Washington State Health Care Authority

Vertebroplasty, Kyphoplasty, Sacroplasty – Rereview

Final Appendix

October 16, 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

Aggregate Analytics, Inc.

Final Appendix

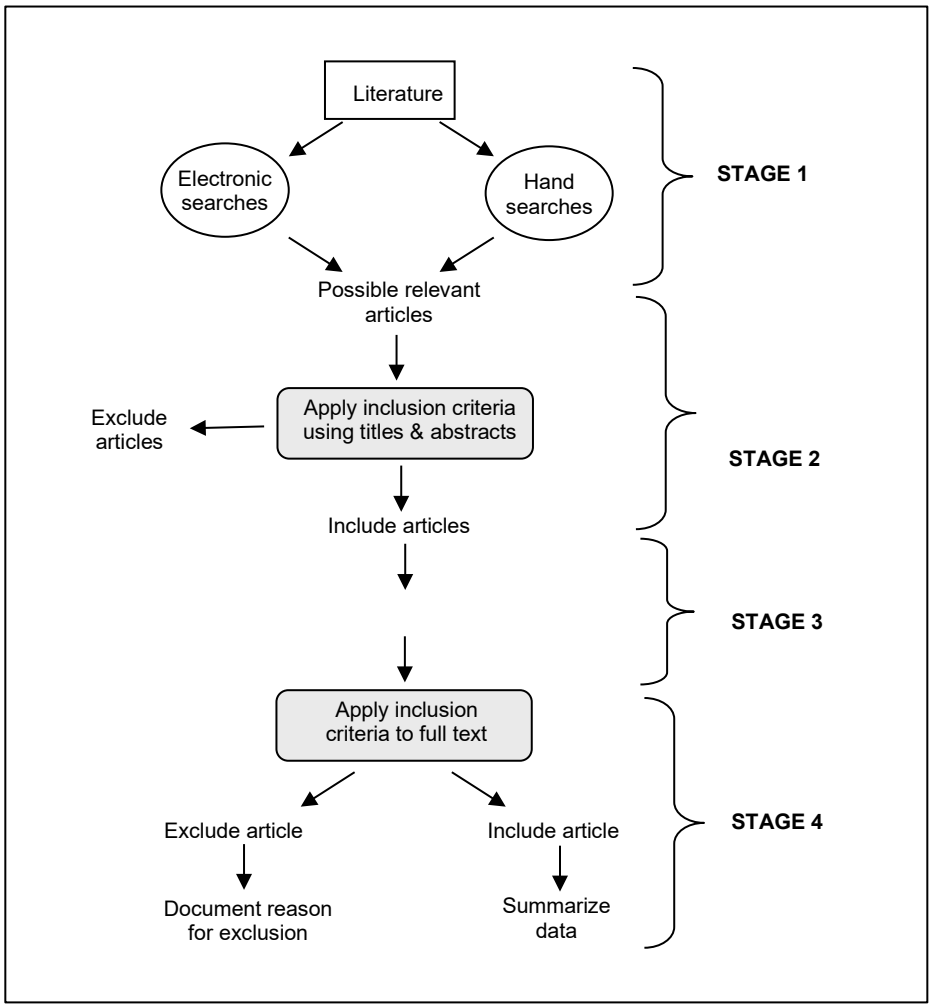
October 16, 2024

Table of Contents

Appendix Tables

Appendix Figures

APPENDIX A. Algorithm for Article Selection



APPENDIX B. Search Strategies

Below is the search strategy for PubMed. Parallel strategies were used to search other electronic databases listed below. Keyword searches were conducted in the other listed resources. In addition, hand-searching of included studies was performed.

Appendix Table B1: PubMed Search Strategy for Lit Search

Search period: January 1, 2010 – January 3, 2024

1.	(vertebroplast* OR kyphoplast* OR sacroplast* OR skyphoplast* OR vertebral augmentation)
2.	Fha[Filter]
3.	#1 and #2
4.	#3 NOT comment[Publication Type]
5.	#4 NOT case reports[Publication Type]
6.	#5 NOT review[Publication Type]
7.	#6 NOT editorial[Publication Type]
8.	#7 NOT cadaver*
9.	#7 Filters: Abstract, from 2010 - 2014

Appendix Table B2: PubMed Search Strategy for Cost Effectiveness Lit Search

Search period: January 1, 2010 – January 3, 2024

1.	(vertebroplast* OR kyphoplast* OR sacroplast* OR vesselplast* OR skyphoplast* OR percutaneous vertebral augmentation OR cement augmentation)
2.	fha[Filter]
3.	#1 AND #2
4.	(economic OR cost OR cost-effectiveness OR cost-analysis)
5.	fha[Filter]
6.	#4 AND #5
7.	#3 AND #6
8.	#7 NOT cadaver
9.	#8 Filters: Abstract, from 2010 - 2024

Appendix Table B3: PubMed Search Strategy for Safety Outcomes

Search period: January 1, 2010 – January 3, 2024

1.	(vertebroplast* OR kyphoplast* OR sacroplast*)
2.	fha[Filter]
3.	(safety OR complication OR complications OR adverse)
4.	fha[Filter]
5.	#1 AND #2 AND #3 AND #4

6.	(case reports[Publication Type] OR review[Publication Type] OR editorial[Publication Type] OR comment[Publication Type])
7.	#6 AND fha[Filter]
8.	#5 NOT #7
9.	#8 NOT cadaver*
10.	#11 Filters: Abstract, from 2010 – 2024

Appendix Table B4: PubMed Search Strategy for Cement Leakage

Search period: January 1, 2010 – January 3, 2024

1.	(vertebroplast* OR kyphoplast* OR sacroplast*)
2.	fha[Filter]
3.	#1 AND #2
4.	(cement leakage)
5.	Fha[Filter]
6.	(case reports[Publication Type] OR review[Publication Type] OR editorial[Publication Type] OR comment[Publication Type])
7.	#6 AND fha[Filter]
8.	#3 AND #4 AND #5 NOT #7
9.	#8 Filters: Abstract, from 2010 – 2024

Appendix Table B5: PubMed Search Strategy for Embolism

Search period: January 1, 2010 – January 3, 2024

1.	(vertebroplast* OR kyphoplast* OR sacroplast*)
2.	fha[Filter]
3.	#1 AND #2
4.	(embolism)
5.	fha[Filter]
6.	(case reports[Publication Type] OR review[Publication Type] OR editorial[Publication Type] OR comment[Publication Type])
7.	#6 AND fha[Filter]
8.	#3 AND #4 AND #5 NOT #7
9.	#8 Filters: Abstract, from 2010 – 2024

Appendix Table B6: PubMed Search Strategy for New Fracture

Search period: January 1, 2010 – January 3, 2024

1.	(vertebroplast* OR kyphoplast* OR sacroplast*)
2.	fha[Filter]

3.	#1 AND #2
4.	(adjacent fracture OR new fracture OR subsequent fracture)
5.	fha[Filter]
6.	(case reports[Publication Type] OR review[Publication Type] OR editorial[Publication Type] OR comment[Publication Type])
7.	#6 AND fha[Filter]
8.	#3 AND #4 AND #5 NOT #7
9.	#8 Filters: Abstract, from 2010 – 2024

Electronic Database Searches

The following databases have been searched for relevant information:

Cochrane Database of Systematic Reviews

Cochrane Registry of Clinical Trials (CENTRAL)

Database of Reviews of Effectiveness (Cochrane Library)

PubMed

ClinicalTrials.gov

Additional Economics, Clinical Guideline and Gray Literature Databases

AHRQ - Healthcare Cost and Utilization Project

Canadian Agency for Drugs and Technologies in Health

Centers for Medicare and Medicaid Services (CMS)

Food and Drug Administration (FDA)

Google

APPENDIX C. Excluded Articles

Articles excluded as primary studies after full text review, with reason for exclusion.

Appendix Table C1. List of Excluded Articles

Citation	Reason for exclusion after full-text review
Aregger FC, Gerber F, Albers C, Oswald K, Knoll C, Benneker L, Heini P, Berlemann U, Hoppe S. Long-term follow-up after vertebroplasty - A mean 10-years follow-up control study. Brain Spine. 2024 Apr 3;4:102783. doi: 10.1016/j.bas.2024.102783. PMID: 38618227; PMCID: PMC11015514.	Ineligible study design
Chabert E, Hugonnet E, Kastler A, Sakka L, Rabbo FA, Zerroug A, et al. Vertebroplasty versus bracing in acute vertebral compression fractures: A prospective randomized trial. Ann Phys Rehabil Med. 2023;66(6):101746.	Ineligible population
Dang SJ, Wei WB, Wei L, Xu J. Vertebroplasty combined with facet joint block vs. vertebroplasty alone in relieving acute pain of osteoporotic vertebral compression fracture: a randomized controlled clinical trial. BMC Musculoskelet Disord. 2022;23(1):807.	Ineligible intervention
D'Oria S, Dibenedetto M, Squillante E, Somma C, Hannan CJ, Giraldi D, et al. Traumatic compression fractures in thoracic-lumbar junction: vertebroplasty vs conservative management in a prospective controlled trial. J Neurointerv Surg. 2022;14(2):202-6.	Ineligible population
Firanescu C, Lohle PN, de Vries J, Klazen CA, Juttmann JR, Clark W, et al. A randomised sham controlled trial of vertebroplasty for painful acute osteoporotic vertebral fractures (VERTOS IV). Trials. 2011;12:93.	Protocol
Gilula L, Persenaire M. Subsequent fractures post-vertebral augmentation: analysis of a prospective randomized trial in osteoporotic vertebral compression fractures. AJNR Am J Neuroradiol. 2013;34(1):221-7.	Ineligible comparator
Klazen CA, Venmans A, de Vries J, van Rooij WJ, Jansen FH, Blonk MC, et al. Percutaneous vertebroplasty is not a risk factor for new osteoporotic compression fractures: results from VERTOS II. AJNR Am J Neuroradiol. 2010;31(8):1447-50.	Included in another publication
Klazen CA, Verhaar HJ, Lohle PN, Lampmann LE, Juttmann JR, Schoemaker MC, et al. Clinical course of pain in acute osteoporotic vertebral compression fractures. J Vasc Interv Radiol. 2010;21(9):1405-9.	Included in another publication
Lin JH, Chien LN, Tsai WL, Chen LY, Chiang YH, Hsieh YC. Early vertebroplasty associated with a lower risk of mortality and respiratory failure in aged patients with painful vertebral compression fractures: a population-based cohort study in Taiwan. Spine J 2017;17:1310-8.	Ineligible Comparator
Longo UG, Loppini M, Denaro L, Brandi ML, Maffulli N, Denaro V. The effectiveness and safety of vertebroplasty for osteoporotic vertebral compression fractures. A double blind, prospective, randomized, controlled study. Clin Cases Miner Bone Metab. 2010;7(2):109-13.	Protocol
Nakano M, Kawaguchi Y, Kimura T, Hirano N. Transpedicular vertebroplasty after intravertebral cavity formation versus conservative treatment for osteoporotic burst fractures. Spine J. 2014;14(1):39-48.	Ineligible study design
Noriega DC, Ramajo RH, Lite IS, Toribio B, Corredera R, Ardura F, et al. Safety and clinical performance of kyphoplasty and SpineJack(®) procedures in the	Ineligible comparator

Citation	Reason for exclusion after full-text review
treatment of osteoporotic vertebral compression fractures: a pilot, monocentric, investigator-initiated study. Osteoporos Int. 2016;27(6):2047-55.	
Noriega D, Marcia S, Theumann N, Blondel B, Simon A, Hassel F, et al. A prospective, international, randomized, noninferiority study comparing an implantable titanium vertebral augmentation device versus balloon kyphoplasty in the reduction of vertebral compression fractures (SAKOS study). Spine J. 2019;19(11):1782-95.	Ineligible comparator
Noriega DC, Rodríguez-Monsalve F, Ramajo R, Sánchez-Lite I, Toribio B, Ardura F. Long-term safety and clinical performance of kyphoplasty and SpineJack® procedures in the treatment of osteoporotic vertebral compression fractures: a pilot, monocentric, investigator-initiated study. Osteoporos Int. 2019;30(3):637-45.	Ineligible comparator
Noriega DC, Rodríguez-Monsalve F, Ramajo R, Sánchez-Lite I, Toribio B, Ardura F. Correction to: Long-term safety and clinical performance of kyphoplasty and SpineJack® procedures in the treatment of osteoporotic vertebral compression fractures: a pilot, monocentric, investigator-initiated study. Osteoporos Int. 2019;30(3):647.	Ineligible comparator
Otten LA, Bornemnn R, Jansen TR, Kabir K, Pennekamp PH, Wirtz DC, et al. Comparison of balloon kyphoplasty with the new Kiva® VCF system for the treatment of vertebral compression fractures. Pain Physician. 2013;16(5):E505-12.	Ineligible intervention
van Berkel D, Ong T, Drummond A, Hendrick P, Leighton P, Jones M, et al. ASSERT (Acute Sacral inSufficiEncy fractuRe augmenTation) randomised controlled, feasibility in older people trial: a study protocol. BMJ Open. 2019;9(7):e032111.	Protocol
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(2):47-51.	Ineligible intervention
Venmans A, Klazen CA, Lohle PN, van Rooij WJ, Verhaar HJ, de Vries J, et al. Percutaneous vertebroplasty and pulmonary cement embolism: results from VERTOS II. AJNR Am J Neuroradiol. 2010;31(8):1451-3.	Included in another publication
Wang D, Cang D, Wu Y, Wang S. Therapeutic effect of percutaneous vertebroplasty and nonoperative treatment on osteoporotic vertebral compression fracture: A randomized controlled trial protocol. Medicine (Baltimore). 2020;99(27):e20770.	Protocol
Werner CM, Osterhoff G, Schlickeiser J, Jenni R, Wanner GA, Ossendorf C, Simmen HP. Vertebral body stenting versus kyphoplasty for the treatment of osteoporotic vertebral compression fractures: a randomized trial. J Bone Joint Surg Am. 2013 Apr 3;95(7):577-84. doi: 10.2106/JBJS.L.00024. PMID: 23553291.	Ineligible comparator
Wickstroem LA, Carreon L, Lund T, Abildgaard N, Lorenzen MD, Andersen M. Vertebroplasty in patients with multiple myeloma with vertebral compression fractures: protocol for a single-blind randomised controlled trial. BMJ Open. 2021;11(9):e045854.	Protocol
Xu JC, Wu GH, Zhou LL, Yang XJ, Liu JT. Two unilateral puncturation comparative analyses of multiple-level fresh osteoporotic vertebral body compression fractures treated with percutaneous vertebroplasty guided by C-arm fluoroscopy or in senile patients. Eur Rev Med Pharmacol Sci. 2017;21(7):1456-61.	Ineligible comparator
Xu JJ, Tang XT, Yang J, Wang YH, Zhu DC, Wu YS, et al. The Effect of Medial Branch Block on Postoperative Residual Pain Relieve After Percutaneous	Ineligible intervention

Citation	Reason for exclusion after full-text review
Kyphoplasty: A Randomized Controlled Trial With 12-Month Follow-up. Pain Physician. 2021;24(7):E1059-e66.	
Yavuz AY, Aydın MV. Long-term Clinical and Radiological Results of Vertebral Augmentation Techniques in Osteoporotic Lumbar Compression Fractures: Vertebroplasty or Kyphoplasty?. J Turk Spinal Surg. 2023 Oct;34(4):180-188. doi:10.4274/jtss.galenos.2023.30502.	Ineligible study design
Yokoyama K, Kawanishi M, Yamada M, Tanaka H, Ito Y, Hirano M, et al. Comparative study of percutaneous vertebral body perforation and vertebroplasty for the treatment of painful vertebral compression fractures. AJNR Am J Neuroradiol. 2012;33(4):685-9.	Ineligible comparator
Anselmetti GC, Marcia S, Saba L, Muto M, Bonaldi G, Carpeggiani P, et al. Percutaneous vertebroplasty: multi-centric results from EVEREST experience in large cohort of patients. Eur J Radiol. 2012;81(12):4083-6.	Ineligible Population
Anselmetti GC, Marcia S, Saba L, Muto M, Bonaldi G, Carpeggiani P, et al. Percutaneous vertebroplasty: multi-centric results from EVEREST experience in large cohort of patients. Eur J Radiol. 2012;81(12):4083-6.	Ineligible study design
Bornemann R, Jansen TR, Kabir K, Pennekamp PH, Stüwe B, Wirtz DC, et al. Comparison of Radiofrequency-targeted Vertebral Augmentation With Balloon Kyphoplasty for the Treatment of Vertebral Compression Fractures: 2-Year Results. Clin Spine Surg. 2017;30(3):E247-e51.	Ineligible comparator
Bozkurt M, Kahilogullari G, Ozdemir M, Ozgural O, Attar A, Caglar S, et al. Comparative analysis of vertebroplasty and kyphoplasty for osteoporotic vertebral compression fractures. Asian Spine J. 2014;8(1):27-34.	SA not focused on harms
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(19):1729-36.	More complete study of database included
Chen F, Xia YH, Cao WZ, Shan W, Gao Y, Feng BO, et al. Percutaneous kyphoplasty for the treatment of spinal metastases. Oncol Lett. 2016;11(3):1799-806.	SA not focused on safety
Cheng Y, Cheng X, Wu H. Risk factors of new vertebral compression fracture after percutaneous vertebroplasty or percutaneous kyphoplasty. Front Endocrinol (Lausanne). 2022;13:964578.	Study design
Clarençon F, Fahed R, Gabrieli J, Guermazi Y, Cormier E, Molet-Benhamou L, et al. Safety and Clinical Effectiveness of Percutaneous Vertebroplasty in the Elderly (≥80 years). Eur Radiol. 2016;26(7):2352-8.	Ineligible Population
Corcos G, Dbjay J, Mastier C, Leon S, Auperin A, De Baere T, et al. Cement leakage in percutaneous vertebroplasty for spinal metastases: a retrospective evaluation of incidence and risk factors. Spine (Phila Pa 1976). 2014;39(5):E332-8.	Ineligible Population
Crouser N, Malik AT, Jain N, Yu E, Kim J, Khan SN. Discharge to Inpatient Care Facility After Vertebroplasty/Kyphoplasty: Incidence, Risk Factors, and Postdischarge Outcomes. World Neurosurg. 2018;118:e483-e8.	More complete study of database included
Delpla A, Tselikas L, De Baere T, Laurent S, Mezaib K, Barat M, et al. Preventive Vertebroplasty for Long-Term Consolidation of Vertebral Metastases. Cardiovasc Intervent Radiol. 2019;42(12):1726-37.	Ineligible Population

Citation	Reason for exclusion after full-text review
Denaro L, Longo UG, Papalia R, De Salvatore S, Ruzzini L, Piergentili I, et al. The burden of percutaneous vertebroplasty: an epidemiological nationwide study in Italy from 2009 to 2015. Eur Spine J. 2021;30(10):3099-106.	Ineligible outcomes
Diel P, Reuss W, Aghayev E, Moulin P, Röder C. SWISSspine-a nationwide health technology assessment registry for balloon kyphoplasty: methodology and first results. Spine J. 2010;10(11):961-71.	SA not focused on safety
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(7):1617-26.	More complete study of database included
Ee GW, Lei J, Guo CM, Yeo W, Tan SB, Tow PB, et al. Comparison of Clinical Outcomes and Radiographic Measurements in 4 Different Treatment Modalities for Osteoporotic Compression Fractures: Retrospective Analysis. J Spinal Disord Tech. 2015;28(6):E328-35.	SA not focused on harms
Fan W, Qiao T, You Y, Zhang J, Gao J. Perioperative prevalence of deep vein thrombosis in patients with percutaneous kyphoplasty: A retrospective study with routine ultrasonography. Medicine (Baltimore). 2020;99(10):e19402.	Ineligible population
Galivanche AR, Toombs C, Adrados M, David WB, Malpani R, Saifi C, et al. Cement Augmentation of Vertebral Compression Fractures May Be Safely Considered in the Very Elderly. Neurospine. 2021;18(1):226-33.	More complete study of database included
He B, Zhao J, Zhang M, Jiang G, Tang K, Quan Z. Effect of Surgical Timing on the Refracture Rate after Percutaneous Vertebroplasty: A Retrospective Analysis of at Least 4-Year Follow-Up. Biomed Res Int. 2021;2021:5503022.	Ineligible design
Hoshino M, Takahashi S, Yasuda H, Terai H, Watanabe K, Hayashi K, et al. Balloon Kyphoplasty Versus Conservative Treatment for Acute Osteoporotic Vertebral Fractures With Poor Prognostic Factors: Propensity Score Matched Analysis Using Data From Two Prospective Multicenter Studies. Spine (Phila Pa 1976). 2019;44(2):110-7.	Ineligible design
Jarrar S, Al Barbarawi MM, S SD, Jaradat A, Alkalbani R, Abu Qayyas L, et al. Cement extravasation as a complication for kyphoplasty and vertebroplasty procedure: a retrospective analysis of 171 cases. Med Glas (Zenica). 2024;21(1).	Ineligible Design
Kasperk C, Haas A, Hillengass J, Weiss C, Neben K, Goldschmidt H, et al. Kyphoplasty in patients with multiple myeloma a retrospective comparative pilot study. J Surg Oncol. 2012;105(7):679-86.	Ineligible Population
Kessler RA, De la Garza Ramos R, Purvis TE, Ahmed AK, Goodwin CR, Sciubba DM, et al. Impact of frailty on complications in patients with thoracic and thoracolumbar spinal fracture. Clin Neurol Neurosurg. 2018;169:161-5.	More complete study of database included
Lee HM, Park SY, Lee SH, Suh SW, Hong JY. Comparative analysis of clinical outcomes in patients with osteoporotic vertebral compression fractures (OVCFs): conservative treatment versus balloon kyphoplasty. Spine J. 2012;12(11):998-1005.	Ineligible design
Lee JK, Jeong HW, Joo IH, Ko YI, Kang CN. Percutaneous balloon kyphoplasty for the treatment of very severe osteoporotic vertebral compression fractures: a case-control study. Spine J. 2018;18(6):962-9.	Ineligible comparator
Lotan R, Smorgick Y, Anekstein Y, Rudik O, Prosso I, Hershkovich O. Kyphoplasty for Elderly Patients With Vertebral Compression Fractures-Do We Save Lives? Mortality Rates Analysis Comparison in a Long-Term Follow-Up Cohort. Global Spine J. 2022;12(7):1443-8.	Ineligible Population/Design

Citation	Reason for exclusion after full-text review
Luetmer MT, Bartholmai BJ, Rad AE, Kallmes DF. Asymptomatic and unrecognized cement pulmonary embolism commonly occurs with vertebroplasty. AJNR Am J Neuroradiol. 2011;32(4):654-7.	Ineligible Population/Design
Nakamae T, Fujimoto Y, Yamada K, Hashimoto T, Olmarker K. Efficacy of Percutaneous Vertebroplasty in the Treatment of Osteoporotic Vertebral Compression Fractures with Intravertebral Cleft. Open Orthop J. 2015;9:107-13.	SA not focused on harms
Pereira LP, Clarençon F, Cormier E, Rose M, Jean B, Le Jean L, et al. Safety and effectiveness of percutaneous sacroplasty: a single-centre experience in 58 consecutive patients with tumours or osteoporotic insufficient fractures treated under fluoroscopic guidance. Eur Radiol. 2013;23(10):2764-72.	Ineligible Population
Pflugmacher R, Bornemann R, Koch EM, Hausmann D, Otten LA, Goost H, et al. [Comparative findings of balloon kyphoplasty in patients with vertebral fractures due to osteoporosis, metastases and myeloma]. Z Orthop Unfall. 2012;150(2):198-204.	No English full text available
Saad A, Botchu R, James S. The Rates of Cement Leakage Following Vertebroplasty in Osteoporotic versus Metastatic Disease. Indian J Radiol Imaging. 2022;32(1):46-50.	Ineligible Population/Design
Saliou G, Kocheida el M, Lehmann P, Depriester C, Paradot G, Le Gars D, et al. Percutaneous vertebroplasty for pain management in malignant fractures of the spine with epidural involvement. Radiology. 2010;254(3):882-90.	Ineligible Population
Shi X, Cui Y, Pan Y, Wang B, Lei M. Epidemiology and detection of cement leakage in patients with spine metastases treated with percutaneous vertebroplasty: A 10-year observational study. J Bone Oncol. 2021;28:100365.	Ineligible Population
Sun G, Li L, Jin P, Liu XW, Li M. Percutaneous vertebroplasty for painful spinal metastasis with epidural encroachment. J Surg Oncol. 2014;110(2):123-8.	Ineligible Population
Sun S, Xu B, Zhang Q, Zhao CS, Ma R, He J, et al. The Early Results of Vertebral Pathological Compression Fracture of Extra- nodal Lymphoma with HIV-positive Patients Treated by Percutaneous Kyphoplasty. Curr HIV Res. 2020;18(4):248-57.	No full text available
Telera S, Gorgoglione N, Raus L, Vidiri A, Villani V, Pace A, et al. Open Kyphoplasty for Metastatic Spine Disease: A Retrospective Clinical Series. World Neurosurg. 2019;127:e751-e60.	Ineligible Population
Tian QH, Liu HF, Wang T, Wu CG, Cheng YS. Fluoroscopy-Guided Percutaneous Sacroplasty for Painful Metastases at the Sacral Ala. J Pain Res. 2020;13:151-6.	Ineligible Population
Toy JO, Basques BA, Grauer JN. Morbidity, mortality, and readmission after vertebral augmentation: analysis of 850 patients from the American College of Surgeons National Surgical Quality Improvement Program database. Spine (Phila Pa 1976). 2014;39(23):1943-9.	More complete study of database included
Tsai YW, Hsiao FY, Wen YW, Kao YH, Chang LC, Huang WF, et al. Clinical outcomes of vertebroplasty or kyphoplasty for patients with vertebral compression fractures: a nationwide cohort study. J Am Med Dir Assoc. 2013;14(1):41-7.	More complete study of database included
Wang C, Zhang X, Liu J, Shan Z, Li S, Zhao F. Percutaneous kyphoplasty: Risk Factors for Recollapse of Cemented Vertebrae. World Neurosurg. 2019;130:e307-e15.	Ineligible design

Citation	Reason for exclusion after full-text review
Wang L, Zhang C, Liang H, Huang T, Zhong W, Zhao Z, et al. Cement leakage in percutaneous vertebroplasty for spinal metastases: a retrospective study of risk factors and clinical outcomes. World J Surg Oncol. 2022;20(1):112.	Ineligible Population
Wu W, Zhang X, Li X, Liu H, Yu S. Clinical Evaluation of Percutaneous Kyphoplasty for the Management of Osteoblastic-Related Metastatic Vertebral Lesions. Acad Radiol. 2022;29 Suppl 3:S183-s7.	More complete study of database included
Wu W, Zhang X, Li X, Liu H, Xu L, Liu T, 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(9):938-41.	More complete study of database included
Yang DH, Cho KH, Chung YS, Kim YR. Effect of vertebroplasty with bone filler device and comparison with balloon kyphoplasty. Eur Spine J. 2014;23(12):2718-25.	Ineligible comparator
Yang JS, Liu JJ, Chu L, Li J, Chen C, Chen H, et al. Causes of Residual Back Pain at Early Stage After Percutaneous Vertebroplasty: A Retrospective Analysis of 1,316 Cases. Pain Physician. 2019;22(5):E495-e503.	Ineligible Outcomes
Yokoyama K, Kawanishi M, Yamada M, Tanaka H, Ito Y, Hirano M, et al. Safety and therapeutic efficacy of the second treatment for new fractures developed after initial vertebroplasty performed for painful vertebral compression fractures. Neurol Res. 2013;35(6):608-13.	Ineligible design
Young C, Munk PL, Heran MK, Lane MD, Le HB, Lee S, et al. Treatment of severe vertebral body compression fractures with percutaneous vertebroplasty. Skeletal Radiol. 2011;40(12):1531-6.	Ineligible Outcomes
Zampini JM, White AP, McGuire KJ. Comparison of 5766 vertebral compression fractures treated with or without kyphoplasty. Clin Orthop Relat Res. 2010;468(7):1773-80.	Ineligible Outcomes
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(1):792.	No English full text available
Zou J, Mei X, Gan M, Yang H. Kyphoplasty for spinal fractures from multiple myeloma. J Surg Oncol. 2010;102(1):43-7.	Accounted for in included systematic review
Chen C, Li DW, Wang Q, Xu XW, Ma YZ, Li Z, et al. [The cost effectiveness analysis of minimally invasive surgery and conservative treatment in elderly osteoporotic spinal fracture]. Zhongguo Gu Shang. 2016;29(7):614-8.	No English full text available
Chew C, O'Dwyer PJ, Edwards R. Health service cost associated with percutaneous vertebroplasty in patients with spinal metastases. Clin Radiol. 2013;68(8):776-9.	Ineligible study design
Itagaki MW, Talenfeld AD, Kwan SW, Brunner JW, Mortell KE, Brunner MC. Percutaneous vertebroplasty and kyphoplasty for pathologic vertebral fractures in the Medicare population: safer and less expensive than open surgery. J Vasc Interv Radiol. 2012;23(11):1423-9.	Ineligible study design
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(4):318-26.	Ineligible outcomes

APPENDIX D. Risk of Bias, Strength of Evidence, QHES, and AMSTAR-2

Each included comparative study is rated against pre-set criteria that resulted in a Risk of Bias (ROB) assessment and presented in a table. Assessment of RCTs followed appropriate criteria based on methods described in *the Cochrane Handbook for Systematic Reviews of Interventions*[5] and guidance from the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[11] In keeping with the AHRQ methods, each study was given a final rating of "good", "fair", or "poor" quality as described below in Table D1. Discrepancies in ratings between reviewers were resolved through discussion and consensus. Where blinding is not possible, studies will automatically be rated as "fair" given the potential for biased assessment of outcomes. The final quality assessments are provided in Appendix E.

Table D2 provides an example of the format used to assess ROB for comparative studies of testing/therapy. Additional criteria for non-randomized studies includes consideration of how patients are selected and appropriate control for confounding. Table D3 provides an example for non-randomized studies of interventions. Table D4 provides an example for evaluating administrative database studies. A "No" indicates that the criterion was not met; an "Unclear" indicates that the criterion could not be determined with the information provided or was not reported by the author. Risk of bias assessments were not conducted for case series; all were considered High risk of bias.

Appendix Table D1. Definition of the risk of bias categories for individual studies of testing

Rating	Description and Criteria
Good	• Least risk of bias; study results generally considered valid • Employ valid methods for selection, inclusion, and allocation of patients to testing; report similar baseline characteristics in different test groups; clearly describe attrition and have low attrition; use appropriate means for preventing bias (e.g., blinding of patients, care providers, and outcomes assessors); and use appropriate analytic methods (e.g., intention-to-treat analysis)
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 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

Appendix Table D2: Assessment of ROB for Individual Randomized Control Trials

Methodological Principle	Author 1, 2023	Author 2 2024	Author 3, 2021
Study design			
Randomized controlled trial	■	■	■
Random sequence generation			
Concealed allocation			
Groups comparable at baseline*			
Outcome assessors independent or blinded			
Care providers blinded			
Patients blinded			
Reporting of attrition			
Complete follow-up of ≥80%			
<10% difference in follow-up between groups			
Intention to treat			
Outcomes prespecified			
Risk of Bias			

Unclear indicates that the study had insufficient detail to determine whether criteria were met
*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table D3: Assessment of ROB for Individual Non-Randomized Studies of Interventions

Methodological Principle	Author 1, 2024	Author 2, 2019	Author 3, 2020
Did the study attempt to enroll a random sample or consecutive patients meeting inclusion criteria (inception cohort) from same underlying population?			
Were the groups comparable at baseline on key prognostic factors?			
Did the article report attrition?			
Overall loss to follow up acceptable? (≤20%) Differential loss to follow up acceptable? (≤10%)			
Were the outcomes investigated prespecified and defined?			
Did the study clearly describe and use accurate methods for ascertaining outcomes, exposures, and potential confounders?			
Were outcome assessors and/or data analysts blinded to treatment?			
Did the study perform appropriate statistical analyses on potential confounders or otherwise control for confounding (e.g. restriction, stratification, matching)?			
Was the duration of follow-up reasonable for investigated events?			
Quality (Risk of Bias)			

NA = not applicable (due to being a case series)
Unclear indicates that the study had insufficient detail to determine whether criteria were met

Appendix Table D4: Assessment of Quality of Administrative Database Studies

Methodological Principle (Interventions)	Author 1, 2023	Author 2, 2020	Author 3, 2021
Study design Administrative database comparative study Administrative database case-control study Administrative database case series	x	x	x
Why database created clearly stated			
Description of database's inclusion/exclusion criteria			
Description of methods for reducing bias in database			
Codes and search algorithms reported			
Rationale for coding algorithm reported			
Code accuracy reported			
Code validity reported			
Clinical significance assessed			
Is the period of data consistent with the outcome data?			
Statement regarding whether data stems from single or multiple hospital admissions			
Statement regarding whether data stems from single or multiple procedures			
Accounting for clustering			
Number of criteria met (maximum: 12)			

Assessment of Economic Studies

Full formal economic analyses evaluate both costs and clinical outcomes of two or more alternative interventions. The four primary types are cost minimization analysis (CMA), cost-utility analysis (CUA), cost-effectiveness analysis (CEA), and cost-benefit analyses (CBA). Each employs different methodologies, potentially complicating critical appraisal, but some common criteria can be assessed across studies.

No standard, universally accepted method of critical appraisal of economic analyses is currently in use. A number of checklists [Canadian, BMJ, AMA] are available to facilitate critique of such studies. The Quality of Health Economic Studies (QHES) instrument developed by Ofman, et al.[7] embodies the primary components relevant for critical appraisal of economic studies. It also incorporates a weighted scoring process which was used as one factor to assess included economic studies. This tool has not yet undergone extensive evaluation for broader use but provides a valuable starting point for critique. Table D4 below provides a template of the instrument.

In addition to assessment of criteria in the QHES, other factors are important in critical appraisal of studies from an epidemiologic perspective to assist in evaluation of generalizability and potential sources of study bias.

Such factors include:

- Are the interventions applied to similar populations (e.g., with respect to age, gender, medical conditions, etc.)? To what extent are the populations for each intervention comparable and are

differences considered or accounted for? To what extent are population characteristics consistent with "real world" applications of the comparators?

- Are the sample sizes adequate so as to provide a reasonable representation of individuals to whom the technology would be applied?
- What types of studies form the basis for the data used in the analyses? Data (e.g., complication rates) from randomized controlled trials or well-conducted, methodologically rigorous cohort studies for data collection are generally of highest quality compared with case series or studies with historical cohorts.
- Were the interventions applied in a comparable manner (e.g., similar protocols, follow-up procedures, evaluation of outcomes, etc.)?
- How were the data and/or patients selected or sampled (e.g., a random selection of claims for the intervention from a given year/source or all claims)? What specific inclusion/exclusion criteria or processes were used?

Were the outcomes and consequences of the interventions being compared comparable for each? (e.g., were all of the relevant consequences/complications for each intervention considered or do they primarily reflect those for one intervention?

Appendix Table D5. Assessment of Quality of Health Economic Studies Criteria

Question	Possible Points*	Criteria For Credit*
1. Was the study **objective** presented in a clear, specific, and measurable manner?	7	Authors must fully describe the objective; is it measurable?
2. Were the **perspective** of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	Authors must state perspective, provide rationale AND have done the correct analysis corresponding to the perspective
3. Were **variable estimates** used in the analysis from the best available source (i.e., randomized controlled trial - best, expert opinion - worst)?	8	No credit if most of estimates are not from the best sources available
4. If estimates came from a **subgroup analysis**, were the groups prespecified at the beginning of the study?	1	-
5. Was **uncertainty** handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	NO credit if they do not give details regarding type of sensitivity analysis, methods (e.g. what assumptions or factors were varied/why), AND the results (what factors are influential, what is the range of ICERs, etc.)
6. Was **incremental analysis** performed between alternatives for resources and costs?	6	-
7. Was the methodology for **data abstraction** (including the value of health states and other benefits) stated?	5	No credit if sources of model inputs and process of choosing model inputs not specified
8. Did the **analytic horizon allow time** for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	No credit if time horizon is too short to allow for important outcomes
9. Was the **measurement of costs** appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	No credit if sources of cost data or methods of estimating costs not clearly described
10. Were the primary **outcome measure(s)** for the economic evaluation clearly stated and did they	6	NO credit if major important outcomes are not included or if time horizon did not allow for important outcomes to be measured

Question	Possible Points*	Criteria For Credit*
include the major short-term, long-term and negative outcomes included?		
11. Were the health outcomes **measures/scales valid** and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	No credit if sources of outcome data or not clearly described or if outcome data is not appropriate for the study population/outcome of interest (i.e. using utility weights from QOL measures that aren't validated or apply to a different population)
12. Were the **economic model** (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	Must provide explicit detail for methods and should be able to trace/identify specific components, how they were derived, etc.
13. Were the choice of economic model, main **assumptions, and limitations** of the study stated and justified?	7	NO credit if insufficient detail of model, assumptions AND limitations are provided (No credit if they do not provide justifications/rationale)
14. Did the author(s) explicitly discuss direction and magnitude of potential **biases?**	6	NO credit if no discussion of direction and magnitude of biases
15. Were the **conclusions/recommendations** of the study justified and based on the study results?	8	NO credit if conclusions/recommendations are stronger than warranted based on findings
16. Was there a statement disclosing the **source of funding** for the study?	3	-
Total	100	

ICER = Incremental Cost-Effectiveness Ratio; QOL = quality of life.
* Study must fit criteria in order to receive full points. Partial credit is not given. If criteria is not met, then the question receives no points.

