-3.47 to -2.06), I^2=0% Conclusion: Large improvement in pain with VP vs. UC.	⊕⊕⊕◯ MODERATE
	≥1 week to ≤2 weeks	5 RCTs (N=557) Blasco 2012 Chen 2014 Klazen 2010 Voormolen 2007 Yang 2016	Yes (-1)	No (excluding outlier)	No	Yes (-1)	**Pooled MD:** -1.22 (-2.80 to 0.21), I^2 = 81.3% Excluding potential outlier (Blasco), N=432: MD -1.99 (95% CI -2.61 to -1.26), I^2=0% Conclusion: Moderate improvement in pain with VP vs. UC after excluding outlier trial.	⊕⊕◯◯ LOW (excluding outlier)

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
	>2 weeks to ≤1 month	3 RCTs (N=398) Chen 2014 Klazen 2010 Yang 2016	Yes (-1)	No	No	Yes (-1)	**Pooled MD:** -2.28 (-3.20 to -1.00), I^2 = 30.5% Conclusion: Large improvement in pain with VP vs. UC.	⊕⊕◯◯ LOW
	>1 month to <6 months	5 RCTs (N=569) Blasco 2012 Chen 2014 Klazen 2010 Rousing 2009 Yang 2016	Yes (-1)	No	No	No	**Pooled MD:** -1.17 (-1.71 to -0.60), I^2 = 0% Conclusion: Moderate improvement in pain with VP vs. UC.	⊕⊕⊕◯ MODERATE
	≥6 months to <12 months	4 RCTs (N=523) Blasco 2012 Chen 2014 Klazen 2010 Yang 2016	Yes (-1)	Yes (-1)	No	Yes (-1)	**Pooled MD:** -0.89 (-2.20 to 0.34), I^2 = 71.9% Conclusion: Similar improvement in pain with VP vs. UC.	⊕⊕◯◯ LOW
	≥12 months	5 RCTs (N=567) Blasco 2012 Chen 2014 Klazen 2010 Rousing 2009 Yang 2016	Yes (-1)	Yes (-1)	No	Yes (-1)	**Pooled MD:** -1.10 (-2.08 to -0.12), I^2 = 63.4% Conclusion: Moderate improvement in pain with VP vs. UC.	⊕⊕◯◯ LOW
Function scores (RDQ [0-24], ODI [0-100] and DPQDA [0-100])	≥1 week to ≤2 weeks	4 RCTs (N=432) Chen 2014 Klazen 2010 Voormolen 2007 Yang 2016	Yes (-1)	Yes (-1)	No	No	**SMD:** -0.37 (-0.61 to -0.17), I^2=11.5% Conclusion: Small improvement in function with VP vs. UC.	⊕⊕◯◯ LOW

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
	>2 weeks to ≤1 month	3 RCTs (N=398) Chen 2014 Klazen 2010 Yang 2016	Yes (-1)	No	No	No	**SMD:** -0.29 (-0.50 to -0.08), I^2=0% Conclusion: Small improvement in function with VP vs. UC.	⊕⊕⊕◯ MODERATE
	>1 month to <6 months	4 RCTs (N=440) Chen 2014 Klazen 2010 Rousing 2009 Yang 2016	Yes (-1)	No	No	No	**SMD:** -0.38 (-0.57 to -0.18), I^2=0% Conclusion: Small improvement in function with VP vs. UC.	⊕⊕⊕◯ MODERATE
	≥6 months to <12 months	3 RCTs (N=398) Chen 2014 Klazen 2010 Yang 2016	Yes (-1)	No	No	No	**SMD:** -0.29 (-0.50 to -0.09), I^2=0% Conclusion: Small improvement in function with VP vs. UC.	⊕⊕⊕◯ MODERATE
	≥12 months	4 RCTs (N=436) Chen 2014 Klazen 2010 Rousing 2009 Yang 2016	Yes (-1)	No	No	No	**SMD:** -0.26 (-0.46 to -0.06), I^2=0% Conclusion: Small improvement in function with VP vs. UC.	⊕⊕⊕◯ MODERATE

CI = confidence interval; DPQDA = Dallas Pain Questionnaire Daily Activities; MD = mean difference; mos = months; ODI = Oswestry Disability Index; RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; RR = risk ratio; SD = standard deviation; SMD = standardized mean difference; SoE = strength of evidence; UC = usual care; VP = vertebroplasty.
*VAS and all function scores are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (VP).

5.1.4 Strength of Evidence Summary: Safety from RCTs of Vertebroplasty versus Usual Care in Patients with Osteoporotic Vertebral Compression Fractures

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	Latest follow-up (6-12 mos.)	5 RCTs (N=844) Blasco 2012 Farrokhi 2011 Klazen 2010 Leali 2016 Rousing 2009 Yang 2016	Yes (-1)	No	No	Yes (-1)	3.1% (13/416) vs. 4.2% (18/428); RR 0.72 (0.35 to 1.50), I^2 = 0% Conclusion: Similar risk of mortality for VP vs. UC at latest follow-up. Results were similar at earlier timepoints, but estimates were more imprecise.	⊕⊕⊕◯ MODERATE
Any new vertebral fracture	Latest follow-up (2 wks. to mean 49 mos.)	9 RCTs (N=830) Blasco 2012 Chen 2014 Farrokhi 2011 Klazen 2010 Leali 2016 Rousing 2010 Voormolen 2007 Yang 2016 Yi 2014	Yes (-1)	Yes (-1)	No	Yes (-1)	10.1% (62/611) vs. 10.5% (67/638); RR 0.96 (0.59 to 1.64), I^2 = 43.4% Conclusion: Similar risk of any new vertebral fracture for VP vs. UC at latest follow-up. Results were similar in sensitivity analyses that excluded poor-quality trials and restricted to trials with ≥12 months follow up.	⊕⊕◯◯ LOW
Any New Symptomatic Vertebral Fracture	Latest follow-up (2 wks. to 12 mos.)	6 RCTs (N=877) Blasco 2012 Farrokhi 2011 Leali 2016 Rousing 2010 Voormolen 2007 Yi 2014	Yes (-1)	Yes (-1)	No	Yes (-1)	6.5% (27/418) vs. 5.9% (27/459); RR 1.24 (0.26 to 6.55), I^2 = 63.5% Conclusion: Similar risk of any new symptomatic vertebral fracture for VP vs. UC at latest follow-up. Results were similar in	⊕⊕◯◯ LOW

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
							sensitivity analyses that excluded poor-quality trials, excluded an outlier trial, and restricted to trials with ≥12 months follow up.	
Serious Adverse Events	Any time	4 RCTs (N=408) Farrokhi, 2011 Rousing, 2009/2010 Yang, 2016 Yi, 2014	Yes (-1)	Unclear	No	Yes (-1)	SAE (unspecified): 2 RCTs (N=261) stated that no SAEs occurred in either treatment group over 12 to a mean 49 months. DVT/thrombophlebitis: 1 RCT, 3.6% (2/56) vs. 7.8% (4/51) at 12 months; RR 0.46 (0.09 to 2.38) Epidural cement leakage causing LE pain and weakness (required reoperation): 2.5% (1/40) at 1 week (NA for UC group) Conclusion: SAEs were poorly reported and occurred with similar frequency between groups were reported.	⊕⊕◯◯ LOW

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
Reoperation	Any time	3 RCTs (N=269) Farrokhi, 2011 Voormolen, 2007 Yi, 2014	Yes (-1)	Unclear	No	Yes (-1)	Range across VP arms, 3 RCTs (n range, 18 to 90): 2.5% to 11.1% Reoperation reasons included symptomatic new fractures or cement leak causing LE pain VP for symptomatic new fractures, 1 RCT, VP vs. UC: 10.0% (9/90) vs. 9.1% (11/121), RR 1.10 (0.48 to 2.54) <u>Conclusion</u>: Similar risk of reoperation for symptomatic new VCF with VP vs. UC in one trial.	⊕⊕◯◯ LOW
Cement Leakage	Any time	7 RCT (n=varies) Blasco, 2012 Chen, 2014 Farrokhi, 2011 Klazen, 2010 Rousing, 2009/2010 Yang, 2016 Yi, 2014	Yes (-1)	No	No	Yes (-1)	Symptomatic cement leakage: range, 0% to 1% across 7 RCTs (range of levels, n=63 to 140 across 6 RCTs; NR by 1 RCT); one symptomatic leakage reported in 1 RCT (1%, 1/100 levels) Asymptomatic cement leakage range, 13.0% to 72.4% across 5 RCTs (range of levels, n=65 to 140); range, 49.3 % to 72.4% across the 3 fair-quality RCTs	⊕⊕⊕◯ MODERATE

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. UC Effect estimate (95% CI) Conclusion	Quality (SoE)
							(range of levels, n=69 to 140); 2 RCTs (63 levels; 90 patients [levels NR]) reported that all fractures were asymptomatic but did not provide clear data Conclusion: Symptomatic cement leakage appears to be rare following VP while asymptomatic leakage is common.	

CI = confidence interval; NR = not reported; RCT = randomized controlled trial; RR = risk ratio; SoE = strength of evidence; UC = usual care; VCF = vertebral compression fracture; VP = vertebroplasty.

Strength of Evidence Summary: Effectiveness for Vertebroplasty versus Kyphoplasty in Patients with Osteoporotic Vertebral Compression Fractures

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Pain Responders (total effective rate)	12 months	1 RCT (N=100) Wang 2023	Yes (-1)	Unknown	No	Yes (-1)	Total effective rate: complete ("cure"), excellent or effective (not defined) improvement in clinical symptoms: 74% (37/50) vs. 94% (47/50), RR 0.79 (0.66 to 0.94) Complete or excellent improvement: 56% (28/50) vs. 74% (37/50); RR 0.76 (0.56 to 1.02) Conclusion: Evidence from one poor-quality trial is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Pain Scores (VAS/ NPRS (0-10))	<1 week	3 RCTs (N=313) Evans 2016 Liu 2010 Wang 2015	Yes (-1)	No	No	No	**Pooled MD:** -0.15 (-0.42 to 0.19), I^2 = 44.6% Conclusion: Similar improvement in pain with VP vs. KP	⊕⊕⊕◯ MODERATE
	≥1 week to ≤2 weeks	1 RCT (N=377) Dohm 2014	Yes (-1)	Unknown	No	Yes (-1)	**MD** -0.25 (-0.81 to 0.31) Conclusion: Evidence from one poor-quality trial is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
	>2 weeks to ≤1 month	3 RCTs (N=560) Dohm 2014 Evans 2016 Wang 2023	Yes (-1)	No (excluding outlier)	No	No (excluding outlier)	**Pooled MD:** 0.35 (-0.60 to 1.24), I^2 = 86.5% Excluding potential outlier (Wang 2023), N=460: MD -0.08 (95% CI -0.58 to 0.41), I^2 = 0% <u>Conclusion</u>: Similar improvement in pain with VP vs. KP	⊕⊕◯◯ LOW (Excluding outlier)
	>1 month to <6 months	3 RCTs (N=519) Dohm 2014 Wang 2015 Wang 2023	Yes (-1)	No (excluding outlier)	No	No (excluding outlier)	**Pooled MD:** 0.46 (-0.43 to 1.26), I^2 = 93.1% Excluding potential outlier (Wang 2023), N=419: MD 0.14 (95% CI -0.29 to 0.45), I^2=0%; N=419 <u>Conclusion</u>: Similar improvement with VP vs. KP	⊕⊕◯◯ LOW (Excluding outlier)
	≥6 months to <12 months	3 RCTs (N=248) Endres 2012 Evans 2016 Liu 2010	Yes (-1)	No	No	Yes (-1)	**Pooled MD:** -0.07 (-0.55 to 0.18), I^2 = 29.0% <u>Conclusion</u>: Similar improvement in pain with VP vs. KP	⊕⊕◯◯ LOW
	≥12 months	5 RCTs (N=673) Dohm 2014 Evans 2016 Griffoni 2020 Liu 2015 Wang 2015	Yes (-1)	No	No	Yes (-1)	**12 months (5 RCTs)** **Pooled MD:** 0.08 (-0.12 to 0.30), I^2 = 0% **24 months (2 RCTs, N=320)** **Pooled MD:** -0.16 (-0.67 to 0.42), I^2 = 0%	⊕⊕◯◯ LOW (12, 24 months ⊕◯◯◯ INSUFFICIENT (60 months)

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
							60 months (1 poor-quality RCT, N=100) **MD**: -0.60 (-1.13 to -0.07) Conclusion: Similar improvement in pain with VP vs. KP at 12 and 24 months. Evidence at 60 months from one poor-quality trial is insufficient to draw conclusions.	
Function scores (RDQ [0-24], ODI [0-100])	<1 week	1 RCT (N=106) Evans 2016	Yes (-1)	Unknown	No	Yes (-1)	**MD:** -0.32 (-3.20 to 2.56), RDQ Conclusion: Similar improvement in function with VP vs. KP	⊕⊕◯◯ LOW
	>2 weeks to ≤1 month	4 RCTs (N=652) Dohm 2014 Evans 2016 Wang 2018 Wang 2023	Yes (-1)	Yes (-1)	No	Yes (-1)	**SMD**: 0.67 (-0.46 to 1.83), I^2=95.8% Excluding potential outlier (Wang 2023), N=552: **SMD** 0.13 (95% CI -0.28 to 0.63), I^2=73.9% Conclusion: Evidence is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
	>1 month to <6 months	3 RCTs (N=499) Dohm 2014 Wang 2015 Wang 2023	Yes (-1)	No [excluding outlier]	No	No [excluding outlier]	**SMD:** 1.20 (-1.27 to 3.70), I^2=97.9% Excluding potential outlier (Wang 2023), N=399: **SMD** 0.14 (-0.11 to 0.38), I^2=0%	⊕⊕◯◯ LOW [excluding outlier]

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
							Conclusion: Similar improvement in function with VP vs. KP	
	≥6 months to <12 months	3 RCTs (N=238) Endres 2012 Evans 2016 Wang 2018	Yes (-1)	No	No	No	**SMD:** 0.26 (-0.10 to 0.63), I^2=25.7% Conclusion: Similar improvement with VP vs. KP	⊕⊕⊕◯ MODERATE
	≥12 months	5 RCTs (N=643) Dohm 2014 Evans 2016 Griffoni 2020 Wang 2015 Wang 2028	Yes (-1)	No (12 months) Unknown (24 months)	No	Yes (-1) (12 months) Yes (-1) (24 months)	12 months: **SMD**: 0.17 (-0.06 to 0.49), I^2=50.2% 24 months (1 RCT, N=201): **MD**: -1.00 (-6.88 to 4.88) Conclusion: Similar improvement in function with VP vs. KP at 12 months. Evidence at 24 months from one poor-quality trial is insufficient to draw conclusions.	⊕⊕◯◯ LOW (12 months) ⊕◯◯◯ INSUFFICIENT (24 months)

CI = confidence interval; MD = mean difference; NPRS = numerical pain rating scale; ODI = Oswestry Disability Index; RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; RR = risk ratio; SD = standard deviation; SMD = standardized mean difference; SoE = strength of evidence; UC = usual care; VP = vertebroplasty.
*VAS and ODI and RDQ are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (VP).

5.1.5 Strength of Evidence Summary: Safety from RCTs of Vertebroplasty versus Kyphoplasty in Patients with Osteoporotic Vertebral Compression Fractures

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	Latest follow-up (12-24 mos.)	4 RCTs (N=631) Dohm 2014 Endres 2012 Vogl 2013 Wang 2015	Yes (-1)	No	No	Yes (-1)	3 RCTs, N=565, 12-24 mos.: 8.9% (24/271) vs. 7.1% (21/294), RR 1.24 (0.56 to 2.38), I^2=0% 1 RCT: 3.0% (2/66), 6 mos.; NR by group [Endres 2012] Conclusion: Similar risk of mortality with VP vs. KP up to 24 months. Results were similar in analysis at 3 months.	⊕⊕◯◯ LOW
Any new vertebral fracture	Latest follow-up (12 to a mean 49 mos.)	6 RCTs (N=781) Dohm 2014 Griffoni 2020 Liu 2015 Vogl, 2013 Wang, 2015 Yi, 2014	Yes (-1)	Yes (-1)	No	Yes (-1)	25.4% (100/393) vs. 20.4% (79/388); RR 1.18 (0.86 to 1.73, I^2=36.4% Conclusion: Similar risk of any new vertebral fracture with VP vs. KP up to a mean of 49 months. Results were similar across sensitivity analyses that stratified by timing and that excluded an outlier trial and for analyses specific to new radiographic fractures and new adjacent level fractures.	⊕⊕◯◯ LOW

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Any New Symptomatic Vertebral Fracture	Latest follow-up (12 mos.)	2 RCTs (N=481) Dohm 2014 Liu 2015	Yes (-1)	Yes (-1)	No	Yes (-1)	11.3% (27/240) vs. 8.7% (21/241), RR 1.23 (0.46 to 3.40), I^2=55.4% Conclusions: Evidence from 2 poor-quality trials is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Refracture or worsening index level fracture	Latest follow-up (12 mos.)	2 RCTs (N=348) Dohm 2014 Vogl 2013	Yes (-1)	Yes (-1)	No	Yes (-1)	6.3% (10/159) vs. 2.6% (5/189), RR 2.24 (0.29 to 8.49), I^2=0% Conclusions: Evidence from 2 poor-quality trials is insufficient to draw conclusions	⊕◯◯◯ INSUFFICIENT
Serious Adverse Events	Various	2 RCTs (N=550) Dohm 2014 Liu 2015	Yes (-1)	Yes (-1)	No	Yes (-1)	1 RCT 30 days: 27.4% (52/190) vs. 26.2% (50/191), RR 1.04 (0.75 to 1.46) 24 months: 65.8% (125/190) vs. 65.5% (125/191), RR 1.00 (0.87 to 1.16) 1 RCT mean 49.4 months: 0% (0/90) vs. 0% (0/79)	⊕◯◯◯ INSUFFICIENT

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
							Conclusions: Evidence from 2 poor-quality trials is insufficient to draw conclusions	
Procedure or device related SAEs (not further defined)	30 days, 24 months	1 RCT (N=381) Dohm 2014	Yes (-1)	Unknown	No	Yes (-1)	30 days: 4.2% (8/190) vs. 4.2% (8/191), RR 1.01 (0.39 to 2.62) 24 months: 5.8% (11/190) vs. 6.3% (12/191), RR 0.92 (95% CI 0.42 to 2.04) Conclusions: Evidence from 1 poor-quality trial is insufficient to draw conclusions	⊕◯◯◯ INSUFFICIENT
Reoperation for any new or repeat fracture	12 – 49.4 months	4 RCTs (N=460) Wang, 2015 Griffoni, 2020 Vogl, 2013 Yi, 2014	Yes (-1)	Yes (-1)	No	Yes (-1)	12 months (3 RCTs): 1 RCT (fair-quality): 2.0% (1/50) vs. 7.8% (4/51), RR 0.26 (0.03 to 2.20) 1 RCT (fair-quality): 23.4% (15/64) vs. 4.1% (2/49), RR 5.74 (1.38 to 23.94) 1 RCT (poor-quality): 3.6% (1/28) vs. 2.0% (1/49), RR 1.75 (0.11 to 26.90)	⊕◯◯◯ INSUFFICIENT

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
							Mean 49.4 months: 1 RCT (poor-quality): 10.0% (9/90) vs. 6.3% (5/79), RR 1.58 (0.55 to 4.52) Conclusions: Evidence is insufficient to draw conclusions. Serious inconsistency across trials and imprecision is present.	
Symptomatic Cement Leakage	Any time	5 RCTs (N=800) Dohm 2014 Endres 2012 Vogl 2013 Wang 2015 Yi 2024	Yes (-1)	No	No	Yes (-1)	VP: range, 0% to 1.1% KP: range, 0% to 1.9% 3 RCTs (N=312) reported no events in either group 1 RCT: 1.1% (2/190) vs. 0.5% (1/191), RR 2.01 (0.18 to 21.99) 1 RCT: 0% (0/53) vs. 1.9% (1/54), p=0.68; required discectomy and fusion Conclusions: Symptomatic cement leakage appears to be rare following VP and KP.	⊕⊕◯◯ LOW
Cement Embolism,	Any time	Symptomatic 1 RCT (N=381) Dohm 2014	Yes (-1)	No	No	Yes (-1)	Symptomatic embolism: 0.5% (1/190) vs. 0.5% (1/191); RR 1.01 (0.06 to 15.96)	⊕⊕◯◯ LOW

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Symptomatic and Asymptomatic		Asymptomatic 1 RCT (N=101) Wang 2015					Asymptomatic embolism: 0% (0/50) vs. 2.0% (1/51) <u>Conclusions</u>: Similar risk of cement embolism with VP and KP. Cement embolism appears to be rare, but studies were likely underpowered to detect rare events.	

CI = confidence interval; KP = kyphoplasty; NR = not reported; RCT = randomized controlled trial; RR = risk ratio; SoE = strength of evidence; VCF = vertebral compression fracture; VP = vertebroplasty.

5.1.6 Strength of Evidence Summary: Effectiveness for Vertebroplasty versus Minimally Invasive Procedures (i.e., Blocks) in Patients with Osteoporotic Vertebral Compression Fractures

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. Nerve/Facet Block Effect estimate (95% CI) Conclusion	Quality (SoE)
Pain **VAS (0-10 scale)**	<1 week	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	**MD:** -1.72 (-1.94 to -1.50) <u>Conclusion:</u> Moderate improvement in pain with VP vs. facet block.	⊕⊕◯◯ LOW
	≥1 week to ≤2 weeks	2 RCTs (N=233) Tan, 2023 Wang, 2016	Yes (-1)	No	No	No	**Pooled MD:** -1.59 (-1.92 to -0.84), I^2 = 35.3% <u>Conclusion:</u> Moderate improvement in pain with VP vs. medial branch nerve or facet block	⊕⊕◯◯ LOW

	>2 weeks to ≤1 month	2 RCTs (N=230) Tan, 2023 Wang, 2016	Yes (-1)	No	No	Yes (-1)	**Pooled MD:** -0.20 (-0.68 to 0.21), I^2 = 0% <u>Conclusion</u>: Similar improvement in pain with VP vs. medial branch nerve or facet blocks	⊕⊕◯◯ LOW
	>1 month to <6 months	2 RCTs (N=227) Tan, 2023 Wang, 2016	Yes (-1)	Unknown (excluding poor-quality RCT)	No	No	**Pooled MD:** 1.16 (-1.92 to 4.59), I^2 = 92.1% Fair-quality RCT (Wang), N=206: **MD** 0.01 (-0.20 to 0.22) <u>Conclusion:</u> Similar improvement in pain with VP vs. facet blocks	⊕⊕◯◯ LOW (large, fair-quality trial)
	≥6 months to <12 months	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	**MD:** 0.03 (-0.18 to 0.24) <u>Conclusion:</u> Similar improvement in pain with VP vs. facet blocks	⊕⊕◯◯ LOW
	≥12 months	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	**MD:** 0.04 (-0.17 to 0.25) <u>Conclusion</u>: Similar improvement in pain with VP vs. facet blocks	⊕⊕◯◯ LOW
Function **RDQ scores (0-24 scale)**	<1 week	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	**MD:** -2.86 (-3.19 to -2.53) <u>Conclusion:</u> Moderate improvement in pain with VP vs. facet block	⊕⊕◯◯ LOW
	≥1 week to ≤2 weeks	2 RCTs (N=233) Tan, 2023 Wang, 2016	Yes (-1)	Unknown (excluding poor-quality RCT)	No	No	**Pooled MD:** -1.69 (-6.54 to 3.98), I^2 = 89.4% Fair-quality RCT (Wang), N=206: **MD** -3.42 (-3.72 to -3.12)	⊕⊕◯◯ LOW (large, fair-quality trial)

							Conclusion: Large improvement in pain with VP vs. facet block after exclusion of poor-quality trial.	
	>2 weeks to ≤1 month	2 RCTs (N=230) Tan, 2023 Wang, 2016	Yes (-1)	No	No	No	**Pooled MD:** 0.15 (-0.45 to 0.19), I^2 = 0% Conclusion: Similar improvement in function with VP vs. medial branch nerve or facet blocks	⊕⊕◯◯ LOW
	>1 month to <6 months	2 RCTS (N=227) Tan, 2023 Wang, 2016	Yes (-1)	Yes (-1)	No	Yes (-1)	**Pooled MD:** -0.11 (-0.59 to 1.94), I^2 = 68.8% Conclusion: Similar improvement in function with VP vs. medial branch nerve or facet blocks	⊕⊕◯◯ LOW
	≥6 months to <12 months	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	**MD:** 0.01 (-0.32 to 0.34) Conclusion: Similar improvement in function with VP vs. facet blocks	⊕⊕◯◯ LOW
	≥12 months	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	**MD:** -0.16 (-0.52 to 0.20) Conclusion: Similar improvement in function with VP vs. facet blocks	⊕⊕◯◯ LOW

CI = confidence interval; MD = mean difference; RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; RR = risk ratio; SD = standard deviation; SoE = strength of evidence; VP = vertebroplasty.
*VAS and RDQ are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (VP).

5.1.7 Strength of Evidence Summary: Safety from RCTs of Vertebroplasty versus Minimally Invasive Procedures (i.e., blocks) in Patients with Osteoporotic Vertebral Compression Fractures

Outcome	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. Blocks Effect estimate (95% CI) Conclusion	Quality (SoE)
New vertebral fractures	12 months	1 RCT (N=206) Wang, 2016	Yes (-1)	Unknown	No	No	13% (13/100) vs. 10.4% (11/106), RR 1.25 (95% CI 0.59 to 2.67) Conclusion: Similar risk of any new vertebral fracture with VP vs. facet block.	⊕⊕◯◯ LOW
Asymptomatic Cement Leakage	1 week, 12 months	1 RCT (N=100, VP arm only) Wang, 2016	Yes (-1)	Unknown	No	Yes (-1)	RCT: 1.0% (1/100), 12 months Conclusion: Evidence is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT

CI = confidence interval; NR = not reported; RCT = randomized controlled trial; RR = risk ratio; SoE = strength of evidence; VCF = vertebral compression fracture; VP = vertebroplasty.

5.1.8 Strength of Evidence Summary: Effectiveness for Kyphoplasty versus Usual Care in Patients with Osteoporotic Vertebral Compression Fractures

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
Pain **VAS scores (0-10 scale)**	<1 week	1 RCT (N=80) Li, 2017	Yes (-1)	Unknown	No	Yes (-1)	**MD:** -6.22 (-7.13 to -5.31) Conclusion: Data from one poor-quality trial is insufficient to draw conclusions	⊕◯◯◯ INSUFFICIENT
	≥1 week to ≤2 weeks	2 RCTs (N=380) Li, 2017 Wardlaw, 2009	Yes (-1)	Unknown (fair-quality trial only)	No	No (fair-quality trial only)	**Pooled MD:** -2.59 (95% CI -3.97 to -1.76), I^2=67.0% Fair-quality trial (Wardlaw), N=300: **MD**: -2.40 (-2.89 to -1.91) Conclusion: Large improvement in pain with KP vs. UC in the large, fair-quality trial.	⊕⊕◯◯ LOW (fair-quality trial)
	>2 weeks to ≤1 month	2 RCTs (N=380) Li, 2017 Van Meirhaeghe 2023	Yes (-1)	Unknown (fair-quality trial only)	No	No (fair-quality trial only)	**Pooled MD:** -1.33 (95% CI -3.02 to 0.57), I^2=85.4% Fair-quality trial (Van Meirhaeghe), N=300: **MD**: -1.96 (-2.50 to -1.42) Conclusion: Moderate improvement in pain with KP vs. UC in the large, fair-quality trial.	⊕⊕◯◯ LOW (fair-quality trial)
	>1 month to <6 months	2 RCTs (N=380) Li, 2017	Yes (-1)	No	No	Yes (-1)	**Pooled MD:** -1.48 (-2.10 to -0.58), I^2=0%	⊕⊕◯◯ LOW