Application of AMSTAR 2 to systematic reviews

Table D6 shows our criteria for RoB assessment based on the AMSTAR-2 tool. AMSTAR-2 is the revised and updated version of AMSTAR[13] published in 2007 used for critical appraisal of systematic reviews (Shea, 2017). It is not intended to provide an overall score, as high scores may hide weaknesses in critical domains. In light of this, we used a modified AMSTAR tool as determined by Dettori et al (2020).[4] Table D7 (adapted from Dettori 2020)[3] describes how overall scores were determined considering critical domains. Bold items in table 1 were considered as critical items. The original AMSTAR-2 guidance suggests grading each item as no or yes, with items 2, 4, 7, 8, and 9 allowing for a 'partial yes'. We considered a 'yes' or 'partial yes' as yes.

Appendix Table D6. Criteria for assessing systematic reviews based on AMSTAR-2.

Item	Criteria
1: Did the research questions and inclusion criteria for the review include the components of PICO?	• **Yes** if all components of PICO are described somewhere in the report. • **No** if any components of PICO are missing.
2: Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?	• **Yes** if the protocol or review methods were established prior to review. • **No** if no protocol or discussion/description of methods decided prior to review.
3: Did the review authors explain their selection of the study designs for inclusion in the review?	• **Yes** if study design inclusion is justified or discussed. No penalty for restricting study designs. • **No** if no discussion of justification for inclusion.

4: Did the review authors use a comprehensive literature search strategy?	• **Yes** if 2 or more electronic databases were searched and key words are available in report or appendices. No penalty for language restrictions. • **No** if less than 2 electronic databases were searched or key words are unavailable.
5: Did the review authors perform study selection in duplicate?	• **Yes** if selection at title/abstract and full text reviews were performed by 2 authors with consensus upon disagreement or single author selecting with a second checking agreement on sample and a kappa reported of ≥0.80. • **No** if no second author involved or no kappa reported.
6: Did the review authors perform data extraction in duplicate?	• **Yes** if abstraction was performed by 2 authors with consensus upon disagreement or single author abstracting with a second checking agreement on sample and a kappa of reported of ≥0.80. • **No** if no second author involved or no kappa reported.
7: Did the review authors provide a list of excluded studies and justify the exclusions?	• **Yes** if a list of potentially relevant studies is reported in appendix or discussed in text with citations with justification for exclusion. List of references must be provided. • **No** if no list of references provided or not potentially relevant but excluded studies are discussed.
8: Did the review authors describe the included studies in adequate detail?	• **Yes** if study characteristics are reported in sufficient detail to determine whether the studies met PICO criteria and provides framework to judge heterogeneity. • **No** if study characteristics are not reported or table 1 does not include age, sex, (and #'s).
9: Did the review authors use a satisfying technique for assessing the RoB in individual studies that were included in the review?	**RCTS** • **Yes** if important domains similar to Cochrane. **Cohort studies** • **Yes** if it addresses all of the following: confounding, selection bias, measurement bias, and selective reporting of outcomes (Newcastle okay if all 8 questions included). **Case series** (study of incidence, no direct comparison) • **Yes** if selection bias, measurement bias, and selective reporting of outcomes met (Newcastle okay IF questions #1, 2, 3, 4, 6, 7, and 8 addressed). **For all studies** • **No** if there is obvious evidence that the authors misapplied an acceptable technique.
10: Did the review authors report on the sources of funding for the studies included in the review?	• **Yes** if authors report funding of individual studies. • **No** if authors do not report funding.

11: If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?	• **Yes** if all the following are present
	  ○ Meta-analysis justified (e.g., studies comparable, direct comparison).
	  ○ Explanation of fixed or random effects (must do more than merely report without explanation).
	  ○ Pooled results reported separately for RCTs and cohort studies.
	  ○ Assessment of heterogeneity (must address I^2).
	• **No** if one or more of the above are not present.

	• If no meta-analysis was done mark as **NM (No meta-analysis)**
12: If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?	• **Yes** if results are stratified by RoB or if the review only included the lowest RoB studies in the analysis. • **No** if results are not stratified by RoB and review includes a range of RoB outcomes in the analysis. No credit if RoB method from item #9 is not acceptable. • If no meta-analysis was done mark as **NM (No meta-analysis**)
13: Did the review authors account for RoB in individual studies when interpreting or discussing the results of the review?	• **Yes** if there is a discussion of the impact of RoB in the interpretation of results and/or accounting for differences between studies. • **No** if there is no discussion of the impact of RoB in the interpretation of results and/or accounting for differences between studies. No credit if method from #9 is not acceptable.
14: Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	• **Yes** if I^2 demonstrates no heterogeneity (<50%) or authors explored reasons for heterogeneity if I^2 is ≥50%. • **No** if I^2 demonstrates heterogeneity (>50%) and authors do not explore reasons for heterogeneity.
15: If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?	• **Yes** if there is an attempt to identify publication bias. Must also show awareness of likely impact of publication bias on results. Credit given if they acknowledge publication bias could be a problem but not enough data given or if they have fewer than 10 studies and show no evidence of publication bias. • **No** if there is no attempt to identify or discuss publication bias. • If no meta-analysis was done mark **as NM (No meta-analysis)**
16: Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	• **Yes** if authors report no competing interests or how they managed potential conflicts of interest. • **No** if there is no discussion or reporting of potential conflicts of interest.

PICO = population, intervention, comparison, outcome; RoB = risk of bias.

Appendix Table D7. Rating overall Confidence in the Results of the Review (Dettori 2020).

High**:** No or 1 noncritical weakness	The systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest.
Moderate**:** More than 1 noncritical weakness*	The systematic review has more than 1 weakness but no critical flaws. It may provide an accurate summary of the results of the available studies that were included in the review.
Low**:** One critical flaw with or without noncritical weaknesses	The review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest.
Critically low**:** More than 1 critical flaw with or without noncritical weaknesses	The review has more than 1 critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies.

* Multiple noncritical weaknesses may diminish confidence in the review, and it may be appropriate to move the overall appraisal down from moderate to low confidence.

Determination of Overall Strength (Quality) of Evidence

Following the assessment of the quality of each individual study included in the report, an *overall* "strength of evidence"/"quality of evidence" for all critical and important *primary* health outcomes and harms based on methods used by GRADE (Grading of Recommendation Assessment, Development and Evaluation) and the Agency for Healthcare Research and Quality (AHRQ)[11] will be reported.

The overall strength of evidence is based on assessment of the following required domains: risk of bias, consistency, directness, and precision. The overall Strength of Evidence (SoE) ranges from high for a body of evidence if new studies are unlikely to change the effect estimates to low if estimates from the currently available body of evidence is very likely to change as new data become available or insufficient if evidence is unavailable or does not permit a conclusion. To evaluate differential efficacy and safety (heterogeneity of effect, interaction), we will focus on RCTs as they have the least potential for bias and confounding thus potentially allowing for causal inference. Further, only RCTs that formally test for interaction between subgroups will be reported. 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 are based on recommendations from Oxman and Guyatt[8] and the Instrument to assess the Credibility of Effect Modification (ICEMAN) criteria.[12] The overall strength of evidence reflects our confidence in the effects estimated in the included studies and how likely new studies are to change the estimates. If only poor-quality studies are available for an outcome, SOE will be graded as insufficient.

The strength of evidence for the overall body of evidence for all *critical health outcomes* was assessed by one researcher following the principles for adapting GRADE (Grades of Recommendation Assessment, Development and Evaluation) as outlined by the Agency for Healthcare Research and Quality (AHRQ). The strength of evidence was based on the highest quality evidence available for a given *primary* outcome. In determining the strength of body of evidence regarding a given *primary* 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 are similar in terms of range and variability.
- **Directness:** describes whether the evidence is directly related to patient health outcomes.
- **Precision:** describes the level of certainty surrounding the effect estimates.
- **Publication bias:** is considered when there is concern of selective publishing.

All AHRQ "required" and "additional" domains (risk of bias, consistency, directness, precision, and if possible, publication bias) were assessed. Bodies of evidence consisting of RCTs were initially considered as High strength of evidence (SoE), while those that comprised nonrandomized studies began as Low strength of evidence. The strength of evidence could be downgraded based on the limitations described above. There could also be situations where the *nonrandomized* studies could be upgraded, including the presence of plausible unmeasured confounding and bias that would decrease an observed effect or increase an effect if none was observed, presence of a dose-response relationship, and large magnitude of effect (strength of association) *if no downgrades for domains above*. Publication and reporting bias are difficult to assess. Publication bias is particularly difficult to assess with fewer than 10 RCTs (AHRQ methods guide). When publication bias was unknown in all studies and this domain is often eliminated

from the strength of evidence tables for our reports. The final strength of evidence for each **primary** outcome 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 probably stable but some doubt remains.

Low— Limited confidence that effect size estimates lie close to the true effect for this outcome; important 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 deficiencies precluding judgment.

Similar 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.

Appendix Table D8. Example methodology outline for determining overall strength of evidence (SoE):

All AHRQ "required" and "additional" domains* are assessed. Only those that influence the baseline grade are listed in table below.
Baseline strength: HIGH = RCTs. LOW = observational, cohort studies, administrative data studies.
DOWNGRADE: Risk of bias for the individual article evaluations (1 or 2); Inconsistency** of results (1 or 2); Indirectness of evidence (1 or 2); Imprecision of effect estimates (1 or 2); Sub-group analyses not stated *a priori* and no test for interaction (2)
UPGRADE (non-randomized studies): Large magnitude of effect (1 or 2); Dose response gradient (1) done for observational studies ***if no downgrade for domains above***

Outcome	Strength of Evidence	Conclusions & Comments	Baseline SOE	DOWNGRADE	UPGRADE
Outcome	**HIGH**	Summary of findings	**HIGH** RCTs	**NO** consistent, direct, and precise estimates	**NO**
Outcome	**MODERATE**	Summary of findings	**LOW** Cohort studies	**NO** consistent, direct, and precise estimates; high quality	**YES** Large effect

				(moderately low ROB)	
Outcome	**LOW**	Summary of findings	**HIGH** RCTs	**YES (2)** Inconsistent Indirect	**NO**

*Required domains: risk of bias, consistency, directness, precision. Plausible confounding that would decrease observed effect is accounted for in our baseline risk of bias assessment through individual article evaluation. Additional domains: dose-response, strength of association, publication bias.
**Single study = "consistency unknown", may or may not be downgraded

APPENDIX E. Study Quality: Risk of Bias evaluation

Appendix Table E1. Risk of Bias Assessment: Osteoporosis Trials Evaluating Vertebroplasty versus Sham

Methodological Principle	Buchbinder, 2009; Kroon 2014; Staples, 2015	Carli, 2023	Clark, 2016; Diamond, 2020	Firanescu, 2018; Firanescu, 2019	Hansen, 2019	Kallmes, 2009; Comstock, 2013
Study design						
Randomized controlled trial	■	■	■	■	■	■
Random sequence generation	Yes	Yes	Yes	Yes	Yes	Yes
Concealed allocation	Yes	Unclear	Yes	Unclear	Unclear	Yes
Groups comparable at baseline*	No	No	No	Yes	No	Yes
Outcome assessors independent or blinded	Yes	Yes	Yes	Yes	Yes	Yes
Care providers blinded	No	No	Yes	Unclear	No	No
Patients blinded	Yes	Yes	Yes	Yes	Yes	Yes
Reporting of attrition	Yes	Yes	Yes	Yes	Yes	Yes
Complete follow-up of ≥80%	Yes	Yes	Yes	Yes	Yes	Yes
<10% difference in follow-up between groups	Yes	Yes	Yes	Yes	Yes	Yes
Intention to treat	Yes	Yes	Yes	Yes	Yes	Yes
Outcomes prespecified	Yes	Yes	Yes	Yes	Yes	Yes
Risk of Bias	**Good**	**Fair**	**Fair**	**Good**	**Fair**	**Good**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E2. Risk of Bias Assessment: Osteoporosis Trials Evaluating Vertebroplasty versus Usual Care

Methodological Principle	Blasco, 2012	Leali, 2016	Chen, 2014	Farrokhi, 2011	Klazen, 2010; Klazen, 2010 (Venmans)	Rousing, 2009; Rousing, 2010	Voormolen, 2007	Yang, 2016
Study design								
Randomized controlled trial	■	■	■	■	■	■	■	■
Random sequence generation	Yes	Unclear	Unclear	Yes	Yes	Unclear	Yes	Yes
Concealed allocation	Unclear	Unclear	Unclear	Unclear	Yes	Unclear	Yes	Unclear
Groups comparable at baseline*	No	Unclear	Yes	No	No	No	No	Yes
Outcome assessors independent or blinded	Unclear	Unclear	Unclear	Yes	No	Unclear	Unclear	No
Care providers blinded	No	No	No	No	No	No	No	No
Patients blinded	No	No	No	No	No	No	No	No
Reporting of attrition	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Complete follow-up of ≥80%	No	Unclear	Yes	Yes	Yes	Yes	Yes	No
<10% difference in follow-up between groups	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	No
Intention to treat	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	No
Outcomes prespecified	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes
Risk of Bias	**Fair**	**Poor**	**Fair**	**Poor**	**Fair**	**Poor**	**Fair**	**Poor**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E3. Risk of Bias Assessment: Osteoporosis Trials Evaluating Vertebroplasty versus Kyphoplasty

Methodological Principle	Dohm, 2014	Endres, 2012	Evans, 2016	Griffoni, 2020	Liu, 2010; Liu, 2015
Study design					
Randomized controlled trial	■	■	■	■	■
Random sequence generation	Yes	Unclear	Yes	Yes	Yes
Concealed allocation	Unclear	Unclear	Yes	Yes	Unclear
Groups comparable at baseline*	Yes	No	No	Yes	Yes
Outcome assessors independent or blinded	Unclear (clinically), yes (radiographs)	Unclear	Yes	Unclear (clinically), yes (radiographs)	Yes (radiographic), Unclear (all others)
Care providers blinded	No	No	No	No	No
Patients blinded	No	Yes	Unclear	No	Unclear
Reporting of attrition	Yes	Yes	Yes	Yes	No
Complete follow-up of ≥80%	Yes (3 mos), No (12-24 mos)	Yes	No	Yes	Unclear
<10% difference in follow-up between groups	Yes	Yes	Yes	Yes	Unclear
Intention to treat	No	Yes	Unclear	Yes	Unclear
Outcomes prespecified	Yes	Yes	Yes	Yes	Yes
Risk of Bias	**Poor**	**Poor**	**Fair**	**Fair**	**Poor**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E3 (continued). Risk of Bias Assessment: Osteoporosis Trials Evaluating VP versus KP

Methodological Principle	Vogl, 2013	Wang, 2015	Wang, 2018	Wang, 2023	Yi, 2014
Study design					
Randomized controlled trial	■	■	■	■	■
Random sequence generation	Unclear	Unclear	Yes	Unclear	Unclear
Concealed allocation	Unclear	Unclear	Unclear	Unclear	Unclear
Groups comparable at baseline*	Yes	Yes	Yes	Yes	Unclear
Outcome assessors independent or blinded	No	Yes (radiographic), Unclear (all others)	Unclear	Unclear	Unclear
Care providers blinded	No	No	Unclear	Unclear	No
Patients blinded	Yes	Yes	Unclear	Unclear	Unclear
Reporting of attrition	Yes	Yes	No	No	Yes
Complete follow-up of ≥80%	No	Yes	Unclear	Unclear	Yes
<10% difference in follow-up between groups	Yes (3 mos), No (2 mos)	Yes	Unclear	Unclear	Yes
Intention to treat	No	Yes	Unclear	Unclear	Yes
Outcomes prespecified	Yes	Yes	Yes	Yes	Yes
Risk of Bias	**Poor**	**Fair**	**Poor**	**Poor**	**Poor**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E4. Risk of Bias Assessment: Osteoporosis Trials Evaluating Vertebroplasty versus Nerve Block

Methodological Principle	Tan, 2023	Wang, 2016
Study design		
Randomized controlled trial	■	■
Random sequence generation	Yes	Yes
Concealed allocation	Yes	Unclear
Groups comparable at baseline*	No	Yes
Outcome assessors independent or blinded	No	Unclear
Care providers blinded	No	Unclear
Patients blinded	No	Unclear
Reporting of attrition	Yes	Yes
Complete follow-up of ≥80%	Yes (1 and 4 wks), No (8 wks)	Yes
<10% difference in follow-up between groups	Yes	Yes
Intention to treat	Yes	Yes
Outcomes prespecified	Yes	Yes
Risk of Bias	**Fair**	**Fair**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E5. Risk of Bias Assessment: Osteoporosis Trials Evaluating Kyphoplasty versus Usual Care

Methodological Principle	Li, 2017	Liu, 2019	Wardlaw, 2009
Study design			
Randomized controlled trial	■	■	■
Random sequence generation	Unclear	Unclear	Yes
Concealed allocation	Unclear	Unclear	Yes
Groups comparable at baseline*	No	Unclear	Yes
Outcome assessors independent or blinded	Unclear	Unclear	No
Care providers blinded	No	No	No
Patients blinded	No	No	No
Reporting of attrition	No	No	Yes
Complete follow-up of ≥80%	Unclear	Unclear	Yes (1, 3, 6 mos), No (12 mos)
<10% difference in follow-up between groups	Unclear	Unclear	Yes (1 and 12 mos), No (3 and 6 mos)
Intention to treat	Unclear	Unclear	Yes
Outcomes prespecified	Yes	Yes	Yes
Risk of Bias	**Poor**	**Poor**	**Fair**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E6. Risk of Bias Assessment: Osteoporosis Trials Evaluating Kyphoplasty versus Vertebral Body Stenting

Methodological Principle	Werner, 2013
Study design	
Randomized controlled trial	■
Random sequence generation	Yes
Concealed allocation	Unclear
Groups comparable at baseline*	No
Outcome assessors independent or blinded	No
Care providers blinded	No
Patients blinded	No
Reporting of attrition	Yes
Complete follow-up of ≥80%	Yes
<10% difference in follow-up between groups	Yes
Intention to treat	Yes
Outcomes prespecified	Yes
Risk of Bias	**Fair**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.

Appendix Table E7. Risk of Bias Assessment: Cancer Trials Evaluating Kyphoplasty versus Usual Care

Methodological Principle	Berenson, 2011†
Study design	
Randomized controlled trial	■
Random sequence generation	Yes
Concealed allocation	Yes
Groups comparable at baseline*	No
Outcome assessors independent or blinded	No
Care providers blinded	No
Patients blinded	No
Reporting of attrition	Yes
Complete follow-up of ≥80%	Yes (1 month), No (3, 6, 12 mos)
<10% difference in follow-up between groups	No (1 month), Yes (3, 6, 12 mos)
Intention to treat	Yes (1 month), No (3, 6, 12 mos)
Outcomes prespecified	Yes
Risk of Bias	**Fair (1 month)**

Unclear indicates that the study had insufficient detail to determine whether criteria were met
*Groups must be comparable on a robust set of baseline characteristics or present evidence that controlling of confounding presented was performed.
† The CAFE Trial allowed for crossover at 1 month. Most patients in the usual care group crossed over and received kyphoplasty at that point.

Appendix Table E8. Risk of Bias Assessment: NRSI evaluating Kyphoplasty versus Usual Care in Patients with Osteoporotic VCFs

Methodological Principle	Faloon, 2015 (retrospective)
Did the study attempt to enroll a random sample or consecutive patients meeting inclusion criteria (inception cohort) from same underlying population?	Yes
Were the groups comparable at baseline on key prognostic factors?	No (age, smoking, diabetes)
Did the article report attrition?	No
Overall loss to follow up acceptable? (≤20%) Differential loss to follow up acceptable? (≤10%)	Unclear
Were the outcomes investigated prespecified and defined?	Yes
Did the study clearly describe and use accurate methods for ascertaining outcomes, exposures, and potential confounders?	Yes
Were outcome assessors and/or data analysts blinded to treatment?	Unclear
Did the study perform appropriate statistical analyses on potential confounders or otherwise control for confounding (e.g. restriction, stratification, matching)?	Yes
Was the duration of follow-up reasonable for investigated events?	Yes
Quality (Risk of Bias)	**Poor (High)**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

Appendix Table E9. Risk of Bias Assessment: NRSI evaluating Vertebroplasty versus Kyphoplasty in Patients with Malignancies (From prior review)

Methodological Principle	Fourney, 2003 (retrospective)	Kose, 2006 (retrospective)
Did the study attempt to enroll a random sample or consecutive patients meeting inclusion criteria (inception cohort) from same underlying population?	Yes	Yes
Were the groups comparable at baseline on key prognostic factors?	No	Unclear
Did the article report attrition?	No	No
Overall loss to follow up acceptable? (≤20%) Differential loss to follow up acceptable? (≤10%)	Unclear	Unclear
Were the outcomes investigated prespecified and defined?	Yes	Yes
Did the study clearly describe and use accurate methods for ascertaining outcomes, exposures, and potential confounders?	No	No
Were outcome assessors and/or data analysts blinded to treatment?	Unclear	No
Did the study perform appropriate statistical analyses on potential confounders or otherwise control for confounding (e.g. restriction, stratification, matching)?	No	No
Was the duration of follow-up reasonable for investigated events?	Unclear*	Yes
Quality (Risk of Bias)	**Poor (High)**	**Poor (High)**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

* Patient follow-up ranged from one day to 19.7 months, and only 14% of the sample was available at 1 year follow-up.

Appendix Table E10. Risk of Bias Assessment: NRSI evaluating Sacroplasty in Patients with Osteoporotic

Methodological Principle	Frey, 2017 (prospective)	Andresen, 2022 (retrospective)
Did the study attempt to enroll a random sample or consecutive patients meeting inclusion criteria (inception cohort) from same underlying population?	No*	Unclear
Were the groups comparable at baseline on key prognostic factors?	Unclear	Unclear
Did the article report attrition?	Yes	No
Overall loss to follow up acceptable? (≤20%) Differential loss to follow up acceptable? (≤10%)	No	Unclear
Were the outcomes investigated prespecified and defined?	Yes	Yes
Did the study clearly describe and use accurate methods for ascertaining outcomes, exposures, and potential confounders?	Yes	Yes
Were outcome assessors and/or data analysts blinded to treatment?	Unclear	Unclear
Did the study perform appropriate statistical analyses on potential confounders or otherwise control for confounding (e.g. restriction, stratification, matching)?	No	No
Was the duration of follow-up reasonable for investigated events?	Yes	Yes
Quality (Risk of Bias)	**Poor (High)**	**Poor (High)**

Unclear indicates that the study had insufficient detail to determine whether criteria were met

* Patients were only referred for sacroplasty if they failed non-surgical management, likely making them more complex cases.

Appendix Table E11. Risk of Bias Assessment: Administrative Database Studies for Patients with Osteoporotic VCFs

Methodological Principle (Interventions)	Spross, 2014 (KP)	Choo, 2018 (KP and VP)	Kim, 2022 (KP and VP)	Ong, 2018 (VP vs. KP vs. non-op)	Wu, 2012 (VP vs. surgery)
Study design Administrative database comparative study Administrative database case-control study Administrative database case series	x	x	x	x	x
Why database created clearly stated	Yes	Yes	No	No	Yes
Description of database's inclusion/exclusion criteria	No	Yes	No	No	Yes
Description of methods for reducing bias in database	No	Yes	Yes	Yes	Yes
Codes and search algorithms reported	No	Yes	Yes	Yes	Yes
Rationale for coding algorithm reported	No	No	No	Yes	Yes
Code accuracy reported	No	Yes	No	No	Yes
Code validity reported	No	No	No	No	Yes
Clinical significance assessed	Yes	Yes	Yes	Yes	Yes
Is the period of data consistent with the outcome data?	Yes	Yes	Yes	Yes	Yes
Statement regarding whether data stems from single or multiple hospital admissions	No	No	No	No	No
Statement regarding whether data stems from single or multiple procedures	No	No	No	Yes	No
Accounting for clustering	No	Yes?	No	Yes	No
Number of criteria met (maximum: 12)	3	8	4	7	9

KP = kyphoplasty; VP = vertebroplasty.

Appendix Table E12. QHES Assessment of U.S. Cost-effectiveness studies

Question	Possible Points*	Edidin, 2012	Hopkins, 2020
1. Was the study **objective** presented in a clear, specific, and measurable manner?	7	7	7
2. Were the **perspective** of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	0	0
3. Were **variable estimates** used in the analysis from the best available source (i.e., randomized controlled trial - best, expert opinion - worst)?	8	0	0
4. If estimates came from a **subgroup analysis**, were the groups prespecified at the beginning of the study?	1	1	1
5. Was **uncertainty** handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	0	9
6. Was **incremental analysis** performed between alternatives for resources and costs?	6	6	6
7. Was the methodology for **data abstraction** (including the value of health states and other benefits) stated?	5	5	5
8. Did the **analytic horizon allow time** for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	7	7
9. Was the **measurement of costs** appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	8	8
10. Were the primary **outcome measure(s**) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	0	6
11. Were the health outcomes **measures/scales valid** and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	0	7
12. Were the **economic model** (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	8	8
13. Were the choice of economic model, main **assumptions, and limitations** of the study stated and justified?	7	0	7
14. Did the author(s) explicitly discuss direction and magnitude of potential **biases?**	6	0	0
15. Were the **conclusions/recommendations** of the study justified and based on the study results?	8	8	8
16. Was there a statement disclosing the **source of funding** for the study?	3	3	3
Total	**100**	**53**	**82**

* Study must fit criteria in order to receive full points. Partial credit is not given. If criteria is not met, then the question receives no points.

Appendix Table E13. QHES Assessment of Non-U.S. Government Cost-effectiveness studies

Question	Possible Points*	Cameron, 2016	Stevenson, 2014
1. Was the study **objective** presented in a clear, specific, and measurable manner?	7	7	7
2. Were the **perspective** of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	4	4
3. Were **variable estimates** used in the analysis from the best available source (i.e., randomized controlled trial - best, expert opinion - worst)?	8	0	8
4. If estimates came from a **subgroup analysis**, were the groups prespecified at the beginning of the study?	1	1	0
5. Was **uncertainty** handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	9	9
6. Was **incremental analysis** performed between alternatives for resources and costs?	6	6	6
7. Was the methodology for **data abstraction** (including the value of health states and other benefits) stated?	5	5	5
8. Did the **analytic horizon allow time** for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	7	7
9. Was the **measurement of costs** appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	8	8
10. Were the primary **outcome measure(s**) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	0	6
11. Were the health outcomes **measures/scales valid** and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	7	7
12. Were the **economic model** (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	8	8
13. Were the choice of economic model, main **assumptions, and limitations** of the study stated and justified?	7	7	7
14. Did the author(s) explicitly discuss direction and magnitude of potential **biases?**	6	0	6
15. Were the **conclusions/recommendations** of the study justified and based on the study results?	8	8	8
16. Was there a statement disclosing the **source of funding** for the study?	3	3	3
Total	**100**	**80**	**99**

* Study must fit criteria in order to receive full points. Partial credit is not given. If criteria is not met, then the question receives no points.

Appendix Table E14. QHES Assessment of Non-U.S. Cost-effectiveness studies

Question	Possible Points*	Svedbom, 2013	Fritzell, 2011	Takahashi, 2019
1. Was the study **objective** presented in a clear, specific, and measurable manner?	7	7	7	7
2. Were the **perspective** of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	4	4	0
3. Were **variable estimates** used in the analysis from the best available source (i.e., randomized controlled trial - best, expert opinion - worst)?	8	0	8	0
4. If estimates came from a **subgroup analysis**, were the groups prespecified at the beginning of the study?	1	1	0	1
5. Was **uncertainty** handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	9	9	9
6. Was **incremental analysis** performed between alternatives for resources and costs?	6	6	6	6
7. Was the methodology for **data abstraction** (including the value of health states and other benefits) stated?	5	5	0	0
8. Did the **analytic horizon allow time** for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	7	0	7
9. Was the **measurement of costs** appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	7	8	8
10. Were the primary **outcome measure(s)** for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	6	6	6
11. Were the health outcomes **measures/scales valid** and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	0	7	7
12. Were the **economic model** (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	8	0	8
13. Were the choice of economic model, main **assumptions, and limitations** of the study stated and justified?	7	7	7	7
14. Did the author(s) explicitly discuss direction and magnitude of potential **biases?**	6	6	6	0
15. Were the **conclusions/recommendations** of the study justified and based on the study results?	8	8	8	8
16. Was there a statement disclosing the **source of funding** for the study?	3	3	3	3
Total	**100**	**84**	**79**	**79**

* Study must fit criteria in order to receive full points. Partial credit is not given. If criteria is not met, then the question receives no points.

APPENDIX F. Detailed Characteristics and Demographic Tables of Randomized Control Trials

Appendix Table F1. Patient Characteristics of Studies Comparing Vertebroplasty to Other Treatments in Patients with Fractures due to Osteoporosis

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty versus Sham						
Carli, 2023 Study period NR; Recruitment period: May 2013 to June 2019 Netherlands	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain (median): 176 days vs. 185 days Duration of symptoms <6 weeks: NR Severity of fracture Mild: 40.7% Moderate: 30.4% Severe: 28.9% Number of vertebral bodies treated: 1: 60.0% 2: 18.8% 3: 15.0% 4: 3.8% 5: 2.5% One or more previous vertebral fractures: NR Fracture appearance	N=80 Mean age (SD): 71 (10) years vs. 69 (10) years Female: 68% Race/Ethnicity: NR	Vertebroplasty (n=40) VP trans- or bipedicular approach using PMMA (mean 1.4 ml). Analgesics allowed during study	Sham (n=40) Sham procedure using periosteal approach with cement mixed but not used to improve blinding. Analgesics allowed during study	1 day 100% (40/40) vs. 100% (40/40) 1 week 100% (40/40) vs. 95.5% (39/40) 1 month 100% (40/40) vs. 97.5% (39/40) 3 months 97.5% (39/40) (97.5%) vs. 97.5% (39/40) 6 months 97.5% (39/40) vs. 92.5% (37/40) 12 months 97.5% (39/40) vs. 90.0% (36/40)	Funding NR Authors report grants and consulting fees from industry

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Wedge: 74.1% Biconcave: 25.9% Crossover interventions: NR					
Clark, 2016; Diamond, 2020 Study period: Nov 4, 2011 to Dec 5, 2014; Recruitment period NR Australia	Fracture type: Osteoporotic VCF Fracture age: 2.6 weeks Duration of back pain (median): NR Duration of symptoms <6 weeks: 100%* Severity of fracture (Genant grade): 1: 9.2% 2: 20.8% 3: 71.7% Number of vertebral bodies treated: KP: 1: 86.7% 2: 13.3% One or more previous vertebral fractures: 56.7% Fracture appearance: NR	N=120 Mean age (SD): 80 (NR) years Female: 73.3% Race/Ethnicity: NR	Vertebroplasty (n=61) VP trans- or bipedicular using PMMA (7.5 ml). Analgesics allowed during study .	Sham (n=59) Sham procedure using periosteal approach with cement mixed but not used to improve blinding. Analgesics allowed during study	3 days 95.1% (58/61) vs. 93.2% (55/59) 14 days 90.2% (55/61) vs. 96.6% (57/59) 1 month 90.2% (55/61) vs. 96.6% (57/59) 3 months 86.9% (53/61) vs. 88.1% (52/59) 6 months 83.6% (51/61) vs. 86.4% (51/59)	CareFusion Corporation No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover interventions: NR					
Diamond, 2020[†] Subgroup analysis of Clark, 2016	Fracture type: Osteoporotic VCF Fracture age: <3 weeks[*] Duration of back pain (median): NR Duration of symptoms <6 weeks: 100% (all within 3 weeks) Severity of fracture (Genant grade): 1: 7.5% 2: 21.5% 3: 71.0% Number of vertebral bodies treated: 1: 88.2% 2: 8.8% One or more previous vertebral fractures: NR Fracture appearance: NR	N=93 Mean age (SD): 82 (8) years Female: 73.1% Race/Ethnicity: NR	Vertebroplasty (n=46) See Clark, 2016	Sham (n=47) See Clark, 2016	3 days 93.5% (87/93) 14 days 92.5% (86/93) 1 month 91.4% (85/93) 3 months 86.0% (80/93) 6 months 83.9% (78/93)	See Clark, 2016

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover interventions: NR					
Firanescu, 2018; Firanescu, 2019 Study period: NR; Recruitment period: Jan 2011 to Jan 2013 The Netherlands	Fracture type: Osteoporotic VCF Fracture age (from initial pain): 40 days Duration of symptoms <6 weeks: NR Severity of fracture (Genant grade): 1: 29.9% 2: 44.6% 3: 25.4% Number of vertebral bodies treated: 1: 77.8% 2: 17.0% 3: 5.1% One or more previous vertebral fractures: NR Fracture appearance: Wedge: 54.0% Biconcave: 46.0% Crossover interventions: NR	N=180‡ randomized N=176 analyzed at baseline Mean age (SD): 75.8 (NR) years Female: 75.6% Race/Ethnicity: NR	Vertebroplasty (n=90) Used PMMA (5.1 ml), approach not specified. Analgesics allowed during study	Sham (n=86) Identical procedure to VP without actual treatment. Analgesics allowed during study	1 month 100% (176/176) 3 months 97.2% (171/176) 6 months 93.8% (165/176) 12 months 86.4% (152/176)	Stryker One author reports consulting fees from industry, as well as serving on a data and safety monitoring board. No other COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Hansen, 2019 Study period: NR Recruitment period: 2011 to 2014 The Netherlands	Fracture type: Osteoporotic VCF New VCFs: 21% vs. 19% Duration of back pain: ≤8 weeks* Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR	N=52 randomized N=46 analyzed at baseline Mean age (SD): 69.9 (NR) years Female: 87.0% Race/Ethnicity: NR	Vertebroplasty (n=24) Procedure used the V-Max Mixing and Delivery system (DePuy Acromed) under fluoroscopy using PMMA (2 to 4 ml).	Sham (n=22) Procedure the same as VP, except 2 mL of Lidocaine was injected into the Sham group.	52 randomized 51 received treatment 46 analyzed at all time points	Danish Rheumatism Society COIs NR
Kallmes, 2009; Comstock, 2013 Study period: NR; Recruitment period: Jun	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain (mean): 18 weeks	N=131 Mean age (SD): 73.8 (NR) years Female: 75.6% Race/Ethnicity: White: 96.9%	Vertebroplasty (n=68) Used PMMA (2.6 ml), via central aspect of target vertebra(e). Filling stopped once cement PMMA	Sham (n=63) During the control intervention, verbal and physical cues, such as pressure on the patient's back, were given, and	1 month 98.5% (67/68) vs. 96.8% (61/63) 3 months 94.1% (64/68) vs. 96.8% (61/63)	National Institute of Arthritis and Musculoskeletal and Skin Diseases Authors report receiving consulting fees and grant support, lecture fees, and having equity interest in industry

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
2004 to Aug 2008 UK, Australia	Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: 1: 67.9% 2: 20.6% 3: 11.5% One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention at 3 months: 11.8% (8/64) vs. 44.3% (27/61)	Not White: 3.1%	reached posterior aspect or entered extraosseous space	the methacrylate monomer was opened to simulate the odor associated with mixing of PMMA, but the needle was not placed and PMMA was not infused.		
Buchbinder, 2009; Kroon 2014; Staples, 2015 Study period: NR; Recruitment period: Jun 2004 to Aug 2008 Australia	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain (mean): 9.3 weeks Duration of symptoms <6 weeks: 32.1% Severity of fracture (total fractures):	N=78 Mean age (SD): 73.8 (NR) years Female: 75.6% Race/Ethnicity: White: 96.9% Not White: 3.1%	Vertebroplasty (n=38) Used PMMA (3 ml). Unipedicular approach preferred, bipedicular approach used only if there was inadequate instillation of cement with the	Sham (n=40) Same procedure as VP until insertion of the need. Needle replaced with a blunt stylet. To simulate vertebroplasty, the vertebral body was gently tapped, and PMMA was prepared so that	1 week 97.4% (37/38) vs. 92.5% (37/40) 1 month 92.1% (35/38) vs. 95.0% (38/40) 3 months 94.7% (36/38) vs. 92.5% (37/40) 6 months	National Health and Medical Research Council of Australia, Arthritis Australia, the Cabrini Education and Research Institute, and Cook Australia One author reports grant support from Cook Australia. No other COIs.

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Mild: 27.2% Moderate: 48.9% Severe: 23.9% Number of vertebral bodies treated: 1: 82.1% 2: 17.9% One or more previous vertebral fractures: 50.0% Fracture appearance: NR Crossover intervention: None		unipedicular approach. All received cephalothin administered intravenously. All patients then received usual care.	its smell permeated the room. All patients then received usual care.	92.1% (35/38) vs. 90.0% (36/40) 12 months 89.5% (34/38) vs. 90.0% (36/40) 24 months 84.2% (32/38) vs. 80.0% (32/40)	
Vertebroplasty versus Usual Care						
Blasco, 2012 Study period NR; recruitment period: April 2006 to January 2010 Spain	Fracture type: Osteoporotic VCF Fracture age: <12 months Duration of back pain (mean): 142 days Patients with symptom onset <6 weeks: 4.8% Patients with symptom onset <4 months: 51.2%	N=125 Mean age (SD): 73.2 (9.3) years Female: 77.6% Race/Ethnicity: NR	Vertebroplasty (n=64) VP using PMMA (volume NR) via bilateral transpedicular approach. After surgery, patients received calcitonin for one month and standard analgesics as necessary. Following one month, patients	Treatment as usual (n=61) Consisted of analgesics with standardized format and nasal calcitonin in the first month. All patients were offered rescue therapy if treatment was ineffective or on intolerance to drug therapy. In case of	2 weeks 79.7% (51/64) Vs. 96.7% (59/61) 2 months 84.4% (54/64) vs. 91.8% (56/61) 6 months 78.2% (50/64) vs. 88.5% (54/61) 12 months 73.4% (47/64) vs. 78.7% (48/61)	Fundació La Marató de TV3, the Spanish Society of Medical Radiology, and the Catalan Society of Rheumatology. No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Severity of fracture: NR Mean number of vertebral bodies treated: 2.46[§] One or more previous vertebral fractures: NR 2 initial fractures: 25% >2 initial fractures: 49% Fracture appearance: NR Crossover interventions: Any patient was offered rescue therapy by intrathecal infusion when treatment was ineffective (VAS ≥7) or on intolerance to drug therapy. Patients in the TAU group were offered vertebroplasty if there was an absence of improvement despite		received bisphosphonates (except for those with intolerance, received teriparatide or strontium ranelate).	no improvement, patient was considered for vertebroplasty and the case was deemed a failure[**]. Following one month, patients received bisphosphonates (except for those with intolerance, received teriparatide or strontium ranelate).		