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
		Van Meirhaeghe 2023					Conclusion: Moderate improvement in pain with KP vs. UC.	
	≥6 months to <12 months	2 RCTs (N=380) Li, 2017 Van Meirhaeghe 2023	Yes (-1)	Unknown (fair-quality trial only)	No	No (fair-quality trial only)	**Pooled MD:** -1.08 (-2.41 to 0.27), I^2=85.0% Fair-quality trial (Van Meirhaeghe), N=300: **MD** -1.62 (-2.18 to -1.06) Conclusion: Moderate improvement in pain with KP vs. UC in the large, fair-quality trial.	⊕⊕◯◯ LOW (fair-quality trial)
	≥12 months	1 RCT (N=300) Van Meirhaeghe 2023	Yes (-1)	Unknown	No	Yes (-1)	12 months: **MD:** -0.98 (-1.56 to -0.40) 24 months: **MD:** -0.83 (-1.41 to -0.25) Conclusion: Small improvement in pain with KP vs. UC at 12 and 24 months.	⊕⊕◯◯ LOW (fair-quality trial)
Function **RDQ scores (0-24 scale)** **ODI scores (0-100)**	<1 week	1 RCT (N=80) Li, 2017	Yes (-1)	Unknown	No	Yes (-1)	**SMD:** -0.49 (-0.94 to -0.05) Conclusion: Evidence from one poor-quality trial is insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
	≥1 week to ≤2 weeks	1 RCT (N=80) Li, 2017	Yes (-1)	Unknown	No	Yes (-1)	**SMD:** -0.05 (-0.49 to 0.39) Conclusion: Evidence from one poor-quality trial is	⊕◯◯◯ INSUFFICIENT

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
							insufficient to draw conclusions.	
	>2 weeks to ≤1 month	2 RCTs (N=380) Li, 2017 Van Meirhaeghe 2023	Yes (-1)	Unknown (fair-quality trial only)	No	No	**Pooled SMD:** -0.48 (-1.13 to 0.27), I^2=81.4% Fair-quality trial (Van Meirhaeghe), N=300: **SMD** -0.71 (-0.95 to -0.48) <u>Conclusion</u>: Moderate improvement in function with KP vs. UC in the large, fair-quality trial.	⊕⊕◯◯ LOW (fair-quality trial)
	>1 month to <6 months	2 RCTs (N=380) Li, 2017 Van Meirhaeghe 2023	Yes (-1)	Unknown (fair-quality trial only)	No	No	**Pooled SMD:** -0.50 (-0.92, 0.16), I^2=69.9% Fair-quality trial (Van Meirhaeghe), N=300: **SMD** -0.60 (-0.83 to -0.37) <u>Conclusion</u>: Moderate improvement in function with KP vs. UC in the large, fair-quality trial.	⊕⊕◯◯ LOW (fair-quality trial)
	≥6 months to <12 months	1 RCT (N=300) Van Meirhaeghe 2023	Yes (-1)	Unknown	No	No	**SMD:** -0.48 (-0.71 to -0.25) <u>Conclusion</u>: Small improvement in function with KP vs. UC.	⊕⊕◯◯ LOW
	≥12 months	1 RCT (N=300) Van Meirhaeghe 2023	Yes (-1)	Unknown	No	No	12 months **SMD:** -0.45 (-0.68 to -0.22) 24 months **MD**: -1.43 (-2.90 to 0.04)	⊕⊕◯◯ LOW

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
							Conclusion: Small improvement in function with KP vs. UC at 12 months but similar improvement between groups at 24 months.	

CI = confidence interval; KP = kyphoplasty; MD = mean difference; ODI = Oswestry Disability Index; RCT = randomized controlled trial; RDQ = Roland-Morris Disability Questionnaire; RCT = randomized controlled trial; RR = risk ratio; SD = standard deviation; SoE = strength of evidence; UC = usual care.
*VAS and RDQ and ODI are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (KP).
†Wardlaw 2009 (index publication) and Van Meirhaeghe 2023 (follow-up publication) are the same trial (FREE trial).

5.1.9 Strength of Evidence Summary: Safety from RCTs of Kyphoplasty versus Usual Care in Patients with Osteoporotic Vertebral Compression Fractures

Outcome	Time	Studies*	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	Latest follow-up (24 months)	1 RCT (N=300) Wardlaw 2009, Boonen 2011	Yes (-1)	Unknown	No	Yes (-1)	8.1% (12/149) vs. 7.2% (11/151); RR 1.11 (0.50 to 2.43) Conclusion: Similar risk of mortality with KP vs. UC. Results were similar at 12 months (6.0% vs. 4.6%).	⊕⊕◯◯ LOW
SAEs (any)	30 days and Latest follow-up	2 RCTs (N=500) Wardlaw 2009, Boonen 2011, Van Meirhaeghe 2013; Yi 2014	Yes (-1)	Yes (-1)	No	Yes (-1)	Fair quality trial, 30 days: 16.1% (24/149) vs. 11.2% (17/151), RR 1.43 (0.80 to 2.55) 24 months: 49.7% (74/149) vs. 48.3% (73/151), RR 1.02 (0.82 to 1.29)	⊕⊕◯◯ LOW

Outcome	Time	Studies*	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
							Poor-quality trial, mean 49 months: 0% (0/79) vs. 0% (0/121) Conclusion: Similar risk of any SAE with KP vs. UC.	
Treatment-related SAEs†	Various	1 RCT (N=300) Boonen 2011, Van Meirhaeghe 2013	Yes (-1)	Unknown	No	Yes (-1)	30 days: 1.3% (2/149) vs. 0.7% (1/151), RR 2.03 (0.19 to 22.12) 12 and 24 months: KP arm only: 1.3% (2/149) and 2.0% (3/149) Conclusion: Similar risk of any treatment-related SAEs with KP vs. UC by 30 days.	⊕⊕◯◯ LOW
Withdrawals due to AEs	Latest follow-up (24 months)	1 RCT (N=300) Wardlaw 2009, Boonen 2011	Yes (-1)	Unknown	No	Yes (-1)	0.6% (1/149) vs. 0.6% (1/151), RR 1.01 (0.06 to 16.05) Conclusion: Similar risk of any withdrawal due to AEs with KP vs. UC.	⊕⊕◯◯ LOW
Cement leakage, symptomatic	Various	2 RCTs (n=379, KP arm only) Boonen 2011, Van Meirhaeghe 2013, Yi 2014	Yes (-1)	Yes (-1)	No	Yes (-1)	Fair-quality trial 30 days: 3.4% (5/149) 24 months: 7.4% (11/149) Poor-quality trial Mean 49 months: 0% (0/79) Conclusion: Similar risk of any withdrawal due to AEs with KP vs. UC.	⊕⊕◯◯ LOW
New clinical/ symptomatic	Latest follow-up (24 to a	2 RCTs (N=500)	Yes (-1)	Yes (-1)	No	Yes (-1)	Fair-quality trial, 24 months:	⊕⊕◯◯ LOW

Outcome	Time	Studies*	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
vertebral fractures	mean 49 months)	Boonen 2011, Van Meirhaeghe 2013, Yi 2014					17.4% (26/149) vs. 11.3% (17/151), RR 1.55 (0.88 to 2.74) Poor-quality trial Mean 49 months: 6.3% (5/79) vs. 14.0% (17/121), RR 0.45 (0.17 to 1.17) <u>Conclusion</u>: Similar risk of new symptomatic vertebral fractures with KP vs. UC.	
New radiographic vertebral fracture (Any, index level, and adjacent level)	Latest follow-up (24 months)	1 RCT (N=300) Boonen 2011, Van Meirhaeghe 2013	Yes (-1)	Unknown	No	Yes (-1)	Any new fracture: 47.5% (56/118) vs. 44.1% (45/102), RR 1.08 (0.81 to 1.44) New index level fractures: 4.2% (5/118) vs. 10.8% (11/102), RR 0.39 (0.14 to 1.09) New adjacent level fractures: 23.7% (28/118) vs. 16.7% (17/102), RR 1.42 (0.83 to 2.45) <u>Conclusion</u>: Similar risk of any new fractures, new index level fractures, and new adjacent level fractures with KP vs. UC.	⊕⊕◯◯ LOW
Reoperation (for new symptomatic fractures)	Latest follow-up (24 months)	1 RCT (N=300) Wardlaw 2009, Boonen 2011	Yes (-1)	Unknown	No	Yes (-1)	8.1% (12/149) vs. 4.0% (6/151), RR 2.03 (0.78 to 5.26)	⊕⊕◯◯ LOW

Outcome	Time	Studies*	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
							Conclusion: Similar risk of reoperation for new symptomatic fractures with KP vs. UC.	

CI = confidence interval; KP = kyphoplasty; NR = not reported; RCT = randomized controlled trial; RR = risk ratio; SoE = strength of evidence; UC = usual care; VCF = vertebral compression fracture.

*FREE trial: Wardlaw 2009 (index publication), Van Meirhaeghe 2013 and Boonen 2011 (24-month follow-up data/subsequent publications).

†At 30 days (and 12 months): 2 serious adverse events were attributed to kyphoplasty – a soft tissue hematoma at the surgical site and a postoperative urinary tract infection that needed intervention; 1 serious event was attributed to UC – back pain. At 24 months, the same patient who had UTI by 12 months developed spondylitis near the cement that required treatment; in another patient there was anterior cement migration that likely caused a recurrent fracture.

Strength of Evidence Summary: Malignant Fractures

5.1.10 Strength of Evidence Summary: Effectiveness of Kyphoplasty versus Usual Care in Patients with Fractures due to Tumors and Malignancies

Outcome*	Time†	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
Pain NRS (0-10)	1 week and 1 month	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	No	1 week N=117, MD -3.50 (95% CI -4.27 to -2.73) 1 month N=114, MD -3.50 (95% CI -4.37 to -2.63) Conclusion: One fair-quality trial found a large improvement in pain with KP compared to usual care.	⊕⊕◯◯ LOW
Function Responders: RDQ (≥2.5-point improvement); KPS (≥5- point improvement)	1 month	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	No	RDQ 80.9% (51/63) vs. 28% (14/50) RR 2.89 (95% CI 1.82 to 4.58) KPS 65.1% (41/63) vs. 26.5% (13/49) RR 2.45 (95% CI 1.49 to 4.04) Conclusion: Large increase in the likelihood of achieving MCIDs on RDQ and KPS for KP compared to usual care.	⊕⊕◯◯ LOW
Function	1 month	1 RCT (N=134)	Yes (-1)	Unknown	No	Yes (-1)	74.6% (47/63) vs. 38.8% (19/49) RR 1.92 (95% CI 1.32 to 2.81)	⊕⊕◯◯ LOW

Outcome*	Time†	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
KPS score ≥70 (ability to care for oneself)		Berenson, 2011 (CAFE trial)					Conclusion: Moderate increase in the likelihood of achieving a score ≥70 on the KPS for KP compared to usual care.	
Function RDQ scores (0-24) KPS scores (0-100)	1 month	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	No	RDQ: N=113, MD -8.9 (95% CI -9.49 to -8.31) KPS: N=112, MD 14.5 (95% CI 12.83 to 16.17) Conclusion: One fair-quality trial found large improvement in function (both measures) with KP compared to usual care.	⊕⊕◯◯ LOW

AE = Adverse event; CI = Confidence interval; KP = Kyphoplasty; KPS = Karnofsky Performance Status; MCS = Mental component scale; MD = Mean difference; NA = Not applicable; NRS = Numerical Rating Scale; PCS = Physical component scale; RCT = Randomized controlled trial; RDQ = Roland Morris Disability Questionnaire; RR = Risk ratio; SF-36 = 36 Item Short-Form Survey; SoE = Strength of evidence.

*VAS and ODI are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (KP); for the KPS, a higher score is better, i.e., a positive score favors the intervention (KP).

†The CAFE Trial reports outcomes beyond 1 month. However, 53% (34/64) of the control group immediately crossed over to KP at 1 month; given the substantial cross-over rate and break in randomization, our efficacy analyses focused on data at 1 month and earlier.

5.1.11 Strength of Evidence Summary: Safety of Kyphoplasty versus Usual Care in Patients with Fractures due to Tumors and Malignancies

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	1 month	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	Yes (-1)	2.8% (2/70) vs. 1.5% (1/64) RR 1.82 (95% CI 0.17 to 19.69) <u>Conclusion:</u> Similar risk of mortality between groups but the estimate was imprecise.	⊕⊕◯◯ LOW
	≥1 month and ≤12 months	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	Yes (-1)	As randomized: 30.0% (21/70) vs. 19.2% (5/26), RR 1.56 (95% CI 0.66 to 3.71) As treated/after crossover: 25.0% (27/108) vs. 19.2% (5/26), RR 1.30 (95% CI 0.55 to 3.05) <u>Conclusion:</u> KP tended to have higher mortality rates compared with usual care but the differences were not statistically significant and the estimates were imprecise.	⊕⊕◯◯ LOW
Serious AEs†	≥1 month and ≤12 months	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	Yes (-1)	As randomized: 52.8% (37/70) vs. 30.7% (8/26) RR 1.72 (95% CI 0.93 to 3.19) As treated/after crossover: 50.9% (55/108) vs. 30.7% (8/26), RR 1.66 (95% CI 0.90 to 3.03) <u>Conclusion:</u> KP tended to have higher mortality rates compared with usual care, but the differences were not statistically different.	⊕⊕◯◯ LOW

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	KP vs. Usual Care Effect estimate (95% CI) Conclusion	Quality (SoE)
Symptomatic Fracture	1 month	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	Yes (-1)	2.8% (2/70) vs. 4.7% (3/64) RR 0.61 (95% CI 0.11 to 3.53) <u>Conclusion:</u> Similar risk of new symptomatic fracture between groups, but the estimate was imprecise	⊕⊕◯◯ LOW
	≥1 month and ≤12 months	1 RCT (N=134) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	Yes (-1)	As randomized: 12.8% (9/70) vs. 0% (0/26), p=0.056 As treated/after crossover: 16.7% (18/108) vs. 0% (0/26), p=0.026 <u>Conclusion:</u> Only patients receiving KP experienced symptomatic fractures in the long term.	⊕⊕◯◯ LOW
Cement Leakage, symptomatic	1 month	1 RCT (N=70) Berenson, 2011 (CAFE trial)	Yes (-1)	Unknown	No	Yes (-1)	1.4% (1/70) <u>Conclusion:</u> One patient receiving KP experienced symptomatic cement leakage and suffered an adjacent fracture the day after the procedure, which was classified as a serious device-related event.	⊕⊕◯◯ LOW

AE = Adverse event; CI = Confidence interval; KP = Kyphoplasty; RCT = Randomized controlled trial; RR = Risk ratio; SoE = Strength of evidence.
*Given the high rate of cross-over (53%) from UC to KP after 1 month, we reported harms for patients both as randomized and as treated.
†Defined as any event that resulted in death, life-threatening injury or permanent impairment, needed intervention to prevent impairment, or resulted in prolonged hospitalizations

5.1.12 Strength of Evidence Summary: Effectiveness of Vertebroplasty versus Kyphoplasty in Patients with Fractures due to Tumors and Malignancies

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Responders (VAS score ≥3)	Discharge or first f/u visit	1 comparative NRSI (N=342) Bae 2016	Yes (-1)	Unknown	No	Yes (-1)	62% (148/238) vs. 57% (59/104), RR 1.10 (95% CI 0.90 to 1.33) <u>Conclusion</u>: Data from one retrospective NRSI are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Complete or improved pain relief	24 hours	1 comparative NRSI (N=49) Fourney 2003	Yes (-1)	Unknown	No	Yes (-1)	86% (30/35 sessions) vs. 80% (12/15 sessions), RR 1.07 (95% CI 0.80 to 1.43) <u>Conclusion</u>: Data from one retrospective NRSI are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
VAS pain scores (0-10)	1, 3, 6, 12, 24 months	2 comparative NRSIs (N=391) Bae 2016 Fourney 2003 1 SR of case series (N=1,445 VP; 1,110 KP) 4 case series (N=261) *VP (N=94)* Moulin 2020 Rocha Romero *KP (N=157)*	Yes (-1)	Unknown	No	Yes (-1)	1 NRSI (N=342), Timing unclear: MD -0.30 (95% CI -0.74 to 0.14) 1 NRSI (N=49) 1 month: median 2 vs. 2.5 3 months: median 2 vs. 2.5 6 months: median 2 vs. 4 12 months: median 1 vs.2 p=NS for all VP (N=1,539), SR and case series: Baseline range, 5.0-7.48 Latest follow-up range (1-24 months): 1.68 to 2.98	⊕◯◯◯ INSUFFICIENT

Outcome*	Time	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
		Wu 2023 Garcia Maroto 2015					KP (N=1,539), SR and case series: Baseline range, 6.3-7.49 Latest follow-up range (9 to ≥12 months): 3.09 to 3.4 Conclusion: Both VP and KP showed improvement in pain from baseline over time. Data from NRSIs and are insufficient to draw conclusions.	
VAS pain (0-50 scale)†	6 weeks, 6 months, 12 months	Köse 2006 (N=34)	Yes (-1)	Unknown	No	Yes (-1)	6 weeks: MD 3.2 (95% CI 0.51 to 5.89) 6 months: MD 3.6 (95% CI 1.74 to 5.46) 12 months: MD 3.8 (95% CI 1.95 to 5.65) Conclusion: Data from one retrospective NRSI are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT

AE = Adverse event; CI = Confidence interval; KP = Kyphoplasty; KPS = Karnofsky Performance Status; MCS = Mental component scale; MD = Mean difference; NA = Not applicable; NRS = Numerical Rating Scale; PCS = Physical component scale; RCT = Randomized controlled trial; RDQ = Roland Morris Disability Questionnaire; RR = Risk ratio; SF-36 = 36 Item Short-Form Survey; SoE = Strength of evidence.

*VAS is scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (VP or KP).

†Average of pain during 5 ADLs: pain at rest, walking, sitting-standing, taking a shower and wearing clothes.

5.1.13 Strength of Evidence Summary: Safety of Vertebroplasty versus Kyphoplasty in Patients with Fractures due to Tumors and Malignancies

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	30 days, 2.5 months	1 comparative NRSI (N=49) Fourney 2003 4 case series (N range, 50-158)	Yes (-1)	Unknown	No	Yes (-1)	**30 days:** 0% (0/34) vs. 0% (0/15) **2.5 months:** 2.9% (1/34) vs. 0% (0/15) 4 case series VP (1 study, N=50): 34.0% at 401 days KP (3 studies, N range 75-158): 0% to 18.8% <u>Conclusion</u>: Data from one retrospective NRSI and case series are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
SAEs	30 days, 12 months	2 comparative NRSI (N=83) Fourney 2003 Kose 2006 4 case series (N range, 50-407)	Yes (-1)	Unknown (different time points)	No	Yes (-1)	**30 days, 1 NRSI:** 2.9% (1/34) vs. 6.7% (1/15), RR 0.44 (95% CI 0.03 to 6.59) VP: Paraplegia due to metastasis KP: Readmission for CHF **12 months, 1 NRSI:** 0% (0/16) vs. 0% (0/18) 4 case series VP (1 study, N=50): 2% KP (2 studies, N=92, 117): 0% VP/KP (1 study, N=407): 0.5%	⊕◯◯◯ INSUFFICIENT

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
							Conclusion: Data from two retrospective NRSI and case series are insufficient to draw conclusions.	
Device- or procedure-related complications	30 days, 12 months	2 comparative NRSI (N=83) Fourney 2003 Kose 2006	Yes (-1)	Unknown (different time points)	No	Yes (-1)	**30 days, 1 NRSI:** 0% (0/34) vs. 0% (0/15) **12 months, 1 NRSI:** NR vs. 5.6% (1/18), asymptomatic balloon rupture Conclusion: Data from two retrospective NRSI are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
New fracture	12 months	1 comparative NRSI (N=34) Kose 2006 2 SRs of case series 4 case series (N range, 44-407)	Yes (-1)	Unknown	No	Yes (-1)	Adjacent level fracture and symptomatic fracture requiring reoperation: 0% (0/16) vs. 0% (0/18) 2 SRs of case series VP: 18% (1 study) KP: range, 10.2%- 17.0% Case series VP (2 studies, N=44, 50): 10%-29.5% KP: (1 study, N=75): 14.7% VP/KP (1 study, N=407): 24.6% Conclusion: Data from one retrospective NRSI and case series are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	VP vs. KP Effect estimate (95% CI) Conclusion	Quality (SoE)
Cement leakage, asymptomatic	30 days	1 comparative NRSI (N=49) Fourney 2003	Yes (-1)	Unknown	No	Yes (-1)	9.2% (6/65 levels) vs. 0% (0/32 levels) <u>Conclusion</u>: Data from one retrospective NRSI are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Symptomatic cement leakage	-----	2 SR of case series 2 case series	-----	------	------	-----	2 SRs of case series: VP: 0.26% (21/760 levels) to 3.1% (3/98) KP: 0% (0/214 levels) and 0% 2 case series, VP/KP: 1.0% (4/407) KP: 0% (0/215)	⊕◯◯◯ INSUFFICIENT
Reoperation	4.5 months	1 comparative NRSI (N=49) Fourney 2003 2 case series (N range, 50-407)	Yes (-1)	Unknown	No	Yes (-1)	**Repeat VP or KP:** 2.9% (1/34) vs. 6.7% (1/15). RR 0.44 (95% CI 0.03 to 6.59) **Subsequent spinal surgery:** 5.9% (2/34) vs. 0% (0/15) Case series VP (1 study, N=50): 6.0% VP/KP (1 study, N=407): 19.2% <u>Conclusion</u>: Data from one retrospective NRSI and case series are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT

AE = Adverse event; CI = Confidence interval; KP = Kyphoplasty; KPS = Karnofsky Performance Status; MCS = Mental component scale; MD = Mean difference; NA = Not applicable; NRS = Numerical Rating Scale; PCS = Physical component scale; RCT = Randomized controlled trial; RDQ = Roland Morris Disability Questionnaire; RR = Risk ratio; SF-36 = 36 Item Short-Form Survey; SoE = Strength of evidence.

5.2 Strength of Evidence Summary: Sacroplasty

5.2.1 Strength of Evidence Summary: Effectiveness of Sacroplasty vs. Nonsurgical Management and vs. Surgery

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	SP vs. NSM Effect estimate (95% CI) Conclusion	Quality (SoE)
Sacroplasty vs. Usual Care								
VAS pain scores (0-10)	Various	3 comparative NRSIs (N=438) Frey, 2017 Sarigul, 2023, Andresen 2022	Yes (-1)	Unknown	No	Yes (-1)	Sacroplasty resulted in significantly greater improvement in pain across most timepoints (follow-up range, 10 days to 2 years) compared with UC and was sustained longer term Latest follow-up 1 NRSI (N=116), 2 years: MD -0.46 (95% CI -0.86 to -0.06); 1 NRSI (N=185), 1 year: mean change scores, -7.67 vs. -1.36, p<0.001; 1 NRSI (N=137), 2 years: data NR, p<0.001 <u>Conclusion</u>: Data from two comparative NRSIs (1, prospective, 1 retrospective) are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Function ODI scores (0-100) HBI scores (0-100)	Various	2 comparative NRSIs (N=382) Sarigul, 2023 Andresen, 2022	Yes (-1)	Unknown	No	Yes (-1)	Sacroplasty resulted in significantly greater improvement in function across all timepoints (follow-up range, 10 days to 2 years) compared with UC and was sustained longer term Latest follow-up ODI 1 NSRI (N=185), 1 year: mean change scores, -70.2 vs. -30.6, p<0.001 HBI	⊕◯◯◯ INSUFFICIENT

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	SP vs. NSM Effect estimate (95% CI) Conclusion	Quality (SoE)
							1 NSRI (N=197), 2 years: MD 7.0 (95% CI 4.24 to 9.76) <u>Conclusion</u>: Data from two retrospective comparative NRSIs are insufficient to draw conclusions.	
Sacroplasty vs. teriparatide								
VAS pain scores (0-10) ODI function scores (0-100)	Various	1 comparative NRSI (N=27) Yang, 2023	Yes (-1)	Unknown	No	Yes (-1)	Sacroplasty resulted in significantly less improvement in pain and function compared with UC at 12 and 24 weeks but not at early timepoints (2, 4 weeks) Latest follow-up, 26 weeks: VAS pain: MD 2.1 (95% CI 1.21 to 3.00) ODI function: MD 9.5 (95% CI 6.14 to 12.86) <u>Conclusion</u>: Data from 1 retrospective comparative NRSI are insufficient to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Sacroplasty vs. surgery								
VAS pain scores (0-10) HBI function scores (0-100)	Various	1 comparative NRSI (N=233) Andresen, 2022	Yes (-1)	Unknown	No	Yes (-1)	Patients in both groups experienced significant improvement in pain (VAS scores) and function (HBI scores) but it was more rapid following sacroplasty versus surgery Latest follow-up, 2 years: VAS pain: data NR, p<0.001 HBI function: MD -1.0 (95% CI -2.89 to 0.89)	⊕◯◯◯ INSUFFICIENT

Outcome*	Time	Studies†	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	SP vs. NSM Effect estimate (95% CI) Conclusion	Quality (SoE)
							Conclusion: Data from 1 retrospective comparative NRSI are insufficient to draw conclusions.	

CI = Confidence interval; MD = Mean difference;NR = not reported; ODI = Oswestry Disability Index;; SoE = Strength of evidence.
*VAS and ODI are scaled such that a lower score means a better outcome, i.e., a negative mean difference favors the intervention (KP); for the KPS, a higher score is better, i.e., a positive score favors the intervention (KP).
†All comparative NRSIs are retrospective except Frey 2017 which is prospective.

5.2.2 Strength of Evidence Summary: Safety of Sacroplasty vs. Nonsurgical or Surgical Management

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	SP vs. NSM or Surgery Effect estimate (95% CI) Conclusion	Quality (SoE)
Mortality	6-12 months	1 comparative NRSI SP vs. UC (N=233) SP vs. Surgery (N=178) Andresen, 2022 1 single arm Registry Beall 2023	Yes (-1)	Unknown (2 different comparators)	No	Yes (-1)	NRSI, 12 months: SP vs. UC 8.4% (10/119) vs. 21.7% (25/114); RR 0.38 (95% CI 0.19 to 0.76) SP vs. Surgery 8.4% (10/119) vs. 13.6% (8/59); RR 0.62 (95% CI 0.23 to 1.49) Registry, 6 months: 0% (0/102) Conclusion: Data from 1 retrospective comparative NRSI and one single arm registry are insufficient to draw conclusions for SP vs. UC and vs. surgery.	⊕◯◯◯ INSUFFICIENT
SAEs	Perioperative	1 comparative NRSI (N=83 in SP arm) Sarigul 2023	Yes (-1)	Unknown	No	Yes (-1)	PE Perioperative: 0% (0/83) vs. NR Conclusion: Insufficient data to draw conclusions.	⊕◯◯◯ INSUFFICIENT

Outcome	Time*	Studies	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	SP vs. NSM or Surgery Effect estimate (95% CI) Conclusion	Quality (SoE)
New fracture, symptomatic	6 months	1 single arm registry (N=102) Beall 2023	Yes (-1)	Unknown	No	Yes (-1)	3% (3/102), all required surgery Conclusion: Insufficient data to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Cement leakage	Perioperative, and 3-18 month	2 comparative NRSI (N=202 in SP arm) Sarigul 2023 Andresen 2022 1 single arm registry (N=102) Beall 2023 1 SR of case series (N=861) Chandra 2019	Yes (-1)	No	No	Yes (-1)	Symptomatic cement leakage: range 0% to 1.0% Asymptomatic cement leakage: range 2.2% to 17.7% Conclusion: Symptomatic cement leakage appears to be rare; asymptomatic cement leakage is common following sacroplasty. However, data is insufficient data to draw conclusions.	⊕◯◯◯ INSUFFICIENT
Reoperation	3-18 months	1 single arm registry (N=102) Beall 2023 1 SR of case series (N=861) Chandra 2019	Yes (-1)	No	No	Yes (-1)	Registry: 3% (3/102), for any new sacral or VCF SR: 0.3% (3/861), radicular pain due to cement leakage Conclusion: Data is insufficient data to draw conclusions.	⊕◯◯◯ INSUFFICIENT

AE = Adverse event; CI = Confidence interval; NA = Not applicable; NRSI = Nonrandomized studies or interventions; RCT = Randomized controlled trial; RR = Risk ratio; SAE = serious adverse event; SoE = Strength of evidence; SP = sacroplasty; SR =systematic review; UC = usual care; VCF = vertebral compression fracture

5.3 Safety Evidence Summary: NRSIs for Osteoporotic Compression Fractures

The following evidence from NRSIs (comparative and single arm) is considered insufficient to draw conclusions due to the methodological flaws/limitations (i.e., high risk of bias) of studies and imprecision of estimates, with many studies not providing any measure of variance. Given the unique analyses across some of the studies, consistency is often unknown. NRSIs start out at Low SOE and given these limitations the studies ended up as insufficient strength of evidence.