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	all standardized measures.					
Yang, 2016 Study period: Jan 2009 to Dec 2011; Recruitment period: NR China	Fracture type: Traumatic Osteoporotic VCF Fracture age (SD): 5.5 (NR) days[††] Duration of back pain (mean): 5.5 days Duration of symptoms <6 weeks: 100% Severity of fracture: NR Number of vertebral bodies treated: 1: 85.0% 2: 15.0% One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention: 15.2% (10/66) conservative	N=135 randomized[‡‡] N=107 analyzed at baseline Mean Age (SD): 76.7 (NR) years Female: 64.5% Race/Ethnicity: NR	Vertebroplasty (n=56) PMMA (4.5 ml) via fluoroscopy-guided transpedicular approach. Injection ceased when cement reached the cortical edge of the vertebral body or leaked into the extraosseous structures or veins, bipedicular approach then used if filling was incomplete. Osteoporotic medication including bisphosphonates, calcium supplementation, and vitamin D prescribed	Conservative Care (n=51) 2 weeks bed rest, then walking with brace and assistance. For pain medication, nonsteroidal anti-inflammatory drugs (NSAIDs) were prescribed for every patient. Additional analgesics, such as tramadol and morphine, would be added in case NSAIDs were not effective. Two weeks after diagnosis, physical therapy was started. Osteoporotic medication including bisphosphonates, calcium	135 randomly assigned, 130 accepted, 107 completed follow up	No funding No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	care crossed over to VP but were excluded from analyses			supplementation, and vitamin D prescribed		
Leali, 2016 Study period: NR; Recruitment period: NR Italy, France, Switzerland	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): NR Duration of symptoms <6 weeks: NR§§ Severity of fracture: NR Number of vertebral bodies treated: 1: 100% One or more previous vertebral fractures: 1: 100%*** Fracture appearance: NR Crossover intervention: None	N=400 Mean Age (SD): NR (range 56 to 82) Female: 100% Race/Ethnicity: NR	Vertebroplasty (n=200) PMMA (4 mL) via transpedicular approach. Patients were treated with acetaminophen, non-steroidal drugs (NSAIDs), or derivatives of morphine as needed after surgery.	Conservative Care (n=200) Pain medication, osteoporosis medication, physiotherapy, or bracing	6 months 96.3% (385/400)	NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Chen, 2014 Study period: NR; Recruitment period: Jan 2007 to Dec 2012 China	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): 30.2 weeks Duration of symptoms <6 weeks: 0% (exclusion criteria) Severity of fracture: NR Number of VCF at baseline: 2.1 One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention: Four patients in the conservative group had VP done at 3 months and were excluded from the study	N=96 randomized N=89 analyzed at baseline Mean Age (SD): 65.5 (9.1) years Female: 70.0% Race/Ethnicity: NR	Vertebroplasty (n=46) PMMA (3.6 ml) via transpedicular approach; Injection was ceased when substantial resistance was met or when the cement reached the cortical edge of the fractured vertebral body; injection was also stopped if cement leaked into extraosseous structures or veins (approximately 3-5 mLs)	Conservative Care (n=43) Hospitalized and offered brace treatment, analgesia, general mobilizing physiotherapy, and osteoporotic medication treatment, including vitamin D, and diphosphonate	12 months: 91.3% (42/46) vs. 86.0% (43/50)[†††]	No funding No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Farrokhi, 2011 Study period: Sep 2004 to Jan 2009; Recruitment period: Sep 2004 to Jan 2006 Iran	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): 28.5 weeks Duration of symptoms <6 weeks: NR Severity of fracture: Mild: 62.2% Moderate: 29.3% Severe: 6.1% Number of vertebral bodies treated: 1: 26.4% 2: 24.8% >2: 48.8% One or more previous vertebral fractures: NR Fracture appearance: Wedge: 84.2% Biconcave: 15.8% Crossover intervention allowed at 1 month: <2 months: NA vs. 9.5% (4/42) <6 months: NA vs. 16.7% (7/42)	N=82 Mean Age (SD): 73 (NR) years Female: 73.2% Race/Ethnicity: NR	Vertebroplasty (n=40) PMMA (3.5 ml) via unilateral parapedicular approach in 87.5% (35/40) patients and bilateral transpedicular in 12.5% (5/40) patients Change in lifestyle and physical treatment were also suggested to patients in both groups	Conservative Care (n=42) 250mg acetaminophen with codeine twice daily, 400mg ibuprofen twice a day, 1000mg calcium daily, 400 IU vitamin D daily, 70mg alendronate orally once weekly, and 200 IU calcitonin daily. Doses of analgesics were a baseline suggestion, and the physician could increase them to achieve an optimum dose. Change in lifestyle and physical treatment were also suggested to patients in both groups	1 week 100% (40/40) vs. 100% (42/42) All other follow-ups had crossover	Vice-chancellor for research affairs of Shiraz University of Medical Sciences and Apadana Tajhizgostar Co. No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	<12 months: NA vs. 23.8% (10/42) <24 months: NA vs. 23.8% (10/42) <36 months: NA vs. 47.6% (20/42)					
Klazen, 2010; Klazen, 2010 (2); Klazen, 2010 (3)‡‡‡ Study period: Sep 2004 to Jan 2009; Recruitment period: Oct 1, 2005 to Jun 30, 2008 The Netherlands	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): 28.1 days Duration of symptoms <6 weeks: 100% (inclusion) Severity of fracture: Mild: 43.8% Moderate: 40.2% Severe: 16.0% Mean number of VCF at baseline: 2.3 One or more previous vertebral fractures: NR Fracture appearance: Wedge: 73.0% Biconcave: 27.0% Crossover intervention allowed at 1 month:	N=202 Mean Age (SD): 75.3 (NR) years Female: 69.3% Race/Ethnicity: NR	Vertebroplasty (n=101) PMMA (4.1 mL) via unilateral or bilateral transpedicular approach Throughout follow-up, analgesia in both groups was individually tailored in a stepwise manner from non-opiates to weak opiate derivatives and strong opiate derivatives. All patients were prescribed bisphosphonates, calcium supplementation, and vitamin D.	Conservative Care (n=101) Throughout follow-up, analgesia in both groups was individually tailored in a stepwise manner from non-opiates to weak opiate derivatives and strong opiate derivatives. All patients were prescribed bisphosphonates, calcium supplementation, and vitamin D.	1 day 97.0% (98/101) vs. 93.1% (94/101) 1 week 96.0% (97/101) vs. 92.1% (93/101) 1 month 95.0% (96/101) vs. 91.1% (92/101) 3 months 91.1% (92/101) vs. 85.1% (86/101) 6 months 88.1% (89/101) vs. 80.2% (81/101) 12 months 85.1% (86/101) vs. 76.2% (77/101)	Netherlands Organization for Health Research and Development and COOK Inc. No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	<2 months: NA vs. 9.5% (4/42) <6 months: 16.7% (7/42) <12 months: 23.8% (10/42) <24 months: 23.8% (10/42) <36 months: 47.6% (20/42)					
Rousing, 2009; Rousing, 2010 Study period: NR; Recruitment period: Jan 2001 to Jan 2008 Denmark	Fracture type: Osteoporotic VCF Fracture age (SD): 7.6 days Duration of back pain (mean): NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: 1: 76% 2: 20% 3: 4%[§§§] One or more previous vertebral fractures: NR Fracture appearance: NR	N=49 Mean Age (SD): 80 (NR) years Female: 81.6% Race/Ethnicity: NR	Vertebroplasty (n=25)[****] PMMA (volume NR) via unilateral or bilateral approach. In case of extravertebral cemental leakage, the injection was terminated. Both groups were offered pain medication and physiotherapy, if necessary, until discharge.	Conservative Care (n=24) Offered brace treatment. Both groups were offered pain medication and physiotherapy, if necessary, until discharge.	3 months 96.0% (24/25) vs. 95.8% (23/24) 12 months 92.0% (23/25) vs. 91.7% (22/24)	Foundation and Danish government funds No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover intervention: None					
Voormolen, 2007 Study period: NR; Recruitment period: Jul 2003 to Jun 2005 The Netherlands	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): 81 days Duration of symptoms <6 weeks: 0%* Severity of fracture: Mild: 12.2% Moderate: 22.4% Severe: 65.3% Number of vertebral bodies treated (mean): 1.4 (range 1 to 3) One or more previous vertebral fractures: NR Fracture appearance: Wedge: 77.6% Biconcave: 22.4%	N=34 Mean Age (SD): 73 (range 55 to 88) years Female: 82.4% Race/Ethnicity: NR	Vertebroplasty (n=18) PMMA (3.2 ml) months frequently via bilateral transpedicular approach. Mean 3.2 mL The pain medication was optimized according to the individual need of patients. In ascending order of anesthesia, the patients were treated with paracetamol (acetaminophen), nonsteroidal anti-inflammatory drugs (NSAIDs), or opiate derivatives. To optimize analgesic use, at first the dose per day of prescribed	Conservative Care (n=16) CC optimized during follow-up by internist and/or orthopedic surgeon but not further described The pain medication was optimized according to the individual need of patients. In ascending order of anesthesia, the patients were treated with paracetamol (acetaminophen), nonsteroidal anti-inflammatory drugs (NSAIDs), or opiate derivatives. To optimize analgesic use, at first the dose per day of prescribed	2 weeks 100% (18/18) vs. 100% (16/16)††††	NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover intervention at 2 weeks: NA vs. 87.5%		analgesics was regulated. Second, the class of pain medication was adjusted. Corrections were made on a daily basis by the endocrinologists on an as-needed basis.	analgesics was regulated. Second, the class of pain medication was adjusted. Corrections were made on a daily basis by the endocrinologists on an as-needed basis.		
Yi, 2014 Study period: Nov 2005 to Jul 2009; Recruitment period: NR China	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated (mean): NR One or more previous vertebral fractures: NR Fracture appearance: NR	N=211 Mean Age (range): NR‡‡‡‡ Female: NR‡‡‡‡ Race/Ethnicity: NR	Vertebroplasty (n=90) PMMA (1.5 to 9 ml) via transpedicular approach	Conservative Care (n=121) CC: pain medication, bed rest, a solf bi-valved body brace, and physiotherapy	Mean 49.4 months: 100% (90/90) vs. 100% (121/121)	Funding NR No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover intervention: None					
Vertebroplasty versus Nerve Block						
Tan, 2023 Study period NR; recruitment period: June 2021 to June 2022 UK	Fracture type: Osteoporotic VCF Fracture age: ≤6 weeks* Duration of back pain (mean): ≤6 weeks* Patients with symptom onset <6 weeks: NR Severity of fracture: NR Mean number of vertebral bodies treated: <3* One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR	N=30 randomized N=27 analyzed at baseline§§§§ Mean age (SD): 82 (NR) years Female: 57.1% vs. 84.6% White: 100% vs. 92.3%	Vertebroplasty (n=14) Bipedicular or unipedicular approach with PMMA (2 to 5 ml). Other details NR. Participants encouraged to mobilize following procedure, and prescribed analgesia as required.	Medial branch spinal nerve block (n=13) Performed targeting facet joints above and below the vertebral fracture using fluoroscopy. Mixed of 0.5% bupivacaine with 40 mg depomedrone used, and each medial branch will be blocked with 1 to 1.5 mL solution.	Randomized 30 IP 90% (27/30) 1 week 90% (27/30) 4 weeks 80% (24/30) 8 weeks 70% (21/30)	National Institute for Health Research No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Wang, 2016 Study period NR; recruitment period: NR China	Fracture type: Osteoporotic VCF Fracture age: ≤6 weeks* Acute fractures (<2 weeks): 87% vs. 84.9% Subacute fractures (2-8 weeks): 13% vs. 15.1% Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR	N=217 Mean age (SD): 63.1 (NR) years Female: 82.5% Race/ethnicity: NR	Vertebroplasty (n=108) Bipedicular or unilateral transpedicular approach under fluoroscopy using PMMA (3 to 9 ml). Patients wore a brace to aid ambulation for 3 months following procedure.	Facet block (n=109) Bilateral posterior approach. Mixture of prednisolone (5 mL) and lidocaine (5 mL) injected under fluoroscopic monitoring. Patients wore a brace to aid ambulation for 3 months following procedure.	12 months 94.9% (206/217)	No funding No COI
Vertebroplasty versus Kyphoplasty						
Wang, 2015 Study period January 1 2012 to	Fracture type: Osteoporotic VCF Fracture age: NR	N=107 Mean age (SD): 69 (NR) years	High-viscosity cement Vertebroplasty (n=53)	kyphoplasty (n=54) KP (Kyphon, Sunnyval) performed using	3 months 100% (53/53) vs. 96.3% (52/54) 12 months	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
February 12 2014, recruitment period NR China	Duration of back pain (mean): ≥4 weeks* Mean duration of symptoms: NR Patients with symptom onset <4 months: 100% Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR	Female: 75.7% Race/ethnicity: NR	VP (Confidence Spinal Cement System, DePuy Spine inc.) performed using unipedicular approach using PMMA (3.31 ml). All patients referred for treatment with calcium, vitamin D supplement, and antiresorptive or anabolic agents.	unipedicular approach using PMMA (4.22 ml). All patients referred for treatment with calcium, vitamin D supplement, and antiresorptive or anabolic agents.	94.3% (50/53) vs. 94.4% (51/54)	
Liu, 2010; Liu 2015 Study period NR, recruitment period NR Taiwan	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain (mean): NR	N=100 Mean age (SD): 73 (NR) years Female: 77% Race/ethnicity: NR	Vertebroplasty (n=50) VP done using bipedicular approach using PMMA (4.91 ml) mixed with an antibiotic	Kyphoplasty (n=50) Balloon kyphoplasty done using bipedicular approach using PMMA (5.56 ml) with an antibiotic (gentamicin) under	6 months: NR*****	Chung-Shan Medical University Hospital No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Mean duration of symptoms: 17.9 days Patients with symptom onset <4 months: 100% Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR		(gentamicin) under mobile C-arm X-ray. All patients undertook an orally administered treatment regimen to protect bone density after surgery (details NR)	mobile C-arm X-ray. All patients undertook an orally administered treatment regimen to protect bone density after surgery (details NR)		
Griffoni, 2020 Study period 2011 to 2015, recruitment period Italy	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain (mean): NR Mean duration of symptoms (SD): NR	N=113 Mean age (SD): 73 (NR) years Female: 82% Race/ethnicity: NR	Vertebroplasty (n=64) VP performed according to standard practice. Further details NR. Used Confidence-DePuy Spine PMMA (volume NR)	Kyphoplasty (n=49) KP performed according to standard practice. Further details NR. Performed with a bilateral approach using Kyphon Osteo Introducer system (Medtronic Spine). Volume NR	12 months 97.3% (110/113)[†††††]	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Patients with symptom onset ≥4 weeks: 100%* Patients with symptom onset <4 months: NR Severity of fracture: NR Fractures in ≥2 levels: 34.5% One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR					
Evans, 2016 Study period NR, recruitment period NR USA	Fracture type: Osteoporotic VCF (41%) Fracture age: NR Duration of back pain (mean): NR Mean duration of symptoms (SD): 17.5 (11.7) days vs. 18 (10.3) days	N=115 Mean age (SD): 75.6 (10) years Female: 71% White: 98% Hispanic/ Latino: 1% Other: 1%	Vertebroplasty (n=56) VP according to standard practice and to each practitioner's preference. Approach, device, and cement (volume NR) used were at operators' discretion	Kyphoplasty (n=59) KP according to standard practice and to each practitioner's preference. Approach, device, and cement (volume NR) used were at operators' discretion	3 days NR 1 month NR 6 months NR 12 months 77.8% (88/113)‡‡‡‡‡	Carefusion, Johnson and Johnson/DePuy Synthes Spine, Cardinal Health, and Stryker Authors report consultancies, grants, payments for lectures, royalties, payment for development of educational presentations, and travel/accommodations/meeting expenses.

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Patients with symptom onset ≤12 months: 100% Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: 13.9% Fracture appearance: NR Crossover interventions: NR					
Endres, 2012 Study period NR; recruitment period: NR Germany	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain (mean): NR Patients with symptom onset ≤6 weeks: 100%*	N=66 Mean age (SD): 68 (NR) years Female: NR Race/ethnicity: NR	Vertebroplasty (n=22) VP (Stryker) performed through a unipedicular transpedicular approach, using liquid and powder PMMA (SpinePlex, Stryker) cement (3.1 ml). All patients also	Balloon kyphoplasty (n=22) KP (Medtronic) performed through a unipedicular approach using PMMA (3.9 ml). All patients also received a daily standard dose of oral amino-bisphosphonate,	6 months 89% (59/66)	Funding NR No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Severity of fracture: NR Mean number of vertebral bodies treated: 1* One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR		received a daily standard dose of oral amino-bisphosphonate, 1000 mg calcium, and 1000 IU vitamin D3. Physiotherapy and pain medication prescribed as needed.	1000 mg calcium, and 1000 IU vitamin D3. Physiotherapy and pain medication prescribed as needed. Shield kyphoplasty (n=22) Sheild KP (Soteira) using unipedicular approach, with the Shield (Soteria) implant and pmma. All patients also received a daily standard dose of oral amino-bisphosphonate, 1000 mg calcium, and 1000 IU vitamin D3. Physiotherapy and pain medication prescribed as needed.		

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Dohm, 2014 Study period NR; recruitment period: October 2006 to May 2011 USA	Fracture type: Osteoporotic VCF (41%)§§§§§ Fracture age: ≤6 months* Duration of back pain (mean): 3.6 weeks Patients with symptom onset ≤6 months: 100% Severity of fracture: NR Mean number of vertebral bodies treated: NR Single fractures treated: 78.5% One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: 3.4% (7/205) vs. 2% (4/199)	N=404****** Mean age (SD): 65.6 (NR) years Female: 77.4% White: 94.3% Black: 2.9% Hispanic/ Latino: 2.7% Asian: 1.1% Other: 0.5%	Vertebroplasty (n=205) Details NR. Cement: (median) 4.0 ml)	Balloon kyphoplasty (n=199) KP (Kyphon Osteo Introducer Systems, Medtronic) and PMMA (HV-R Bone Cement, Medtronic) Performed using bilateral approach according to local practices. Details NR. Cement: (median) 4.6 ml)	Baseline†††††† 92.6% (190/205) vs. 96% (191/199) 1 week‡‡‡‡‡‡ 91.7% (188/205) vs. 95% (189/199) 1 month‡‡‡‡‡‡ 88.3% (181/205) vs. 90.5% (180/199) 3 months‡‡‡‡‡‡ 79.5% (163/205) vs. 81.4% (162/199) 12 months‡‡‡‡‡‡ 63.4% (130/205) vs. 71.8% (143/199) 24 months‡‡‡‡‡‡ 44.4% (91/205) Vs. 50.3% (100/199) Modified ITT†††††† 92.6% (190/205) vs. 96% (191/199)	Medtronic Authors report receiving fees for participating in review activities. Some authors report consultancies, stock options, and fees being paid to their institutions as compensation.

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Vogl, 2013 Study period March 2008 to September 2009; recruitment period: NR Germany	Fracture type: Osteoporotic VCF Fracture age: ≤6 weeks* Duration of back pain (mean): ≤6 weeks* Patients with symptom onset <6 weeks: NR Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR	N=77 Mean age (SD): 73 (NR) years Female: 71.4% Race/ethnicity: NR	Vertebroplasty (n=28) Performed using bipedicular cement injection and cement (3.99 ml)	Cement Directed Kyphoplasty System (n=49) Used lateral intrapedicular or extrapedicular approach, with a 10 mm implant and cement (3.77 mL)	3 months 82.1% (23/28) vs. 75.5% (37/49) 12 months 67.9% (19/28) vs. 57.1% (28/49)	Soteira Inc. Authors report board membership, consultancies, expert testimony, payment for lectures, stock/stock options, travel/accommodations/meeting expenses, and grants.
Yi, 2014 Study period: Nov 2005 to Jul 2009;	Fracture type: Osteoporotic VCF Fracture age (SD): NR	N=169 Mean Age (range): NR‡‡‡‡ Female: NR‡‡‡‡	Vertebroplasty (n=90) PMMA (1.5 to 9 ml) via	Kyphoplasty (n=79) KP: PMMA with inflatable bone tamps via	Mean 49.4 months: 100% (90/90) vs. 100% (79/79)	Funding NR No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Recruitment period: NR China	Duration of back pain (mean): NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated (mean): NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention: None	Race/Ethnicity: NR‡‡‡‡	transpedicular approach	transpedicular approach		
Wang, 2018 Study period NR, recruitment period September 2015 to August 2016 China	Fracture type: VCF, bilateral resection of ovarian cancer§§§§§§ Fracture age: NR Duration of back pain (mean): NR Mean duration of symptoms: NR	N=86 Mean age (SD): 42.2 (NR) years Female: NR Race/ethnicity: NR	High-viscosity cement Vertebroplasty (n=43) VP (Haraeus Medical HmbH) using transpedicular approach and cement (3.97 ml). Antibiotic prophylaxis administered	kyphoplasty (n=43) KP not detailed. Cement: 3.89 ml. Antibiotic prophylaxis administered within 24 hours, and out of bed activities after 24 hours of bedrest. All patients received antibiotics 2 hours before,	NR	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Patients with symptom onset <4 months: NR Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR		within 24 hours, and out of bed activities after 24 hours of bedrest. All patients received antibiotics 2 hours before, and fasted for 8 hours before operation	and fasted for 8 hours before operation		
Wang, 2023 Study period: NR; Recruitment period: Jan 2021 to Jan 2022 China	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): NR Duration of symptoms ≤3 weeks: 100%* Severity of fracture: NR	N=100 Mean Age (SD): 81.7 (NR) years Female: 44.0% Race/Ethnicity: NR	Vertebroplasty (n=50) PMMA via posterior approach (cement volume NR); The injection action is stopped before the bone cement is pushed to the posterior edge of the vertebra.	Kyphoplasty (n=50) Balloon kyphoplasty using unspecified bone cement (volume NR)	NR	Science and Technology Program of Health commission of Jiangxi Province and Shangrao 2021 Annual Science and Technology Project No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral bodies treated (mean): NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention at 2 weeks: NA vs. 87.5% (14/16)					

COI = conflict of interest; KP = kyphoplasty; NA = not applicable; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; VCF = vertebral compression fracture; VP = vertebroplasty.
* Inclusion criteria.
† Diamond 2020 is a subgroup analysis of patients receiving surgery within 3 weeks of fracture in Clark 2016.
‡ A total of 180 patients were randomized, 4 did not receive the intervention and were excluded at baseline.
§ Vertebroplasty only.
** Timing of failure was not reported.
†† Vertebroplasty performed on average 8.4 days after onset but presentation time appeared to be 5.5 days after onset.
‡‡ 135 were initially randomized, but authors only retained those that completed one year follow-up (n=107). All others were excluded from the study.
§§ Authors report that patients were acute, but do not define timing.
*** Authors report that they only treated one fracture in each participant
††† Vertebroplasty group initially consisted of 46 patients, but four were lost to follow-up. The conservative group initially consisted of 50 patients, but three were lost to follow-up and four had VP done. The four that crossed over were excluded from the study.
‡‡‡ All Klazen publications were included as part of a dissertation under Venmans 2010/2011.
§§§ All in conservative group
**** Initial randomized included N=50 and n=25 in the VP group. However, one patient refused to attend the hospital and to have a visit at 3-month follow-up, and was therefore excluded from the study.
†††† Authors report that, because the majority of patients in the conservative care group elected to have PV 2 weeks after the beginning of treatment, they stopped the study early.
‡‡‡‡ Study reports patients receiving VP vs. KP vs. conservative care. Authors do not report by group, only for the whole group, and then those that received surgery once, surgery twice, and then conservative therapy once and twice respectively.
§§§§ 30 patients were randomized, but several dropped out before the procedure.
***** Attrition assumed to be 100% as authors do not report otherwise.
††††† Not reported at earlier timepoints or by intervention group.

‡‡‡‡‡ 2 patients did not have pain rating scales before surgery, and were not included in the loss to follow-up.
§§§§§ 45% of vertebroplasty patients and 37.2% of kyphoplasty 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.
****** Used a modified ITT, only 23 patients were enrolled but withdrew before surgery and were therefore not analyzed.
†††††† Patients dropped out after enrollment but before surgery.
‡‡‡‡‡‡ Authors report that the sponsor (Medtronic) terminated the study in 14.1% (29/205 vs. 14.6% (29/199) of patients. They do not report details.
§§§§§§ Fractures due to osteoporosis, but all patients had ovarian cancer.

Appendix Table F2. Patient Characteristics of Studies Comparing Kyphoplasty to Other Treatments in Patients with Fractures due to Osteoporosis

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Kyphoplasty versus Usual Care						
Li, 2017 Study period NR, recruitment period January 2013 to June 2015 China	Fracture type: Osteoporotic VCF Median Symptomatic Fracture age: NR Duration of back pain (mean): NR Mean duration of symptoms: NR Patients with symptom onset <4 months: NR Severity of fracture: Fracture ratings I: 11.25% II: 37.5% III: 51.25%	N=80 Mean age (SD): 74.03 (6.21) years Female: 30% Race/Ethnicity: NR	Kyphoplasty (n=40) KP with PMMA (volume NR), but not further detailed	Conservative treatment (n=40) Confined to rests on platform beds for 8-10 weeks. Exercise interventions also required, but not detailed. Supplementation of antiosteoporosis drugs including vitamin D3 and calcium carbonate for symptomatic treatment and nutrition and nursing	1 month 92.8% (65/70) vs. 81.3% (52/64)	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	ASA Grading 1: 65% 2: 25% 3: 10% Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR					
Yi, 2014 Study period: Nov 2005 to Jul 2009; Recruitment period: NR China	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain (mean): NR Duration of symptoms <6 weeks: NR Severity of fracture: NR	N=200 Mean Age (SD): NR* Female: NR* Race/Ethnicity: NR*	Kyphoplasty (n=79) KP: PMMA (volume NR) with inflatable bone tamps via transpedicular approach	Conservative treatment (n=121) CC: pain medication, bed rest, a solf bi-valved body brace, and physiotherapy	Mean 49.4 months 100% (79/79) vs. 100% (121/121)	Funding NR No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral bodies treated (mean): NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention: None					
Liu, 2019 Study period: NR Recruitment period: Jan 2016 to Jun 2017 China	Fracture type: Osteoporotic VCF Fracture age (SD): NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR	N=116 Mean Age (SD): 65.6 (NR) years Female: Unclear[†] Race/Ethnicity: NR	Kyphoplasty (n=58) Balloon kyphoplasty using bipedicular approach; type of bone cement not specified (volume NR). Length of time NR	Conservative treatment (n=58) Analgesia using drugs, physical treatment, and fixation with waist orthosis, and maintained in bed for 3 months	NR	Funding NR No COIs

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	One or more previous vertebral fractures: NR Fracture appearance: NR Crossover intervention: None					
Wardlaw, 2009, Boonen, 2011, Van Meirhaeghe, 2013 Study period February 2004 to December 2005, recruitment period Austria, Belgium, France, Germany, Italy, UK, USA, Sweden	Fracture type: Osteoporotic VCF, Cancer‡ Duration of back pain (mean): <3 months§ Mean fracture age: 6 weeks Patients with symptom onset <4 months: 100%§ Severity of fracture: 70% Grade ≥2. 29% Grade 3 on Genant assessment Mean number of vertebral bodies treated: NR	N=300 Mean age (SD): 73.2 (NR) years Female: 77% Race/ethnicity: NR	Kyphoplasty (n=149) KP with PMMA (Medtronic; volume NR) by a percutaneous, bilateral, transpedicular, or extrapedicular approach. All patients received analgesics, bed rest, back braces, physiotherapy, rehabilitation programs, and walking aids according to standard practices of participating hospitals. Calcium and Vitamin D supplements and antiresorptive or anabolic agents given as needed.	Non-surgical care (n=151) Details NR.	1 month 92.6% (138/149) vs. 84.7% (128/151) 3 months 89.9% (134/149) vs. 77.5% (117/151) 6 months 87.9% (131/149) vs. 76.2% (115/151) 12 months 83.2% (124/149) vs. 73.5% (111/151) 24 months	Medtronic Authors report receiving honoraria for consulting, research funding, research support, being employed by the funder, and owning stock and stock options.

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	One or more previous vertebral fractures: NR Fracture appearance: NR Fractures treated (KP only) 0 Fractures: 7%** 1 fracture: 67% 2 fractures: 19% 3 fractures: 7% Crossover interventions: NR				80.5% (120/149) vs. 74.2% (112/151)	
KP versus Other Surgical Procedures						
Werner, 2013 Study period NR; recruitment period: NR Switzerland	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain: NR Patients with symptom onset <6 weeks: NR Patients with symptom onset <4 months: NR	N=65 Mean age (SD): 70 (NR) years Female: 61.5% Race/ethnicity: NR	Kyphoplasty (n=32) KP with KyphX HV-R using transpedicular approach with use of Jamshidi needles and working cannulas, cement volume NR	Vertebral body stenting (n=33) VBS with Verecem V+ Cement Kit	Post-tx (timing NR) 100% (32/32) vs. 100% (33/33) Total: 100% (65/65 patients)	No funding COI NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Severity of fracture: NR Mean number of vertebral bodies treated: 1.5 levels One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR					

ASA = American Society of Anesthesiologists; COI = conflict of interest; KP = kyphoplasty; NA = not applicable; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; VCF = vertebral compression fracture.

* Study reports patients receiving VP vs. KP vs. conservative care. Authors do not report by group, only for the whole group, and then those that received surgery once, surgery twice, and then conservative therapy once and twice respectively.

† Authors report 28 males and 39 females in the KP group (n=67) and 29 males and 29 females in the control care group. These do not add up to the total sample size.

‡ Only 1% of patients had cancer as the primary cause of VCF.

§ Inclusion criteria.

** Ten kyphoplasty patients did not receive surgery, but were still included in the ITT analysis.

Appendix Table F3. Patient Characteristics of Studies Comparing Kyphoplasty to Usual Care in Patients with Fractures due to Malignancies (Berenson, 2011)

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Kyphoplasty versus Usual Care						
Berenson, 2011 Study period NR, recruitment period May 16 2005 to March 11 2008 Australia, Canada, Europe, USA	Fracture type: Fractures due to malignancies Median Symptomatic Fracture age: 3.5 months Duration of back pain (mean): NR Mean duration of symptoms: NR Patients with symptom onset <4 months: NR Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR 1 fracture: 35% vs. 44% 2 fractures: 26% vs. 33% 3 fractures: 38% vs. 23%	N=129[*] Mean age (SD): 63.9 (NR) years Female: 58% White: 91% vs. 85% Black: 3% vs. 11% Hispanic/ Latino: 1% vs. 0% Asian: 1% vs. 0% Other: 3% vs. 2%	kyphoplasty (n=68) KP using PMMA (Medtronic; volume NR)) by percutaneous, bilateral, transpedicular, or extrapedicular method. All patients could receive analgesics, bed rest, bracing, physiotherapy, rehabilitation programs, walking aids, radiation treatment, and other antitumor therapy at the discretion of the treating physician. Patients with concurrent osteoporosis or bone metastasis could also receive treatment with calcium, vitamin D supplements, and antiresorptive or anabolic agents as necessary.	Non-surgical fracture management (n=61) Protocol not described in detail. All patients could receive analgesics, bed rest, bracing, physiotherapy, rehabilitation programs, walking aids, radiation treatment, and other antitumor therapy at the discretion of the treating physician. Patients with concurrent osteoporosis or bone metastasis could also receive treatment with calcium, vitamin D supplements, and antiresorptive or anabolic agents as necessary. Control group patients	1 month 95.6% (65/68) vs. 85.2% (52/61) After crossover[‡] 3 months KP: 83.8% (57/68) control: 14/61 (23.0%) Crossover: 51.6% (33/64) (34 crossed over, 1 dropped out) 6 months KP: KP: 72.1% (49/68) CMT: 14.8% (9/61) Crossover: 33 (3 new	Medtronic Spine LLC. Authors report receiving honoraria, consulting fees, research funding, employment, owning stock, stock options, and providing expert testimony.

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Previous treatments Radiation (spine): 24% vs. 18% Radiation (bone): 10% vs. 23% Surgery: 50% vs. 52% Chemotherapy/hormonal: 66% vs. 67% Steroids: 29% vs. 41% Status of cancer at baseline† No evidence: 15% vs. 16% Remission: 6% vs. 11% Stable: 40% vs. 36% Progressive: 38% vs. 34% Crossover interventions: Crossover was offered to control group after 1-month follow-up.			were offered kyphoplasty after the 1-month follow-up.	crossover, 3 deaths) 12 months KP: 58.8% (40/68) CMT: 9.8% (6/61) Crossover: 28 (1 new crossover, 6 lost to follow-up)	

COI = conflict of interest; KP = kyphoplasty; NA = not applicable; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; VCF = vertebral compression fracture.

* 134 were enrolled (64 vs. 70), but 2 withdrew and 3 did not reach 1 month in the KP group and 3 withdrew before assignment and 9 did not reach 1 month assessment in the non-surgical management group.

† Data were unknown for one patient in each group.

‡ Berenson 2011 allowed patients receiving non-surgical management to receive kyphoplasty following 1 month. For the purpose of this report, we do not report on efficacy results beyond 1 month. Safety results are included for all patients that remained in their initially randomized group; we do not report safety outcomes for crossover patients.

APPENDIX G. Outcome Data Abstraction of Randomized Control Trials

Appendix Table G1. Efficacy Results of Studies Comparing Vertebroplasty to Other Treatments in Patients with Fractures due to Osteoporosis

Author (year)	F/U	Function	Pain	Quality of Life	Other
Vertebroplasty versus Sham					
Carli, 2023	1 day 1 week 1 month 3 months 6 months 12 months	VP vs. sham *ITT analysis* **RDQ (0-100), mean (95% CI)** Baseline (n=40 vs. n=40) 64.7 (58.5 to 70.9) vs. 63.8 (57.6 to 70.0) 1 week (n=40 vs. n=40) 51.3 (48.9 to 53.6) vs. 52.7 (50.4 to 55.0) 1 month (n=40 vs. n=40) 44.6 (38.2 to 51.1) vs. 52.3 (45.8 to 58.8) 3 months (n=40 vs. n=40) 42.6 (35.8 to 49.4) vs. 52.8 (46.2 to 59.5) 6 months (n=40 vs. n=40) 45.2 (37.7 to 52.6) vs. 48.7 (41.4 to 56.0) 12 months (n=40 vs. n=40) 42.0 (34.8 to 49.2) vs. 49.0 (41.7 to 56.3) Adjusted MD from baseline 7.1 (95% CI -3.3 to 17.5)	VP vs. sham *ITT analysis* **VAS (0-10), mean (95% CI)** Baseline (n=40 vs. n=40) 7.6 (7.0 to 8.2) vs. 7.3 (6.9 to 7.8) 1 day (n=40 vs. n=40) 5.1 (4.3 to 5.9) vs. 4.7 (3.9 to 5.5) 1 week (n=40 vs. n=40) 4.5 (3.8 to 5.2) vs. 5.0 (4.3 to 5.8) 1 month (n=40 vs. n=40) 4.0 (3.3 to 4.8) vs. 4.9 (4.1 to 5.7) 3 months (n=40 vs. n=40) 3.5 (2.7 to 4.4) vs. 4.9 (4.1 to 5.7) 6 months (n=40 vs. n=40) 3.9 (3.1 to 4.7) vs. 4.9 (4.1 to 5.6) 12 months (n=40 vs. n=40) 3.9 (3.1 to 4.8) vs. 5.1 (4.3 to 6.0)	VP vs. sham *ITT analysis* **QUALEFFO (0-100), mean (95% CI)** Baseline (n=40 vs. n=40) 56.3 (53.1 to 59.5) vs. 55.3 (52.1 to 58.4) 1 week (n=40 vs. n=40) 51.3 (48.9 to 53.6) vs. 52.7 (50.4 to 55.0) 1 month (n=40 vs. n=40) 48.6 (46.2 to 51.0) vs. 51.5 (49.1 to 53.1) 3 months (n=40 vs. n=40) 48.0 (44.7 to 51.3) vs. 52.1 (48.9 to 55.4) 6 months (n=40 vs. n=40) 48.6 (45.9 to 51.4) vs. 51.4 (48.7 to 54.2) 12 months (n=40 vs. n=40) 47.9 (44.9 to 50.9) vs. 53.1 (50.2 to 56.0) Adjusted MD from baseline 5.2 (95% CI 0.9 to 9.4)	VP vs. sham *ITT analysis* **Analgesic use, % (n/N)** Baseline Strong opioids: 40.0% (16/40) vs. 22.5% (9/40) Weak opioids: 25.0% (10/40) vs. 12.5% (5/40) Nonopioids: 80.0% (32/40) vs. 70.0% (28/40) 12 months Strong opioids: 17.1% (6/35) vs. 14.3% (5/35) Weak opioids: 5.9% (2/34) vs. 8.6% (3/35) Nonopioids: 54.3% (19/35) vs. 60.0% (21/35) **Progressive Height Loss** 12 months: 0% (0/40) vs. 10% (4/40)

Author (year)	F/U	Function	Pain	Quality of Life	Other
			Adjusted MD from baseline 1.3 (95% CI 0.1 to 2.6)		
Clark, 2016	3 days 2 weeks 1 month 3 months 6 months	VP vs. Sham **Reduction in RDQ (0-24), mean reduction (SD)** Baseline score (n=61 vs. n=59) 19.5 (3.5) vs. 19.8 (3.7) 3 days (n=58 vs. n=55) -4.5 (6.2) vs. -2.9 (4.4) MD -1.6 (95% CI -3.6 to 0.4) 2 weeks (n=53 vs. n=56) -5.9 (5.8) vs. -4.1 (6.3) MD -1.8 (95% CI -4.1 to 0.5) 1 month (n=55 vs. n=54) -6.9 (6.0) vs. -4.3 (5.6) MD -2.6 (95% CI -4.8 to -0.4) 3 months (n=53 vs. n=50) -9.6 (7.7) vs. -6.4 (7.0) MD -3.2 (95% CI -6.1 to -0.3) 6 months (n=49 vs. n=51) -11.7 (6.5) vs. -7.4 (6.9) MD -4.2 (95% CI -6.9 to -1.6)	VP vs. Sham **Mean Reduction NRS Pain (0-10), mean reduction (SD)** Baseline Score (n=61 vs. n=59) 8.6 (1.3) vs. 8.6 (1.2) 3 days (n=58 vs. n=55) -3.5 (2.6) vs. -1.8 (2.3) MD -1.8 (95% CI -2.7 to -0.8) 2 weeks (n=55 vs. n=57) -4.2 (2.7) vs. -3.0 (3.0) MD -1.2 (95% CI -2.3 to -0.1) 1 month (n=55 vs. n=57) -4.6 (3.0) vs. -3.2 (2.7) MD -1.4 (95% CI -2.5 to -0.4) 3 months (n=53 vs. n=52) -5.4 (3.5) vs. -4.1 (3.1) MD -1.3 (95% CI -2.6 to 0) 6 months (n=51 vs. n=51) -6.1 (3.3) vs. -4.8 (3.1) MD -1.3 (95% CI -2.6 to 0)	VP vs. Sham **QUALEFFO (0-100), mean (SD)** Baseline (n=61 vs. n=59) 65.4 (11.4) vs. 67.7 (11.2) 2 weeks (n=48 vs. n=54) MD -6 (95% CI -11 to -1) 49 (13) vs. 55 (14) 1 month (n=48 vs. n=52) 49 (17) vs. 52 (15) MD -4 (95% CI -10 to 3) 6 months (n=46 vs. n=48) 38 (15) vs. 45 (16) MD -7 (95% CI -13 to -1) **EQ-5D (0-1)** Baseline (n=61 vs. n=59) 0.60 (0.07) vs. 0.59 (0.06) 3 days (n=58 vs. n=52) 0.69 (0.11) vs. 0.65 (0.09) MD 0.03 (95% CI -0.05 to 0.07) 2 weeks (n=49 vs. n=56) 0.69 (0.10) vs. 0.68 (0.11) MD 0.01 (95% CI -0.03 to 0.06) 1 month (n=47 vs. n=51) 0.75 (0.11) vs. 0.70 (0.11) MD 0.05 (95% CI 0 to 0.09) 3 months (n=51 vs. n=49)	VP vs. Sham **Analgesic use, % (n/N)** Baseline NR 3 days 96.6% (57/59) vs. 98.2% (56/57) 14 days 87.5% (49/56) vs. 91.2% (52/57) 1 month 74.5% (41/55) vs. 87.7% (50/57) 3 months 64.2% (34/53) vs. 83.0% (44/53) 6 months 58.0% (29/50) vs. 76.5% (39/51) **Vertebral Height Loss, mean % (SD)** 6 months: 27% (12) vs. 63% (17) **Opioid use:** p>0.05

Author (year)	F/U	Function	Pain	Quality of Life	Other
			Proportion with NRS <4, % (n/N) Baseline: NR 3 days: 31% (18/51) vs. 9% (5/55) 2 weeks: 44% (24/55) vs. 21% (12/57) 1 month: 51% (28/55) vs. 18% (10/57) 3 months: 55% (29/53) vs. 33% (17/57) 6 months: 69% (35/51) vs. 47% (24/51) **VAS Pain (patient observed) (0-100), mean (SD)** Baseline (n=61 vs. n=59) 81 (18) vs. 82 (15) 2 weeks (n=41 vs. n=47) 39.0 (28.0) vs. 49.0 (28.0) 6 months (n=42 vs. n=46) 23.0 (26.0) vs. 34.0 (27.0)	0.75 (0.12) vs. 0.71 (0.11) MD 0.03 (95% CI -0.01 to 0.08) 6 months (n=47 vs. n=50) 0.80 (0.11) vs. 0.74 (0.12) MD 0.06 (95% CI 0.01 to 0.10)	
Diamond, 2020 Subgroup analysis of Clark, 2016	See Clark, 2016 Above	VP vs. Sham **RDQ (0-24), mean (SD)** Baseline Score (n=46 vs. n=47) 19.7 (2.8) vs. 19.9 (4.1) 3 days (n=44 vs. n=43) 14.0 (6.5) vs. 17.1 (4.2) MD -3.1 (95% CI -5.4 to -0.7) 2 weeks (n=39 vs. n=44) 13.1 (6.2) vs. 16.0 (6.3)	VP vs. Sham **NRS Pain (0-10)** Baseline (n=46 vs. n=47) 8.7 (1.3) vs. 8.6 (1.2) 3 days (n=43 vs. n=43) 4.8 (2.4) vs. 7.2 (2.0) MD -2.4 (95% CI -3.4 to -1.5) 2 weeks (n=41 vs. n=45)	VP vs. Sham **QUALEFFO (0-100), mean (SD)** Baseline Score (n=46 vs. n=47) 67.0 (11.0) vs. 68.8 (11.7) Additional time points NR **EQ-5D (0-1), mean (SD)**	VP vs. Sham **Opioid use:** Not different between groups. Data NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		MD -2.9 (95% CI -5.6 to -0.2) 1 month (n=40 vs. n=43) 12.9 (5.9) vs. 15.4 (5.9) MD -2.5 (95% CI -5.1 to 0.1) 3 months (n=39 vs. n=39) 10.2 (7.5) vs. 13.6 (6.2) MD -3.4 (95% CI -6.5 to -0.3) 6 months (n=37 vs. n=40) 9.0 (6.4) vs. 12.5 (6.5) MD -3.4 (95% CI -6.4 to -0.5)	3.8 (2.6) vs. 5.6 (2.8) MD -1.9 (95% CI -3.0 to -0.7) 1 month (n=40 vs. n=45) 3.7 (2.7) vs. 5.5 (2.5) MD -1.9 (95% CI -3.0 to -0.7) 3 months (n=39 vs. n=41) 3.1 (3.1) vs. 4.5 (3.0) MD -1.4 (95% CI -2.8 to -0.1) 6 months (n=38 vs. n=40) 2.1 (2.6) vs. 3.5 (2.6) MD -1.4 (95% CI -2.6 to -0.3) **Proportion with NRS<4** 3 days: 14/43 (33%) vs. 3/43 (7%) 2 weeks: 21/41 (51%) vs. 9/45 (20%) 1 month: 22/40 (55%) vs. 7/45 (16%) 3 months: 21/39 (54%) vs. 12/41 (29%) 6 months: 28/38 (74%) vs. 19/40 (48%)	Baseline Score (n=46 vs. n=47) 0.59 (0.06) vs. 0.59 (0.06) Additional time points NR	
Firanescu, 2018	1 month 3 months 6 months 12 months	VP vs. Sham *ITT analysis* **RDQ (0-24), mean (95% CI)** Baseline (n=90 vs. n=86)	VP vs. Sham *ITT analysis* **VAS Pain (mean, 95% CI) (0-10)**	VP vs. Sham *ITT analysis* **QUALEFFO (0-100), mean (95% CI)**	VP vs. Sham *ITT analysis* **Opioid use, % (n/N)**

Author (year)	F/U	Function	Pain	Quality of Life	Other
		18.02 (95% CI 16.75 to 19.29) vs. 17.79 (95% CI 16.49 to 19.09) 1 week (n=90 vs. n=86) 14.83 (95% CI 13.55 to 16.10) vs. 14.01 (95% CI 12.71 to 15.31) Adjusted MD 0.81 (95% CI -1.01 to 2.630 1 month (n=90 vs. n=86) 11.86 (95% CI 10.56 to 13.14) vs. 12.98 (95% CI 11.67 to 14.29) Adjusted MD -1.12 (95% CI -2.95 to 0.71) 3 months (n=90 vs. n=86) 10.90 (95% CI 9.62 to 12.20) vs. 11.51 (95% CI 10.18 to 12.84) Adjusted MD -0.60 (95% CI -2.46 to 1.25) 6 months (n=90 vs. n=86) 10.09 (95% CI 8.79 to 11.39) vs. 10.97 (95% CI 9.62 to 12.33) Adjusted MD -0.88 (95% CI -2.76 to 1.00) 12 months (n=90 vs. n=86) 10.31 (95% CI 8.98 to 11.63) vs. 10.32 (95% CI 8.92 to 11.72) Adjusted MD -0.01 (95% CI -1.94 to 1.92) MD from baseline -0.12 (95% CI -1.35 to 1.11)	Baseline (n=90 vs. n=86) 7.72 (95% CI 7.21 to 8.24) vs. 7.92 (95% CI 7.40 to 8.45) 1 day (n=90 vs. n=86) 5.24 (95% CI 4.73 to 5.67) vs. 4.82 (95% CI 4.29 to 5.34) Adjusted MD 0.43 (95% CI -0.31 to 1.17) 1 week (n=90 vs. n=86) 4.38 (95% CI 3.86 to 4.90) vs. 4.27 (95% CI 3.74 to 4.79) Adjusted MD 0.11 (95% CI -0.63 to 0.85) 1 month (n=90 vs. n=86) 3.32 (95% CI 2.80 to 3.84) vs. 3.73 (95% CI 3.20 to 4.26) Adjusted MD -0.41 (95% CI -1.15 to 0.33) 3 months (n=90 vs. n=86) 2.69 (95% CI 2.16 to 3.21) vs. 2.90 (95% CI 2.35 to 3.44) Adjusted MD -0.21 (95% CI 0.96 to 0.54) 6 months (n=90 vs. n=86) 3.02 (95% CI 2.48 to 3.55) vs. 3.41 (95% CI 2.86 to 3.96) Adjusted MD -0.39 (95% CI -1.15 to 0.37)	Baseline (n=90 vs. n=86) 59.73 (95% CI 55.96 to 63.51) vs. 60.70 (95% CI 56.84 to 64.56) 1 week (n=90 vs. n=86) 53.07 (95% CI 49.29 to 56.85) vs. 51.84 (95% CI 47.97 to 55.70) Adjusted MD 1.23 (95% CI -4.17 to 6.64) 1 month (n=90 vs. n=86) 47.77 (95% CI 43.99 to 51.56) vs. 49.32 (95% CI 45.45 to 53.19) Adjusted MD -1.55 (95% CI -6.96 to 3.87) 3 months (n=90 vs. n=86) 44.24 (95% CI 40.44 to 48.04) vs. 44.97 (95% CI 41.07 to 48.87) Adjusted MD -0.73 (95% CI -6.17 to 4.72) 6 months: 43.56 (95% CI 39.73 to 47.38) vs. 42.90 (95% CI 38.95 to 46.84) Adjusted MD 0.66 (95% CI -4.83 to 6.16) 12 months (n=90 vs. n=86) 41.41 (95% CI 37.54 to 45.28) vs. 42.09 (95% CI 38.05 to 46.13) Adjusted MD 0.14 (95% CI -2.76 to 3.04)	Strong (morphine, fentanyl): Baseline 47% (41/90) vs. 29% (25/86) 1 day 34% (30/89) vs. 28% (24/86) 1 week 35% (31/88) vs. 19% (16/85) 1 month 21% (18/86) vs. 22% (19/85) 3 months 20% (17/85) vs. 16% (13/80) 6 months 14% (12/83) vs. 17% (13/78) 12 months 16% (13/79) vs. 16% (11/70) Weak (codeine, tramadol): Baseline 14% (13/90) vs. 20% (17/86) 1 day 8% (7/89) vs. 12% (10/86) 1 week 7% (6/88) vs. 8% (7/85) 1 month

Author (year)	F/U	Function	Pain	Quality of Life	Other
			12 months (n=90 vs. n=86) 2.72 (95% CI 2.18 to 3.26) vs. 3.17 (95% CI 2.60 to 3.75) Adjusted MD -0.45 (95% CI -1.24 to 0.36) Adjusted MD from baseline -0.13 (95% CI -0.66 to 0.41)		7% (6/86) vs. 5% (4/85) 3 months 6% (5/85) vs. 4% (3/80) 6 months 5% (4/83) vs. 5% (4/78) 12 months 2% (2/79) vs. 0% (0/70) Non-opiates Baseline 87% (78/90) vs. 76% (65/86) 1 day 67% (60/89) vs. 54% (46/86) 1 week 76% (67/88) vs. 71% (60/85) 1 month 50% (43/86) vs. 56% (48/85) 3 months 53% (45/85) vs. 52% (42/80) 6 months 45% (37/83) vs. 51% (40/78) 12 months 44% (35/79) vs. 46% (32/70)

Author (year)	F/U	Function	Pain	Quality of Life	Other
					Loss of Disc Height (≥4 mm), % (n/N) 12 months: 7.8% (7/90) vs. 48.8% (39/86)
Firanescu, 2019 Follow-up to Frianescu, 2018	NR	NR	NR	NR	NR
Hansen, 2019	1 week 2 weeks 3 weeks 1 month 5 weeks 6 weeks 7 weeks 2 months 9 weeks 10 weeks 11 weeks 3 months 12 months	NR	VP vs. Sham *ITT analysis* **VAS Pain (Forward Bending) (0-10), mean (reported SE; likely SD)** Baseline (n=22 vs. n=24) 74.68 (4.55) vs. 76.08 (4.35) 1 week (n=22 vs. n=24) 26.80 (4.77) vs. 41.83 (4.45) 2 weeks (n=22 vs. n=24) 28.52 (4.65) vs. 34.83 (4.45) 3 weeks (n=22 vs. n=24) 17.81 (4.65) vs. 28.83 (4.45) 1 month (n=22 vs. n=24) 17.33 (4.65) vs. 26.27 (4.55) 5 weeks (n=22 vs. n=24) 14.33 (4.65) vs. 27.14 (4.55) 6 weeks (n=22 vs. n=24)	VP vs. Sham *ITT analysis* **EQ-5D (0-1), mean (reported SE; likely SD)** Baseline (n=22 vs. n=24) 0.44 (NR) vs. 0.49 (NR) 3 months (n=22 vs. n=24) 0.68 (0.23) vs. 0.71 (0.23) 12 months (n=22 vs. n=24) 0.67 (0.27) vs. 0.74 (0.22) **SF-36 PCS (0-100), mean (reported SE; likely SD)** Baseline (n=22 vs. n=24) 25.12 (6.86) vs. 25.53 (4.64) 3 months (n=22 vs. n=24) 31.44 (10.03) vs. 33.93 (10.56) 12 months (n=22 vs. n=24) 31.90 (9.19) vs. 35.15 (11.92) **SF-36 MCS (0-100), mean (reported SE; likely SD)** Baseline (n=22 vs. n=24)	VP vs. Sham **Opioid use:** Similar at baseline, 12 wks, 12 months. Data NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
			15.27 (4.55) vs. 21.09 (4.45) 7 weeks (n=22 vs. n=24) 13.62 (4.65) vs. 19.26 (4.45) 2 months (n=22 vs. n=24) 13.24 (4.65) vs. 19.77 (4.55) 9 weeks (n=22 vs. n=24) 10.00 (4.55) vs. 15.87 (4.45) 10 weeks (n=22 vs. n=24) 10.50 (4.77) vs. 14.00 (4.65) 11 weeks (n=22 vs. n=24) 9.50 (5.03) vs. 16.48 (4.45) 3 months (n=22 vs. n=24) 16.09 (4.55) vs. 18.70 (4.45) 12 months (n=22 vs. n=24) 28.35 (5.16) vs. 30.67 (4.65)	42.00 (9.75) vs. 44.29 (13.10) 3 months (n=22 vs. n=24) 49.70 (12.02) vs. 51.40 (10.98) 12 months (n=22 vs. n=24) 48.60 (10.75) vs. 53.60 (10.29)	
Kallmes, 2009	3 days 2 weeks 1 month	VP vs. Sham *ITT analysis* **Modified RDQ (0-23), mean (SD)** Baseline (n=68 vs. n=63) 16.6 (3.8) vs. 17.5 (4.1)	VP vs. Sham *ITT analysis* **Pain (0-10 worsening pain), mean (SD)** Baseline (n=68 vs. n=63) 6.9 (2.0) vs. 7.2 (1.8)	VP vs. Sham *ITT analysis* **EQ-5D (0-1), mean (SD)** Baseline (n=68 vs. n=63) 0.57 (0.18) vs. 0.54 (0.23) 1 month (n=67 vs. n=61)	VP vs. Sham **Opioid use, % (n/N)** 1 months 53.7% (36/67) vs. 42.6% (26/61)*

Author (year)	F/U	Function	Pain	Quality of Life	Other
		3 days (n=68 vs. n=63) 13.0 (5.2) vs. 12.5 (5.5) 2 weeks (n=67 vs. n=61) 12.4 (5.8) vs. 12.3 (5.9) 1 month (n=67 vs. n=61) 12.0 (6.3) vs. 13.0 (6.4)	3 days (n=68 vs. n=63) 4.2 (2.8) vs. 3.9 (2.9) 2 weeks (n=67 vs. n=61) 4.3 (2.9) vs. 4.5 (2.8) 1 month (n=67 vs. n=61) 3.9 (2.9) vs. 4.6 (3.0) **Pain improvement >=30%** 1 month: 43/67 (64.2%) vs. 29/61 (47.5%)*	0.70 (0.18) vs. 0.64 (0.20) **SOF-ADL (0-18), mean (SD)** Baseline (n=68 vs. n=63) 10.0 (3.6) vs. 10.3 (2.8) 1 month (n=67 vs. n=61) 7.7 (3.7) vs. 8.2 (3.6) **SF-36 MCS (0-100), mean (SD)** Baseline (n=68 vs. n=63) 44.8 (11.8) vs. 41.5 (14.1) 1 month (n=67 vs. n=61) 46.9 (12.0) vs. 45.6 (14.8) **SF-36 PCS (0-100), mean (SD)** Baseline (n=68 vs. n=63) 25.3 (7.8) vs. 25.3 (7.3) 1 month (n=67 vs. n=61) 29.7 (9.6) vs. 28.7 (8.0)	
Comstock, 2013 See Kallmes, 2009 Above	3 months 6 months 12 months	VP vs. Sham **Modified RDQ (0-23), mean (SD)‡** 3 months (n=64 vs. n=61) 10.5 (1.5) vs. 9.5 (2.5) 6 months (n=63 vs. n=58) 9.0 (2.0) vs. 8.8 (2.5) 12 months (n=63 vs. n=56) 9.5 (2.0) vs. 8.5 (2.8)	VP vs. Sham **Pain (0-10 worsening pain), mean (SD)‡** 3 months (n=64 vs. n=61) 3.4 (0.7) vs. 3.3 (1.1) 6 months (n=63 vs. n=58) 3.5 (0.8) vs. 4.0 (1.2) 12 months (n=63 vs. n=56) 3.2 (0.8) vs. 3.1 (1.2)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
			Pain improvement ≥30%, % (n/N) 12 months: 69.8% (44/63) vs. 44.6% (25/56)		
Buchbinder, 2009	1 week 1 month 3 months 6 months	VP vs. Sham *ITT analysis* **Modified RDQ change from baseline (0-23), mean (SD)** Baseline (n=38 vs. n=40) 17.3 (2.8) vs. 17.3 (2.9) 1 week (n=37 vs. n=37) -1.8 (5.0) vs. -4.0 (6.8) Adjusted MD 2.1 (95% CI -0.9 to 5.2) 1 month (n=35 vs. n=38) -4.4 (6.6) vs. -3.1 (6.8) Adjusted MD -1.7 (95% CI -5.2 to 1.8) 3 months (n=36 vs. n=37) -3.7 (5.4) vs. -5.3 (7.2) Adjusted MD 1.5 (95% CI -1.7 to 4.8) 6 months (n=35 vs. n=36) -4.1 (5.8) vs. -3.7 (5.8) Adjusted MD 0.0 (95% CI -2.9 to 3.0)	VP vs. Sham *ITT analysis* **Overall Pain (0-10 worsening pain) change from baseline, mean (SD)** Baseline (n=38 vs. n=40) 7.4 (2.1) vs. 7.1 (2.3) 1 week (n=37 vs. n=37) -1.5 (2.5) vs. -2.1 (2.8) Adjusted MD 0.7 (95% CI -0.4 to 1.8) 1 month (n=37 vs. n=37) -2.3 (2.6) vs. -1.7 (2.3) Adjusted MD -0.5 (95% CI -1.7 to 0.8) 3 months (n=36 vs. n=37) -2.6 (2.9) vs. -1.9 (3.3) Adjusted MD -0.6 (95% CI -1.8 to 0.7) 6 months (n=35 vs. n=36) -2.4 (3.3) vs. -2.1 (3.3) Adjusted MD -0.1 (95% CI -1.4 to 1.2) **Perceived Pain, % (n/N)** 1 week	VP vs. Sham *ITT analysis* **QUALEFFO change from baseline (0-100), mean (SD)** Baseline (n=38 vs. n=40) 56.9 (13.4) vs. 59.6 (17.1) 1 week (n=37 vs. n=37) 0.5 (7.4) vs. -3.6 (9.2) Adjusted MD 4.0 (95% CI 0.2 to 7.8) 1 month (n=35 vs. n=38) -2.8 (9.3) vs. -2.4 (12.3) Adjusted MD -0.9 (95% CI -6.0 to 4.2) 3 months (n=36 vs. n=37) -6.0 (9.6) vs. -6.1 (13.7) Adjusted MD -0.7 (95% CI -5.7 to 4.4) 6 months (n=35 vs. n=36) -6.4 (13.4) vs. -6.1 (13.4) Adjusted MD -0.6 (95% CI -6.2 to 5.1) **EQ-5D change from baseline (0-1), mean (SD)** Baseline (n=38 vs. n=40) 0.30 (0.32) vs. 0.28 (0.33) 1 week (37 vs. n=37) 0.1 (0.3) vs. 0.1 (0.3)	VP vs. Sham **Discontinued Opioids (of those taking opioids at baseline), % (n/N)** 1 week: 10.0% (3/30) vs. 20.6% (7/34) 1 month: 13.3% (4/30) vs. 26.5% (9/34) 3 months: 36.7% (11/30) vs. 32.4% (11/34) 6 months: 56.7% (17/30) vs. 52.9% (18/34)

Author (year)	F/U	Function	Pain	Quality of Life	Other
			Better: 16.2% (6/37) vs. 35.1% (13/37) No Change: 70.3% (26/37) vs. 62.2% (23/37) Worse: 13.5% (5/37) vs. 2.7% (1/37) 1 month Better: 34.3% (12/35) vs. 24.7% (9/38) No Change: 60.0% (21/35) vs. 52.6% (20/38) Worse: 5.7% (2/35) vs. 23.7% (9/38) 3 months Better: 38.9% (14/36) vs. 32.4% (12/37) No Change: 52.8% (19/36) vs. 48.6% (18/37) Worse: 8.3% (3/36) vs. 18.9% (7/37) 6 months Better: 45.7% (16/35) vs. 41.7% (15/36) No Change: 34.3% (12/35) vs. 44.4% (16/36) Worse: 20.0% (7/35) vs. 13.9% (5/36)	Adjusted MD 0.0 (95% CI -0.2 to 0.1) 1 month (n=35 vs. n=38) 0.1 (0.3) vs. 0.1 (0.3) Adjusted MD 0.0 (95% CI -0.1 to 0.1) 3 months (n=36 vs. n=37) 0.2 (0.3) vs. 0.2 (0.4) Adjusted MD 0.0 (95 CI -0.2 to 0.1) 6 months (n=35 vs. n=36) 0.2 (0.4) vs. 0.2 (0.4) Adjusted MD 0.0 (95% CI -0.2 to 0.1) **AQoL change from baseline, mean (SD)** Baseline (n=38 vs. n=40) 0.33 (0.25) vs. 0.27 (0.26) 1 week (37 vs. n=37) 0.0 (0.2) vs. 0.0 (0.2) Adjusted MD 0.0 (95% CI -0.2 to 0.1) 1 month (n=35 vs. n=38) 0.0 (0.2) vs. 0.1 (0.3) Adjusted MD 0.0 (-0.1 to 0.1) 3 months (n=36 vs. n=37) 0.0 (0.2) vs. 0.1 (0.3) Adjusted MD 0.0 (95% CI -0.1 to 0.1) 6 months (n=35 vs. n=36) 0.0 (0.3) vs. 0.1 (0.3) Adjusted MD 0.1 (95% CI -0.2 to 0.1)	

Author (year)	F/U	Function	Pain	Quality of Life	Other
Staples, 2015 Follow-up to Buchbinder, 2009	NR	NR	NR	NR	NR
Kroon, 2014 Follow-up to Buchbinder, 2009	12 months 24 months	VP vs. Sham *ITT analysis* **RDQ (0-24) change from baseline, mean (SD)** 12 months (n=33 vs. n=34) -2.0 (5.7) vs. -2.6 (6.9) Adjusted MD 0.5 (95% CI -3.2 to 4.3) 24 months (n=29 vs. n=28) -2.6 (7.0) vs. -2.7 (5.6) Adjusted MD -0.3 (95% CI -4.1 to 3.5)	VP vs. Sham *ITT analysis* **Overall Pain (0-10 worsening pain) change from baseline, mean (SD)** 12 months (n=33 vs. n=34) -2.4 (2.7) vs. -1.9 (2.8) Adjusted MD -0.3 (95% CI -1.5 to 0.9) 24 months (n=29 vs. n=28) -3.0 (3.1) vs. -1.9 (3.0) Adjusted MD -1.1 (95% CI -2.4 to 0.3) **Perceived Pain, % (n/N)** 12 months Better: 45.5% (15/33) vs. 44.1% (15/34) No Change: 45.5% (15/33) vs. 50.0% (17/34) Worse: 9.0% (3/33) vs. 5.9% (2/34) 24 months Better: 41.4% (12/29) vs. 35.7% (10/28)	VP vs. Sham *ITT analysis* **QUALEFFO change from baseline (0-100), mean (SD)** 12 months (n=33 vs. n=34) -6.7 (12.2) vs. -8.8 (13.3) Adjusted MD 1.3 -4.3 to 7.0) 24 months (n=29 vs. n=28) -5.9 (10.7) vs. -4.6 (15.0) Adjusted MD -2.1 (95% CI -8.5 to 4.4) **EQ-5D change from baseline (0-1), mean (SD)** 12 months (n=33 vs. n=34) 0.2 (0.4) vs. 0.2 (0.4) Adjusted MD 0.0 (95% CI -0.2 to 0.2) 24 months (n=29 vs. n=28) 0.2 (0.4) vs. 0.2 (0.4) Adjusted MD 0.0 (95% CI -0.2 to 0.2) **AQoL change from baseline, mean (SD)** 12 months (n=33 vs. n=34) 0.1 (0.3) vs. 0.2 (0.3) Adjusted MD -0.1 (95% CI -0.2 to 0.0)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
			No Change: 48.3% (14/29) vs. 42.9% (12/28) Worse: 10.3% (3/29) vs. 21.4% (6/28) **Participants who improved, % (n/N) (n/N)** ≥2.5 units (VAS) 12 months 45% (15/33) vs. 38% (13/34) 24 months 34% (10/29) vs. 50% (14/28) ≥2.5 units (RDQ) 12 months 64% (16/25) vs. 50% (12/24) 24 months 55% (12/22) vs. 58% (11/19) Overall pain (≥30%) 12 months 48% (16/33) vs. 47% (16/34) 24 months 69% (20/29) vs. 61% (17/28) RDQ (≥30%) 12 months 45% (15/33) vs. 57% (20/35) 24 months 50% (15/30) vs. 63% (19/30)	24 months (n=29 vs. n=28) 0.1 (0.3) vs. 0.1 (0.3) Adjusted MD 0.1 (95% CI -0.1 to 0.2)	

Author (year)	F/U	Function	Pain	Quality of Life	Other
Vertebroplasty versus Usual Care					
Blasco, 2012	2 weeks 2 months 6 months 12 months	NR	VP vs. Usual Care *ITT analysis* **VAS pain (0-10), mean (SD)** Baseline (n=64 vs. n=61) 7.21 (2.8) vs. 6.31 (2.7) 2 weeks (n=64 vs. n=61) 5.8 (3.6) vs. 4.7 (3.3) 2 months (n=64 vs. n=61) 4.1 (3.4) vs. 4.8 (3.3) 6 months (n=64 vs. n=61) 4.7 (3.0) vs. 4.2 (2.9) 12 months (n=64 vs. n=61) 4.4 (3.0) vs. 4.2 (2.9) **Vertebral pain ≤4 on VAS, % (n/N)** 12 months 56.1% (23/41) vs. 52.4% (22/42) RR 1.07 (95% CI 0.72 to 1.59)	VP vs. Usual Care *ITT analysis* **Qualeffo-41 Total Score (0-100), mean (SD)** Baseline (n=64 vs. n=61) 65.2 (95% CI 60.9 to 69.6) vs. 59.2 (95% CI 54.9 to 63.4) 2 weeks (n=64 vs. n=61) 62 (18) vs. 57 (18) 2 months (n=64 vs. n=61) 57 (18) vs. 55 (18) 6 months (n=64 vs. n=61) 54 (18) vs. 52 (18) 12 months (n=64 vs. n=61) 54 (18) vs. 52 (18)	VP vs. Usual Care *ITT analysis* **Minor opioid use, % (n/N)** 2 weeks 23.2% (13/56) vs. 32.8% (19/58), RR 0.71 (95% CI 0.39 to 1.29) 2 months 26.9% (14.52) vs. 28.6% (16.56), RR 0.94 (95% CI 0.51 to 1.73) 6 months 16.3% (8/49) vs. 26.9% (14/52), RR 0.61 (95% CI 0.28 to 1.32) 12 months 17.1% (7/41) vs. 23.8% (10/42), RR 0.72 (95% CI 0.30 to 1.70) **Major opioid use, % (n/N)** 2 weeks

Author (year)	F/U	Function	Pain	Quality of Life	Other
					35.7% (20/56) vs. 29.3% (17/58), RR 1.22 (95% CI 0.72 to 2.07) 2 months 30.1% (16/52) vs. 30.4% (17/56), RR 1.01 (95% CI 0.57 to 1.79) 6 months 36.7% (18/49) vs. 32.7% (17/52), RR 1.12 (95% CI 0.66 to 1.92) 12 months 36.6% (15/41) vs. 16.7% (7/42), RR 2.19 (95% CI 0.99 to 4.82) **Height reduction from baseline, mean (SD)** 12 months -0.28 cm (0.15) vs. -0.13 cm (0.17), p>0.05
Yang, 2016	1 day 1 week 1 month 3 months 6 months 12 months	VP vs. Conservative Care **ODI (0-100), mean (SD)** Baseline (n=56 vs. n=51) 80.2 (9.9) vs. 81.5 (9.7) 1 week (n=56 vs. n=51) 62.5 (10) vs. 80 (7) 1 month (n=56 vs. n=51) 47 (10) vs. 71.5 (6.5) 3 months (n=56 vs. n=51) 30.5 (8) vs. 56.5 (8.5)	VP vs. Conservative Care **VAS (0-10), mean (SD)** Baseline (n=56 vs. n=51) 7.5 (1.1) vs. 7.7 (1.1) 1 day (n=56 vs. n=51) 4.3 (1.3) vs. 7.3 (1.2) 1 week (n=56 vs. n=51) 3.4 (1.0) vs. 6.4 (1.3) 1 month (n=56 vs. n=51) 2.4 (0.7) vs. 4.9 (0.9)	VP vs. Conservative Care **QUALEFFO (0-100), mean (SD)** Baseline (n=56 vs. n=51) 78.1 (8.1) vs. 77.5 (8.6) 1 week (n=56 vs. n=51) 65 (6.5) vs. 75 (5.5) 1 month (n=56 vs. n=51) 49.5 (6.0) vs. 66 (5) 3 months (n=56 vs. n=51)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		6 months (n=56 vs. n=51) 29.5 (5.5) vs. 48 (7) 12 months (n=56 vs. n=51) 30 (7) vs. 40 (7)	3 months (n=56 vs. n=51) 2.1 (0.6) vs. 3.9 (0.8) 6 months (n=56 vs. n=51) 2.3 (0.7) vs. 3.6 (0.7) 12 months (n=56 vs. n=51) 2.0 (0.5) vs. 3.3 (0.7)	43 (5.5) vs. 56 (5.5) 6 months (n=56 vs. n=51) 40 (5) vs. 53 (5) 12 months (n=56 vs. n=51) 42.5 (5) vs. 49 (5)	
Leali, 2016	1 day 2 days 6 weeks 3 months 6 months	VP vs. Conservative Care **ODI (0-100), mean (SD)** Baseline (n=NR vs. n=NR) 53.6 (NR) vs. NR[§] Day 1 (n=NR vs. n=NR) 31.7 (NR) vs. NR[§] Clinical results for 6 weeks, 3 months, and 6 months were similar in both groups. Authors report that control group patients had no change at any timepoint.	VP vs. Conservative Care **VAS (0-10), mean (SD)** Baseline (n=NR vs. n=NR) 4.8 (NR) vs. NR[§] Day 1 (n=NR vs. n=NR) 2.3 (NR) vs. NR[§] Clinical results for 6 weeks, 3 months, and 6 months were similar in both groups. Authors report that control group patients had no change at any timepoint.	NR	VP vs. Conservative Care **Discontinued Analgesics, % (n/N)** 2 days: 65.0% (120/200) vs. NR
Chen, 2014	1 week 1 month 3 months 6 months 12 months	VP vs. Conservative Care **ODI (0-100), mean (SD)** Baseline (n=46 vs. n=43) 59.9 (2.2) vs. 57.9 (1.9) 1 week (n=46 vs. n=43) 30.3 (3.2) vs. 44.5 (3.9) 1 month (n=46 vs. n=43) 20.4 (3.1) vs. 35.4 (2.9) 3 months (n=46 vs. n=43) 16.6 (1.6) vs. 30.0 (2.4)	VP vs. Conservative Care **VAS (0-10), mean (SD)** Baseline (n=46 vs. n=43) 6.5 (0.9) vs. 6.4 (0.9) 1 week (n=46 vs. n=43) 3.4 (0.5) vs. 5.0 (0.7) 1 month (n=46 vs. n=43) 2.8 (0.4) vs. 4.0 (0.6) 3 months (n=46 vs. n=43)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		6 months (n=46 vs. n=43) 15.5 (1.1) vs. 31.3 (3.5) 12 months (n=46 vs. n=43) 15.0 (1.3) vs. 32.1 (4.5) **RDQ (0-24), mean (SD)** Baseline (n=46 vs. n=43) 18.6 (1.8) vs. 16.7 (1.3) 1 week (n=46 vs. n=43) 13.2 (1.5) vs. 15.7 (1.6) 1 month (n=46 vs. n=43) 11.7 (1.0) vs. 13.8 (1.5) 3 months (n=46 vs. n=43) 9.9 (1.2) vs. 12.5 (1.0) 6 months (n=46 vs. n=43) 9.3 (0.9) vs. 11.1 (0.9) 12 months (n=46 vs. n=43) 8.1 (0.7) vs. 10.7 (1.1)	2.5 (0.5) vs. 3.9 (0.7) 6 months (n=46 vs. n=43) 2.5 (0.6) vs. 4.0 (0.8) 12 months (n=46 vs. n=43) 2.5 (0.5) vs. 4.1 (0.8) **Complete Pain Relief, % (n/N)**[**] 12 months 84.8% (39/46) vs. 34.9% (15/43) **Receiving Pain Treatment, % (n/N)**[**] Baseline 100% vs. 100% 1 week 37.0% (17/46) vs. 100% (43/43) 1 month 28.3% (13/46) vs. 76.7% (33/43) 3 months 15.2% (7/46) vs. 60.5% (26/43) 6 months 13.0% (6/46) vs. 55.8% (24/43) 12 months 15.2% (7/46) vs. 65.1% (28/43)		