5.3.1 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSI of VP vs. UC and case series of VP

Outcome	VP vs. UC Effect estimate (95% CI) Conclusion
Mortality	Admin data: 3 studies VP vs. UC Ong: HR 0.926 (0.926, 0.917) Lin: adj HR 1.39 (1.09–1.78) Huang: Adj HR 0.87 (0.77–0.99) Case series (all VP): 2 case series range: 0%(N=1512) to 1.2% (N=485) Mortality due to embolism: 0% (N=1512)
New fracture	Admin data: VP vs. UC Huang: <0.3% of 1389 vs. <0.1% of 6017), p=NS Case series (all VP): Any new, range: 11.6% (N=293) to 22.1% (N= 1090) Adjacent: 6.6% (N=361) to 7.8%(N=293)
Pulmonary embolism (PE)	Admin data: VP vs. UC Edidin 2015: 4 years: Adj. HR 1.07 (95% CI 0.98 to 1.18 Case series (all VP): Asymptomatic pulmonary cement embolism: 3.7% (11/299)
Deep vein thrombosis (DVT)	Admin data: VP vs. UC Edidin 2015, 4 years: Adj. HR 1.05 (95% CI 0.96 to 1.15) Case series (all VP): NR
Cardiac complications (to include MI) **Other cardiac complications**	Admin data: VP vs. UC Edidin 2015, Adj. HR 1.05 (95% CI 0.94 to 1.16) Case series (all VP): All from 1 Case series (N=1512) Cardiorespiratory arrest: <0.1% Any intracardiac cement embolism 4.8% Intracardiac cement embolism with associated PCE 4.1%

Outcome	VP vs. UC Effect estimate (95% CI) Conclusion
	Symptomatic intracardiac cement embolism: 0.3% (none resulted in death)
Respiratory (including resp failure)	Admin data: VP vs. UC Edidin 2015, 4 years: Adj. HR 1.05 (95% CI 1.01 to 1.09 Choo, 2018: 30 days: Adj. OR 3.28 (95% CI 1.56-6.88) Case series (all VP): NR
Infection, osteomyelitis or infection	Admin data: VP vs. UC Edidin 2015, adj HR 1.05 (95% CI 0.87 to 1.27) Case series (all VP): NR
Additional procedures, reoperation, subsequent augmentation	Admin data: VP vs. UC Edidin 2015: Any subsequent augmentation: Adj. HR 11.1 (95% CI 11.1 to 12.5) Subsequent augmentation or fusion: Adj. HR 11.1 (95% CI 11.1 to 12.5) Case series (all VP): New fracture: 22.1% (241/1090)

Adj = adjusted; CI = confidence interval; HR = hazard ratio; MI = myocardial infarction; NRSI = nonrandomized study of interventions; OR = odds ratio; RR = risk ratio; UC = usual care; VP = vertebroplasty.

5.3.2 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSI of VP vs. KP

Outcome	VP vs. KP Effect estimate (95% CI) Conclusion
Mortality	Admin data, 2 studies: Kim 2022, Adj OR 0.94 (0.27-3.24), 30 days, KP vs. VP >30 days, 1KP vs. VP Edidin 2015, Adj HR 0.83 (0.81-0.85), >30 days, KP vs. VP
Pulmonary embolism (PE)	Admin data, 1 study: Edidin 2015, Adj HR 1.16 (1.01-1.35), 4 years, VP vs. KP
Deep vein thrombosis (DVT)	Admin data, 1 study: Edidin 2015, Adj HR 1.16 (1.01-1.35), 4 years, VP vs. KP
Cardiac complications (due to include MI) **Other cardiac complications**	Admin data, 1 study: Edidin 2015, Adj HR 1.05 (0.94-1.16), 4 years, VP vs. KP

Outcome	VP vs. KP Effect estimate (95% CI) Conclusion
Pulmonary/ respiratory complications	Admin data, 2 studies: Edidin 2015, Adj HR 1.05 (1.01-1.16), 4 years, VP vs. KP Choo 2018, Adj. OR 3.28 (1.56-6.88), 30 days, VP vs. KP
Infection	Admin data, 1 study: Edidin 2015, Adj HR 1.05 (0.87 to 1.27), 4 years, VP vs. KP
Any SAE	Admin data, 1 study: Kim 2022, Adj OR 1.93 (0.58 to 6.41), 30 days, VP vs. KP
Subsequent reoperation	Admin data, 1 study: Edidin 2015, 4 years, VP vs. KP: Any subsequent augmentation: Adj. HR 1.03 (0.97-1.09) Subsequent augmentation or fusion: Adj. HR 1.03 (0.97-1.09) Subsequent VCF with repair: Adj. HR 1.03 (0.97-1.09)

Adj = adjusted; CI = confidence interval; HR = hazard ratio; KP = kyphoplasty; MI = myocardial infarction; NRSI = nonrandomized study of interventions; OR = odds ratio; RR = risk ratio; SAE = serious adverse event; VCF = vertebral compression fracture; VP = vertebroplasty.

5.3.3 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSI of VP vs. Other surgery

Outcome	VP vs. Other surgery Effect estimate (95% CI) Conclusion
Mortality	Admin data, 1 study: Huang 2019, mean 4.5 years, VP vs. Open surgery, 19.2% (267/1389) vs. 17.4% (212/1219), p>0.05
New fracture	Admin data, 1 study: Huang 2019, mean 4.5 years, VP vs. Open surgery, <0.3% (NR/1389) vs. <0.3% (NR/1219), p>0.05
Stroke	Admin data, 1 study: Wu 2012, ≤5 years, VP vs. other surgery: Any stroke: Adj. HR 1.22 (95% CI 0.67 to 2.24) Hemorrhagic Stroke: Adj. HR 3.17 (95% CI 0.97 to 10.3) Ischemic stroke: Adj. HR 0.96 (95% CI 0.49 to 1.91)
Pulmonary embolism (PE)	Admin data, 1 study: Huang 2019, mean 4.5 years, VP vs. Open surgery, 0.4% (6/1389) vs. ≤0.3% (NR/1219), p>0.05
Vertebral osteomyelitis or infection	Admin data, 1 study: Huang 2019, mean 4.5 years, VP vs. Open surgery, 1.0% (14/1389) vs. 1.0% (12/1219), p>0.05

Adj = adjusted; CI = confidence interval; HR = hazard ratio; NRSI = nonrandomized study of interventions; VP = vertebroplasty

5.3.4 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSIs of VP vs. Minimally Invasive Procedures (Blocks)

Outcome*	VP vs. Blocks Effect estimate (95% CI) Conclusion
New vertebral fractures	1 comparative NRSI (N=164) Bae 2019, 24 months, VP vs. medial branch block 15.2% (14/92) vs. 4.2% (3/72), RR 3.65 (1.09-12.23), remained significant after adjustment for confounding factors (data NR)
Symptomatic Cement Leakage	1 comparative NRSI (N=92 VP arm only) Bae 2019, 1 week, 1.1% (1/92), subjective leg weakness, resolved
Asymptomatic Cement Leakage	1 comparative NRSI (N=92 VP arm only) Bae 2019, 1 week, 4.3% (4/92)

CI = confidence interval; NRSI = nonrandomized study of interventions; VP = vertebroplasty

5.3.5 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSI of KP vs. UC and case series of KP

Outcome	KP vs. UC Effect estimate (95% CI) Conclusion
Mortality	Admin data, 2 studies Zampini 2010, 30 days, KP vs. UC, 0.3% vs. 1.6%, Adj OR 0.52, p=0.003 (CI NR); Edidin 2015, 4 years, UC vs. KP, Adj HR 1.62 (1.60–1.64) Case series (all KP): 1 case series (Bergmann 2012): 0.3% (1/297)
Pulmonary embolism (PE)	Admin data, 1 study: Edidin 2015, 4 years, KP vs. UC, Adj HR 0.99 (0.92-1.08) Case series (all KP): No studies
Deep vein thrombosis (DVT)	Admin data, 2 studies: Edidin 2015, 4 years, KP vs. UC, Adj HR 0.92 (0.0.87-0.96) Zampini 2010, inpatient, KP vs. UC, 0.2% (n=882 fractures) vs. 0.2% (n=4884 fractures), p=0.899 Case series (all KP): No studies

Outcome	KP vs. UC Effect estimate (95% CI) Conclusion
Cardiac complications (including MI)	Admin data, 1 study: Edidin 2015, 4 years, KP vs. UC, Adj HR 0.88 (0.83-0.93) Case series (all KP): No studies
Pulmonary/respiratory complications	Admin data, 1 study: Edidin 2015, 4 years, KP vs. UC, Adj HR 1.00 (0.98-1.02) Case series (all KP): No studies
Infection	Admin data, 2 studies: Edidin 2015, 4 years, KP vs. UC, Adj HR 1.00 (0.90-1.10) Zampini 2010, inpatient, 0.1% (n=882 fractures) vs. 0.1% (n=4884 fractures), p=0.929 Case series (all KP): No studies
Any SAE	Admin data, 1 study: Zampini 2010, inpatient, 1.7% (n=882 fractures) vs. 1.0% (n=4884 fractures), p=0.061 Case series (all kyphoplasty): 1 case series (Bergmann 2012): <1.0% (2/297) allergic reaction to the balloon (severe hypotension and tachycardia) and subcutaneous hematoma requiring release
New fracture (various)	Case series (all kyphoplasty): **Any new fracture** 3 studies, range 12.1% to 22.2%; **Any new symptomatic fracture**, 2 studies, range 8.1% to 10.6%; **Any new adjacent level fracture**, 4 studies, range 4.6% to 10.5%; **Symptomatic adjacent level fracture**, 2 studies, range 0.3% to 6.6%; **Refracture at index level,** 1 study, 0.7%
Symptomatic cement leakage	Case series (all kyphoplasty): 2 studies, range 0% to 2.3%
Subsequent augmentation	Admin data, 1 study: Edidin 2015, 4 years, KP vs. UC, Subsequent augmentation: Adj HR 12.5 (12.5-14.3); Subsequent augmentation or Fusion: Adj HR 12.5 (12.5-14.3) Case series (all kyphoplasty): Repeat KP for symptomatic fracture, 2 studies, range 8% to 10.6%

Adj = adjusted; CI = confidence interval; HR = hazard ratio; KP = kyphoplasty; MI = myocardial infarction; NRSI = nonrandomized study of interventions; SAE = serious adverse event; UC = usual care.

5.3.6 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSI of KP vs. Pedicle Screw Fixation

Outcome	KP vs. Surgery Effect estimate (95% CI) Conclusion
Mortality	1 comparative NRSI Wen 2021, 3 years, KP vs. screw fixation, 0% (0/376) vs. 0% (0/121)
Deep vein thrombosis (DVT)	1 comparative NRSI Wen 2021, 3 years, KP vs. screw fixation, 0% (0/376) vs. 2.5% (3/121)
Adjacent or distant new vertebral fracture	1 comparative NRSI Wen 2021, 3 years, KP vs. screw fixation, 7.7% (29/376) vs. 5.8% (7/121), unadjusted RR 1.33 (0.60-2.96)
Any reoperation	1 comparative NRSI Wen 2021, 3 years, KP vs. screw fixation, 7.7% (29/376) vs. 5.8% (7/121), unadjusted RR 1.33 (0.60-2.96), all for new fractures

CI = confidence interval; KP = kyphoplasty; NRSI = nonrandomized study of interventions.

5.3.7 Adverse events reported in nonrandomized studies in patients with osteoporotic vertebral compression fractures: Comparative NRSIs for any vertebral augmentation (VP or KP) vs. UC

Outcome	Any VA (VP or KP) vs. UC Effect estimate (95% CI) Conclusion
Mortality	Admin data: VP/KP vs. UC **<30 days, 1 study** McCullough 2013, 0.3% (31/9017) vs. 0.6% (51/9017), Adj OR 0.61 (95% CI 0.39-0.95) **>30 days, 3 studies** McCullough 2013, 5.2% (469/9017) vs. 5.6% (505/9017), Adj HR 0.92 (95% CI, 0.81-1.04); Levy 2012, Adj HR: 0.81 (0.42, 1.59) Lange 2014, Survival rates, 69.9% vs. 53.8%, Adj HR 0.58 (0.48-0.70); Case series (all VP/KP): 2 database studies, 2%
Any SAE	Admin data: VP/KP vs. UC **1 study, Purvis 2018, post-op** Post-op, 8.1% (900/11116) vs. 8.7% (4086/46962), Adj OR 0.95 (0.87-1.03) **1 study, McCullough 2013, 30 days and 1 year** 30 days, 9.5% (860/9017) vs. 10.5% (947/9017), Adj OR 0.90 (0.81-0.99) 1 year, 29.8% (2691/9017) vs. 30.0% (2709/9017), Adj HR 1.00 (0.94-1.06)

Outcome	Any VA (VP or KP) vs. UC Effect estimate (95% CI) Conclusion
	Case series (all VP/KP): 2 database studies (N=1932-2433), range, 4.9% to 5.8% 1 case series, no SAEs occurred in any patient (N=358)
Specific SAE	Admin data: VP/KP vs. UC **1 study, Purvis 2018, post-op** Stroke: 0.1% (11/11116) vs. 0% (0/46962) MI: 0.6% (67/11116) vs. 0.8% (376/46962) PE: 0.2% (22/11116) vs. 0.3% (141/46962) Shock: 0.2% (22/11116) vs. 0.2% (94/46962) All p>0.05 Case series (all VP/KP): 2 database studies (N=1932-2433), the incidence of individual SAEs – thromboembolic events, cardiac events, cerebrovascular events – was very low ≤1.0%
Recurrent/new fracture	Admin data: VP/KP vs. UC 1 study, Levy 2012, 17.5% (10/57) vs. 25.9% (7/27), unadjusted RR 0.68 (0.29 to 1.58); p>0.05 in adjusted analyses Case series (all VP/KP): 1 study, Wang 2014 Any new fracture: 12.6% (45/358) Symptomatic Adjacent level: 7.3% (26/358)
Deep infection, Sepsis	Case series (all VP/KP): Deep infection, 1 database, Kim 2022, 0% (0/1932) Sepsis/septic complication, 2 databases, Choo 2018, Kim 2022, range, 0.5%-0.8% (N=1932-2433)
Cement Embolism	Case series (all VP/KP): 2 studies, no symptomatic cases in either study. 1 study, Sun 2023, 17.2% (64/373) 1 study, Venmans 2008, 3.7% (11/299)
Symptomatic Cement leakage	Case series (all VP/KP): 1 study, Wang 2014, 0% (0/358)
Reoperation	Case series (all VP/KP): Any reoperation 2 database studies, Choo 2018, Kim 2022, range 3.2%-3.6% Repeat VP or KP for symptomatic new fracture 1 study, Wang 2014, 7.3%

Adj = adjusted; CI = confidence interval; HR = hazard ratio; KP = kyphoplasty; MI = myocardial infarction; NRSI = nonrandomized study of interventions; OR = odds ratio; PE = pulmonary embolism; RR = risk ratio; SAE = serious adverse event; VP = vertebroplasty.