Author (year)	F/U	Function	Pain	Quality of Life	Other
Farrokhi, 2011	1 week 2 months 6 months 12 months 24 months 36 months	VP vs. Conservative Care *ITT Analysis* **ODI (0-100), mean (SD)** Baseline (n=40 vs. n=42) 52.2 (2.4) vs. 50.4 (2.8) 1 week (n=40 vs. n=42) 30.1 (3.0) vs. 44.0 (2.5) 2 months (n=40 vs. n=42) 15.0 (2.2) vs. 30.0 (3.1) 6 months (n=40 vs. n=42) 10.0 (2.0) vs. 21.0 (2.5) 12 months (n=38 vs. n=39) 8.0 (3.2) vs. 20.0 (1.7) 24 months (n=38 vs. n=39) 8.0 (2.2) vs. 20.0 (2.0) 36 months (n=37 vs. n=39) 8.0 (1.7) vs. 22.0 (1.2)	VP vs. Conservative Care *ITT Analysis* **VAS Pain (0-10), mean (SD)** Baseline (n=40 vs. n=42) 8.4 (1.6) vs. 7.2 (1.7) 1 week (n=40 vs. n=42) 3.3 (1.5) vs. 6.4 (2.1) 2 months (n=40 vs. n=42) 3.2 (2.2) vs. 6.1 (2.1) 6 months (n=40 vs. n=42) 2.2 (2.1) vs. 4.1 (1.5) 12 months (n=38 vs. n=39) 2.2 (2.1) vs. 4.1 (1.8) 24 months (n=38 vs. n=39) 2.8 (2.0) vs. 3.7 (2.0) 36 months (n=37 vs. n=39) 1.8 (1.7) vs. 3.7 (2.5)	NR	VP vs. Conservative Care *ITT Analysis* **Vertebral Body Height, mean (SD)** Baseline (n=40 vs. n=42) 2.8 cm (1.5) vs. 2.5 cm (1.3) 1 week (n=40 vs. n=42) 3.2 cm (1.1) vs. 2.0 cm (1.0) 6 months (n=40 vs. n=42) 3.2 cm (1.1) vs. 1.9 cm (1.4) 12 months (n=38 vs. n=39) 3.2 (1.5) vs. 2.0 (1.2) 24 months (n=38 vs. n=39) 3.0 (1.5) vs. 2.1 (1.2) 36 months (n=37 vs. n=39) 3.0 (1.2) vs. 2.0 (1.0)
Klazen, 2010; Klazen, 2010 (2)	1 day 1 week 1 month 3 months 6 months 12 months	VP vs. Conservative Care *ITT Analysis* **RDQ (0-24), mean (SD)** Baseline (n=101 vs. n=101) 18.6 (3.6) vs. 17.2 (4.2) 1 week (n=101 vs. n=101)	VP vs. Conservative Care *ITT Analysis* **VAS (0-10), mean (SD)** Baseline (n=101 vs. n=101) 7.85 (NR) vs. 7.50 (NR)	VP vs. Conservative Care *ITT Analysis* **QUALEFFO (0-100), mean (SD)** Baseline (n=101 vs. n=101) 58.7 (13.5) vs. 54.7 (14.4)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		13.7 (5.4) vs. 15.7 (4.7) 1 month (n=101 vs. n=101) 12.5 (6.3) vs. 14.0 (5.7) 3 months (n=101 vs. n=101) 10.5 (6.8) vs. 12.9 (6.0) 6 months (n=101 vs. n=101) 10.0 (6.6) vs. 11.7 (6.6) 12 months (n=101 vs. n=101) 9.6 (6.8) vs. 11.5 (6.9)	1 day (n=101 vs. n=101) 3.7 (2.4) vs. 6.7 (2.1) 1 week (n=101 vs. n=101) 3.5 (2.5) vs. 5.6 (2.5) 1 month (n=101 vs. n=101) 2.5 (2.5) vs. 4.9 (2.6) 3 months (n=101 vs. n=101) 2.5 (2.7) vs. 3.9 (2.8) 6 months (n=101 vs. n=101) 2.3 (2.7) vs. 3.9 (2.9) 12 months (n=101 vs. n=101) 2.2 (2.7) vs. 3.8 (2.8)	1 week (n=101 vs. n=101) 45.6 (14.5) vs. 49.5 (15.5) 1 month (n=101 vs. n=101) 42.9 (15.8) vs. 47.1 (16.1) 3 months (n=101 vs. n=101) 39.6 (17.1) vs. 44.2 (16.6) 6 months (n=101 vs. n=101) 38.9 (17.8) vs. 42.3 (18.3) 12 months (n=101 vs. n=101) 39.7 (18.3) vs. 42.2 (17.9)	
Rousing, 2009	3 months	VP vs. Conservative Care **Barthel Index (0-20), mean (95% CI)[††]** Baseline (n=12 vs. 15) 17.7 (95% CI 15.6 to 19.8) vs. 17.0 (95% CI 14.2 to 19.8) 3 months (n=11 vs. 16) 19.6 (95% CI 19.0 to 20.3) vs. 18.1 (95% CI 16.8 to 19.4) Dallas Pain Questionnaire (activities of daily living) (0-100), mean (95% CI) Baseline (n=16 vs. n=19) 47.8 (95% CI 22.5 to 73.1) vs. 68.5 (95% CI 47.0 to 90.1) 3 months (n=21 vs. n=21_	VP vs. Conservative Care **VAS Pain (0-10), mean (95% CI)** Baseline (n=19 vs. 17): 7.5 (95% CI 6.6 to 8.4) vs. 8.8 (95% CI 8.2 to 9.3) 3 months (n= 23 vs. 23) 1.8 (95% CI 0.8 to 2.8) vs. 2.6 (95% CI 1.2 to 4.0)	VP vs. Conservative Care **EQ-5D (0-1), mean (95% CI)[††]** Baseline (n=17 vs. 16) 0.356 95% CI (0.196 to 0.516) vs. 0.083 (95% CI -0.151 to 0.317) 3 months (n=15 vs. 17) 0.731 (95% CI 0.653 to 0.809) vs. 0.543 (95% CI 0.387 to 0.699) **SF-36 MCS (0-100), mean (95% CI)[‡‡]** Baseline (n=17 vs. 17) 49.7 (95% CI 43.6 to 55.8) vs. 49.6 (95% CI 41.9 to 57.3) 3 months (n=23 vs. 20)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		47.1 (95% CI 32.9 to 61.4) vs. 68.5 (95% CI 47.0 to 90.1)		48.9 (95% CI 43.8 to 54.0) vs. 46.2 (95% CI 39.2 to 53.2) **SF-36 PCS (0-100), mean (95% CI)** †††‡‡‡ Baseline (n=17 vs. 17) 36.7 (95% CI 30.0 to 43.4) vs. 33.4 (95% CI 26.2 to 40.7) 3 months (n=23 vs. 20) 34.0 (95% CI 30.1 to 37.9) vs. 29.3 (95% CI 24.5 to 34.1)	
Rousing, 2010 Follow-up to Rousing, 2009	12 months	VP vs. Conservative Care **Barthel (0-20), mean (95% CI)**†† 12 months (n=12 vs. 17) 19.8 (95% CI 19.5 to 20.0) vs. 18.5 (95% CI 17.6 to 19.3)	VP vs. Conservative Care **Pain VAS (0-10), mean (95% CI)** 12 months (n=22 vs. 22) 2.0 (95% CI 1.1 to 3.0) vs. 2.9 (95% CI 1.6 to 4.1)	VP vs. Conservative Care **EQ-5D (0-1), mean (95% CI)**†† 12 months (n=14 vs. 18) 0.675 (95% CI 0.576 to 0.775) vs. 0.571 (95% CI 0.448 to 0.694) **SF-36 MCS (0-100), mean (95% CI)**†††‡‡‡ 12 months (n=20 vs. 21) 48.7 (95% CI 42.7 to 54.6) vs. 49.0 (95% CI 43.9 to 54.1) **SF-36 PCS (0-100), mean (95% CI)**†††‡‡‡ 12 months (n=20 vs. 21) 32.1 (95% CI 27.8 to 36.3) vs. 30.5 (95% CI 25.2 to 35.7) Dallas Pain Questionnaire (activities of daily living) (0-100), mean (95% CI)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
				Baseline (n=16 vs. n=19) 47.8 (95% CI 22.5 to 73.1) vs. 68.5 (95% CI 47.0 to 90.1) 12 months (n=21 vs. n=17) 53.0 (95% CI 38.3 to 67.7) vs. 53.6 (95% CI 34.8 to 72.5)	
Voormolen, 2007	2 weeks	VP vs. Conservative Care *ITT Analysis* **RDQ (0-24), mean (range)** Baseline (n=18 vs. n=16) 15.7 (8 to 22) vs. 17.8 (9 to 24) 2 weeks (n=18 vs. n=16) 13 (3 to 22) vs. 18 (9 to 23)	VP vs. Conservative Care *ITT Analysis* **VAS Pain (0-10), mean (range)** Baseline (n=18 vs. n=16) 7.1 (5 to 9) vs. 7.6 (5 to 10) 1 day (n=18 vs. n=16) 4.7 (1 to 8) vs. 7.1 (5 to 10) 2 weeks (n=18 vs. n=16) 4.9 (0 to 10) vs. 6.4 (3 to 9)	VP vs. Conservative Care *ITT Analysis* **QUALEFFO (0-100), mean (SD)** Baseline (n=18 vs. n=16) 60 (37 to 86) vs. 67 (38 to 86) 2 weeks (n=18 vs. n=16) 53 (28 to 79) vs. 67 (40 to 88)	VP vs. Conservative Care **Analgesic Use (0-3), mean (range)** Baseline: 1.9 (0 to 3) vs. 1.7 (0 to 3) 1 day: 1.1 (0 to 3) vs. 2.5 (1 to 3) 2 weeks: 1.2 (0 to 3) vs. 2.6 (2 to 3)
Yi, 2014	NR	NR	NR	NR	NR
Vertebroplasty versus Nerve Block					
Tan, 2023	1 week 1 month 2 months	VP vs. Minimally Invasive Surgeries *ITT analysis*	VP vs. Nerve Block *ITT analysis*	VP vs. Nerve Block *ITT analysis*	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		RDQ (0-24), median (IQR) Baseline (n=14 vs. n=13) 20.6 (16 to 22) vs. 19 (17 to 21) 1 week (n=14 vs. n=13) 19 (16 to 21) vs. 18 (15 to 20) 1 month (n=12 vs. n=12) 18 (15.5 to 20) vs. 17 (13 to 21.5) 2 months (n=11 vs. n=10) 12.5 (7 to 13.5) vs. 9 (6 to 12) Authors report no significance at any timepoints **NEADL (0-66), median (IQR)** Baseline (n=14 vs. n=13) 14 (10 to 16) vs. 11 (7 to 19) 1 week (n=14 vs. n=13) 15 (11 to 22) vs. 11 (6 to 16) 1 month (n=12 vs. n=12) 8 (4.5 to 11) vs. 9 (6 to 14.5) 2 months (n=11 vs. n=10) 13.5 (7 to 17.5) vs. 9 (4 to 9) Authors report no significance at any timepoints	**NRS (0-11), median (IQR)** Baseline (n=14 vs. n=13) 9 (8 to 10) vs. 10 (9 to 10) 1 week (n=14 vs. n=13) 6.5 (5 to 8) vs. 7 (4 to 8) 1 month (n=12 vs. n=12) 7 (5.5 to 8) vs. 7.5 (3.5 to 8) 2 months (n=11 vs. n=10) 6 (4 to 7) vs. 3 (2 to 5) Authors report no significance at any timepoints	**EQ-5D (Unclear), median (IQR)**[§§] Baseline (n=14 vs. n=13) 44496.5 (42343 to 53351) vs. 43541 (43441 to 45553) 1 week (n=14 vs. n=13) 32381.5 (22322 to 33532) vs. 33432 (22532 to 44532) 1 month (n=12 vs. n=12) 32987 (22271 to 44546.5) vs. 27826.5 (12217 to 43937) 2 months (n=11 vs. n=10) 27276.5 (22221 to 33926.5) vs. 31441 (22511 to 33311) Authors report no significance at any timepoints	
Wang, 2016	1 day 1 week 1 month 3 months 6 months 12 months	VP vs. Facet block **ODI (0-100), mean (SD)** Baseline (n=100 vs. 106) 46.03 (2.13) vs. 46.46 (1.86) 1 day (n=100 vs. 106) 34.64 (2.57) vs. 42.99 (3.35) MD -8.35 (95% CI -9.15 to -7.55) 1 week (n=100 vs. 106)	VP vs. Facet block **VAS pain score (0-10), mean (SD)** Baseline (n=100 vs. 106) 7.65 (1.11) vs. 7.76 (1.06) 1 day (n=100 vs. 106)	VP vs. Facet block **SF-36 PCS (0-100), mean (SD)** Baseline (n=100 vs. 106) 36.42 (1.55) vs. 36.74 (1.31) 1 month (n=100 vs. n=106) 37.06 (1.64) vs. 36.98 (2.28) MD 0.08 (95% CI –0.46 to 0.62)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		32.37 (1.71) vs. 40.16 (2.29) MD -7.79 (95% CI -8.34 to -7.24) 1 month (n=100 vs. 106) 30.71 (1.73) vs. 30.49 (2.12) MD 0.22 (95% CI -0.30 to 0.74) 3 months (n=100 vs. 106) 24.27 (1.94) vs. 23.82 (2.12) MD 0.45 (95% CI -0.10 to 1.00) 6 months (n=100 vs. 106) 20.16 (2.06) vs. 20.23 (2.16) MD -0.07 (95% CI -0.64 to 0.50) 12 months (n=100 vs. 106) 18.64 (1.77) vs. 18.87 (1.77) MD -0.23 (95% CI -0.71 to 0.25) **RDQ (0-24), mean (SD)** Baseline (n=100 vs. 106) 18.30 (0.99) vs. 18.45 (0.98) 1 day (n=100 vs. 106) 13.35 (1.43) vs. 16.21 (0.96) MD –2.86 (95% CI –3.19 to –2.53) 1 week (n=100 vs. 106) 12.52 (1.25) vs. 15.94 (0.92) MD –3.42 (95% CI –3.82 to –3.12) 1 month (n=100 vs. 106) 12.38 (1.25) vs. 12.24 (1.21) MD 0.15 (95% CI –0.45 to 0.87) 3 months (n=100 vs. 106) 10.99 (1.14) vs. 11.12 (1.19) MD –0.13 (95% CI –0.45 to 0.19) 6 months (n=100 vs. 106) 10.49 (1.14) vs. 10.48 (1.24) MD 0.01 (95% CI –0.32 to 0.34) 12 months (n=100 vs. 106)	1.47 (0.80) vs. 3.19 (0.83) MD –1.72 (95% CI –1.94 to –1.50) 1 week (n=100 vs. 106) 1.62 (0.83) vs. 3.23 (0.82) MD –1.59 (95% CI –1.92 to –0.84) 1 month (n=100 vs. 106) 1.63 (0.88) vs. 1.83 (0.91) MD –0.20 (95% CI –0.44 to 0.04 3 months (n=100 vs. 106) 1.45 (0.77) vs. 1.44 (0.73) MD 0.01 (95% CI –0.20 to 0.22) 6 months (n=100 vs. 106) 1.31 (0.79) vs. 1.28 (0.74) MD 0.03 (95% CI –0.18 to 0.24) 12 months (n=100 vs. 106) 1.19 (0.80) vs. 1.15 (0.75) MD 0.04 (95% CI –0.17 to 0.25)	3 months (n=100 vs. n=106) 38.75 (1.79) vs. 38.32 (2.23) MD 0.43 (95% CI –0.12 to 0.98) 6 months (n=100 vs. n=106) 38.84 (2.14) vs. 38.83 (2.20) MD 0.01 (95% CI –0.58 to 0.60) 12 months (n=100 vs. 106) 39.01 (2.12) vs. 39.04 (2.29) MD –0.03 (95% CI –0.63 to 0.57) **SF-36 MCS (0-100), mean (SD)** Baseline (n=100 vs. 106) 49.97 (2.29) vs. 50.17 (2.35) 1 month (n=100 vs. 106) 48.86 (2.47) vs. 48.28 (2.39) MD 0.58 (95% CI –0.08 to 1.24) 3 months (n=100 vs. 106) 49.10 (2.04) vs. 48.44 (3.35) MD 0.66 (95% CI –0.10 to 1.42) 6 months (n=100 vs. 106) 49.43 (1.70) vs. 49.41 (1.79) MD 0.02 (95% CI. -0.45 to 0.49) 12 months (n=100 vs. 106) 50.26 (1.86) vs. 50.60 (1.98) MD –0.34 (95% CI –0.86 to 0.18)	

Author (year)	F/U	Function	Pain	Quality of Life	Other
		9.42 (1.35) vs. 9.58 (1.31) MD –0.16 (95% CI –0.52 to 0.20)			
Vertebroplasty versus Kyphoplasty					
Wang, 2015	3 months 12 months	VP vs. KP **ODI (0-100), mean (SD)** Baseline (n=53 vs. n=54) 71.22 (10.56) vs. 71.30 (10.22) 3 months (n=53 vs. n=52) 19.74 (6.44) vs. 19.18 (5.89) 12 months (n=50 vs. n=51) 17.04 (6.43) vs. 16.20 (6.70)	VP vs. KP **VAS (0-10), mean (SD)** Baseline (n=53 vs. n=54) 8.10 (1.23) vs. 8.04 (1.13) Post-op (n=53 vs. n=54) 2.59 (0.76) vs. 2.54 (0.81) 3 months (n=53 vs. n=52) 1.24 (0.72) vs. 1.06 (0.68) 12 months (n=50 vs. n=51) 1.24 (0.95) vs. 1.02 (0.80)	NR	VP vs. KP **Mean vertebral height restoration rate, mean % (SD)** 12 months (n=50 vs. n=51) 30.04% (17.38) vs. 42.65% (20.11), p>0.05
Liu, 2010	Post-op 6 months	NR	VP vs. KP **VAS Pain (0-10), mean (SD)** Baseline (n=NR vs. n=NR) 7.9 (0.7) vs. 8.0 (0.8) 3 days (n=NR vs. n=NR) 2.3 (0.5) vs. 2.6 (0.6) 6 months (n=NR vs. n=NR) 2.6 (0.6) vs. 2.6 (0.7)	NR	VP vs. KP **Vertebral body height, mean (SD)** Baseline (n=NR vs. n=NR) 1.01 cm (0.22) vs. 1.13 cm (0.34) Post-op (n=NR vs. n=NR) 1.32 cm (0.26) vs. 2.04 cm (0.41) MD NR, p<0.001

Author (year)	F/U	Function	Pain	Quality of Life	Other
					Kyphotic wedge angle, mean (SD) Baseline (n=NR vs. n=NR) 15.5º (4.2) vs. 17.0º (7.3) Post-op (n=NR vs. n=NR) 12.2º (3.6) vs. 9.0º (5.7)
Liu, 2015 Follow-up to Liu, 2010	12 months 24 months 5 years	NR	VP vs. KP **VAS Pain (0-10), mean (SD)** Baseline (n=NR vs. n=NR) 7.9 (0.7) vs. 8.0 (0.8) 12 months (n=NR vs. n=NR) 2.5 (1.0) vs. 2.7 (0.7) 24 months (n=NR vs. n=NR) 2.6 (1.1) vs. 2.8 (1.3) 5 years (n=NR vs. n=NR) 2.4 (1.5) vs. 3.0 (1.2)	NR	VP vs. KP **Vertebral body height, mean (SD)** Baseline (n=NR vs. n=NR) 1.01 cm (0.22) vs. 1.13 cm (0.34) 12 months (n=NR vs. n=NR) 1.3 cm (0.2) vs. 2.0 cm (0.4) 24 months (n=NR vs. n=NR) 1.9 cm (0.4) vs. 1.3 cm (0.3) 5 years (n=NR vs. n=NR) 1.3 cm (0.2) vs. 1.9 cm (0.5) **Kyphotic wedge angle, mean (SD)** Baseline (n=NR vs. n=NR)

Author (year)	F/U	Function	Pain	Quality of Life	Other
					15.5º (4.2) vs. 17.0º (7.3) 12 months (n=NR vs. n=NR) 12.1º (3.3) vs. 8.7º (5.5) 24 months (n=NR vs. n=NR) 12.2º (3.2) vs. 8.5º (5.6) 5 years (n=NR vs. n=NR) 12.1º (3.3) vs. 8.3º (5.2)
Griffoni, 2020	12 months	VP vs. KP **ODI (0-100), mean (95% CI)** Baseline (n=64 vs. n=49) 54.5 (95% CI 37.4 to 71.7) vs. 55.2 (95% CI 35.5 to 74.9) 12 months (n=NR) 33.6 (95% CI 12.25 to 55.47) vs. 28.3 (95% CI 10.4 to 46.4)	VP vs. KP **VAS pain (0-10), mean (95% CI)** Baseline (n=64 vs. n=49) 7.8 (95% CI 6.0 to 9.8) vs. 8.1 (95% CI 6.7 to 9.9) 12 months (n=NR) 4.7 (95% CI 2.2 to 7.6) vs. 4.4 (95% CI 1.9 to 7.4)	VP vs. KP **EQ-5D (0-100), mean (95% CI)** Baseline (n=64 vs. n=49) 35.8 (95% CI 18.9 to 53.5) vs. 35.7 (95% CI 16.5 to 55.2) 12 months (n=NR) 53.0 (95% CI 29.4 to 76.9) vs. 55.2 (95% CI 34.8 to 76.3)	VP vs. KP **Kyphotic wedge angle, mean (SD)** Baseline (n=64 vs. n=49) 10.8 (5.8) vs. 9.3 (6.5) 12 months (n=NR) 8.8 (5.1) vs. 7.5 (4.3), p=0.202 **Sagittal index, mean (SD)** Baseline (n=64 vs. n=49) 0.63 (0.18) vs. 0.66 (0.20) 12 months (n=NR) 0.71 (0.16) vs. 0.73 (0.18)

Author (year)	F/U	Function	Pain	Quality of Life	Other
Evans, 2016	3 days 1 month 6 months 12 months	VP vs. KP *ITT analysis* **SOF-ADL6 (Scale unclear) mean (SD)** Baseline (n=56 vs. n=59) 17.4 (3.1) vs. 17.7 (4.0) 3 days (n=56 vs. n=59) 12.8 (NR) vs. 14.7 (NR), p=0.49 1 month (n=56 vs. n=59) 11.2 (NR) vs. 11.8 (NR), p=0.90 6 months (n=56 vs. n=59) 13.7 (NR) vs. 12.7 (NR), p=0.89 12 months (n=56 vs. n=59) 13.1 (NR) vs. 12.8 (NR), p=0.82 **RDQ (0-24), mean (SD)** Baseline (n=56 vs. n=59) 16.3 (7.4) vs. 17.3 (6.6) 3 days (n=56 vs. n=59) 11.0 (NR) vs. 11.6 (NR), p=0.86 1 month (n=56 vs. n=59) 9.0 (NR) vs. 9.0 (NR), p=0.99 6 months (n=56 vs. n=59) 7.2 (NR) vs. 8.0), p=93 12 months (n=56 vs. n=59) 6.8 (NR) vs. 7.5 (NR), p=0.85	VP vs. KP *ITT analysis* **VAS pain (0-10), mean (SD)** Baseline (n=56 vs. n=59) 7.9 (2.0) vs. 7.4 (1.9) 3 days (total n=107) 3.7 (NR) vs. 4.0 (NR), p=0.83 1 month (total n=100) 3.7 (NR) vs. 3.4 (NR), p=0.74 6 months (total n=89) 3.2 (NR) vs. 3.7 (NR), p=0.59 12 months (total n=84) 2.3 (NR) vs. 2.9 (NR), p=0.72 <u>Subgroup analyses</u> VAS Pain 12 months Men, p=0.51 Women, p=0.27 Age <75 years, p=0.09 Age ≥75 years, p=0.14 Preoperative average pain score <7, p=0.40 Preoperative average pain score ≥7, p=0.69	VP vs. KP *ITT analysis* **SF-36 PCS (0-100) mean (SD)** Baseline (n=56 vs. n=59) 26.6 (7.6) vs. 26.1 (6.9) 3 days NR 1 month (n=56 vs. n=59) 32 (NR) vs. 31 (NR), p=0.69 6 months (n=56 vs. n=59) 33 (NR) vs. 32 (NR), p=0.80 12 months (n=56 vs. n=59) 33 (NR) vs. 36 (NR), p=0.90 **SF-36 MCS (0-100) mean (SD)** Baseline (n=56 vs. n=59) 42.4 (12.7) vs. 45.4 (14.2) 3 days NR 1 month (n=56 vs. n=59) 49 (NR) vs. 51 (NR), p=0.78 6 months (n=56 vs. n=59) 52 (NR) vs. 53 (NR), p=0.38 12 months (n=56 vs. n=59) 54 (NR) vs. 51 (NR), p=0.92 **EQ-5D (0-100), mean (SD)** Baseline (n=56 vs. n=59) 10.1 (1.6) vs. 10.4 (1.9) 3 days NR 1 month (n=56 vs. n=59)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
				8.5 (NR) vs. 8.1 (NR), p=0.05 6 months (n=56 vs. n=59) 8.3 (NR) vs. 8.3 (NR), p=0.99 12 months (n=56 vs. n=59) 8.2 (NR) vs. 8.0 (NR), p=0.39	
Endres, 2012	6 months	VP vs. Balloon KP vs. Shield KP **ODI (0-100), mean (SD)** Baseline (n=22 vs. n=22 vs. n=22) 68.2 (5.7) vs. 77 (4.2) vs. 75.7 (9.1) 6 months (n=21 vs. n=20 vs. n=18) 53.1 (8.5) vs. 43.1 (19.5) vs. 56.1 (7.6)	VP vs. Balloon KP vs. Shield KP **VAS pain (0-100), mean (SD)** Baseline (n=22 vs. n=22 vs. n=22) 78.2 (9.36) vs. 90 (7.07) vs. 88.16 (15.06) 6 months (n=21 vs. n=20 vs. n=18) 32.4 (14.04) vs. 3.65 (6.36) vs. 40.16 (7.44)	NR	NR
Dohm, 2014	Post-tx 3 months 12 months 24 months	VP vs. KP *Modified ITT analysis* **ODI (0-100), mean (95% CI)** Baseline (n=189 vs. n=191) 57.8 (NR) vs. 59.3 (NR) 1 month (n=NR) 34.3 (NR) vs. 35.8 (NR) 3 months (n=141 vs. n=153) 31.2 (NR) vs. 29.9 (NR) MD from baseline -25.2 (95% CI -28.5 to -22.0) vs. -28.4 (95% CI -31.5 to -25.3) 12 months (n=119 vs. n=138) 27.9 (NR) vs. 29.1 (NR)	VP vs. KP *Modified ITT analysis* **VAS pain (0-10), mean (95% CI)** Baseline (n=190 vs. n=191) 7.6 (NR) vs. 7.8 (NR) 1 week (n=NR) 3.9 (NR) vs. 4.2 (NR) 3 months (n=156 vs. n=158) 3.2 (NR) vs. 3.4 (NR)	VP vs. KP *Modified ITT analysis* **EQ-5D (0-1), mean (95% CI)** Baseline (n=189 vs. n=189) 0.42 (NR) vs. 0.45 (NR) 1 month (n=NR) 0.71 (NR) vs. 0.70 (NR) 3 months (n=140 vs. n=152) 0.75 (NR) vs. 0.74 (NR) MD from baseline 0.32 (95% CI 0.27 to 0.36) vs. 0.29 (95% CI 0.25 to 0.33) 12 months (n=119 vs. n=137) 0.77 (NR) vs. 0.76 (NR)	VP vs. KP *Modified ITT analysis* **Opioid use concomitant with pain relief, % (n/N)** Baseline 74.6% (126/169) vs. 23.9% (34/142) 6 months 73.9% (122/165) vs. 17.6% (25/142)

Author (year)	F/U	Function	Pain	Quality of Life	Other
		MD from baseline -28.0 (95% CI -31.6 to -24.5) vs. -28.8 (95% CI -32.2 to -25.4) 24 months (n=93 vs. n=108) 30.5 (NR) vs. 31.7 (NR) MD from baseline -25.9 (95% CI -30.2 to -21.6) vs. -26.9 (95% CI 30.9 to -22.8)	MD from baseline -4.6 (95% CI -5.1 to -4.1) vs. -4.5 (95% CI -5.0 to -4.0) 12 months (n=133 vs. n=142) 3.5 (NR) vs. 3.2 (NR) MD from baseline -4.3 (95% CI -4.9 to -3.7) vs. -4.5 (95% CI -5.0 to -4.0) 24 months (n=108 vs. n=112) 3.8 (NR) vs. 3.6 (NR) MD from baseline -4.0 (95% CI -4.7 to -3.4) vs. -4.0 (95% CI -4.7 to -3.3)	MD from baseline 0.32 (95% CI 0.28 to 0.37) vs. 0.30 (95% CI 0.25 to 0.35) 24 months (n=94 vs. n=108) 0.75 (NR) vs. 0.72 (NR) MD from baseline 0.31 (95% CI 0.26 to 0.36) vs. 0.28 (95% CI 0.22 to 0.34) **SF36 PCS (0-100), mean (95% CI)** Baseline (n=190 vs. n=189) 27.9 (NR) vs. 27.3 (NR) 1 month (n=NR) 34.4 (NR) vs. 32.7 (NR) 3 months (n=138 vs. n=153) 36.2 (NR) vs. 35.6 (NR) MD from baseline 8.3 (95% CI 6.4 to 10.1) 8.0 (95% CI 6.3 to 9.7) 12 months (n=118 vs. n=138) 37.0 (NR) vs. 35.7 (NR) MD from baseline 9.6 (95% CI 7.6 to 11.6) vs. 8.1 (95% CI 6.4 to 9.9) 24 months (n=92 vs. n=108) 35.0 (NR) vs. 34.6 (NR) MD from baseline 7.5 (95% CI 5.3 to 9.8) vs. 7.6 (95% CI 5.4 to 9.8)	**Kyphotic angulation correction, mean (95% CI)** 1 month: 3.41º (95% CI 2.61 to 4.21) vs. 3.10º (95% CI 2.39 to 3.80) ANCOVA MD 0.21º (95% CI -0.73 to 1.14) 3 months: 2.28º (95% CI 1.37 to 3.19) vs. 1.78º (95% CI 0.98 to 2.58) ANCOVA MD -0.04º (95% CI -1.10 to 1.01) 12 months: 1.51º (95% CI 0.58 to 2.44) vs. 1.97º (95% CI 1.11 to 2.82) ANCOVA MD 0.92º (95% CI -0.14 to 1.98) 24 months: 1.43º (95% CI 0.39 to 2.47) vs. 2.09º (95% CI 0.90 to 3.28) ANCOVA MD 1.42º (95% CI 0.10 to 2.74) **Patients with Perioperative postural reduction, % (n/N)** 12 months 75.1% (142/189) vs. 80.6% (154/191)

Author (year)	F/U	Function	Pain	Quality of Life	Other
Vogl, 2013	3 months 12 months	NR	NR	NR	VP vs. CDKS **Changes in vertebral body height, mean change (SD)** 3 months -9.5% (8.3) (n=10 levels) vs. -4.0% (8.5) (n=11 levels)
Yi, 2014	NR	NR	NR	NR	NR
Wang, 2018	1 month 6 months 12 months	VP vs. KP **ODI (0-100), mean (SD)** Baseline (n=43 vs. n=43) 31.25 (3.34) vs. 30.89 (3.26) 1 month (n=43 vs. n=43) 13.59 (3.37) vs. 11.47 (3.63) 6 months (n=43 vs. n=43) 6.93 (2.36) vs. 5.75 (2.26) 12 months (n=43 vs. n=43) 5.78 (2.37) vs. 4.12 (2.23)	VP vs. KP **JOA low back pain score (0-18), mean (SD)** Baseline (n=43 vs. n=43) 2.78 (0.36) v. 2.82 (0.35) 1 month (n=43 vs. n=43) 3.32 (0.34) vs. 4.57 (0.36) 6 months (n=43 vs. n=43) 3.33 (0.32) vs. 4.57 (0.35) 12 months (n=43 vs. n=43) 4.87 (0.34) vs. 6.25 (0.36) **SF-MPQ (0-78), mean (SD)** Baseline (n=43 vs. n=43)	VP vs. KP **OQOLS - Disease (Scale unclear), mean (SD)** Baseline NR 12 months (n=43 vs. n=43) 52.78 (3.32) vs. 63.82 (3.34), p<0.001 **OQOLS - Physiology (Scale unclear), mean (SD)** Baseline NR 12 months (n=43 vs. n=43) 45.34 (3.36) vs. 53.56 (3.35), p<0.001 **OQOLS - Society (Scale unclear), mean (SD)** Baseline	VP vs. KP **Change in vertebral body height, mean (SD)** Baseline (n=43 vs. n=43) 22.74 (2.36) vs. 22.62 (2.34) 1 month (n=43 vs. n=43) 24.34 (2.38) vs. 25.56 (2.37), p=0.019 3 months (n=43 vs. n=43) 25.89 (2.43) vs. 29.24 (2.47), p<0.001 **Change in cobb angle, mean (SD)** Baseline (n=43 vs. n=43)

Author (year)	F/U	Function	Pain	Quality of Life	Other
			57.16 (3.26) vs. 56.95 (3.15) 1 month (n=43 vs. n=43) 48.06 (3.24) vs. 36.85 (3.16) 6 months (n=43 vs. n=43) 24.63 (3.22) vs. 18.56 (3.18) 12 months (n=43 vs. n=43) 16.28 (3.14) vs. 9.16 (3.15)	NR 12 months (n=43 vs. n=43) 46.79 (3.44) vs. 54.26 (3.56), p<0.001 **OQOLS - Psychology (Scale unclear), mean (SD)** Baseline NR 12 months (n=43 vs. n=43) 21.89 (3.34) vs. 28.24 (3.36), p<0.001 **OQOLS – Degree of Satisfaction (Scale unclear), mean (SD)** Baseline NR 12 months (n=43 vs. n=43) 30.69 (3.25) vs. 38.26 (3.26), p<0.001	22.31 (1.38) vs. 22.25 (1.37) 1 month (n=43 vs. n=43) 18.32 (1.03) vs. 20.76 (1.05), p<0.001 3 months (n=43 vs. n=43) 13.49 (0.84) vs. 17.34 (0.76), p<0.001
Wang, 2023	1 month 3 months	VP vs. KP **ODI (0-45), mean (SD)** Baseline (n=50 vs. n=50) 38.36 (4.19) vs. 38.39 (4.22) 1 month (n=NR vs. n=NR) 26.40 (3.13) vs. 19.51 (3.08) 3 months (n=NR vs. n=NR) 18.69 (1.86) vs. 12.68 (1.62)	VP vs. KP **VAS (0-10), mean (SD)** Baseline (n=50 vs. n=50) 7.35 (1.17) vs. 7.38 (1.20) 1 month (n=NR vs. n=NR) 5.39 (1.11) vs. 4.30 (1.02) 3 months (n=NR vs. n=NR) 3.68 (0.75) vs. 2.57 (0.51)	NR	VP vs. KP **Cobb angle, mean (SD)** Baseline (n=50 vs. n=50) 23.35º (4.49) vs. 23.38º (4.53) 1 month (n=NR vs. n=NR) 19.10º (3.21) vs. 15.41º (3.12) 3 months (n=NR vs. n=NR)

Author (year)	F/U	Function	Pain	Quality of Life	Other
					13.39º (2.21) vs. 8.48º (2.02)

CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; F/U = follow-up; ITT = intention-to-treat; JOA = Japanese Orthopedic Association; KP = kyphoplasty; MCS = mental component score; MD = mean difference; NEADL = Nottingham Extended Activities of Daily Living; NR = not reported; NRS = numerical rating scale; ODI = Oswestry Disability Index; OQOLS = Osteoporosis Quality of Life Scale; PCS = physical component score; QUALEFFO = Quality of Life Questionnaire of the European Foundation for Osteoporosis; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SE = standard error; SF-36 = 36-item Short-Form Questionnaire; SF-MPQ = Osteoporosis Quality of Life Scale; SOF-ADL = Study of Osteoporotic Fractures Activities of Daily Living; VAS = visual analogue scale; VP = vertebroplasty.
* n's back-calculated.
† Numerators back-calculated using percentages given in Figure 2G (VP 54% vs. Sham 43%); OR 1.15 (95% CI 0.98 to 1.35), from a logistic regression model with adjustment for baseline opioid use and study center.
‡ High crossover in control group at 6 (n=30) and 12 (n=33) months
§ Reports patients in usual care group had no change in pain or disability, but data not provided
** Undefined.
†† A PhD project became affiliated with the project in November 2004, and added questionnaires for the EuroQoL (EQ5D), Barthel, Modified mini-mental state examination (MMSE), and 3 physical tests. The available sample size was smaller for these.
‡‡ Only patients with acute fracture answered questionnaires concerning SF-36 and DPQ at inclusion as the authors wanted to register the health state before the fracture, and patients with subacute fractures may not recall the before fracture condition.
§§ Data is reported exactly as is published. Appears to be transformed.