5.3.8 Adverse events reported in nonrandomized studies of patients with osteoporotic vertebral compression fractures: Comparative NRSIs for any vertebral augmentation (VP or KP) vs. operative treatment

Outcome	Any VA (VP or KP) vs. Operative Effect estimate (95% CI) Conclusion
Any and specific SAEs	Admin data: VP/KP vs. Open surgery **1 study, Purvis 2018, post-op** Stroke: 0.1% (11/11116) vs. 0.3% (4/1487), p<0.001 MI: 0.6% (67/11116) vs. 2.2% (33/1487), p<0.001 PE: 0.2% (22/11116) vs. 1.2% (18/1487), p<0.001 Shock: 0.2% (22/11116) vs. 1.0% (15/1487), p<0.001 Any SAE: 8.1% (900/11116) vs. 16.3% (242/1487); Adj. OR 0.48 (95% CI 0.41-0.56) All p>0.05

CI = confidence interval; KP = kyphoplasty; MI = myocardial infarction; NRSI = nonrandomized study of interventions; OR = odds ratio; PE = pulmonary embolism; SAE = serious adverse event; VCF = vertebral compression fracture; VP = vertebroplasty.

6 References

1. WebPlotDigitizer.
2. Alexandru D, So W. Evaluation and management of vertebral compression fractures. Perm J 2012;16:46-51.
3. Alfonso M, Llombart R, Gil L, et al. [Translated article] Tumor ablation and vertebral augmentation in the treatment of vertebral metastases: A multicenter study. Rev Esp Cir Ortop Traumatol 2023;67:S480-s6.
4. Andresen JR, Radmer S, Andresen R, et al. Comparative outcome of different treatment options for fragility fractures of the sacrum. BMC Musculoskelet Disord 2022;23:1106.
5. Anselmetti GC, Manca A, Ortega C, Grignani G, Debernardi F, Regge D. Treatment of extraspinal painful bone metastases with percutaneous cementoplasty: a prospective study of 50 patients. Cardiovasc Intervent Radiol 2008;31:1165-73.
6. Ardakani AHG, Faimali M, Nystrom L, et al. Metastatic bone disease: Early referral for multidisciplinary care. Cleve Clin J Med 2022;89:393-9.
7. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. Bmj 2004;328:1490.
8. Bae IS, Chun HJ, Bak KH, Yi HJ, Choi KS, Kim KD. Medial Branch Block Versus Vertebroplasty for 1-Level Osteoporotic Vertebral Compression Fracture: 2-Year Retrospective Study. World Neurosurg 2019;122:e1599-e605.
9. Bae JS, Park JH, Kim KJ, Kim HS, Jang IT. Analysis of Risk Factors for Secondary New Vertebral Compression Fracture Following Percutaneous Vertebroplasty in Patients with Osteoporosis. World Neurosurg 2017;99:387-94.
10. Bae JW, Gwak HS, Kim S, et al. Percutaneous vertebroplasty for patients with metastatic compression fractures of the thoracolumbar spine: clinical and radiological factors affecting functional outcomes. Spine J 2016;16:355-64.
11. Basile A, Tsetis D, Cavalli M, et al. Sacroplasty for local or massive localization of multiple myeloma. Cardiovasc Intervent Radiol 2010;33:1270-7.
12. Bayley E, Srinivas S, Boszczyk BM. Clinical outcomes of sacroplasty in sacral insufficiency fractures: a review of the literature. Eur Spine J 2009;18:1266-71.
13. Beall DP, Shonnard NH, Shonnard MC, et al. An Interim Analysis of the First 102 Patients Treated in the Prospective Vertebral Augmentation Sacroplasty Fracture Registry. J Vasc Interv Radiol 2023;34:1477-84.
14. Berenson J, Pflugmacher R, Jarzem P, et al. Balloon kyphoplasty versus non-surgical fracture management for treatment of painful vertebral body compression fractures in patients with cancer: a multicentre, randomised controlled trial. Lancet Oncol 2011;12:225-35.
15. Bergmann M, Oberkircher L, Bliemel C, Frangen TM, Ruchholtz S, Krüger A. Early clinical outcome and complications related to balloon kyphoplasty. Orthop Rev (Pavia) 2012;4:e25.
16. Berkman ND, Lohr KN, Ansari M, et al. AHRQ Methods for Effective Health Care Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008.
17. Bian F, Bian G, Zhao L, Huang S, Fang J, An Y. Risk factors for recollapse of new vertebral compression fractures after percutaneous kyphoplasty in geriatric patients: establishment of a nomogram. BMC Musculoskelet Disord 2022;23:458.

18. Blasco J, Martinez-Ferrer A, Macho J, et al. 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;27:1159-66.
19. Boonen S, Van Meirhaeghe J, Bastian L, et al. Balloon kyphoplasty for the treatment of acute vertebral compression fractures: 2-year results from a randomized trial. J Bone Miner Res 2011;26:1627-37.
20. Borgström F, Beall DP, Berven S, et al. Health economic aspects of vertebral augmentation procedures. Osteoporos Int 2015;26:1239-49.
21. Bouza C, Lopez-Cuadrado T, Cediel P, Saz-Parkinson Z, Amate JM. Balloon kyphoplasty in malignant spinal fractures: a systematic review and meta-analysis. BMC Palliat Care 2009;8:12.
22. Briggs P, King SW, Staniland T, Gopal S, Shah R, Chimutengwende-Gordon M. A Systematic Review of Sacral Insufficiency Fractures: Treatment Modalities and Outcomes. Cureus 2023;15:e41745.
23. Buchbinder R, Johnston RV, Rischin KJ, et al. Percutaneous vertebroplasty for osteoporotic vertebral compression fracture. Cochrane Database Syst Rev 2018;11:Cd006349.
24. Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. N Engl J Med 2009;361:557-68.
25. Bureau USC. 2023 Population Projections Summary Tables. 2023.
26. Bureau USC. U.S. Older Population Grew From 2010 to 2020 at Fastest Rate Since 1880 to 1890 [Internet]. 2023.
27. Burton AW, Mendoza T, Gebhardt R, et al. Vertebral compression fracture treatment with vertebroplasty and kyphoplasty: experience in 407 patients with 1,156 fractures in a tertiary cancer center. Pain Med 2011;12:1750-7.
28. Butler CL, Given CA, 2nd, Michel SJ, Tibbs PA. Percutaneous sacroplasty for the treatment of sacral insufficiency fractures. AJR Am J Roentgenol 2005;184:1956-9.
29. Carli D, Venmans A, Lodder P, et al. Vertebroplasty versus Active Control Intervention for Chronic Osteoporotic Vertebral Compression Fractures: The VERTOS V Randomized Controlled Trial. Radiology 2023;308:e222535.
30. Carlson BC, Robinson WA, Wanderman NR, et al. A Review and Clinical Perspective of the Impact of Osteoporosis on the Spine. Geriatr Orthop Surg Rehabil 2019;10:2151459319861591.
31. Chandra V, Wajswol E, Shukla P, Contractor S, Kumar A. Safety and Efficacy of Sacroplasty for Sacral Fractures: A Systematic Review and Meta-Analysis. J Vasc Interv Radiol 2019;30:1845-54.
32. Chen AT, Cohen DB, Skolasky RL. Impact of nonoperative treatment, vertebroplasty, and kyphoplasty on survival and morbidity after vertebral compression fracture in the medicare population. J Bone Joint Surg Am 2013;95:1729-36.
33. 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;21:473-7.
34. Cheng J, Muheremu A, Zeng X, Liu L, Liu Y, Chen Y. Percutaneous vertebroplasty vs balloon kyphoplasty in the treatment of newly onset osteoporotic vertebral compression fractures: A retrospective cohort study. Medicine (Baltimore) 2019;98:e14793.
35. Choo S, Malik AT, Jain N, Yu E, Kim J, Khan SN. 30-day adverse outcomes, re-admissions and mortality following vertebroplasty/kyphoplasty. Clin Neurol Neurosurg 2018;174:129-33.
36. Chou R, Deyo R, Friedly J, et al. Noninvasive Treatments for Low Back Pain. 2016.
37. Chou R, Fu R, Dana T, Pappas M, Hart E, Mauer KM. Interventional Treatments for Acute and Chronic Pain: Systematic Review. Comparative Effectiveness Review No. 247. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 75Q80120D00006.) AHRQ Publication No. 21-EHC030. Rockville, MD: Agency for Healthcare Research and Quality; September 2021. DOI: https://doi.org/10.23970/AHRQEPCCER247. Posted final reports are located on the Effective Health Care Program search page.

38. Chou R, Fu R, Dana T, Pappas M, Hart E, Mauer KM. AHRQ Comparative Effectiveness Reviews. Interventional Treatments for Acute and Chronic Pain: Systematic Review. Rockville (MD): Agency for Healthcare Research and Quality (US); 2021.
39. Chou R, Hartung D, Turner J, et al. Opioid Treatments for Chronic Pain. 2020.
40. Chou R, Wagner J, Ahmed AY, et al. Treatments for Acute Pain: A Systematic Review. 2020.
41. Clark W, Bird P, Gonski P, et al. Safety and efficacy of vertebroplasty for acute painful osteoporotic fractures (VAPOUR): a multicentre, randomised, double-blind, placebo-controlled trial. Lancet 2016;388:1408-16.
42. Comstock BA, Sitlani CM, Jarvik JG, Heagerty PJ, Turner JA, Kallmes DF. Investigational vertebroplasty safety and efficacy trial (INVEST): patient-reported outcomes through 1 year. Radiology 2013;269:224-31.
43. Cui Y, Pan Y, Lin Y, Mi C, Wang B, Shi X. Risk factors for predicting cement leakage in percutaneous vertebroplasty for spinal metastases. J Orthop Sci 2022;27:79-83.
44. Daher M, Kreichati G, Kharrat K, Sebaaly A. Vertebroplasty versus Kyphoplasty in the Treatment of Osteoporotic Vertebral Compression Fractures: A Meta-Analysis. World Neurosurg 2023;171:65-71.
45. Damron TA, Mann KA. Fracture risk assessment and clinical decision making for patients with metastatic bone disease. J Orthop Res 2020;38:1175-90.
46. De Leacy R, Chandra RV, Barr JD, et al. The evidentiary basis of vertebral augmentation: a 2019 update. J Neurointerv Surg 2020;12:442-7.
47. Deibert CP, Gandhoke GS, Paschel EE, Gerszten PC. A Longitudinal Cohort Investigation of the Development of Symptomatic Adjacent Level Compression Fractures Following Balloon-assisted Kyphoplasty in a Series of 726 Patients. Pain Physician 2016;19:E1167-e72.
48. Dettori JR, Norvell DC. Discordant Systematic Reviews: Which to Believe? Global Spine J 2020;10:237-9.
49. Dettori JR, Skelly AC, Brodt ED. Critically Low Confidence in the Results Produced by Spine Surgery Systematic Reviews: An AMSTAR-2 Evaluation From 4 Spine Journals. Global Spine J 2020;10:667-73.
50. Diamond T, Clark W, Bird P, Gonski P, Barnes E, Gebski V. Early vertebroplasty within 3 weeks of fracture for acute painful vertebral osteoporotic fractures: subgroup analysis of the VAPOUR trial and review of the literature. Eur Spine J 2020;29:1606-13.
51. Ding J, Zhang Q, Zhu J, et al. Risk factors for predicting cement leakage following percutaneous vertebroplasty for osteoporotic vertebral compression fractures. Eur Spine J 2016;25:3411-7.
52. Dohm M, Black CM, Dacre A, Tillman JB, Fueredi G. A randomized trial comparing balloon kyphoplasty and vertebroplasty for vertebral compression fractures due to osteoporosis. AJNR Am J Neuroradiol 2014;35:2227-36.
53. Donnally IC, DiPompeo CM, Varacallo M. Vertebral Compression Fractures. 2024.
54. Ebeling PR, Akesson K, Bauer DC, et al. The Efficacy and Safety of Vertebral Augmentation: A Second ASBMR Task Force Report. J Bone Miner Res 2019;34:3-21.
55. Edidin AA, Ong KL, Lau E, Kurtz SM. Mortality risk for operated and nonoperated vertebral fracture patients in the medicare population. J Bone Miner Res 2011;26:1617-26.
56. Edidin AA, Ong KL, Lau E, Kurtz SM. Morbidity and Mortality After Vertebral Fractures: Comparison of Vertebral Augmentation and Nonoperative Management in the Medicare Population. Spine (Phila Pa 1976) 2015;40:1228-41.
57. Edidin AA, Ong KL, Lau E, Schmier JK, Kemner JE, Kurtz SM. Cost-effectiveness analysis of treatments for vertebral compression fractures. Appl Health Econ Health Policy 2012;10:273-84.
58. Ender SA, Gradl G, Ender M, Langner S, Merk HR, Kayser R. Osseofix® system for percutaneous stabilization of osteoporotic and tumorous vertebral compression fractures - clinical and radiological results after 12 months. Rofo 2014;186:380-7.