Appendix Table G2. Efficacy Results of Studies comparing Kyphoplasty to Other Treatments in Patients with Fractures due to Osteoporosis

Author (year)	F/U	Function	Pain	Quality of Life	Other
Kyphoplasty versus Usual Care					
Li, 2017	3 days 1 week 1 month 3 months 6 months	KP vs. Conservative treatment **ODI (0-100), mean (SD)** Baseline (n=40 vs. n=40) 42.3 (6.7) vs. 41.3 (6.2), p<0.05 3 days (n=40 vs. n=40)	KP vs. Conservative treatment **VAS pain (0-10), mean (SD)** Baseline (n=40 vs. n=40) 8.60 (0.46) vs. 8.43 (0.60) 3 days (n=40 vs. n=40) 2.10 (0.28) vs. 8.32 (0.37), p<0.05 1 week (n=40 vs. n=40) 3.80 (0.35) vs. 7.20 (0.38), p<0.05 1 month (n=40 vs. n=40) 2.64 (0.22) vs. 3.10 (0.45), p<0.05	NR	KP vs. Conservative treatment **Vertebral body height, mean (SD*)** Baseline (n=40 vs. n=40) 9.8 (2.1) vs. 9.6 (1.9) 1 weeks (n=40 vs. n=40) 14.2 (3.1) vs. 10.4 (2.0), 1 month (n=40 vs. n=40) 14.5 (4.2) vs. 10.5 (3.2), 3 months (n=40 vs. n=40)

Author (year)	F/U	Function	Pain	Quality of Life	Other
		20.2 (5.4) vs. 36.5 (5.1), p<0.05 1 week (n=40 vs. n=40) 18.5 (4.3) vs. 19.7 (3.4), p<0.05 1 month (n=40 vs. n=40) 15.1 (3.6) vs. 18.7 (5.3), p<0.05 3 months (n=40 vs. n=40) 14.2 (4.2) vs. 18.2 (5.0), p<0.05	3 months (n=40 vs. n=40) 1.42 (0.34) vs. 2.38 (0.52), p<0.05 6 months (n=40 vs. n=40) 1.02 (0.24) vs. 1.53 (0.21), p<0.05		14.5 (1.3) vs. 11.5 (2.3), 6 months (n=40 vs. n=40) 14.1 (2.6) vs. 11.2 (2.7), **Cobbs angle, mean (SD)** Baseline (n=40 vs. n=40) 26.31º (2.1) vs. 26.24º (2.4) 3 days (n=40 vs. n=40) 13.20º (1.2) vs. 25.63º (1.27) 1 week (n=40 vs. n=40) 13.45º (1.24) vs. 16.86º (2.12) 1 month (n=40 vs. n=40) 13.80º (1.24) vs. 17.62º (1.29) 3 months (n=40 vs. n=40) 14.31º (1.63) vs. 18.27º (1.55) 6 months (n=40 vs. n=40) 14.47º (1.20) vs. 18.97º (1.46)
Yi, 2014	NR	NR	NR	NR	NR
Liu, 2019	"After Treatment"	KP vs. Non-KP treatment **Barthel Index Daily Life Disturbance (0-100), mean (SD)** Baseline (n=58 vs. n=58) 89.76 (5.27) vs. 89.83 (4.37) After treatment (n=58 vs. n=58) 24.34 (4.53) vs. 31.57 (4.25)	KP vs. Non-KP treatment **VAS (0-10), mean (SD)** Baseline (n=58 vs. n=58) 8.56 (0.39) vs. 8.58 (0.36) After treatment (n=58 vs. n=58) 2.25 (0.21) vs. 4.54 (0.28)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
Wardlaw, 2009[†]	1 week 1 month 3 months 6 months 12 months	KP vs. Non-surgical care *ITT analysis* **RDQ (0-24), Mean (95% CI)** Baseline (n=149 vs. n=151) 16.90 (95% CI 16.00 to 17.80) vs. 17.00 (95% CI 16.10 to 18.00) 1 month (n=149 vs. n=151) 10.90 (95% CI 9.90 to 11.80) vs. 15.10 (95% CI 14.10 to 16.00) 3 months (n=149 vs. n=151) 9.21 (95% CI 8.22 to 10.20) vs. 12.90 (95% CI 11.90 to 13.90) 6 months (n=149 vs. n=151) 8.45 (95% CI 7.44 to 9.45) vs. 11.50 (95% CI 10.40 to 12.50) 12 months (n=149 vs. n=151) 8.60 (95% CI 7.57 to 9.63) vs. 11.50 (95% CI 10.40 to 12.50)	KP vs. Non-surgical care *ITT analysis* **NRS back pain (0-10), Mean (95% CI)** Baseline (n=149 vs. n=151) 6.79 (95% CI 6.42 to 7.16) vs. 6.93 (95% CI 6.56 to 7.30) 1 week (n=149 vs. n=151) 3.60 (95% CI 3.30 to 4.00) vs. 6.00 (95% CI 5.60 to 6.30) 1 month (n=149 vs. n=151) 3.52 (95% CI 3.14 to 3.90) vs. 5.48 (95% CI 5.08 to 5.87) 3 months (n=149 vs. n=151) 2.93 (95% CI 2.55 to 3.32) vs. 4.52 (95% CI 4.11 to 4.93) 6 months (n=149 vs. n=151) 2.73 (95% CI 2.34 to 3.12) vs. 4.35 (95% CI 3.93 to 4.76) 12 months (n=149 vs. n=151) 2.81 (95% CI 2.40 to 3.21) vs. 3.79 (95% CI 3.37 to 4.21)	KP vs. Non-surgical care *ITT analysis* **SF-36 PCS (0-100), Mean (95% CI)** Baseline (n=149 vs. n=151) 26.00 (95% CI 24.40 to 27.50) vs. 25.50 (95% CI 24.00 to 27.10) 1 months (n=149 vs. n=151) 33.40 (95% CI 31.80 to 35.00) vs. 27.50 (95% CI 25.90 to 29.10) 3 months (n=149 vs. n=151) 35.60 (95% CI 34.00 to 37.20) vs. 31.10 (95% CI 29.40 to 32.80) 6 months (n=149 vs. n=151) 36.40 (95% CI 34.80 to 38.00) vs. 32.60 (95% CI 31.00 to 34.30) 12 months (n=149 vs. n=151) 35.90 (95% CI 34.30 to 37.50) vs. 33.80	KP vs. Non-surgical care **Analgesic use**, % (n/N) Non-opioid Baseline 21% (29/140) vs. 25% (36/146) 1 month 25% (28/114) vs. 27% (31/115) 12 months 24% (28/117) vs. 35% (35/101) Combination (non-opioid + opioid) Baseline 58% (81/140) vs. 56% (82/146) 1 month 41% (47/114) vs. 57% (65/115) 12 months 24% (28/117) vs. 29% (29/101) Strong opioid Baseline 16% (22/140) vs. 12% (17/146) 1 month 5% (6/114) vs. 8% (9/115) 12 months 4% (5/117) vs. 5% (5/101)

Author (year)	F/U	Function	Pain	Quality of Life	Other
				(95% CI 32.10 to 35.50) **EQ-5D (0-1), Mean (95% CI)** Baseline (n=149 vs. n=151) 0.16 (95% CI 0.11 to 0.22) vs. 0.17 (95% CI 0.12 to 0.22) 1 month (n=149 vs. n=151) 0.54 (95% CI 0.49 to 0.60) vs. 0.37 (95% CI 0.31 to 0.42) Baseline to 1 month MD 0.17 (95% CI 0.08 to 0.28), p=0.0003 3 months (n=149 vs. n=151) 0.59 (95% CI 0.53 to 0.65) vs. 0.49 (95% CI 0.44 to 0.55) 6 months (n=149 vs. n=151) 0.63 (95% CI 0.57 to 0.68) vs. 0.50 (95% CI 0.45 to 0.56) 12 months (n=149 vs. n=151) 0.61 (95% CI 0.56 to 0.67) vs. 0.51 (95% CI 0.45 to 0.57)	

Author (year)	F/U	Function	Pain	Quality of Life	Other
Boonen, 2011 Follow-up to Wardlaw, 2009	24 months	KP vs. Non-surgical care *ITT analysis* **RDQ (0-24), Mean (95% CI)** 24 months (n=149 vs. n=151) 8.87 (95% CI 7.82 to 9.91) vs. 10.30 (95% CI 9.30 to 11.40) Baseline to 24 months MD -3.01 (95% CI -4.14 to -1.89), p<0.0001	KP vs. Non-surgical care *ITT analysis* **NRS back pain (0-10), Mean (95% CI)** 24 months (n=149 vs. n=151) 2.82 (95% CI 2.41 to 3.22) vs. 3.65 (95% CI 3.23 to 4.07) Baseline to 24 months MD -1.49 (95% CI -1.88 to -1.10), p<0.0001 Treatment-by-visit interaction, p<0.0001	NR	NR
Van Meirhaeghe, 2013 Follow-up to Wardlaw, 2009	24 months	NR	NR	KP vs. Non-surgical care *ITT analysis* **SF-36 PCS (0-100), Mean (95% CI)** 24 months (n=149 vs. n=151) 35.80 (95% CI 34.20 to 37.40) vs. 33.80 (95% CI 32.10 to 35.50) Baseline to 24 months MD 3.24 (95% CI 1.47 to 5.01), p=0.0004	KP vs. Non-surgical care *ITT analysis* **Kyphotic angle, mean (SD)** 24 months MD from baseline 3.13º (NR) vs. 0.82º (NR), p=0.003 **Anterior vertebral body height, mean (SD)** Baseline 62.6% (23.0%) vs. 61.1% (21.4%) 24 months MD 6.7% (95% CI NR) vs. 1.1% (95% CI NR)

Author (year)	F/U	Function	Pain	Quality of Life	Other
				Treatment-by-visit interaction, p=0.004‡ **EQ-5D (0-1), Mean (95% CI)** 24 months (n=149 vs. n=151) 0.61 (95% CI 0.56 to 0.67) vs. 0.53 (95% CI 0.47 to 0.59) Baseline to 24 months MD 0.12 (95% CI 0.06 to 0.18), p=0.0002 Treatment-by-visit interaction, p=0.16	**Midvertebral body height, mean (SD)** Baseline 65.8% (19.5%) vs. 64.5% (19.2%) 24 months 10.0% (14.1%) vs. 8.3% (12.6%)
KP versus Surgical Procedures					
Werner, 2013 Study period NR; recruitment period: NR Switzerland	Post-tx (timing NR)	NR	NR	NR	KP vs. VBS *ITT Analysis* **Mean reduction of the kyphosis, mean (SD)** 4.5º (3.6) vs. 4.7º (4.2) **Radiation exposure time** Data NR, p>0.05

CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; F/U = follow-up; ITT = intention-to-treat; KP = kyphoplasty; MCS = mental component score; MD = mean difference; NR = not reported; NRS = numerical rating scale; ODI = Oswestry Disability Index; PCS = physical component score; RDQ = Roland Morris Disability Questionnaire; SF-36 = 36-Item Short-Form Questionnaire; VAS = visual analogue scale.

* Assumed to be standard deviations.

† All data abstracted from table 3 in Van Meirhaeghe except for data at 1 week, which was estimated from figures in Wardlaw 2009. Sample sizes abstracted from Wardlaw 2009.

‡ There was a significant treatment-by-visit interaction (p=0.004), indicating that the treatment effect is not uniform across follow-up, a result from earlier improvement in the kyphoplasty group. The treatment difference remained statistically significant (3.39 points, 95% CI 1.13–5.64, p=0.003) at 6 months but not at 12 months (1.70 points, 95% CI 0.59 to 3.98, p=0.15) or 24 months (1.68 points, 95% CI 0.63 to 3.99, p=0..15).

Appendix Table G3. Efficacy Results of Studies comparing Kyphoplasty to Other Treatments in Patients with Fractures due Malignancies (Berenson, 2011)

Author (year)	F/U	Function	Pain	Quality of Life	Other
Kyphoplasty versus Usual Care					
Berenson, 2011	1 month Crossover 3 month 6 month 12 month	KP vs. Non-surgical fracture management *Modified ITT* **RDQ (0-24), Mean (SD)** Baseline (n=68 vs. n=60) 17.6 (1.0) vs. 18.2 (0.8) 1 month (n=63 vs. n=50) 9.1 (1.9) vs. 18.0 (1.0) MD -8.9 (95% CI -9.49 to -8.31) MD from baseline -8.3 (95% CI -6.4 to -10.2) vs. 0.1 (95% CI -0.8 to 1.0)* **Minimally clinically important difference on RDQ (≥2 points), % (n/N)** 1 month 80.9% (51/63) vs. 28% (14/50) RR 2.89 (95% CI 1.82 to 4.58) **KPS score (0-100), mean (SD)***	KP vs. Non-surgical fracture management *Modified ITT* **NRS (0-10), Mean (95% CI)*†** Baseline (n=68 vs. n=60) 7.3 (95% CI 6.9 to 7.6) vs. 7.3 (95% CI 6.9 to 7.6) 1 week (n=63 vs. n=54) 3.5 (2.4) vs. 7.0 (1.7) MD -3.50 (95% CI -4.27 to -2.73) 1 month (n=64 vs. n=50) 3.3 (2.9) vs. 6.8 (1.4) MD -3.50 (95% CI -4.37 to -2.63) MD from baseline -3.3 (95% CI -3.6 to -3.0), p<0.0001 **Minimally clinically important difference on NRS (≥1.0 to ≥2.5 points), % (n/N)** 1 week Patients in the kyphoplasty group improved by a mean 3.8 points vs. non-surgical management showed no	KP vs. Non-surgical fracture management *Modified ITT* **SF-36 PCS (0-100), mean (SD)*** Baseline (n=68 vs. n=59) 25.5 (1.5) vs. 25.5 (2.0) 1 month (n=58 vs. n=47) 35 (2.5) vs. 27 (1.5) MD 8.0 (95% CI 7.18 to 8.82) MD from baseline 8.4 (95% CI 7.7 to 9.1) p<0.0001 **Minimally clinically important difference on SF-36 PCS (Improvement ≥3.5 to ≥4.3 points)** 1 month Patients in kyphoplasty group improved 9.4 points vs. non-surgical management showed no significant improvement. Data NR.	KP vs. Non-surgical fracture management *Modified ITT* **Analgesic use** Fewer patients in KP group used analgesics to manage pain relief than in the control group at 1 month (data NR), p=0.0018 **Change in Vertebral body height, mean (SD)** 1 month 2.4 mm (NR) vs. -0.7 mm (NR)

Author (year)	F/U	Function	Pain	Quality of Life	Other
		Baseline (n=68 vs. n=59) 57.0 (0.5) vs. 57.5 (4.5) 1 month (n=63 vs. n=49) 73.0 (4.5) vs. 58.5 (4.5) MD 14.5 (95% CI 12.83 to 16.17) MD from baseline 15.3 (95% CI 13.5 to 17.1), p<0.0001 **Minimally clinically important difference on KPS (≥5 points), % (n/N)** 1 month 65.1% (41/63) vs. 26.5% (13/49) RR 2.45 (95% CI 1.49 to 4.04)	significant improvement. Data NR. 1 month NR	**SF-36 MCS (0-100), mean (SD)** Baseline (n=68 vs. n=59) 38.0 (3.5) vs. 37.5 (3.0) 1 month (n=58 vs. n=47) 46.5 (3.0) vs. 36.5 (3.5) MD 10.0 (95% CI 8.74 to 11.26) MD from baseline 11.1 (95% CI 10.7 to 11.5), p<0.0001	

CI = confidence interval; F/U = follow-up; ITT = intention-to-treat; KP = kyphoplasty; KPS = Karnofsky Performance Status; MCS = mental component score; MD = mean difference; NR = not reported; NRS = numerical rating scale; ODI = Oswestry Disability Index; PCS = physical component score; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SF-36 = 36-Item Short-Form Questionnaire.

* Means not reported for most data. Figures for everything other than NRS is MD from baseline. Includes data past 1 month, but not included here 65% (34/52) of the control group immediately crossed over to KP at 1 month. Final analysis included 54% (28/52) of the original control group sample.

† Standard deviations calculated from 95% confidence intervals.

Appendix Table G4. Safety Results of Studies comparing Vertebroplasty to Other Treatments in Patients with Fractures due to Osteoporosis

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Vertebroplasty versus Sham					
Carli, 2023	12 months	VP vs. Sham **New fractures (remote or adjacent), % (n/N)** 12 months 17.5% (7/40) vs. 15.0% (6/40)	VP only Cement leakage was detected at CT in 70% of treated vertebrae; no specifications whether they were symptomatic or not.	VP vs. Sham **Mortality, % (n/N)** 12 months 0% (0/40) vs. 5% (2/40)*	VP vs. Sham **Serious AEs, % (n/N)** **Spinal cord compression** 12 months 0% (0/40) vs. 2.5 % (1/40)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Clark, 2016	6 months	VP vs. Sham **New Fracture, % (n/N)**[†] 6 months 7.3% (3/41) vs. 4.7% (2/43)	NR	VP vs. Sham **Mortality, % (n/N)**[*] 6 months 4.9% (3/61) vs. 6.8% (4/59)	VP vs. Sham **Serious AEs related to procedure, % (n/N)** 6 months 3.3% (2/61)[‡] vs. 3.4% (2/59)[§]
Diamond, 2020 Subgroup analysis of Clark, 2016	6 months	NR	NR	NR	VP vs. Sham **Serious AEs, % (n/N)**[**] 6 months 4.6% (2/43) vs. 4.6% (2.43)
Firanescu, 2018	12 months	VP vs. Sham **New fractures, % (n/N)** 12 months: 16.7% (15/90) vs. 22.1% (19/86) Twelve participants in each group underwent re-intervention during follow-up for one or more new symptomatic fractures	VP vs. Sham **Cement leakage, % (n/N)** 12 months: 91.3% (105/115 levels) vs. NA Threshold for leakage as anything perceptible on computed tomography Type of Leakage: •Type 3=perivertebral tissue •4= perivertebral veins •5 = pulmonary •6= spinal canal	VP vs. Sham **Mortality, % (n/N)**[*] 8.8% (8/90) vs. 5.8% (5/86)	VP vs. Sham **Adverse events, % (n/N)** Respiratory insufficiency: 12 months 1.1% (1/90) vs. 0% (0/86) Vasovagal reaction: 1.1% (1/90)[††] vs. 0% (0/86)
Firanescu, 2019 Follow-up to Firanescu, 2018	See Franescu 2018	See Franescu 2018[‡‡]	See Franescu 2018[‡‡]	See Franescu 2018[‡‡]	See Franescu 2018[‡‡]

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Hansen, 2019	12 months	NR	VP vs. Sham **Cement leakage, % (n/N)** 12 months 0% (0/24) vs. 0% (0/22)	NR	VP vs. Sham **Infection, % (n/N)** 12 months 0% (0/24) vs. 0% (0/22)
Kallmes, 2009	3 months	NR	NR	None	VP vs. Sham **Severe AEs, % (n/N)** 3 months Thecal sac injury: 1.5% (1/68) vs. 0% (0/63) Tachycardia and rigors requiring hospitalization: 0% (0/68) vs. 1.6% (1/63)
Comstock, 2013 Follow-up to Kallmes, 2009	NR	NR	NR	VP vs. Sham **Mortality, % (n/N)*** 3% (2/68) vs. 5% (3/63)	NR
Buchbinder, 2009	6 months	VP vs. Sham **New fractures, % (n/N)** 1 week: 2.6% (1/38) vs. 0% (0/40) 1 month: 5.3% (2/38) vs. 7.5% (3/40) 3 months: 5.3% (2/38) vs. 7.5% (3/40) 6 months: 7.9% (3/38) vs. 10% (4/40)	VP vs. Sham **Cement Leakage, % (n/N)** 6 months 36.8% (14/38) vs. NA	VP vs. Sham **Mortality, % (n/N)*** 6 months 5.3% (2/38) vs. 2.5% (1/40)	VP vs. Sham **Adverse events, % (n/N)** 6 months Osteomyelitis: 2.6% (1/38) vs. 0% (0/40) Tightness in back/rib cage: 2.6% (1/38) vs. 5.0% (2/40) Pain/burning in thigh/leg: 10.5% (4/38) vs. 5.0% (2/40) Stomach pain: 5.3% (2/38) vs. 2.5% (1/40) Increased pain/cramping around puncture site: 5.3% (2/38) vs. 2.5% (1/40) Chest pain: 7.9% (3/38) vs. 0% (0/40)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Staples, 2015 Follow-up to Buchbinder, 2009	24 months	VP vs. Sham **New fractures, % (n/N)** 12 months: 31.6% (12/38) vs. 27.5% (11/40) 24 months: 36.8% (14/38) vs. 32.5% (13/40)	VP vs. Sham **Cement leakage, % (n/N)** 40.0% (18/45 levels) vs. NA	NR	NR
Kroon, 2014 Follow-up to Buchbinder, 2009	24 months	NR	NR	VP vs. Sham **Mortality, % (n/N)**[*] 24 months 13.2% (5/38) vs. 17.5% (7/40)	VP vs. Sham **Withdrawal, % (n/N)** 2.6% (1/38) vs. 2.5% (1/40)
Vertebroplasty versus Usual Care					
Blasco, 2012	2 weeks 2 months 6 months 12 months	VP vs. Usual Care **New radiological vertebral fracture, % (n/N)** 12 months 26% (17/64) vs. 13% (8/61) OR 2.78 (95% CI 1.02 to 7.62)	NR	VP vs. Usual Care **Mortality, % (n/N)**[*] 12 months 2 weeks: 0% (0/64) vs. 3.3% (2/61) 2 months: 1.6% (1/64) vs. 3.3% (2/61) 6 months: 1.6% (1/64) vs. 6.6% (4/61) 12 months: 4.7% (3/64) vs. 9.8% (6/61)	NR
Yang, 2016	12 months	VP vs. Conservative Care **New fractures, % (n/N)** 12 months 8.9% (5/56) vs. 7.8% (4/51)	VP vs. Conservative Care **Cement leakage, % (n/N)** 12 months 33.8% (22/65 levels) vs. NA	NR	VP vs. Conservative Care **Other complications, % (n/N)**[§§] 12 months 16.1% (9/56) vs. 35.3% (18/51)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Leali, 2016	6 months	VP vs. Conservative Care **New fractures, % (n/N)** 6 weeks 1.6% (3/185) vs. 0% (0/200)	NR	VP vs. Conservative Care **Fracture related mortality, % (n/N)** 6 months 0.5% (1/185) vs. 1.5% (3/200)	VP vs. Conservative Care **Adverse events, % (n/N)** 6 months Transverse process fracture 0.5% (1/185) vs. 0% (0/200) Psoas muscle bleeding: 0.5% (1/185) vs. 0% (0/200)
Chen, 2014	12 months	VP vs. Conservative care **New fractures, % (n/N)** 12 months 8.7% (4/46) vs. 16.3% (7/43)	VP vs. Conservative care **Cement Leakage, % (n/N)** 12 months: 52.2% (36/69 levels) vs. NA	NR	NR
Farrokhi, 2011	3 years	VP vs. Conservative care **New fractures (symptomatic, adjacent level), % (n/N)** 24 months: 2.5% (1/38 vs. 15.4% (6/39)	VP vs. Conservative care **Leakage causing severe lower-extremity pain and weakness, % (n/N)** 1 weeks 2.5% (1/40) vs. NA	VP vs. Conservative care **Mortality, % (n/N)**[*] 12 months: 5.0% (2/40) vs. 2.4% (1/42)	NR
Klazen, 2010 (3)	12 months Mean 22 months	VP vs. Conservative Care **New fractures, % (n/N)** 12 months 16.5% (15/91) vs. 24.7% (21/85)	VP vs. Conservative Care **Cement Leakage, % (n/N)** 12 months 72.4% (97/134 levels) vs. NA Mean 22 months **Perivertebral cement leakage**: 80.0% (64/80 vertebrae) **Discal leakage**: 17.5% (14/80 levels)	VP vs. Conservative Care **Morality, % (n/N)**[*] 12 months 5.0% (5/101) vs. 5.9% (6/101)	VP vs. Conservative Care **Adverse events, % (n/N)** 12 months Urinary tract infection: 1.0% (1/101) vs. 0% (0/101) Cement deposition in segmental pulmonary artery: 1.0% (1/101) vs. 0% (0/101)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
			Perivertebral cement and discal leakage: 10.0% (8/80 vertebrae)		
Rousing, 2009	3 months	VP vs. Conservative Care **New fractures, % (n/N)** 3 months 12.0% (3/25) vs. 4.2% (1/24)	NR	VP vs. Conservative Care **Mortality, % (n/N)*** 3 months 4.0% (1/25) vs. 4.2% (1/24)	NR
Rousing, 2010 Follow-up to Rousing, 2009	12 months	VP vs. Conservative Care **New fractures, % (n/N)** 12 months: 28.0% (7/25) vs. 16.7% (4/24)	NR	VP vs. Conservative Care **Mortality, % (n/N)*** 12 months 7.7% (2/26) vs. 8.3% (2/24)	NR
Voormolen, 2007	2 weeks	VP vs. Conservative Care **New fractures, % (n/N)** 2 weeks 11.1% (2/18) vs. 0% (0/16)	NR	NR	VP vs. Conservative Care One PV procedure-related complication occurred in a patient initially randomized in the control arm but who requested to be treated by PV after 2 weeks of control treatment.
Yi, 2014	Mean 49.4 months	VP vs. Conservative Care **New fractures, % (n/N)** Follow-up 10.0% (9/90) vs. 14.0% (17/121)	VP/KP vs. Conservative Care **Cement leakage, % (n/N)** Follow-up 1.8% (4/169) vs. NA	NR	VP vs. Conservative Care **Major AEs, % (n/N)***** Follow-up 0% (0/90) vs. 0% (0/121)
Vertebroplasty Versus Nerve Block					

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Tan, 2023	1 week 4 weeks 8 weeks	NR	NR	NR	Total only **Overall AEs, % (n/N)** 8 weeks: 48.1% (13/27) VP vs. Nerve block **AEs related to the trial (readmission due to back pain)** 8 weeks: 0% (0/14) vs. 15.4% (2/13)
Wang, 2016	12 months	NR	VP vs. Facet block **Cement leakage, % (n/N)** 12 months 1% (1/100) vs. NA	NR	VP vs. Facet block **Treatment related AEs, % (n/N)** 12 months 1% (1/100)[†††] vs. 0% (0/106)
Vertebroplasty versus Kyphoplasty					
Wang, 2015	3 months 12 months	VP vs. KP **New adjacent vertebral fractures, % (n/N)** 12 months 2% (1/50) vs. 7.8% (4/51)	VP vs. KP **Asymptomatic cement leakage, % (n/N)** 12 months 13.2% (9/68 levels) vs. 30.6% (22/72 levels)	VP vs. KP **Mortality, % (n/N)[*]** 3 months 0% (0/53) vs. 1.9% (1/52) 12 months 2% (1/50) vs. 0% (0/51)	VP vs. KP **Neurological deficit, % (n/N)** 12 months 0% (0/50) vs. 0% (0/51) **Cement embolism, % (n/N)** 12 months 0% (0/50) vs. 1.9% (1/51) **Severe discogenic back pain related to disc leak, % (n/N)** 12 months 0% (0/50) vs. 1.9% (1/51) **Infections, % (n/N)** 12 months 0% (0/50) vs. 0% (0/51)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Liu, 2010	6 months	VP vs. KP **Adjacent segment fractures, % (n/N)** 6 months 0% (0/50) vs. 4% (2/50)	NR	NR	NR
Liu, 2015 Follow-up to Liu, 2010	5 years	VP vs. KP **Adjacent segment fractures, % (n/N)** 5 years‡‡‡ 14%% (7/50) vs.16% (8/50)	NR	NR	NR
Griffoni, 2020	12 months	VP vs. KP **New radiographic fractures, (≥20% height reduction) % (n/N)** 12 months 23.4% (15/64) vs. 4.1% (2/49) **Adjacent level fractures, % (n/N)** 12 months 17.2% (11/64) vs. 2% (1/49)	VP vs. KP **Cement leakage, % (n/N)** 12 months 4.7% (3/64) vs. 4.1% (2/49)	NR	NR
Evans, 2016	NR	NR	NR	NR	NR
Endres, 2012	6 months	VP vs. balloon KP vs. Shield KP	VP vs. balloon KP vs. Shield KP **Cement Leakage, % (n/N)§§§**	Total only **Mortality, % (n/N)** 6 months	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
		Adjacent fractures (≥20% height reduction), % (n/N) 6 months 0% (0/21) vs. 0% (0/20) vs. 0% (0/18)	6 months 0% (0/21) vs. 0% (0/20) vs. 0% (0/18) **Lateral leakages, % (n/N)** 6 months 19% (4/21) vs. 0% (0/20) vs. 0% (0/18) **Disk leakages, % (n/N)** 6 months 19% (4/21) vs. 5% (1/20) vs. 6% (1/18) **Anterior leakages, % (n/N)** 6 months 0% (0/21) vs. 15% (3/20) vs. 0% (0/18)	3% (2/66)****	
Dohm, 2014	Post-tx 3 months 12 months 24 months	VP vs. KP **New (all subsequent) radiographic fractures, % (n/N)** 0-3 months 27.4% (40/146) vs. 23.3% (35/150) 0-12 months 43.5% (57/131) vs. 35.7% (50/140) 0-24 months 57.7% (64/111) vs. 49.1% (54/110)	NR	VP vs. KP **Mortality, % (n/N)*** 24 months 10% (21/205) vs. 8% (16/199)	VP vs. KP **AEs attributable to procedure, device, or anesthesia, % (n/N)††††** 24 months Bone marrow edema: 0.5% 1/190 vs. 0% (0/191) Constipation: 0% (0/190) vs. 0.5% (1/191) Hypersensitivity: 0.5% 1/190 vs. 0% (0/191) Cement embolism: 0.5% (1/190) vs. 0.5% (1/191) Implant site extravasation to the disc: 0.5% (1/190) vs. 0.5% (1/191) Mental status changes postoperatively: 0.5% (1/190) vs. 0% (0/191)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Procedural hypotension: 0% (0/190) vs. 0.5% (1/191) Procedural nausea/vomiting: 0% (0/190) vs. 0.5% (1/191) Procedural pain: 1.6% (3/190) vs. 1.6% (3/191) Spinal fracture: 0% (0/190) vs. 0.5% (1/191) Arthralgia: 0% (0/190) vs. 0.5% (1/191) Back pain: 1.6% (3/190) vs. 1% (2/191) Muscle spasm: 0% (0/190) vs. 0.5% (1/191) Symptomatic vertebral fracture: 1% (2/190) vs. 0.5% (1/191) Hallucination: 0% (0/190) vs. 0.5% (1/191) Chronic obstructive pulmonary disease: 0% (0/190) vs. 0.5% (1/191) Hypoxia: 0.5% (1/190) vs. 0% (0/191) Respiratory failure: 0.5% (1/190) vs. 0% (0/191) Hematoma: 0.5% (1/190) vs. 0% (0/191)
Vogl, 2013	Post-tx 12 months	VP vs. KP **Refracture of the treated level, % (n/N)** Post-tx 0% (0/28) vs. 2% (1/49)‡‡‡‡	VP vs. KP **Cement leakage, % (n/N)** 12 months 12 levels with multiple leaks (total 54 leaks, n=39 levels) vs. 6 levels with multiple leaks (total 42 leaks, n=65 levels), Overall number of	VP vs. KP **Mortality, % (n/N)*** 12 months 7.1% (2/28) vs. 8.1% (4/49)	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
		Adjacent level fractures, % (n/N) Post-tx 3.6% (1/28) vs. 4.1% (2/49)	levels with leaks, p=p=0.0132 Total number of leaks per level, p=0.0012 **Leaks requiring additional intervention, % (n/N)** 12 months 0% (0/28) vs. 0% (0/49)		
Yi, 2014	Mean 49.4 months	VP vs. KP **New fractures, % (n/N)** Follow-up 10.0% (9/90) vs. 6.3% (5/79)	NR§§§§	NR	VP vs. KP **Major***** AEs, % (n/N)** Follow-up 0% (0/90) vs. 0% (0/79)
Wang, 2018	NR	NR	NR	NR	NR
Wang, 2023	3 months	NR	VP vs. KP **Cement leakage, % (n/N)** 3 months 24.0% (12/50) vs. 8.0% (4/50)	NR	NR

AE = adverse event; F/U = follow-up; KP = kyphoplasty; NR = not reported; VP = vertebroplasty;
* Deaths were unrelated to the device/procedure.
† Fracture data only available for patients with radiographs available at baseline and 6 months.
‡ One patient experienced respiratory arrest after sedation, and fully recovered following resuscitation. They went on to receive the procedure without incident. Another patient experienced supracondylar humerus fracture during transfer onto the radiology table. This fracture healed with a plaster cast.
§ Both patients developed spinal cord compression due to interval collapse and retropulsion of the fracture several weeks after enrolment. One patient underwent spinal decompressive surgery with resolution of the neurological deficit, the other became paraplegic.
** Adverse events reported in Diamond 2020 are the same as reported in Clark, 2016.
†† Vasovagal reaction occurred during the procedure, but spontaneously resolved.
‡‡ Franescu, 2019 reports the same data as Franescu, 2018, with no additional results.

§§ Ten complications in nine patients in the VP group: UTI in two, deep vein thrombosis in two, depression in two, and sleep disorder in four. 24 complications in 18 patients in the conservative care group: pneumonia in two patients, UTI in five, deep vein thrombosis/thrombophlebitis requiring treatment in four, depression in five, and sleep disorder in eight. Some patients had multiple AEs.
*** AEs not defined. Authors report that no major adverse events were observed during the follow-up period.
††† Treatment related complication was cement leakage.
‡‡‡ All but 1 patient in the kyphoplasty group experienced fractures within 12 months. Outlier occurred 16 months after treatment.
§§§ Authors report in the results that there was no cement leakage. However, in the discussion they reference that 36%, 13%, and 4% for KP, balloon KP, and shield KP occurred. It is unclear if this is in reference to another study, as the article is reported to have been was translated from German into English in the acknowledgements section.
**** Intervention group was not reported. None were reported as associated with the interventions.
†††† Authors report numerous other AEs that were not attributable to the device, procedure, or anesthesia.
‡‡‡‡ Patient went on to receive follow up vertebroplasty.
§§§§ Authors report this with VP and KP combined. Results in table for VP vs conservative care)
***** Not defined. Authors report that no major adverse events were observed during the follow-up period.

Appendix Table G5. Safety Results of Studies comparing Kyphoplasty to Other Treatments in Patients with Fractures due to Osteoporosis

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Kyphoplasty versus Usual Care					
Li, 2017	NR	NR	NR	NR	NR
Yi, 2014	Mean 49.4 months	KP vs. Conservative Care **New fractures, % (n/N)** Follow-up 6.3% (5/79) vs. 14.0% (17/121)	NR*	NR	KP vs. Conservative Care **Major† AEs, % (n/N)** Follow-up 0% (0/79) vs. 0% (0/121)
Liu, 2019	"After Treatment"	NR	KP vs. Non-KP Treatment **Cement leakage, % (n/N)** After treatment 1.7% (1/58) vs. NA	NR	KP vs. Non-KP Treatment **Adverse events, % (n/N)** After treatment Venous embolism: 0% (0/58) vs. 1.7% (1/58) Decubitus: 0% (0/58) vs. 6.9% (4/58) Infection: 0% (0/58) vs. 6.9% (4/58)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Wardlaw, 2009	1 week 1 month 3 months 6 months 12 months	KP vs. Non-surgical care **New or worsening radiographic vertebral fractures, % (n/N)** 12 months 33% (38/115) vs. 25% (24/95)[‡]	NR	KP vs. Non-surgical care **Mortality, % (n/N)** 12 months 6% (9/149) vs. 4.6% (7/151)	KP vs. Non-surgical care **Overall AEs, % (n/N)** 12 months 87.2% (130/149) vs. 80.8% (122/151) **Withdrew because of AE, % (n/N)** 12 months 0.6% (1/149) vs. 0.6% (1/151) **Overall serious[§] AEs[§], % (n/N)[**]** 12 months 38.9% (58/149) vs. 35.7% (54/151) 12 months Anemia: 2% (3/149) vs. 0.6% (1/151) Back pain: 6.7% (10/149) vs. 6.7% (10/151) Coronary heart disease: 4.7% (7/149) vs. 2.6% (4/151) Arrhythmia: 1.3% (2/149) vs. 1.3% (2/151) Pulmonary embolism: 2% (3/149) vs. 0% (0/151) Stroke: 0.6% (1/149) vs. 0.6% (1/151) Haematoma: 0.6% (1/149)[††] vs. 0% (0/151)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Other (NR): 4% (6/149) vs. 3.3% (5/151) Clostridium infection: 0.6% (1/149) vs. 0.6% (1/151) Sepsis: 0.6% (1/149) vs. 1.3% (2/151) Urinary tract infection: 0.6% (1/149)[††] vs. 1.3% (2/151) Neoplasms/cancer: 4% (6/149) vs. 4% (6/149) Nervous system disorders: 2% (3/149) vs. 1.3% (2/151) Psychiatric disorders: 2% (3/149) vs. 0% (0/151) Pneumonia: 4% (6/149) vs. 3.3% (5/151) Other respiratory disorders: 3.4% (5/149) vs. 0.6% (1/151) Cardiovascular events that resulted in death: 3.4% (5/149) vs. 2% (3/151) Pneumonia that resulted in death: 0% (0/149) vs. 0.6% (1/151) Cancer that resulted in death: 1.3% (2/149) vs. 0.6% (1/151) Other (NR) AEs that resulted in death: 1.3% (2/149) vs. 1.3% (2/151)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Boonen, 2011 follow-up to Wardlaw, 2009	24 months	KP vs. Non-surgical care **New radiographic vertebral fractures, % (n/N)** 24 months 47.5% (56/118) vs. 44.1% (45/102) **New clinical fractures, % (n/N)** 24 months 20.8% (31/149) vs. 17.9% (27/151) **New radiographic adjacent fractures, % (n/N)** 24 months 95.8% (113/118) vs. 16.7% (17/102)	NR	KP vs. Non-surgical care **Mortality, % (n/N)‡‡** 24 months 8.1% (12/149) vs. 7.2% (11/151)	KP vs. Non-surgical care **Further kyphoplasty following new vertebral fractures, % (n/N)** 24 months 46% (12/26) vs. 2% (3/151) **Vertebroplasty following new vertebral fractures, % (n/N)** 24 months NR vs. 2% (3/151) **All AEs, % (n/N)** 24 months 89.9% (134/149) vs. 88.7% (134/151) **Withdrawal due to AE, % (n/N)** 24 months 0.6% (1/149) vs. 0.6% (1/151) **Serious AEs, % (n/N)**** 24 months 49.7% (74/149) vs. 48.3% (73/151) Blood and lymphatic systems disorders (anemia): 2% (3/149) vs. 1.3% (2/151) Angina pectoris: 1.3% (2/149) vs. 3.3% (5/151) Arrhythmia: 1.3% (2/149) vs. 1.9% (3/151)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Myocardial infarction: 3.3% (5/149) vs. 1.9% (3/151) Rectal hemorrhage: 0% (0/149) vs. 1.9% (3/151) Cholelithiasis: 0% (0/149) vs. 3/151) Sepsis/septic shock: 1.3% (2/149) vs. 1.9% (3/151) Urinary tract infection: 1.3% (2/149)[§§] vs. 1.9% (3/151) Hematoma: 0.6% (1/149)[††] vs. 0.6% (1/151) Back pain: 3.4% (5/149) vs. 7.9% (12/151) Spondylitis: 0.6% (1/149)[§§] vs. 0% (0/151) Neoplasms/cancer: 4.6% (7/149) vs. 5.9% (9/151) Psychiatric disorders (depression): 2% (3/149) vs. 0.6% (1/151) Dyspnea: 0.6% (1/149) vs. 2.6% (4/151) Pneumonia: 5.3% (8/149) vs. 3.9% (6/151) Pulmonary embolism: 2.7% (4/149) vs. 0.6% (1/151) Cardiovascular event resulting in death: 3.3% (5/149)[***] vs. 3.3% (5/151)[***] Respiratory event resulting in death: 0.6% (1/149)[***] vs. 1.3% (2/151)[***]

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Cancer resulting in death: 2% (3/149)*** vs. 1.3% (2/151)*** Other (NR) event resulting in death: 2% (3/149)*** vs. 1.3% (2/151)***
Van Meirhaeghe, 2013 follow-up to Wardlaw, 2009	24 months	NR	KP vs. Control **Cement leakage, % (n/N)** 24 months Total sample 27.1% (51/188 vertebral bodies)		NR
Kyphoplasty versus Surgical Procedures					
Werner, 2013 Study period NR; recruitment period: NR Switzerland	Post-tx (Timing NR)	NR	KP vs. VBS **Minor cement leakage, % (n/N)†††** **Post-tx** 12% (6/50 levels) vs. 20% (10/50 levels) **Major cement leakage, % (n/N)†††** 8% (4/50 levels) vs. 10% (5/50 levels)	NR	KP vs. VBS **Revision surgery Post-tx, % (n/N)** 0% (0/50) vs. 0% (0/50) **Total complications, % (n/N)** 22% (11/50 levels) vs. 48% (24/50 levels), p=0.013 **Post-operative sequelae, % (n/N)** Post-tx 0% (0/50) vs. 0% (0/50) **Intraoperative material-related complications**

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Cannula, % (n/N) 0% (0/50 levels) vs. 10% (5/50) **Balloon rupture, % (n/N)** 2% (1/50 levels) vs. 2% (1/50 levels) **Stent related, % (n/N)** NR vs. 6% (3/50 levels)

AE = adverse event; F/U = follow-up; KP = kyphoplasty; NR = not reported; VBS = vertebral body scenting.
* Authors report this with VP and KP combined. Results in table for VP vs conservative care)
† Not defined. Authors report that no major adverse events were observed during the follow-up period.
‡ Plain radiographs were only available for 115 and 95 patients in the VP and TAU groups respectively at 12 months.
§ Defined as any event which resulted in death, life-threatening injury or permanent impairment, or if it required extended hospital stay or intervention to prevent impairment.
** Authors report that some participants might have had multiple serious AEs. But does not elaborate.
†† Event was deemed related to kyphoplasty.
‡‡ Deaths were unrelated to the device/procedure.
§§ One urinary traction infection was considered procedure-related; the same patient had subsequent spondylitis at the treated level that was considered possibly cement-related.
*** All events were deemed unrelated to the procedures.
††† Cement leakage was defined as “minor” (paravertebral) or “major” (into the venous plexus, into the spinal canal, behind the anterior longitudinal ligament, or into the intervertebral disc space).