59. Endres S, Badura A. Shield kyphoplasty through a unipedicular approach compared to vertebroplasty and balloon kyphoplasty in osteoporotic thoracolumbar fracture: a prospective randomized study. Orthop Traumatol Surg Res 2012;98:334-40.
60. England RW, Gong A, Li T, et al. Clinical outcomes and safety of the SpineJack vertebral augmentation system for the treatment of vertebral compression fractures in a United States patient population. J Clin Neurosci 2021;89:237-42.
61. Evans AJ, Kip KE, Brinjikji W, et al. Randomized controlled trial of vertebroplasty versus kyphoplasty in the treatment of vertebral compression fractures. J Neurointerv Surg 2016;8:756-63.
62. Fadili Hassani S, Cormier E, Shotar E, et al. Intracardiac cement embolism during percutaneous vertebroplasty: incidence, risk factors and clinical management. Eur Radiol 2019;29:663-73.
63. Faloon MJ, Ruoff M, Deshpande C, et al. Risk Factors Associated with Adjacent and Remote-Level Pathologic Vertebral Compression Fracture Following Balloon Kyphoplasty: 2-Year Follow-Up Comparison Versus Conservative Treatment. J Long Term Eff Med Implants 2015;25:313-9.
64. 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;14:561-9.
65. Ferreira ML, March L. Vertebral fragility fractures - How to treat them? Best Pract Res Clin Rheumatol 2019;33:227-35.
66. Firanescu CE, de Vries J, Lodder P, et al. Percutaneous Vertebroplasty is no Risk Factor for New Vertebral Fractures and Protects Against Further Height Loss (VERTOS IV). Cardiovasc Intervent Radiol 2019;42:991-1000.
67. Firanescu CE, de Vries J, Lodder P, et al. Vertebroplasty versus sham procedure for painful acute osteoporotic vertebral compression fractures (VERTOS IV): randomised sham controlled clinical trial. Bmj 2018;361:k1551.
68. Food US, Drug A. Clinical Trial Considerations: Vertebral Augmentation Devices to Treat Spinal Insufficiency Fractures - Guidance for Industry and FDA Staff. 2004.
69. Fourney DR, Schomer DF, Nader R, et al. Percutaneous vertebroplasty and kyphoplasty for painful vertebral body fractures in cancer patients. J Neurosurg 2003;98:21-30.
70. Frey ME, Depalma MJ, Cifu DX, Bhagia SM, Carne W, Daitch JS. Percutaneous sacroplasty for osteoporotic sacral insufficiency fractures: a prospective, multicenter, observational pilot study. Spine J 2008;8:367-73.
71. Frey ME, Warner C, Thomas SM, et al. Sacroplasty: A Ten-Year Analysis of Prospective Patients Treated with Percutaneous Sacroplasty: Literature Review and Technical Considerations. Pain Physician 2017;20:E1063-e72.
72. Fritzell P, Ohlin A, Borgström F. Cost-effectiveness of balloon kyphoplasty versus standard medical treatment in patients with osteoporotic vertebral compression fracture: a Swedish multicenter randomized controlled trial with 2-year follow-up. Spine (Phila Pa 1976) 2011;36:2243-51.
73. Galibert P, Deramond H, Rosat P, Le Gars D. Note préliminaire sur le traitement des angiomes vertébraux par vertébroplastie acrylique percutanée [Preliminary note on the treatment of vertebral angioma by percutaneous acrylic vertebroplasty]. Neurochirurgie 1987;33:166-8.
74. García-Maroto R, García-Coiradas J, Milano G, Cebrián JL, Marco F, López-Durán L. [Safety and efficacy of kyphoplasty in the treatment of tumoral disease of the spine]. Rev Esp Cir Ortop Traumatol 2015;59:406-12.
75. Georgy BA. Interventional techniques in managing persistent pain after vertebral augmentation procedures: a retrospective evaluation. Pain Physician 2007;10:673-6.
76. GitHub I. A collection of R functions supporting the text book Modern Epidemiology, Second Edition, by Kenneth J. Rothman and Sander Greenland. 2017.

77. Griffoni C, Lukassen JNM, Babbi L, et al. Percutaneous vertebroplasty and balloon kyphoplasty in the treatment of osteoporotic vertebral fractures: a prospective randomized comparison. Eur Spine J 2020;29:1614-20.
78. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383-94.
79. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 2011;64:380-2.
80. Hansen EJ, Rousing, R., Tropp, H., Carreon, L.Y., Osterhaden Andersen, M. Vertebroplasty vs. SHAM for treating osteoporotic vertebral compression fractures: a double blind RCT (VOPE). Integrative Journal of Orthopaedics ad Traumatology 2019;2:2-6.
81. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. Stat Med 1996;15:619-29.
82. Heo DH, Park CK. Percutaneous Sacroplasty for Non-neoplastic Osteoporotic Sacral Insufficiency Fractures. Pain Physician 2017;20:89-94.
83. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj 2003;327:557-60.
84. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. 2011.
85. Hinde K, Maingard J, Hirsch JA, Phan K, Asadi H, Chandra RV. Mortality Outcomes of Vertebral Augmentation (Vertebroplasty and/or Balloon Kyphoplasty) for Osteoporotic Vertebral Compression Fractures: A Systematic Review and Meta-Analysis. Radiology 2020;295:96-103.
86. Hopkins TJ, Eggington S, Quinn M, Nichols-Ricker CI. Cost-effectiveness of balloon kyphoplasty and vertebroplasty versus conservative medical management in the USA. Osteoporos Int 2020;31:2461-71.
87. Huang KY, Lee SC, Liu WL, Wang JD. Long Term Survival of Pathological Thoracolumbar Fractures Treated with Vertebroplasty: Analysis Using a Nationwide Insurance Claim Database. J Clin Med 2019;9.
88. International Osteoporosis Foundation. Facts & Statistics [Internet]. Nyon (Switzerland): International Osteoporosis Foundation 2022.
89. Kallmes DF, Comstock BA, Heagerty PJ, et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. N Engl J Med 2009;361:569-79.
90. Kamel EM, Binaghi S, Guntern D, Mouhsine E, Schnyder P, Theumann N. Outcome of long-axis percutaneous sacroplasty for the treatment of sacral insufficiency fractures. Eur Radiol 2009;19:3002-7.
91. Kang SE, Lee JW, Kim JH, Park KW, Yeom JS, Kang HS. Percutaneous sacroplasty with the use of C-arm flat-panel detector CT: technical feasibility and clinical outcome. Skeletal Radiol 2011;40:453-60.
92. Kim HJ, Zuckerman SL, Cerpa M, Yeom JS, Lehman RA, Jr., Lenke LG. Incidence and Risk Factors for Complications and Mortality After Vertebroplasty or Kyphoplasty in the Osteoporotic Vertebral Compression Fracture-Analysis of 1,932 Cases From the American College of Surgeons National Surgical Quality Improvement. Global Spine J 2022;12:1125-34.
93. Klazen CA, Lohle PN, de Vries J, et al. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. Lancet 2010;376:1085-92.
94. Klazen CA, Venmans A, de Vries J, et al. Percutaneous vertebroplasty is not a risk factor for new osteoporotic compression fractures: results from VERTOS II. AJNR Am J Neuroradiol 2010;31:1447-50.
95. Klazen CA, Verhaar HJ, Lohle PN, et al. Clinical course of pain in acute osteoporotic vertebral compression fractures. J Vasc Interv Radiol 2010;21:1405-9.
96. Kobayashi N, Noguchi T, Kobayashi D, et al. Safety and Efficacy of Percutaneous Vertebroplasty for Osteoporotic Vertebral Compression Fractures: A Multicenter Retrospective Study in Japan. Interv Radiol (Higashimatsuyama) 2021;6:21-8.

97. Kortman K, Ortiz O, Miller T, et al. Multicenter study to assess the efficacy and safety of sacroplasty in patients with osteoporotic sacral insufficiency fractures or pathologic sacral lesions. J Neurointerv Surg 2013;5:461-6.
98. Kose KC, Cebesoy O, Akan B, Altinel L, Dincer D, Yazar T. Functional results of vertebral augmentation techniques in pathological vertebral fractures of myelomatous patients. J Natl Med Assoc 2006;98:1654-8.
99. Kroon F, Staples M, Ebeling PR, et al. Two-year results of a randomized placebo-controlled trial of vertebroplasty for acute osteoporotic vertebral fractures. J Bone Miner Res 2014;29:1346-55.
100. Lange A, Kasperk C, Alvares L, Sauermann S, Braun S. Survival and cost comparison of kyphoplasty and percutaneous vertebroplasty using German claims data. Spine (Phila Pa 1976) 2014;39:318-26.
101. Leali PT, Solla F, Maestretti G, Balsano M, Doria C. Safety and efficacy of vertebroplasty in the treatment of osteoporotic vertebral compression fractures: a prospective multicenter international randomized controlled study. Clin Cases Miner Bone Metab 2016;13:234-6.
102. Lee IJ, Choi AL, Yie MY, et al. CT evaluation of local leakage of bone cement after percutaneous kyphoplasty and vertebroplasty. Acta Radiol 2010;51:649-54.
103. Levy H, Seydafkan S, Rice JD, Easley KA, Tangpricha V. Comparative efficacy of vertebroplasty, kyphoplasty, and medical therapy for vertebral fractures on survival and prevention of recurrent fractures. Endocr Pract 2012;18:499-507.
104. Lewiecki EM, Ortendahl JD, Vanderpuye-Orgle J, et al. Healthcare Policy Changes in Osteoporosis Can Improve Outcomes and Reduce Costs in the United States. JBMR Plus 2019;3:e10192.
105. Li Y, Zhu, J, Xie, C. A comparative study of percutanous kyphoplasty and conservative therapy on vertebral osteoporotic compression fractures in elderly patients. Int J Clin Exp Med 2017;10:8139-45.
106. Lin D, Hao J, Li L, et al. Effect of Bone Cement Volume Fraction on Adjacent Vertebral Fractures After Unilateral Percutaneous Kyphoplasty. Clin Spine Surg 2017;30:E270-e5.
107. Liu J, Li X, Tang D, et al. Comparing pain reduction following vertebroplasty and conservative treatment for osteoporotic vertebral compression fractures: a meta-analysis of randomized controlled trials. Pain Physician 2013;16:455-64.
108. Liu JT, Li CS, Chang CS, Liao WJ. Long-term follow-up study of osteoporotic vertebral compression fracture treated using balloon kyphoplasty and vertebroplasty. J Neurosurg Spine 2015;23:94-8.
109. Liu JT, Liao WJ, Tan WC, et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. Osteoporos Int 2010;21:359-64.
110. Liu Q, Cao J, Kong JJ. Clinical effect of balloon kyphoplasty in elderly patients with multiple osteoporotic vertebral fracture. Niger J Clin Pract 2019;22:289-92.
111. Liu Y, Liu J, Suvithayasiri S, Han I, Kim JS. Comparative Efficacy of Surgical Interventions for Osteoporotic Vertebral Compression Fractures: A Systematic Review and Network Meta-analysis. Neurospine 2023;20:1142-58.
112. Long Y, Yi W, Yang D. Advances in Vertebral Augmentation Systems for Osteoporotic Vertebral Compression Fractures. Pain Res Manag 2020;2020:3947368.
113. Mahmood B, Pasternack J, Razi A, Saleh A. Safety and efficacy of percutaneous sacroplasty for treatment of sacral insufficiency fractures: a systematic review. J Spine Surg 2019;5:365-71.
114. Masala S, Ciarrapico AM, Konda D, Vinicola V, Mammucari M, Simonetti G. Cost-effectiveness of percutaneous vertebroplasty in osteoporotic vertebral fractures. Eur Spine J 2008;17:1242-50.
115. Matsumoto T, Hoshino M, Tsujio T, et al. Prognostic factors for reduction of activities of daily living following osteoporotic vertebral fractures. Spine (Phila Pa 1976) 2012;37:1115-21.

116. Mattie R, Brar N, Tram JT, et al. Vertebral Augmentation of Cancer-Related Spinal Compression Fractures: A Systematic Review and Meta-Analysis. Spine (Phila Pa 1976) 2021;46:1729-37.
117. McCarthy J, Davis A. Diagnosis and Management of Vertebral Compression Fractures. Am Fam Physician 2016;94:44-50.
118. McCullough BJ, Comstock BA, Deyo RA, Kreuter W, Jarvik JG. Major medical outcomes with spinal augmentation vs conservative therapy. JAMA Intern Med 2013;173:1514-21.
119. Mendel E, Bourekas E, Gerszten P, Golan JD. Percutaneous techniques in the treatment of spine tumors: what are the diagnostic and therapeutic indications and outcomes? Spine (Phila Pa 1976) 2009;34:S93-100.
120. Molloy S, Sewell MD, Platinum J, et al. Is balloon kyphoplasty safe and effective for cancer-related vertebral compression fractures with posterior vertebral body wall defects? J Surg Oncol 2016;113:835-42.
121. Moulin B, Tselikas L, Gravel G, et al. Safety and Efficacy of Multilevel Thoracolumbar Vertebroplasty in the Simultaneous Treatment of Six or More Pathologic Compression Fractures. J Vasc Interv Radiol 2020;31:1683-9.e1.
122. Nguyễn MV, Carlier C, Nich C, Gouin F, Crenn V. Fracture Risk of Long Bone Metastases: A Review of Current and New Decision-Making Tools for Prophylactic Surgery. Cancers (Basel) 2021;13.
123. Nieuwenhuijse MJ, van Rijswijk CS, van Erkel AR, Dijkstra SP. The intravertebral cleft in painful long-standing osteoporotic vertebral compression fractures treated with percutaneous vertebroplasty: diagnostic assessment and clinical significance. Spine (Phila Pa 1976) 2012;37:974-81.
124. Ning L, Zhu J, Tian S, et al. Correlation Analysis Between Basic Diseases and Subsequent Vertebral Fractures After Percutaneous Kyphoplasty (PKP) for Osteoporotic Vertebral Compression Fractures. Pain Physician 2021;24:E803-e10.
125. Nuti R, Brandi ML, Checchia G, et al. Guidelines for the management of osteoporosis and fragility fractures. Intern Emerg Med 2019;14:85-102.
126. O'Neill BE, Godil JA, Smith S, et al. Utilization of Vertebroplasty/Kyphoplasty in the Management of Compression Fractures: National Trends and Predictors of Vertebroplasty/Kyphoplasty. Neurospine 2023;20:1132-9.
127. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. J Manag Care Pharm 2003;9:53-61.
128. Olivarez LM, Dipp JM, Escamilla RF, et al. Vertebral augmentation treatment of painful osteoporotic compression fractures with the Kiva VCF Treatment System. Sas j 2011;5:114-9.
129. Ong KL, Beall DP, Frohbergh M, Lau E, Hirsch JA. Were VCF patients at higher risk of mortality following the 2009 publication of the vertebroplasty "sham" trials? Osteoporos Int 2018;29:375-83.
130. Ontario HQ. Vertebral Augmentation Involving Vertebroplasty or Kyphoplasty for Cancer-Related Vertebral Compression Fractures: A Systematic Review. Ont Health Technol Assess Ser 2016;16:1-202.
131. Ontario HQ. Vertebral Augmentation Involving Vertebroplasty or Kyphoplasty for Cancer-Related Vertebral Compression Fractures: An Economic Analysis. Ont Health Technol Assess Ser 2016;16:1-34.
132. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992;116:78-84.
133. Parreira PCS, Maher CG, Megale RZ, March L, Ferreira ML. An overview of clinical guidelines for the management of vertebral compression fracture: a systematic review. Spine J 2017;17:1932-8.
134. Pron G, Hwang M, Nasralla M, Smith R, Cheung A, Murphy K. Cost-effectiveness and willing-to-pay thresholds for vertebral augmentation of osteoporotic vertebral fractures, what are they based on: a systematic review. BMJ Open 2023;13:e062832.