Appendix Table G6. Safety Results of Studies comparing Kyphoplasty to Other Treatments in Patients with Fractures due to Malignancies (Berenson, 2011)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Kyphoplasty versus Usual Care					
Berenson, 2011	1 month Crossover 3 month 6 month 12 month	KP vs. Non-surgical fracture management **Symptomatic fracture, % (n/N)** 1 month 2.8% (2/70) vs. 4.7% (3/64) RR 0.61 (95% CI 0.11 to 3.53) Between 1 and 12 months 12.8% (9/70) vs. 0% (0/26)[*] RR NC	KP vs. non-surgical management **Cement leakage, % (n/N)** 1 month 1.5% (1/68) vs. NA	KP vs. non-surgical fracture management **Any event resulting in death, % (n/N)[†]** 1 month 2.8% (2/70) vs. 1.5% (1/64) RR 1.82 (95% CI 0.17 to 19.69) Between 1 and 12 months 30% (21/70) vs. 19.2% ((5/26) RR 1.56 (95% CI 0.66 to 3.71)	KP vs. Non-surgical fracture management **Any AEs** **1 month** 37.1% (26/70) vs. 29.7% (19/64) RR 1.25 (95% CI 0.77 to 2.03) **All other AEs, % (n/N)[‡]** All within 1-month Blood and lymphatic disorders: 0% (0/70) vs. 1.6% (1/64) Cardiac disorders: 4.3% (3/70) vs. 4.7% (3/64) Eye disorders: 0% (0/70) vs. 1.6% (1/64) Gastrointestinal disorders 5.7% (4/70) vs. 1.6% (1/64) General disorders: 7.1% (5/70)[§] vs. 4.7% (3/64) Infections: 8.6% (6/70)[**] vs. 3.1% (2/64)[††] Balloon rupture (asymptomatic): 1.4% (1/70)[‡‡] vs. 0% (0/64) Myocardial infarction: 1.4% (1/70)[‡‡] vs. 0% (0/64) Procedure related pain: 1.4% (1/70)[‡‡] vs. 0% (0/64)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Postoperative urine retention: 1.4% (1/70) vs. 0% (0/64) Metabolic/nutritional disorder: 0% (0/70) vs. 1.6% (1/64) Musculoskeletal disorders: 11.4% (8/70) vs. 12.5% (8/64) Neoplasms: 0% (0/70) vs. 1.6% (1/64) Nervous system disorders: 2.8% (2/70) vs. 3.1% (2/64) Psychiatric disorders: 0% (0/70) vs. 1.6% (1/64) Respiratory disorders: 2.8% (2/70) vs. 1.6% (1/64) Vascular disorders: 0% (0/70) vs. 3.1% (2/64) MI resulting in death: 1.4% (1/70) vs. 0% (0/64) Cardiac failure resulting in death: 0% (0/70) vs. 1.6% (1/64) General disorders resulting in death: 1.4% (1/70) vs. 0% (0/64) After 1 month (excluding crossover patients)[§§] **Serious[***] AEs, % (n/N)** 52.8% (37/70) vs. 30.7% (8/26) RR 1.72 (95% CI 0.93 to 3.19)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Blood and lymphatic disorders: 2.8% (2/70) vs. 0% (0/26) Cardiac disorders: 7.1% (5/70) vs. 3.8% (1/26) Gastrointestinal disorders: 5.8% (4/70) vs. 0% (0/26) General disorders (2.9% (2/70) vs. 3.8% (1/26) Hepatobiliary disorders: 0% (0/70) vs. 0% (0/26) All infections: 7.1% (5/70) vs. 7.7% (2/26) Osteomyelitis: 1.4% (1/70) vs. 0% (0/26) Pneumonia: 4.3% (3/70) vs. 3.8% (1/26) Respiratory infection: 0% (0/70) vs. 0% (0/26) Sepsis: 0% (0/70) vs. 3.8% (1/26) Urinary tract infection: 1.4% (1/70) vs. 0% (0/26) Wound infection: 1.4% (1/70) vs. 0% (0/26) Other (NR): 2.8% (2/70) vs. 0% (0/26) All injury or procedural complications: 7.1% (5/70) vs. 3.8% (1/26) Airway complication: 0% (0/70) vs. 0% (0/26) Traumatic chest injury 1.4% (1/70) vs. 0% (0/64) Limb fracture: 4.3% (3/70) vs. 3.8% (1/26)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Nerve injury: 1.4% (1/70) vs. 0% (0/26) Procedure related pain: 0% (0/70) vs. 0% (0/26) Metabolic or nutritional disorders: 1.4% (1/70) vs. 0% (0/26) Musculoskeletal disorders: 14.3% (10/70) vs. 3.8% (1/26) Neoplasms: 25.7% (18/70) vs. 7.7% (2/26) All nervous system disorders: 2.8% (2/70) vs. 3.8% (1/26) Stroke: 1.4% (1/70) vs. 0% (0/26) Paraparesis: 1.4% (1/70) vs. 0% (0/26) Transient ischaemic attack: 0% (0/70) vs. 3.8% (1/26) Renal/urinary disorders: 2.8% (2/70) vs. 0% (0/26) Reproductive/breast disorders: 0% (0/70) vs. 3.8% (1/26) Respiratory disorders: 7.1% (5/70) vs. 0% (0/26) Vascular disorders: 2.8% (2/70) vs. 0% (0/26) Cardiac failure resulting in death: 1.4% (1/70) vs. 3.8% (1/26) General disorders resulting in death: 1.4% (1/70) vs. 3.8% (1/26)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Pneumonia resulting in death: 2.8% (2/70) vs. 3.8% (1/26) Traumatic chest injury resulting in death: 1.4% (1/70) vs. 0% (0/26) Neoplasms resulting in death: 18.5% (13/70) vs. 7.7% (2/26) Respiratory failure resulting in death: 4.3% (3/70) vs. 0% (0/26)

AE = adverse event; CI = confidence interval; F/U = follow-up; KP = kyphoplasty; NA = not applicable; RR = risk ratio;
* 9 patients that crossed over to KP from usual care experienced symptomatic fractures between 1 and 12 months; one patient 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.
† Deaths were unrelated to the device/procedure.
‡ Tables include all patients cumulatively up to 1 month follow-up. KP and control group include all randomized.
§ 2 extravasations to the disc were considered device related but not serious.
** Only one was deemed possibly device related. Others included 2 urinary tract infection and 3 labeled 'other'.
†† No events were deemed related to the device or procedure.
‡‡ Device related.
§§ Included cumulatively until no longer in group. Authors reported all data from 1 month to 12 months together. Control group includes all patients up to the point they experienced an AE, but not after they crossed over.
*** Serious AEs were 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.

APPENDIX H. Demographic Data of Non-Randomized Comparative Studies

Appendix Table H1. Patient Characteristics of Non-Randomized Studies of Interventions (Included for Efficacy and Harms)

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty versus Minimally Invasive Surgeries						
Bae, 2019 retrospective cohort Study period NR, recruitment period March 2011 to February 2016 South Korea	Fracture type: Osteoporotic VCF New fracture*: 15.2% vs. 4.2% Severity of fracture: NR Mean number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: NR	N=164 Mean age (SD): 75.1 (range 39 to 96) years Female: 26.2% Race/ethnicity: NR	Vertebroplasty (n=92) Performed from a bilateral or unilateral transpedicular approach. 5 to 10 ml of PMMA injected under fluoroscopic monitoring. All patients were given absolute bed rest before they began the intervention.	Medial branch block (n=72) Performed with a spinal needle using a bilateral posterior approach under fluoroscopic guidance. Mixture included 20 ml 2% lidocaine, 20 ml saline solution, and 1 ml dexamethasone phosphate; 2 ml of the mixture was injected. All patients were given absolute bed rest before they began the intervention.	24 months: 100% Retrospective analysis	No funding No COI
Vertebroplasty versus Kyphoplasty						
Bae, 2016 Retrospective cohort Study period NR; March 2002 to	Fracture type: Malignancy Fracture age: NR Duration of back pain: NR	N=342† Median age (range): 61 (22 to 89) years Female: 51%	Vertebroplasty (n=238) Details NR. For patients that underwent multiple rounds of VP, only the first round was analyzed.	Kyphoplasty (n=104) Details NR	NR	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
September 2013 South Korea	Duration of symptoms <6 weeks: NR Severity of fracture: Severe VCFs: 65.8% Number of vertebral bodies treated: single: 58% Multiple: 42% Fracture appearance: NR Crossover interventions: None	Race/Ethnicity: NR				
Sacroplasty versus Usual Care						
Sarigul, 2023 Retrospective cohort Study period: 2005 to 2017; Recruitment period NR Turkey	Fracture type: Osteopenia: 20.5% Osteoporosis: 38.9% Malignancy: 2.2% Fracture age: NR Duration of back pain: 11.2 weeks Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR	N=185 Mean age (range) 69.2 (46-93) years Female: 83.2% Race/Ethnicity: NR	Sacroplasty (n=83) PMMA via lateral approach; For patients who had bilateral SIF, the procedure was repeated on the contralateral side	Usual Care (n=102) Analgesic drugs, muscle relaxants, and bed rest for 2 weeks	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Fracture appearance: NR Crossover interventions: None					
Andresen, 2022 Retrospective cohort Study period: NR Recruitment period: Jan 2014 to Jun 2019 Austria	Fracture type: Osteopenia: 20.5% Osteoporosis: 38.9% Malignancy: 2.2% Fracture age: NR Duration of back pain: 11.2 weeks Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR Fracture appearance: Unilateral: 31.2% Bilateral: 68.8% Crossover interventions: During follow-up, 26 patients receiving initially receiving conservative therapy were referred for screw fixation due to increasing fracture extension, increased	N=233 Mean age‡ (range): 69.2 (46-93) years Female‡: 83.2% Race/Ethnicity‡: NR	Sacroplasty (n=119) Included balloon sacroplasty, radiofrequency sacroplasty, vertebrosacroplasty, and cement sacroplasty	Conservative treatment (n=114) Divided by the Study into VAS≤5 (n=50) and VAS>5 (n=100) groups. Included bed rest, adjuvant medicinal pain therapy according to the WHO schedule, and mobilization using a walker or on forearm crutches with pain-adapted weight-bearing	24 months: 82.9% (242/292)	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	pain, and pronounced immobility. An additional 10 patients were referred for sacroplasty. All crossover patients were included in the group they ultimately received for outcome analyses.					
Sacroplasty versus Other Surgical Interventions						
Andresen, 2022 Retrospective cohort Study period: NR Recruitment period: Jan 2014 to Jun 2019 Austria	Fracture type: Osteopenia: 20.5% Osteoporosis: 38.9% Malignancy: 2.2% Fracture age: NR Duration of back pain: 11.2 weeks Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR Fracture appearance: Unilateral: 31.2% Bilateral: 68.8% Crossover interventions: During follow-up, 26 patients receiving initially	N=178 Mean age‡ (range): 69.2 (46-93) years Female‡: 83.2% Race/Ethnicity‡: NR	Sacroplasty (n=119) Included balloon sacroplasty, radiofrequency sacroplasty, vertebrosacroplasty, and cement sacroplasty 8.4%% (10/119) were initially receiving conservative treatment, but were referred to receive sacroplasty due to increasing fracture extension, pain >7, and pronounced immobility, and were included in all screw fixation outcome analyses.	Screw Fixation (n=59) Included iliosacral screw fixation (n=38, 32/38 also had cement augmentation), transsacral screw fixation (n=8), transsacral positioning rod (n=3), percutaneous plate osteosynthesis (n=1), lumbopelvic stabilization (n=8), and internal fixator with additional transiliac screw fixation (n=1) 44.1% (26/59) were initially receiving conservative treatment, but were referred to receive screw fixation due to increasing fracture extension, pain >7, and pronounced	24 months: 82.9% (242/292)	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	receiving conservative therapy were referred for screw fixation due to increasing fracture extension, increased pain, and pronounced immobility. An additional 10 patients were referred for sacroplasty. All crossover patients were included in the group they ultimately received for outcome analyses.			immobility, and were included in all screw fixation outcome analyses.		
Sacroplasty versus Non-surgical Management						
Yang, 2023 Retrospective cohort Study period: NR Recruitment period: Jan 2014 to Jun 2019 Taiwan	Fracture type: SIF, osteoporosis Fracture age: NR Duration of back pain: 6.8 weeks[§] Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR Fracture appearance: NR Crossover interventions: None	N=27 Mean age (SD): 77 (NR) years Female: 100%[**] Race/Ethnicity: NR	Sacroplasty (n=13) Cement injected via long-axis approach All patients did not receive osteoporotic treatment prior to their SIF treatment. They were prescribed 1000mg calcium, 400iu vitamin D supplements after SIF diagnosis. Compliance was ensured at each follow-up by the doctor. Denosumab were prescribed to both group of patients after bone reunion after 6 months of SIF treatment.	Non-surgical Management (n=14) 20mcg teriparatide once per day for 6 months All patients did not receive osteoporotic treatment prior to their SIF treatment. They were prescribed 1000mg calcium, 400iu vitamin D supplements after SIF diagnosis. Compliance was ensured at each follow-up by the doctor. Denosumab were prescribed to both group of patients after bone reunion after 6	84.4% (27/32)	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
				months of SIF treatment.		
Frey, 2017 Prospective cohort Study period: December 2003 to August 2015 Recruitment period: Jan 2004 to Jan 2014 United States	Fracture type: SIF, osteoporosis Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR Fracture appearance: NR Crossover interventions: None	N=244 Mean age (SD): 74 (NR) years Female: 81.2% Race/Ethnicity: NR	Sacroplasty (n=210) Patients received sacroplasty if they initially failed to improve while receiving non-surgical management. The procedures was performed with light conscious sedation and fluoroscopic guidance. Antibiotics were administered before the procedure. Procedures were performed according to different protocols depending on the interventional physician.	Non-surgical Management (n=34) Analgesics including opioid analgesics, corsets, and/or bed rest for at least 3 weeks. Patients were included in the non-surgical management group at baseline if their treatment was successful, otherwise they received sacroplasty.	10 years 55.7 (117/210) vs. NR††	Funding NR No COI

COI = conflict of interest; KP = kyphoplasty; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; SIF = sacral insufficiency fracture; VCF = vertebral compression fracture; VP = vertebroplasty.
* Not defined.
† Authors appear to describe the entire population as having received 'vertebroplasty', with n=238 receiving 'Simple Vertebroplasty' and n=104 receiving Balloon kyphoplasty.
‡ Authors report by whole study; includes sacroplasty, usual care, and screw fixation.
§ Yang, 2023: Considerable difference between groups for duration of back pain at baseline: 8.7 vs. 5.0 weeks
** Inclusion criteria.
†† All 34 non-surgical management patients were followed up to 2 years, but none were contacted at 10-year follow-up.

Appendix Table H2. Patient Characteristics of Non-Randomized Studies of Interventions (Included for Harms only)

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty/Kyphoplasty versus Other Surgical Management						
Purvis, 2018 Database: NIS Study period NR; Recruitment period NR USA	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR	N=12603 Mean age (SD): 83.4 (NR) years Female: 74.0% Race/Ethnicity: White: 90.6% Black: 1.4% Hispanic: 3.8% Other: 4.2%	Vertebroplasty/Kyphoplasty (n=11116) VP/KP, methods NR	Other Surgical Management (n=1487) Spinal decompression and/or fusion	NR	No funding, Authors report industry relations

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover interventions: None					
Vertebroplasty/Kyphoplasty versus Usual Care						
Purvis, 2018 Database: NIS Study period NR; Recruitment period NR USA	Fracture type: OVCF Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR	N=58078 Mean age (SD): 84.1 (NR) years Female: 73.4% Race/Ethnicity: White: 89.6% Black: 1.7% Hispanic: 4.3% Other: 4.4%	Vertebroplasty/Kyphoplasty (n=11116) VP/KP, methods NR	Usual Care (n=46962) Non-surgical treatment, methods NR	NR	No funding, Authors report industry relations

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Crossover interventions: None					
Vertebroplasty versus Other Surgical Management						
Huang, 2019 Database: NIS Study period 2003 to 2013; Recruitment period NR Taiwan	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: None	N=2608 Mean age (SD): >59: 30.0% 60-64: 7.5% 65-69: 10.3% 70-74: 14.4% 75-79: 16.9% >79: 20.9% Female: 65.9% Race/Ethnicity: NR	Vertebroplasty (n=1389) sedation	Conventional Open Surgery (n=1219)	Mean 3.2 years vs. 4.7 years	Funding: University No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Wu, 2012 Database: NHIRD Taiwan (for controls) Retrospective case series Study period NR; Recruitment period: Control: 1996 to 2008, VP: 2000 to 2008 Taiwan	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: NR	N=1989 Mean age (SD): 75 (NR) years Female: 78.1% Race/Ethnicity: NR	Vertebroplasty (n=334) sedation	Other surgery (n=1655) Patients receiving other surgical interventions were matched to VP patients for age, sex, and propensity score. Details NR	≤5 years	No funding No COI
Vertebroplasty versus Non-Operative Management						
Ong, 2018 Database: Medicare Study period NR; Recruitment period: 2005 to 2014‡ United States	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: NR	N=1816188 Age 65-69: 12% 70-74: 15% 75-79: 20% 80-84: 23% ≥85: 29% Female: 71% Race/Ethnicity: White: 93% Black: 2.4% Other: 4.4%	Vertebroplasty (n=117232) Details NR	Non-operative Management (n=1698956) Patients did not receive KP or VP.	≥12 months*	Medtronic Authors report being employees of a scientific and engineering consulting firm, as well as being paid fees by companies and suppliers for consulting services, and other consulting fees, stock, and company membership

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty versus Kyphoplasty						
Cheng, 2019 Retrospective cohort Study period Jan 2008 to Dec 2016 Recruitment period NR China	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain: 15.0 days Duration of symptoms <6 weeks: 100% † Severity of fracture: NR Number of vertebral bodies treated: 1: 100%† One or more previous vertebral fractures: NR	N=338 Mean age (SD): 67.3 (NR) years Female: 79.3% Race/Ethnicity: NR	Vertebroplasty (n=215) VP with mean 4.2ml (0.8) PMMA under imaging guidance	Kyphoplasty (n=123) KP with mean 4.9ml (0.7) PMMA under imaging guidance	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Fracture appearance: NR Crossover interventions: None					
Ong, 2018 Retrospective database Study period NR; Recruitment period: Medicare, 2005 to 2014‡ United States	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: Severity of fracture: NR Number of vertebral levels treated: 1:	N=2077944 Age 65-69: 11% 70-74: 15% 75-79: 21% 80-84: 25% ≥85: 27% Female: 73% Race/Ethnicity: NR	Vertebroplasty (n=117232) Details NR	Kyphoplasty (n=261756) Details NR	≥12 months*	Medtronic Authors report being employees of a scientific and engineering consulting firm, as well as being paid fees by companies and suppliers for consulting services, and other consulting fees, stock, and company membership
Kyphoplasty versus Usual Care						
Faloon, 2015 Retrospective cohort Study period NR;	Fracture type: Osteoporotic VCFs Fracture age: NR	N=258 Mean age (range): 76.4 (47 to 96) years Female: 77.1%	Kyphoplasty (n=137) Percutaneous balloon KP using PMMA (KyphX; volume NR)	Usual care (n=121) Details NR	≥2 years	Funding NR COI NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Recruitment period: 1999 to 2007 United States	Duration of back pain: Acute or subacute (details NR) Severity of fracture: NR Number of vertebral levels treated: NR	Race/Ethnicity: NR				
Kyphoplasty versus Non-Operative Management						
Ong, 2018 Retrospective database Study period NR; Recruitment period: Medicare, 2005 to 2014‡ United States	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: NR	N=1960712 Age 65-69: 11% 70-74: 15% 75-79: 20% 80-84: 23% ≥85: 30% Female: 73% Race/Ethnicity: NR	Kyphoplasty (n=261756) Details NR	Non-operative Management (n=1698956) Patients did not receive KP or VP.	≥12 months*	Medtronic Authors report being employees of a scientific and engineering consulting firm, as well as being paid fees by companies and suppliers for consulting services, and other consulting fees, stock, and company membership
Kyphoplasty versus Other Surgeries						

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Wen, 2021 Retrospective cohort Study period Jan 2013 to Jan 2018 Recruitment period NR USA	Fracture type: Osteoporotic VCF Fracture age: NR Duration of back pain: 5.9 days Duration of symptoms <6 weeks: NR Severity of fracture: Severe: 100%[†] Number of vertebral bodies treated: NR One or more previous vertebral fractures: NR Fracture appearance: NR Crossover interventions: None	N=497 Mean age (SD): 72 (NR) years Female: 67.4% Race/Ethnicity: NR	Kyphoplasty (n=376) KP with PMMA (volume NR) under fluoroscopic guidance via classic unilateral or bilateral transpedicular approach Rehabilitation was performed in accordance with the standard rehabilitation protocol formulated by the hospital, including raising the leg, quadriceps active contractions, ankle pump exercises, and lumbar dorsal muscle strengthening. Generally, a brace was required for 1 month after surgery.	Pedicle Screw Fixation (n=121) Screw fixation via standard posterior midline approach; A longitudinal incision was made over the fractured vertebra and extended 1 or 2 levels above and below for the short- or long-segment fixation technique, respectively. Locating probes were then inserted into each pedicle channel of the fractured and adjacent vertebrae. This was followed by suitable pedicle screw insertion 1 or 2 levels above and below the fractured vertebra, with or without cement reinforcement of the pedicle screws, as appropriate Rehabilitation was performed in accordance with the standard rehabilitation protocol formulated by the hospital, including raising the leg, quadriceps active contractions, ankle pump exercises, and lumbar dorsal muscle strengthening.	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
				Generally, a brace was required for 3 months after surgery.		

COI = conflict of interest; KP = kyphoplasty; NHIRS = National Health Insurance Research Database; NIS = Nationwide Inpatient Sample; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; VCF = vertebral compression fracture; VP = vertebroplasty.
* Up to 10 years for all outcomes other than reoperation in order to limit the effects from other unrelated interventions.
† Inclusion criteria.
‡ There was significant overlap in Medicare search between Ong 2018 and another retrospective database study, Edidin 2015. One 2018 includes all of the same years as Edidin 2015, and extends the search up to 2014. Edidin 2015 was excluded because of this.

Appendix Table H3. Patient Characteristics of Non-Randomized Studies of Interventions (From Prior Review)

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty versus Kyphoplasty						
Fourney, 2003 Retrospective cohort Study period NR; Recruitment period: October 2000 to February 2002 USA	Fracture type: Malignancies Fracture age: NR Duration of back pain: NR Duration of symptoms (median): 3.2 months Severity of fracture: NR	N=56 Median age (range): 64 (30 to 82) years Female: 45% Race/Ethnicity: NR	Vertebroplasty (n=34) VP performed as standard operating procedure by radiologists	Kyphoplasty (n=15) KP performed as standard operating procedure by radiologists VP and KP (n=7) Seven patients underwent both procedures at separate levels	Median 4.5 months 1 month 73% (41/56) 3 months 66% (37/56) 6 months 38% (21/56) 1 year 14% (8/56)	NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral bodies treated: 1.7 One or more previous vertebral fractures: NR Fracture appearance: NR					
Kose, 2006 Retrospective cohort Study period NR; Recruitment period: June 2003 to June 2005 Turkey	Fracture type: Malignancies (multiple myeloma) Fracture age: NR Duration of back pain: NR Duration of symptoms: NR Severity of fracture: NR Number of vertebral bodies treated: NR One or more previous vertebral fractures: NR	N=34 Mean age (range): 63 (45 to 82) years Female: NR Race/Ethnicity: NR	Vertebroplasty (n=16) Vertebroplasty performed under fluoroscopy with mean 3.3 ml PMMA. If patient underwent two levels of application, the other levels were augmented after two days interval.	Kyphoplasty (n=18) Kyphoplasty (15 mm balloons) with mean 4.3 ml PMMA. If patient underwent two levels of application, the other levels were augmented after two days interval.	12 months: 100%	NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Comparator Intervention	Follow-up (% followed)	Funding/COI
	Fracture appearance: NR					

APPENDIX I. Outcome Data Abstraction of Included Non-Randomized Comparative Studies

Appendix Table I1. Efficacy Results of Non-Randomized Studies of Interventions

Author (year)	F/U	Function	Pain	Quality of Life	Other
Vertebroplasty versus Nerve Block					
Bae, 2019 Retrospective cohort	1 week 2 weeks 1 month 3 months 1 year 2 years	VP vs. NB **ODI ((0-100), mean (SD)** Baseline (n= 92 vs. n=72) 24.3 (4.7) vs. 25.7 (4.4) 1 month (n= 92 vs. n=72) 16.7 (3.7) vs. 18.8 (3.3) 3 months (n= 92 vs. n=72) 14.1 (3.6) vs. 17.2 (3.5) 1 year (n= 92 vs. n=72) 11.0 (2.6) vs. 13.3 (3.1) 2 years (n= 92 vs. n=72) 10.1 (2.4) vs. 12.5 (3.2)	VP vs. NB **VAS pain (0-10), mean (SD)** Baseline (n= 92 vs. n=72) 7.6 (1.1) vs. 7.4 (1.3) 1 week (n= 92 vs. n=72) 7.4 (1.2) vs. 6.6 (1.4) 2 weeks (n= 92 vs. n=72) 5.7 (1.1) vs. 5.6 (1.8) 1 month (n= 92 vs. n=72) 3.8 (1.0) vs. 4.3 (1.4) 3 months (n= 92 vs. n=72) 2.7 (1.0) vs. 3.7 (1.2) 1 year (n= 92 vs. n=72) 1.9 (1.1) vs. 2.2 (1.0) 2 years (n= 92 vs. n=72) 2.0 (1.3) vs. 2.0 (0.9)	NR	NR
Vertebroplasty versus Kyphoplasty					
Bae, 2016 Retrospective cohort	NR	NR	VP vs. KP **VAS (0-10), mean (SD)** Baseline (n=238 vs. n=104) 5.8 (2.3) vs. 5.7 (2.4) Post-op (timing NR) (n=238 vs. N=104) 2.5 (1.8) vs. 2.8 (2.1) MD –0.3 (95% CI –0.74 to 0.14)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
			Treatment success (VAS improvement ≥3 points), % (n/N) 62% (148/238) vs. 57% (59/104) RR 1.10 (95% CI 0.90 to 1.33) Subgroup analyses, including treated segment (thoracic vs lumbar), severity of degree of compression, and symmetry of six-column involvement showed no difference (data NR)		
Sacroplasty versus Usual Care					
Sarigul, 2023 Retrospective cohort	10 days 3 months 1 year	SP vs. UC **ODI (0-100), mean (SD)** Baseline (n=83 vs n=102) 78.64 (NR) vs. 51.79 (NR) 10 days (n=83 vs n=102) 24.31 (NR) vs. 48.76 (NR) 3 months (n=83 vs n=102) 14.28 (NR) vs. 42.94 (NR) 1 year (n=83 vs n=102) 8.44 (NR) vs. 21.16 (NR)	SP vs. UC **VAS pain (0-10), mean (SD)** Baseline (n=83 vs n=102) 8.82 (NR) vs. 4.18(NR) 10 days (n=83 vs n=102) 5.91 (NR) vs. 1.48(NR) 3 months (n=83 vs n=102) 4.22 (NR) vs. 1.36 (NR) 1 year (n=83 vs n=102) 1.15 (NR) vs. 2.82 (NR)	NR	NR
Andresen, 2022 Retrospective cohort	2 years	SP vs. Conservative treatment* Overall: Patients receiving sacroplasty experienced pain reductions rapidly and	SP vs. Conservative treatment* **HBI (0 to 100), mean (SD)** Baseline (n=119 vs. n=114) 37 (6) vs. 55 (15) 2 years (n=109 vs. n=88)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		significantly (p<0.001), while patients receiving conservative therapy benefited if baseline pain levels were below ≤5 on VAS, while those with >5 at baseline experienced delayed gratification (generally after 6 months).	83 (6) vs. 76 (13)		
Sacroplasty versus Other Surgical Management					
Andresen, 2022 Retrospective cohort	2 years	SP vs. Screw Fixation* Overall: Patients receiving sacroplasty experienced pain reductions rapidly and significantly (p<0.001), all patients receiving screw fixation benefited after 6 months with sustained benefits.	SP vs. Screw fixation* **HBI (0 to 100), mean (SD)** Baseline (n=119 vs. n=59) 37 (6) vs. 35 (4) 2 years (n=109 vs. n=45) 83 (6) vs. 84 (6)	NR	NR
Sacroplasty versus Non-surgical Management					
Yang, 2023 Retrospective cohort	2 weeks 1 month 3 months 6 months	SP vs. NSM **ODI (0-100), mean (SD)** Baseline (n=13 vs. n=14) 82.6 (9.1) vs. 82.7 (9.7) 2 weeks (n=13 vs. n=14) 68.3 (3.5) vs. 64.6 (8.2) 1 month (n=13 vs. n=14) 56.9 (4.1) vs. 48.8 (8.0) 3 months (n=13 vs. n=14) 32.4 (4.8) vs. 22.6 (9.4)	SP vs. NSM **VAS pain (0-10), mean (SD)** Baseline (n=13 vs. n=14) 7.7 (0.8) vs. 8.0 (1.0) 2 weeks (n=13 vs. n=14) 4.7 (1.3) vs. 5.0 (0.8) 1 month (n=13 vs. n=14) 4.6 (1.2) vs. 3.8 (1.1) 3 months (n=13 vs. n=14) 3.8 (1.5) vs. 1.8 (0.6) 6 months (n=13 vs. n=14) 2.7 (1.4) vs. 0.6 (0.8)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		6 months (n=13 vs. n=14) 20.7 (4.9) vs. 11.2 (3.5)			
Frey, 2017 Prospective cohort	Post-op 2 weeks 4 weeks 12 weeks 24 weeks 1 year 2 years 10 years	NR	SP vs. Non-surgical management **VAS (0-10), mean (SE)** Baseline (n=210 vs. n=34) 8.29 (0.13) vs. 7.47 (0.38) Post-op (n=210 vs. n=NA[†]) 3.63 (0.17) vs. NA 2 weeks (n=NR vs. n=34) 2.82 (0.17) vs. 5.44 (0.44) 4 weeks (n=NR vs. n=34) 2.39 (0.15) vs. 4.24 (0.42) 12 weeks (n=NR vs. n=34) 1.93 (0.14) vs. 3.47 (0.46) 24 weeks (n=NR vs. n=34) 1.45 (0.13) vs. 2.47 (0.42) 1 year (n=NR vs. n=34) 0.89 (0.10) vs. 1.44 (0.28) 2 years (n=82 vs. n=34) 0.66 (0.08) vs. 1.12 (0.25) 10 years (n=117 vs. n=NA[†]) 0.50 (0.08) vs. NA	NR	SP vs. Non-surgical management **Opioid users, % (n/N)** Baseline 77.1% (162/210) vs. 70.6% (24/34) Post-op 32.9% (69/210) vs. NR 10 years 0% (0/117) vs. NR **Non-opioid pharmaceutical users, % (n/N)** Baseline 31% (65/210) vs. 38.2% (13/34) Post-op 0.005% (1/210) vs. NR 10 years 0% (0/117) vs. NR **Over the counter drug users, % (n/N)** Baseline 20.5% (43/210) vs. 41.2% (14/34) Post-op 0.7% (15/210) vs. NR 10 years 0% (0/117) vs. NR

CI = confidence interval; F/U = follow-up; FFP = classification of Fragility Fractures of the Pelvis; HBI = modified Hamburg Barthel Index; KP = kyphoplasty; MD = mean difference; NB = nerve block; NR = not reported; NSM = non-surgical management; ODI = Oswestry Disability Index; RR = risk ratio; SD = standard deviation; SP = sacroplasty; UC = usual care; VAS = visual analogue scale; VP = vertebroplasty.
* Andresen 2022 divides patients into several subgroups for VAS. Conservative treatment patients were split into those above (n=46) and below 5 (n=68) on VAS at baseline. Screw fixation patients were split into FFP type II (n=14), FFP type III (n=14), and FFP type IV (n=31). Sacroplasty patients were split into vertebrosacroplasty (n=20), balloon sacroplasty (n=25), radiofrequency sacroplasty (n=25), and cement sacroplasty (n=49). n at final follow-up is impossible to determine for each subgroup. 150 patients were initially included at baseline for conservative treatment, but 36 of them were referred to screw fixation (n=26) or sacroplasty (n=10) and were therefore included in those groups for all analyses.
† Non-surgical management patients were not assessed at post-operation, nor were they contacted at 10-year follow-up.

Appendix Table I2. Safety Results of Non-Randomized Studies of Interventions

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Vertebroplasty versus Nerve Block					
Bae, 2019 Retrospective cohort	2 years	VP vs. NB **New fractures, % (n/N)** 2 years: 15.2% (14/92) vs. 4.2% (3/72)	VP vs. Nerve Block **Cement leakage, % (n/N)** 2 years: 5.4% (5/92) vs. NA	NR	VP vs. Nerve Block **Left leg weakness, % (n/N)** 2 years: 1.1% (1/92) vs. 0% (0/72) NB group had no reported complications over 2-year follow-up
Vertebroplasty versus Kyphoplasty					
Bae, 2016 Retrospective cohort	NR	NR	NR	NR	NR
Sacroplasty versus Usual Care					
Sarigul, 2023 Retrospective cohort	Mean 7.23 years	NR	SP vs. UC **Cement leakage, % (n/N)** 2.4% (2/83) vs. N/A	NR	NR
Andresen, 2022 Retrospective cohort	2 years	NR	SP vs. Conservative care **Asymptomatic Leakage, % (n/N)** 2 years	SP vs. Conservative care **Mortality, % (n/N)** 2 years	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
			8.4% (10/119) vs. N/A	8.4% (10/119) vs. 21,7% (25/114)	
Sacroplasty versus Other Surgical Management					
Andresen, 2022 Retrospective cohort	2 years	NR	SP vs. Other surgery **Asymptomatic Leakage, % (n/N)** 2 years 8.4% (10/119) vs. N/A	SP vs. Other surgery **Mortality, % (n/N)** 2 years 8.4% (10/119) vs. 13.6% (8/59)	NR
Sacroplasty versus Non-surgical Management					
Yang, 2023 Retrospective cohort	6 months	NR	NR	Reported by whole group **Mortality, % (n/N)** 3.1% (1/32)	NR
Frey, 2017 Retrospective cohort	NR	NR	NR	NR	NR

F/U = follow-up; NA = not applicable; NB = nerve block; NR = not reported; SP = sacroplasty; UC = usual care; VP = vertebroplasty.