135. Pron G, Hwang M, Smith R, Cheung A, Murphy K. Cost-effectiveness studies of vertebral augmentation for osteoporotic vertebral fractures: a systematic review. Spine J 2022;22:1356-71.
136. Purvis TE, De la Garza Ramos R, Sankey EW, Karikari IO, Goodwin CR, Sciubba DM. In-hospital medical complications after non-operative and operative treatment of thoracolumbar fractures in patients over 75 years of age. J Clin Neurosci 2018;50:83-7.
137. Qi J, Hu Y, Yang Z, et al. Incidence, Risk Factors, and Outcomes of Symptomatic Bone Cement Displacement following Percutaneous Kyphoplasty for Osteoporotic Vertebral Compression Fracture: A Single Center Study. J Clin Med 2022;11.
138. Qiu Z, Wang P, Chao Y, Yu Y. The risk of new vertebral fracture after percutaneous vertebral augmentation in patients suffering from single-level osteoporotic vertebral compression fractures: A meta-analysis and systematic review. Medicine (Baltimore) 2023;102:e35749.
139. Quality AfHRa. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. 2014.
140. Rizzo SE, Kenan S. Pathologic Fractures. 2024.
141. Rocha Romero A, Hernández-Porras BC, Plancarte-Sanchez R, et al. Risk of New Fractures in Vertebroplasty for Multiple Myeloma. A Retrospective Study. Pain Med 2020;21:3018-23.
142. Röllinghoff M, Zarghooni K, Schlüter-Brust K, et al. Indications and contraindications for vertebroplasty and kyphoplasty. Arch Orthop Trauma Surg 2010;130:765-74.
143. Rose LD, Bateman G, Ahmed A. Clinical significance of cement leakage in kyphoplasty and vertebroplasty: a systematic review. Eur Spine J 2023.
144. 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 (Phila Pa 1976) 2009;34:1349-54.
145. Rousing R, Hansen KL, Andersen MO, Jespersen SM, Thomsen K, Lauritsen JM. Twelve-months follow-up in forty-nine patients with acute/semiacute osteoporotic vertebral fractures treated conservatively or with percutaneous vertebroplasty: a clinical randomized study. Spine (Phila Pa 1976) 2010;35:478-82.
146. Sarafrazi N, Wambogo EA, Shepherd JA. Osteoporosis or Low Bone Mass in Older Adults: United States, 2017-2018. NCHS Data Brief 2021:1-8.
147. Sarigul B, Ogrenci A, Yilmaz M, Koban O, Mammadov M, Dalbayrak S. Sacral insufficiency fracture: a single-center experience of 185 patients with a minimum 5-year follow-up. Eur Spine J 2023.
148. Savage JW, Schroeder GD, Anderson PA. Vertebroplasty and kyphoplasty for the treatment of osteoporotic vertebral compression fractures. J Am Acad Orthop Surg 2014;22:653-64.
149. Schünemann HJ, Tugwell P, Reeves BC, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Res Synth Methods 2013;4:49-62.
150. Shah LM, Jennings JW, Kirsch CFE, et al. ACR Appropriateness Criteria(®) Management of Vertebral Compression Fractures. J Am Coll Radiol 2018;15:S347-s64.
151. Shin MH, Zmuda JM, Barrett-Connor E, et al. Race/ethnic differences in associations between bone mineral density and fracture history in older men. Osteoporos Int 2014;25:837-45.
152. Skelly AC, Chou R, Dettori JR, et al. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review. 2018.
153. Skelly AC, Chou R, Dettori JR, et al. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review Update. 2020.
154. Sørensen ST, Kirkegaard AO, Carreon L, Rousing R, Andersen M. Vertebroplasty or kyphoplasty as palliative treatment for cancer-related vertebral compression fractures: a systematic review. Spine J 2019;19:1067-75.
155. Spine Market G. Kiva VCF Treatment System Receives FDA Clearance for Vertebral Compression Fractures.

156. Spross C, Aghayev E, Kocher R, Röder C, Forster T, Kuelling FA. Incidence and risk factors for early adjacent vertebral fractures after balloon kyphoplasty for osteoporotic fractures: analysis of the SWISSspine registry. Eur Spine J 2014;23:1332-8.
157. Staples MP, Howe BM, Ringler MD, et al. New vertebral fractures after vertebroplasty: 2-year results from a randomised controlled trial. Arch Osteoporos 2015;10:229.
158. StataCorp. Stata Statistical Software: Release 14. 2015.
159. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Bmj 2011;343:d4002.
160. Stevenson M, Gomersall T, Lloyd Jones M, et al. Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: a systematic review and cost-effectiveness analysis. Health Technol Assess 2014;18:1-290.
161. Ström O, Leonard C, Marsh D, Cooper C. Cost-effectiveness of balloon kyphoplasty in patients with symptomatic vertebral compression fractures in a UK setting. Osteoporos Int 2010;21:1599-608.
162. Strub WM, Hoffmann M, Ernst RJ, Bulas RV. Sacroplasty by CT and fluoroscopic guidance: is the procedure right for your patient? AJNR Am J Neuroradiol 2007;28:38-41.
163. Sun X, Deng M, Xu W, et al. CT features and risk factors of pulmonary cement embolism after vertebroplasty or kyphoplasty in patients with vertebral compression fracture: a retrospective cohort study. Quant Imaging Med Surg 2023;13:2397-407.
164. Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. Jama 2014;311:405-11.
165. Svedbom A, Alvares L, Cooper C, Marsh D, Ström O. Balloon kyphoplasty compared to vertebroplasty and nonsurgical management in patients hospitalised with acute osteoporotic vertebral compression fracture: a UK cost-effectiveness analysis. Osteoporos Int 2013;24:355-67.
166. Svensson HK, Olsson LE, Hansson T, Karlsson J, Hansson-Olofsson E. The effects of person-centered or other supportive interventions in older women with osteoporotic vertebral compression fractures-a systematic review of the literature. Osteoporos Int 2017;28:2521-40.
167. Takahashi S, Hoshino M, Yasuda H, et al. Cost-effectiveness of Balloon Kyphoplasty for Patients With Acute/Subacute Osteoporotic Vertebral Fractures in the Super-Aging Japanese Society. Spine (Phila Pa 1976) 2019;44:E298-e305.
168. Tan CW, Arlachov Y, Czernicki M, et al. Spinal medial branch nerve root block (MBNB) intervention compared to standard care-vertebroplasty (VP) for the treatment of painful osteoporotic vertebral fractures in frail, older hospitalised patients: a feasibility study. Arch Osteoporos 2023;18:126.
169. Tang B, Cui L, Chen X, Liu Y. Risk Factors for Cement Leakage in Percutaneous Vertebroplasty for Osteoporotic Vertebral Compression Fractures: An Analysis of 1456 Vertebrae Augmented by Low-Viscosity Bone Cement. Spine (Phila Pa 1976) 2021;46:216-22.
170. Tian J, Xiang L, Zhou D, Fan Q, Ma B. The clinical efficacy of vertebroplasty on osteoporotic vertebral compression fracture: a meta-analysis. Int J Surg 2014;12:1249-53.
171. Tomasian A, Khan MA, Jennings JW. Percutaneous Treatment of Spinal Metastases. Neuroimaging Clin N Am 2023;33:499-506.
172. Tsuda T. Epidemiology of fragility fractures and fall prevention in the elderly: a systematic review of the literature. Curr Orthop Pract 2017;28:580-5.
173. Tsujio T, Nakamura H, Terai H, et al. Characteristic radiographic or magnetic resonance images of fresh osteoporotic vertebral fractures predicting potential risk for nonunion: a prospective multicenter study. Spine (Phila Pa 1976) 2011;36:1229-35.
174. 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 (Phila Pa 1976) 2013;38:971-83.

175. Vanni D, Galzio R, Kazakova A, et al. Third-generation percutaneous vertebral augmentation systems. J Spine Surg 2016;2:13-20.
176. Vanni D, Pantalone A, Bigossi F, Pineto F, Lucantoni D, Salini V. New perspective for third generation percutaneous vertebral augmentation procedures: Preliminary results at 12 months. J Craniovertebr Junction Spine 2012;3:47-51.
177. Venmans A, Lohle PN, van Rooij WJ, Verhaar HJ, Mali WP. Frequency and outcome of pulmonary polymethylmethacrylate embolism during percutaneous vertebroplasty. AJNR Am J Neuroradiol 2008;29:1983-5.
178. Viswanathan M, Ansari MT, Berkman ND, et al. Chapter 9: Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. 2014.
179. Vogl TJ, Pflugmacher R, Hierholzer J, et al. Cement directed kyphoplasty reduces cement leakage as compared with vertebroplasty: results of a controlled, randomized trial. Spine (Phila Pa 1976) 2013;38:1730-6.
180. Voormolen MH, Mali WP, Lohle PN, et al. Percutaneous vertebroplasty compared with optimal pain medication treatment: short-term clinical outcome of patients with subacute or chronic painful osteoporotic vertebral compression fractures. The VERTOS study. AJNR Am J Neuroradiol 2007;28:555-60.
181. 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;25:3486-94.
182. 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:E187-94.
183. Wang S, Wang H, Niu L. Clinical efficacy of PVP and PKP in the treatment of OVCFs after bilateral resection of ovarian cancer. Oncol Lett 2018;16:151-6.
184. Wang W, Liu Y, Wan H, et al. Effectiveness and prognostic factors of different minimally invasive surgeries for vertebral compression fractures. BMC Musculoskelet Disord 2023;24:11.
185. Wang YT, Wu XT, Chen H, Wang C, Mao ZB. Adjacent-level symptomatic fracture after percutaneous vertebral augmentation of osteoporotic vertebral compression fracture: a retrospective analysis. J Orthop Sci 2014;19:868-76.
186. Wang Z, Wang R, Xiang S, et al. Assessment of the effectiveness and satisfaction of platelet-rich plasma compared with hyaluronic acid in knee osteoarthritis at minimum 7-year follow-up: A post hoc analysis of a randomized controlled trial. Front Bioeng Biotechnol 2022;10:1062371.
187. Wardlaw D, Cummings SR, Van Meirhaeghe J, et al. Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet 2009;373:1016-24.
188. Washington State Health Care Authority HTACC. Vertebroplasty, Kyphoplasty and Sacroplasty. 2010.
189. Wen Z, Mo X, Zhao S, et al. Comparison of Percutaneous Kyphoplasty and Pedicle Screw Fixation for Treatment of Thoracolumbar Severe Osteoporotic Vertebral Compression Fracture with Kyphosis. World Neurosurg 2021;152:e589-e96.
190. Werner CM, Osterhoff G, Schlickeiser J, et al. Vertebral body stenting versus kyphoplasty for the treatment of osteoporotic vertebral compression fractures: a randomized trial. J Bone Joint Surg Am 2013;95:577-84.
191. Whitlow CT, Mussat-Whitlow BJ, Mattern CW, Baker MD, Morris PP. Sacroplasty versus vertebroplasty: comparable clinical outcomes for the treatment of fracture-related pain. AJNR Am J Neuroradiol 2007;28:1266-70.

192. Wu CL, Wu JC, Huang WC, et al. The risk of stroke after percutaneous vertebroplasty for osteoporosis: a population-based cohort study. PLoS One 2012;7:e31405.
193. Wu W, Zhang X, Li X, et al. Comparison of the clinical outcomes of percutaneous kyphoplasty for the management of osteolytic and osteoblastic-related metastatic vertebral lesions. J Neurointerv Surg 2022;14:938-41.
194. Wu W, Zhang X, Li X, Yu S. Clinical analysis of percutaneous kyphoplasty for spinal metastases in older adults with comorbidities. J Geriatr Oncol 2023;14:101410.
195. Yang EZ, Xu JG, Huang GZ, et al. Percutaneous Vertebroplasty Versus Conservative Treatment in Aged Patients With Acute Osteoporotic Vertebral Compression Fractures: A Prospective Randomized Controlled Clinical Study. Spine (Phila Pa 1976) 2016;41:653-60.
196. Yang YC, Hsieh MH, Chien JT, Liu KC, Yang CC. Teriparatide treatment shows faster healing than sacroplasty for postmenopausal women with sacral insufficiency fracture. Osteoporos Sarcopenia 2023;9:27-31.
197. Yi X, Lu H, Tian F, et al. Recompression in new levels after percutaneous vertebroplasty and kyphoplasty compared with conservative treatment. Arch Orthop Trauma Surg 2014;134:21-30.
198. Zampini JM, White AP, McGuire KJ. Comparison of 5766 vertebral compression fractures treated with or without kyphoplasty. Clin Orthop Relat Res 2010;468:1773-80.
199. Zhang B, Chen G, Yang X, Fan T, Chen Z. Percutaneous Kyphoplasty Versus Percutaneous Vertebroplasty for Neurologically Intact Osteoporotic Kümmell's Disease: A Systematic Review and Meta-Analysis. Global Spine J 2022;12:308-22.
200. Zhang T, Peng Y, Li J. Comparison of clinical and radiological outcomes of vertebral body stenting versus percutaneous kyphoplasty for the treatment of osteoporotic vertebral compression fracture: A systematic review and meta-analysis. Jt Dis Relat Surg 2024;35:218-30.
201. Zhang TY, Zhang PX, Xue F, Zhang DY, Jiang BG. Risk factors for cement leakage and nomogram for predicting the intradiscal cement leakage after the vertebra augmented surgery. BMC Musculoskelet Disord 2020;21:792.
202. Zhao C, Liu X, Wang Y, et al. The effects of biomechanical factors on adjacent vertebral compression fractures after percutaneous kyphoplasty: a propensity score matching analysis. Osteoporos Int 2022;33:1795-806.
203. Zou J, Mei X, Gan M, Yang H. Kyphoplasty for spinal fractures from multiple myeloma. J Surg Oncol 2010;102:43-7.
204. Zuo XH, Zhu XP, Bao HG, et al. Network meta-analysis of percutaneous vertebroplasty, percutaneous kyphoplasty, nerve block, and conservative treatment for nonsurgery options of acute/subacute and chronic osteoporotic vertebral compression fractures (OVCFs) in short-term and long-term effects. Medicine (Baltimore) 2018;97:e11544.