Appendix Table I3. Safety Results of Non-Randomized Studies of Interventions (Included for Harms only)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Vertebroplasty/Kyphoplasty versus Other Surgical Management					
Purvis, 2018 Database: NIS	NR	NR	NR	NR	VP/KP vs. Other surgery **Any AE, % (n/N)*** 8.1% (900/11116) vs. 16.3% (242/1487) **Stroke, % (n/N)*** 0.1% (11/11116) vs. 0.3% (4/1487)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Myocardial infarction, % (n/N)* 0.6% (67/11116) vs. 2.2% (33/1487) **Pulmonary embolism, % (n/N)*** 0.2% (22/11116) vs. 1.2% (18/1487) **Shock, % (n/N)*** 0.2% (22/11116) vs. 1.0% (15/1487)
Vertebroplasty/Kyphoplasty versus Usual Care					
Purvis, 2018 Database: NIS	NR	NR	NR	NR	VP/KP vs. UC **Any AE, % (n/N)*** 8.1% (900/11116) vs. 8.7% (4086/46962) **Stroke, % (n/N)*** 0.1% (11/11116) vs. 0% (0/46962) **Myocardial infarction, % (n/N)*** 0.6% (67/11116) vs. 0.8% (376/46962) **Pulmonary embolism, % (n/N)*** 0.2% (22/11116) vs. 0.3% (141/46962) **Shock, % (n/N)***

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					0.2% (22/11116) vs. 0.2% (94/46962)
Vertebroplasty versus Other Surgical Management					
Huang, 2019 Database: NIS	Mean 3.2 (2.5) vs. 4.7 4.7 (3.1) years	VP vs. Conventional Open Surgery **New fracture, % (n/N)** <0.3% (NR/1389) vs. <0.3% (NR/1219)	NR	VP vs. Conventional Open Surgery **Mortality, % (n/N)** 19.2% (267/1389) vs. 17.4% (212/1219)	VP vs. Conventional Open Surgery Pulmonary Embolism: 0.4% (6/1389) vs. ≤0.3% (NR/1219) Other AE: 1.0% (14/1389) vs. 1.0% (12/1219)
Wu, 2012 Database: NHIRD Taiwan, Retrospective case series	≤5 years	NR	NR	NR	VP vs. other surgery **Any Stroke, incidence rate (95% CI)** Within 5 years 25.9 (95% CI 15.0 to 44.6) vs. 22.1 (95% CI 17.5 to 27.8) per 1000 person years Adjusted hazard ratio after adjusted for demographic characteristics, co-morbidities, and medications, adjusted HR = 1.22 (95% CI 0.67 to 2.24) **Hemorrhagic Stroke, incidence rate (95% CI)** Within 5 years 7.7 (95% CI 2.9 to 20.6) vs. 3.6 (95% CI 2.0 to 6.3) per 1000 person years

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Adjusted hazard ratio after adjusted for demographic characteristics, co-morbidities, and medications, adjusted HR = 3.17 (95% CI 0.97 to 10.3) **Ischemic Stroke, incidence rate (95% CI)** Within 5 years 19.7 (95% CI 10.6 to 36.7) vs. 19.6 (95% CI 15.4 to 25.1) per 1000 person years Adjusted hazard ratio after adjusted for demographic characteristics, co-morbidities, and medications, adjusted HR = 0.96 (95% CI 0.49 to 1.91)
Vertebroplasty versus Non-Operative Management					
Ong, 2018 Database: Medicare	10 years	NR	NR	VP vs. non-operation Mortality Non-operative group had 8% (95% CI 8 to 9%) higher propensity-adjusted 10-year mortality risk than VP patients	NR
Vertebroplasty versus Kyphoplasty					

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Cheng, 2019 Retrospective cohort	1 year	VP vs. KP **Adjacent level fracture, % (n/N)** 1 year 3.3% (7/215) vs. 9.8% (12/123)	VP vs. KP **Asymptomatic leakage, % (n/N)** 1 year 7.0% (15/215) vs. 0% (0/123)	NR	NR
Ong, 2018 Database: Medicare	10 years	NR	NR	VP vs. KP **Mortality** KP group had a 13% (95% CI 12 to 13%) higher propensity-adjusted 10-year mortality risk than VP patients	NR
Kyphoplasty versus Usual Care					
Faloon, 2015 Retrospective cohort	≥2 years	KP vs usual care **New vertebral fractures, % (n/N)** Within 2 years 26.3% (36/137) vs. 47.1% (57/121) Multivariate logistic regression for new vertebral fractures with age as a consistent second variable found no variables to be statistically significant.	NR	NR	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Kyphoplasty versus Non-Operative Management					
Ong, 2018 Retrospective database	10 years	NR	NR	KP vs. non-operation **Mortality** Non-operative group had a 24% (95% CI 23 to 24%) higher propensity-adjusted 10-year mortality risk than KP patients	NR
Kyphoplasty versus Other Surgeries					
Wen, 2021 Retrospective cohort	3 years	KP vs. Screw Fixation **New fracture, % (n/N)** 3 years 7.7% (29/376) vs. 5.8% (7/121)	KP vs. Screw Fixation **Asymptomatic leakage, % (n/N)** 3 years 30.1% (113/376)	KP vs. Screw Fixation **Mortality, % (n/N)†** 0% (0/376) vs. 0% (0/121)	KP vs. Screw Fixation **Reoperation, % (n/N)‡** 7.7% (29/376) vs. 5.8% (7/121) **Removal of device, % (n/N)** NA vs. 0.8% (1/121) **Deep vein thrombosis, % (n/N)** 0% vs. 2.5% (3/121) **Other AE (screw fixation only), % (n/N)** 5.0% (6/121)

AE = adverse event; CI = confidence interval; F/U = follow-up; HR = hazard ratio; KP = kyphoplasty; NHIRS = National Health Insurance Research Database; NIS = Nationwide Inpatient Sample; NR = not reported; VP = vertebroplasty.

* n/N's back-calculated.

† Patients were retrospectively excluded if they died.

‡ All re-operations were due to new fractures; no revision surgeries occurred.

Appendix Table H3. Efficacy Results of Non-Randomized Studies of Interventions (From Prior Review)

Author (year)	F/U	Function	Pain	Quality of Life	Other
Vertebroplasty versus Kyphoplasty					
Fourney, 2003 Retrospective cohort	24 hours 1 month 3 months 6 months 12 months	VP vs. KP **ODI ((0-100), mean (SD)**	VP vs. KP **VAS pain (0-10), Median (SD)** Baseline (n=34) 8.0 (NR) vs. 8.0 (NR) 1 month (n=34 vs. n=15) 2.0 (NR) vs. 2.5 (NR) 3 months (n=34 vs. n=15) 2.0 (NR) vs. 2.5 (NR) 6 months (n=34 vs. n=15) 2.0 (NR) vs. 4 (NR) 12 months (n=34 vs. n=15) 1.0 (NR) vs. 2.0 (NR) **Complete or Improved pain relief*, % (n/N)** 24 hours 86% (30/35 sessions) vs. 80% (12/15 sessions), RR 1.07 (95% CI 0.80 to 1.43) **Complete pain relief, % (n/N)** 24 hours 23% (8/35 sessions) vs. 7.0% (1/15 sessions), RR 3.43 (95% CI 0.47 to 25.06) **Improved pain relief, % (n/N)** 24 hours 63% (22/35 sessions) vs. 73% (11/15 sessions), RR 0.86 (95% CI 0.58 to 1.28)	NR	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
Kose,			VP vs. KP VAS (0-50)†, mean (SD) Baseline (n=16 vs. n=18)_ 37.83 (3.25) vs. 36.0 (4.50) 6 weeks (n=16 vs. n=18) 15.3 (4.1) vs. 12.1 (3.6) MD 3.20 (95% CI 0.51 to 5.89) 6 months (n=16 vs. n=18) 12.2 (3.0) vs. 8.6 (2.30) MD 3.60 (95% CI 1.74 to 5.46) 12 months (n=16 vs. n=18) 13.5 (2.9) vs. 9.7 (2.4) MD 3.80 (95% CI 1.95 vs. 5.65)		

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

Appendix Table H4. Safety Results of Non-Randomized Studies of Interventions (From Prior Review)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Vertebroplasty versus Kyphoplasty					
Fourney, 2003 Retrospective cohort	30 days	NR	VP vs. KP **Asymptomatic cement leakage, % (n/N)** 30 days 9.2% (6/65 levels) vs. 0% (0/32 levels)	VP vs. KP **Mortality, % (n/N)** 30 days 0% (0/34) vs. 0% (0/15) 2.5 months 2.9% (1/34) vs. 0% (0/15)	VP vs. KP **Serious AE, % (n/N)** Paraplegia due to metastasis: 30 days 2.9% (1/34) vs. 0% (0/15) Readmission for CHF 30 days 0% (0/34) vs. 6.7% (1/15)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Any device or procedures related complications, % (n/N) 30 days 0% (0/34) vs. 0% (0/15)
Kose, 2016	12 months	VP vs. KP **Adjacent level fracture, % (n/N)** 12 months 0% (0/16) vs. 0% (0/18) **Symptomatic fracture requiring reoperation, % (n/N)** 12 months 0% (0/16) vs. 0% (0/18)	NR	NR	VP vs. KP **Serious AE, % (n/N)** Neurological complications 12 months 0% (0/16) vs. 0% (0/18) Pulmonary complications 12 months 0% (0/16) vs. 0% (0/18) **Device related balloon rupture (asymptomatic), % (n/N)** 12 months NA vs. 5.6% (1/18)

AE = adverse events; F/U = follow-up; KP = kyphoplasty; NR = not reported; VP = vertebroplasty.

Appendix Table I6. Information for Chandra 2019 (Systematic Review Included for Efficacy and Harms related to Sacroplasty)

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Systematic Reviews						

SR, Search dates Database	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Chandra, 2019 Medline, Web of Science, SCOPUS 1980 to February 2018	Sacroplasty (not comparative) Sacral insufficiency fractures	<u>Pain</u> VAS Clinical success <u>Harms</u> Cement Leakage, surgical decompression	1 NRSI 13 case series	No	No	<u>Pain</u> **VAS:** Sacroplasty was associated with an improvement in pain at 24-48 hours post-procedure (MD from baseline 2.70, 95% CI 2.19 to 3.20) and 12 months post-procedure (MD from baseline 2.01, 95% CI 1.35 to 2.67) **Clinical success*:** 95.7% of patients experienced clinical success. <u>Harms</u> **Cement** leakage: 2.2% of procedures resulted in cement leakage. **Surgical decompression:** 0.3% of patients required surgical decompression.

CI = confidence interval; MD = mean difference; NRSI = non-randomized study of intervention; SR = systematic review; VAS = visual analogue scale.
* Patients were considered to be clinically successful if the patient's pain improved, stayed the same, or if remobilization was achieved after the sacroplasty procedure.

APPENDIX J. Demographic Data of Included Case Series

Appendix Table J1. Patient Characteristics of Case Series in Patients with Fractures due to Osteoporosis

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty					
Bae, 2017 Retrospective case series Study period NR; Recruitment period: October 2009 to September 2014 South Korea	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: minimum 2 weeks Severity of fracture: NR Number of vertebral levels treated: 1: 87% 2: 11% 3: 1% 4: 0.3%	N=293 Mean age (SD): 71.9 (8.9) years Female: 79.5% Race/Ethnicity: NR	Vertebroplasty (n=293) Patients received midazolam and fentanyl for low level sedation	36 ± 18.6 months	Funding NR No COI
Ding, 2016 Retrospective case series Study period NR; Recruitment period: January 2009 to March 2011	Fracture type: Osteoporotic VCFs Fracture age (range): 5.6 (1.9 to 16.4) months Duration of back pain: NR	N=292 Mean age (range): 69 (52 to 89) years Female: 75% Race/Ethnicity: NR	Vertebroplasty (n=292) Local anesthesia and conscious sedation under fluoroscopic guidance. Two types of PMMA, low-viscosity (OsteoPal-V) and medium-viscosity (Spineplex), volume NR	≥2 years	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
China	Severity of fracture: Mild: 19.9% Moderate: 39.3% Severe: 40.8% Number of vertebral levels treated: 1: 100%*				
Fadili Hassani, 2019 Retrospective case series Study period Jan 2004 to Dec 2011; Recruitment period NR France	Fracture type: OVCF: 34.0% Malignancy: 39.9% Other: 16.0% Trauma: 10.1% Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: 1.9 (1.4) One or more previous vertebral fractures: NR	N=1512 Mean age (SD): 68.8 (13.9) Female: 66% Race/Ethnicity: NR	Vertebroplasty (n=1512) VP with PMMA under imaging guidance via bilateral or parapedicular approach	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Fracture appearance: NR				
Kobayashi, 2021 Retrospective Case Series Study period NR; Recruitment period: April 2017 to March 2018 Japan	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: 69.7 days Severity of fracture: NR Number of vertebral levels treated (SD): 1.4 (0.8)	N=485 Mean age (SD): 81.4 (8.0) years Female: 74.6% Race/Ethnicity: NR	Vertebroplasty (n=485) Performed under local anesthesia using PMMA (Vertaplex or Simplex P; volume NR)	1 year 74.4% (361/485)	No funding No COI
Tang, 2021 Retrospective Case Series Study period NR; Recruitment period: January 2016 and June 2019 China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: 15.6 days Severity of fracture: Mild: 75% Moderate: 17%	N=1090 Mean age (SD): 72 (8.6) years Female: 75.3% Race/Ethnicity: NR	Vertebroplasty (n=1090) Bipedicular approach using PMMA (Medec; 5.2 ml)	NR	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Severe: 8% Number of vertebral levels treated: 1: 73.8% 2: 26.2%				
Kyphoplasty					
Bergmann, 2012 Prospective case series Study period NR; Recruitment period: February 2002 and February 2011 Germany	Fracture type: Osteoporotic VCFs[†] Fracture age: NR Duration of back pain: Mean ASA score: 3.02 Classification Endplate impressions: 66 Wedge compression: 178 Burst: 181 Number of vertebral levels treated: 1: 78% 2: 24% 3: 6% 4: 1%	N=297 Mean age (SD): 76 (10.7) years Female: 73% Race/Ethnicity: NR	Kyphoplasty (n=297) Percutaneous balloon kyphoplasty with mean 7.7 ml PMMA	NR	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	5: 0.3%				
Bian, 2022 Retrospective case series Study period NR; Recruitment period: January 2017 to June 2020 China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: Severity of fracture: NR Number of vertebral levels treated: 1: 100%*	N=371 Mean age (SD): 72 (7.29) years Female: 76.8% Race/Ethnicity: NR	Kyphoplasty (n=371) All patients treated with local anesthesia. Bilateral balloons individually inflated to restore vertebral height using mean 4.7 ml PMMA All patients received 5 mg zoledronic acid dissolved in 100 ml saline infused intravenously once a year after procedure.	≥1 year. Details NR	No funding No COI
Deibert, 2016 Retrospective case series Study period NR; Recruitment period: 2001 to 2014 United States	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR	N=726 Mean age (SD): NR Female: NR Race/Ethnicity: NR	Kyphoplasty (n=726) Kyphon KP system (Medtronic) used in all cases with PMMA (volume NR), under general anesthesia using fluoroscopic guidance.	≥1 year	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral levels treated: 922 total				
Lin, 2017 Retrospective case series Study period NR; Recruitment period: 2006 to 2011 China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: 1: 100%*	N=495 Mean age (range): 72 (56 to 88) years Female: 57.6% Race/Ethnicity: NR	Kyphoplasty (n=495) Unilateral KP with PMMA. Details NR	1 year	Affiliated Southeast Hospital of Xiamen University No COI
Ning, 2021 Retrospective case series Study period NR; Recruitment period: # China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR	N=921 Mean age (SD): 72.1 (8.95) years Female: 82.8% Race/Ethnicity: NR	Kyphoplasty (n=921) All procedures performed under general anesthesia using fluoroscopic guidance with PMMA (volume NR)	Mean 42.63 ± 22.18 months	Natural Science Fund of China No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral levels treated: NR				
Qi, 2022 Retrospective case series Study period NR; Recruitment period: January 2012 to December 2020 China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: NR	N=896 Mean age (SD): 72 (9.22) years Female: 76.5% Race/Ethnicity: NR	Kyphoplasty (n=896) Details NR	6 months	Peking University Third Hospital Clinical Cohort Project No COI
Spross, 2014 Database: SWISSspine Study period NR; Recruitment period: 2005 to 2012 Switzerland	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR	N=375 Mean age (SD): 73 (NR) years Female: 75.7% Race/Ethnicity: NR	Kyphoplasty (n=375) Details NR	≥6 months	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral levels treated: 1: 100%*				
Zhao, 2022 Retrospective case series Study period NR; Recruitment period: September 2013 to March 2020 China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR Number of vertebral levels treated: NR	N=1752 Mean age (SD): 73.5(NR) years Female: 100% Race/Ethnicity: NR	Kyphoplasty (n=1752) All procedures performed with aid of fluoroscopy, using transpedicular approach and PMMA (volume NR). All patients were treated with standard PTH therapy after surgery (or bisphosphonate when contraindicated)	1 year	National Natural Science Foundation of China and the National Key Research and Development Project No COI
Mixed Vertebroplasty and Kyphoplasty					
Choo, 2018 Database: ACS-NSQIP Study period NR; Recruitment period: 2012 to 2014‡ United States	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: Severity of fracture: NR	N=2433 Age 18 to 65: 18.3% 66 to 75: 24.9% 76 to 89: 48.2% 90+: 8.6% Female: 70.4% Race/Ethnicity: White: 88.5% Black/AA: 2.1%	Vertebroplasty/Kyphoplasty (n=2433) Details NR. 90% underwent kyphoplasty, 10% underwent vertebroplasty.	30 days	No funding COI NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral levels treated: 1: 80.8% >1: 19.2%	Asian: 3.8% Native Hawaiian/PI: 0.2% Amer Indian/ Alask Nat: 0.2% NR: 5.3%			
Kim, 2022 Database: ACS-NSQIP Study period NR; Recruitment period: 2011 to 2013‡ United States	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: ASA 1 and 2: 26% 3: 63.9% 4: 10.1% Number of vertebral levels treated: NR	N=1932 Mean age (SD): 74.9 (11.9) years Female: 71% Race/Ethnicity: White: 87.9% Black: 2% Hispanic: 5.6% Asian: 4.3% Unknown: 0.2%	Vertebroplasty/Kyphoplasty (n=1932) Details NR. 90% underwent kyphoplasty, 10% underwent vertebroplasty.	1 month	No funding No COI
Sun, 2023 Retrospective case series Study period NR; Recruitment period: January	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR	N=373 Mean age (SD): 76 (9.4) years Female: 74.3% Race/Ethnicity: NR	Vertebroplasty/Kyphoplasty (n=373) Details NR	NR	National High Level Hospital Clinical Research, Elite Medical Professionals Project of China-Japan Friendship Hospital, Medical and health Science and Technology Innovation Project of Chinese Academic of Medical Science, and the National Natural Science Foundation of China

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
2017 to December 2020 China	Severity of fracture: NR Number of vertebral levels treated: NR				No COI
Wang, 2014 Retrospective case series Study period NR; Recruitment period: # China	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: >2 weeks* Severity of fracture: NR Number of vertebral levels treated: 1: 100%*	N=358 Mean age (SD): 71 (9.1) years Female: 75.7% Race/Ethnicity: NR	Vertebroplasty/Kyphoplasty (n=358) All procedures done using bipedicular approach under local anesthesia. Cement volume ranged from 2 ml to 10 ml (mean 5.7) 37.4% underwent kyphoplasty, 62.6% underwent vertebroplasty.	≥18 months*	National Natural Science Foundation of China No COI
Zhang, 2020 Retrospective case series Study period NR; Recruitment period: January 2015 to March 2019	Fracture type: Osteoporotic VCFs Fracture age: NR Duration of back pain: NR Severity of fracture: NR	N=268 Mean age (SD): 74 (NR) years Female: 79.2% Race/Ethnicity: NR	Vertebroplasty/Kyphoplasty (n=268) Unilateral VP or KP under local anesthesia. Cement volume NR 67.1% underwent kyphoplasty, 32.9% underwent vertebroplasty.	NR	Beijing Science and Technology Planning Project, Ministry of Education Key Laboratory of Trauma Treatment and Nerve Regeneration No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
China	Number of vertebral levels treated: 1: 89.6% 2: 10.1% 3: 0.4%				
Sacroplasty					
Beall, 2022 Database: The Sacroplasty Registry Study period Feb 2013 to Jan 2020; Recruitment period NR USA	Fracture type: OVCF: 98.0% Malignancies: 1.0% Unknown: 1.0% Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: NR One or more previous vertebral fractures: NR	N=102 Mean age (SD): 74.1 (10.1) years Female: 68.6% Race/Ethnicity: NR	Sacroplasty (n=102) Clinical sites were not limited to a specific sacral approach technique, and data regarding the technical approach to SP were not collected.	48.0% (49/102)	Society of Interventional Radiology Foundation Authors report industry relations

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Fracture appearance: NR				
Heo, 2017 Retrospective case series Study period Mar 2009 to Early 2016; Recruitment period NR Korea	Fracture type: SIF (osteoporosis) Fracture age: NR Duration of back pain: >3 weeks: 100%* Duration of symptoms <6 weeks: NR Severity of fractures: NR Number of vertebral bodies treated: 1: 100% One or more previous vertebral fractures: NR Fracture appearance (Total fractures): Bilateral: 69.1%	N=68 Mean age (years, (SD)): 76.8 (6.2) Female: 94.1% Race/Ethnicity: NR	Sacroplasty (n=68) SP with low pressure–high viscosity SP with mean 1.61 ml PMMA via the short axis approach under fluoroscopic guidance	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Unilateral: 30.9% Crossover interventions: None				
Kortman, 2013 Retrospective case series Study period Mar 2009 to Early 2016 Recruitment period NR USA	Fracture type: SIF (osteoporosis) Fracture age: NR Duration of back pain: 30.5 months Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: 1: 100%* One or more previous vertebral fractures: NR Fracture appearance: Bilateral: 82.8% Unilateral: 17.2%	N=204 Mean age (SD): 77.2 (NR) years Female: 86.3% Race/Ethnicity: NR	Sacroplasty (n=204) SP with mean 4.1ml cement under imaging guidance via short- or long-axis approach; Acrylic cement used in 202 procedures and bio-ceramic cement used in remaining 2 procedures	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI

ACS-NSQIP = American College of Surgeons National Surgical Quality; ASA = American Society of Anesthesiologists; COI = conflict of interest; KP = kyphoplasty; NA = not applicable; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; SIF = sacral insufficiency fracture; SP = sacroplasty; VCF = vertebral compression fracture; VP = vertebroplasty;
* Inclusion criteria.
† 91.6% of patients had osteoporotic VCFs, 8.4% had fractures from malignancies.
‡ There were 2 years overlap between Choo and Kim in their search terms with the ACS-NSQIP.

Appendix Table J2. Patient Characteristics of Case Series in Patients with Fractures due to Malignancies

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
Vertebroplasty					
Cui, 2022 Retrospective case series Study period Jan 2010 to Aug 2017; Recruitment period NR China	Fracture type: Malignant Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated (mean): 2.3	N=230 Mean age (range): 65 (24 to 93) years Female: 43.0% Race/Ethnicity: NR	Vertebroplasty (n=230) VP with cement via unilateral or bilateral transpedicular approach under x-ray guidance. Injection was stopped if blood oxygen saturation decreased, spinal canal leakage was detected, or the patient complained of neurological symptoms. Mean cement volume 2.9 ml.	100%	No funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	One or more previous vertebral fractures: NR Fracture appearance: NR				
Moulin, 2020 Retrospective case series Study period Dec 2015 to Jun 2019; Recruitment period NR France	Fracture type: Malignant Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: 6: 26% 7: 26% 8: 16% 9: 14% 10: 6% 11: 4% 12: 6% 13: 2%	N=50 Mean age years (SD): 66 (10) years Female: 48.0% Race/Ethnicity: NR	Vertebroplasty (n=50) VP with PMMA (volume NR) under fluoroscopic guidance via a unilateral transpedicular or intercostovertebral approach. Injection was stopped when the anterior two thirds of the vertebra was filled.	100%	Funding NR COI include industry relations

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	One or more previous vertebral fractures: NR Fracture appearance: NR				
Rocha Romero, 2020 Retrospective case series Study period Dec 2015 to Jun 2019 Recruitment period NR Costa Rica	Fracture type: Malignant Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated (mean): 4.3 One or more previous vertebral fractures: NR Fracture appearance: NR	N=44 Mean age (range): 57 (30 to 75) years Female: 47.3% Race/Ethnicity: NR	Vertebroplasty (n=44) Details NR	100%	No funding No COI
Kyphoplasty					

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
Garcia-Maroto, 2015 Retrospective case series Study period Jan 2010 to Aug 2017; Recruitment period NR Spain	Fracture type: Malignant Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated: 1: 55% 2: 33% 3: 8.0% 4+: 4% One or more previous vertebral fractures: NR Fracture appearance: NR	N=75 Mean age (range): 68 (42 to 86) years Female: 66.7% Race/Ethnicity: NR	Kyphoplasty (n=75) KP via bilateral transpedicular approach	9.3% (7/75)	No funding No COI
Molloy, 2016 Retrospective case series	Fracture type: Malignant Fracture age: NR	N=158 Median age (range):	Kyphoplasty (n=158) KP with PMMA via unilateral pedicular approach under fluoroscopy guidance. 2 to 3 mL in thoracic spine and 3 to 4 mL in lumbar spine	100%	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
Study period 2007 to 2014; Recruitment period NR UK	Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture (Total N fractures): NR Number of vertebral bodies treated (mean): 1: 67% 2+: 33% One or more previous vertebral fractures: NR Fracture appearance (Total fractures): Burst: 100% Crossover interventions: None	62.6 (16 to 92) years Female: 37.3% Race/Ethnicity: NR			
Wu, 2022 Retrospective case series Study period Jan 2017 to Dec 2019;	Fracture type: Malignant Fracture age: NR Duration of back pain: NR	N=117 Mean age (SD): 59 (NR) years Female: 53.8%	Kyphoplasty (n=117) KP with mean 2.8 ml PMMA under fluoroscopic guidance via bipedicular approach. PMMA max volume 4ml for thoracic vertebrae and 6ml for lumbar vertebrae.	100%	Funding NR No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
Recruitment period NR China	Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated (mean): 1.8 One or more previous vertebral fractures: NR Fracture appearance: NR	Race/Ethnicity: NR			
Wu, 2023 Retrospective case series Study period Feb 2013 to Jan 2020; Recruitment period NR China	Fracture type: Malignant Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR	N=92 Mean age (SD): 66.6 (4.7) years Female: 46.7% Race/Ethnicity: NR	Kyphoplasty (n=92) KP with mean 2.7 ml PMMA under fluoroscopic guidance via bilateral approach. Injection was stopped when high resistance was obviously felt, leakage of PMMA was observed or the PMMA reached the posterior margin of the vertebra.	100%	Government funding No COI

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	Number of vertebral bodies treated: 1: 55% 2: 27% 3: 10% 4: 3% 5: 3% 6: 2% One or more previous vertebral fractures: NR Fracture appearance: NR				
Zou, 2010 Retrospective case series Study period Jan 2003 to Jan 2008; Recruitment period NR China	Fracture type: Malignant Fracture age: NR Duration of back pain: 5.2 months Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated:	N=21 Mean age (range): 65.9 (47 to 81) Female: 57.1% Race/Ethnicity: NR	Kyphoplasty (n=21) KP with PMMA under fluoroscopic guidance via posterior transpedicular approach. PMMA volume NR	100%	Funding NR COI NR

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	1: 10% 2: 76% 3: 14% One or more previous vertebral fractures: NR Fracture appearance: NR				
Mixed Vertebroplasty and Kyphoplasty					
Burton, 2011 Retrospective case series Study period Jan 1, 2001 to May 31, 2008 Recruitment period NR USA	Fracture type: Malignancy: 65% Osteoporotic VCF: 35% Fracture age: NR Duration of back pain: NR Duration of symptoms <6 weeks: NR Severity of fracture: NR Number of vertebral bodies treated (mean): 2.84	N=407 Mean age (SD): 62.9 (NR) years Female: 52% Race/Ethnicity: NR	Mixed Vertebroplasty/Kyphoplasty (n=407) VP only (methods NR): 64.3% (262/407) VP only (methods NR): 38.3% (156/407) Mixed VP and KP (methods NR): 27.3% (111/407)	Mean 25 (13.9) days 2 months: 41.8% (170/407)	Medtronic Authors report industry relations

Author (year), Study Period, Country	Characteristics	Population	Intervention	Follow-up (% followed)	Funding/COI
	One or more previous vertebral fractures: NR Fracture appearance: NR				

COI = conflict of interest; KP = kyphoplasty; NA = not applicable; NR = not reported; PMMA = polymethylmethacrylate; SD = standard deviation; SIF = sacral insufficiency fracture; VCF = vertebral compression fracture; VP = vertebroplasty.

APPENDIX K. Outcome Data Abstraction of Included Case Series

Appendix Table K1. Efficacy Results of Case Series of Osteoporotic Fractures

Author (year)	F/U	Function	Pain	Quality of Life	Other
Sacroplasty					
Beall, 2022 Database: The Sacroplasty Registry	6 months	SP only **RDQ (0-24), mean (SD)** Baseline (n=102) 17.7 (6.4) 1 month (n=51) 8.4 (4.9), p<0.001) 3 months (n=52) 6.9 (4.9), p<0.001) 6 months (n=49) 5.2 (5.2), p<0.001) **RDQ Success (≥5 points), % (n/N)** 1 month 76.5% (39/51) 3 months 78.8% (41/52) 6 months 83.7% (41/49)	SP only **NRS (0-10), mean (SD)** Baseline (n=102) 7.8 (2.4) 1 month (n=51) 2.4 (3.3), p<0.001) 3 months (n=52) 1.2 (2.5), p<0.001) 6 months (n=49) 0.9 (2.2), p<0.001) **NRS Success (≥2 points), % (n/N)** 1 month 72.6% (37/51) 3 months 90.4% (47/52) 6 months 91.8% (45/49)	NR	NR

F/U = follow-up, NR = not reported; NRS = numerical rating scale; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SP = sacroplasty.

Appendix Table K2. Safety Results of Case Series of Osteoporotic Fractures

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Vertebroplasty					
Bae, 2017 Retrospective case series	36 ± 18.6 months	VP only	NR	NR	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
		Secondary new VCFs, % (n/N) Timing NR: 9.8% (25/256)* Patients with bone cement leakage during procedure had a higher incidence of new fractures than patients without leakage, p=0.039			
Ding, 2016 Retrospective case series	≥2 years	NR	VP only **Cement leakage, % (n/N)** 77.7% (227/292)	NR	NR
Fadili Hassani, 2019 Retrospective case series	Mean 8.1 months	NR	NR	VP only **Death due to embolism, % (n/N)** 0% (0/1512)	VP only **Any Intracardiac cement embolism, % (n/N)** 3.9% (72/1512) **Multiple Intracardiac cement embolism, % (n/N)** 1.2% (18/1512) **Intracardiac cement embolism with associated pulmonary cement embolism, % (n/N)** 4.1% (62/1512) **Symptomatic Intracardiac cement embolism, % (n/N)** 0.3% (6/1512)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Cardiorespiratory arrest: <0.1% (1/1512)
Kobayashi, 2021 Retrospective Case Series	1 year	VP only **New adjacent vertebral compression fractures, % (n/N)** 1 year 6.6% (24/361) **New distant fractures, % (n/N)** 1 year 12.7% (46/361) **Total new fractures, % (n/N)** 1 year 18.6% (67/485)	VP only **Symptomatic cement leakage, % (n/N)** 1 year 0% (0/361)[†] **Asymptomatic cement leakage, % (n/N)** 1 year 35.7% (173/485)[†]	VP only **Mortality, % (n/N)** 1 year 1.2% (6/485)[‡]	VP only **Adverse events related to VP, % (n/N)** 1 year 0% (0/485)
Tang, 2021 Retrospective Case Series	NR	NR	VP only **Cortical leakage, % (n/N)** Timing NR 20.3% (295/1456 levels) Multivariate regression analysis showed risk factors as age (OR 1.03, 95% CI 1.01 to 1.05), cause (trauma vs. non-trauma; OR 1.73, 95% CI 1.23 to 2.44), Cortical disruption (OR 23.22, 95% CI 13.21 to	NR	VP only **Reoperation due to new fractures, % (n/N)** Timing NR 22.1% (241/1090)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
			40.83), intravertebral cleft (OR 1.63, 95% CI 1.17 to 2.26), Cement distribution (OR 1.68, 95% C 1.21 to 2.33) and cement volume (OR 1.15, 95% CI 1.05 to 1.25). **Venous leakage, % (n/N)** 56.2% (819/1456 levels) Multivariate regression analysis showed risk factors as gender (male vs female; OR 1.42, 95% CI 1.09 to 1.85), fracture severity (OR 0.64, 95% CI 0.51 to 0.79), intravertebral cleft (OR 0.51, 95% CI 0.38 to 0.67), Basivertebral foramen (OR 1.80, 95% CI 1.37 to 2.35), cement distribution pattern (OR 0.32, (95% CI 0.25 to 0.42), and cement volume (OR 1.08, 95% CI 1.01 to 1.17).		
Kyphoplasty					
Bergmann, 2012 Prospective case series	NR	KP only **Symptomatic fractures, % (n/N)** Timing NR 8% (23/293)	KP only **Cement leakage, % (n/N)** Timing NR 40.1% (129/293)		KP only **Other adverse events, % (n/N)** Timing NR Pain persisting at same level as KP: 1% (3/293)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Subcutaneous hematoma: 1% (3/293) Urinary tract infection: 4.4% (13/293)[§] Cardiac arrest resulting in death: <1% (1/293) Return due to recurring pain: 8.1% (24/293)
Bian, 2022 Retrospective case series	≥1 year	KP only **New osteoporotic fracture** ≥1 year Multivariate logistic regression of risk factors for new VCFs showed HU value (OR 0.96, 95% CI 0.94 to 0.97), having cement leakage (OR 2.96, 95% CI 1.49 to 5.88) and having thoracolumbar junction (OR 3.11, 95% CI 1.41 to 6.89).	KP only **Cement leakage, % (n/N)** ≥1 year 28.3% (105/371)	NR	NR
Deibert, 2016 Retrospective case series	≥1 year	KP only **Symptomatic new fractures, % (n/N)** Average 350 days 10.6% (77/726)	NR	NR	KP only **Re-operation, % (n/N)** Average 350 days 10.6% (77/726)[**]
Lin, 2017 Retrospective case series	1 year	KP only **New fractures, % (n/N)**	KP only **Cement leakage, % (n/N)** 20% (99/495)[††]	NR	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
		1 year 22.2% (110/495)			
Ning, 2021 Retrospective case series	NR	KP only **New fractures, % (n/N)** 12.05% (111/921) Multivariate logistic regression for new fractures, adjusted OR, 95% CI Gender: 2.67, 95% CI 1.25 to 5.73 BMD: 0.79, 95% CI 0.64 to 0.96 Calcium + vitamin D: 0.31, 95% CI 0.19 to 0.53 Gallstone disease: 1.42, 95% CI 0.67 to 3.01 Stomach diseases: 7.92, 95% CI 3.28 to 19.15 History of previous fracture: 4.83, 95% CI 2.43 to 9.58 Ovariectomy: 3.01, 95% CI 1.05 to 13.76 Zoledronic acid: 0.13, 95% CI 0.07 to 0.25	NR	NR	NR
Qi, 2022 Retrospective case series	6 months	NR	KP only	NR	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
			Symptomatic bone cement displacement, % (n/N) 2.3% (21/896)‡‡ Multivariate logistic regression for bone cement displacement, adjusted OR, 95% CI Anterior leakage: 1.74, 95% CI 1.22 to 3.30 Intravertebral vacuum cleft: 3.36, 95% CI 1.61 to 13.04 Bone cement distribution score: 0.47, 95% CI 0.23 to 0.90 Paraspinal muscle relative cross-sectional area: 0.95, 95% CI 0.92 to 0.99 Paraspinal muscle fatty degeneration: 1.06, 95% CI 1.01 to 1.12		
Spross, 2014 Database: SWISSspine	Mean 3.6 months	KP only **Adjacent vertebral fracture, % (n/N)** Mean 3.6 months 9.9% (37/375) Multivariate logistic regression for new adjacent vertebral fracture, adjusted OR, 95% CI	NR	NR	NR

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
		Preoperative segmental kyphosis: 8.36, 95% CI 1.61 to 43.5 Rheumatoid arthritis: 2.96, 95% CI 1.07 to 8.21 Cardiovascular disease: 2.66, 95% CI 1.01 to 7.00			
Zhao, 2022 Retrospective case series	1 year	KP only **Adjacent vertebral fracture, % (n/N)** 1 year 4.6% (80/1752)	KP only **Cement leakage, % (n/N)** 1 year 11.5% (202/1752)[§§]	NR	NR
Mixed Vertebroplasty and Kyphoplasty					
Choo, 2018 Database: ACS-NSQIP	1 month	NR	NR	VP/KP only **30-day mortality, % (n/N)** 30 days 2.0% (49/2433) Analysis adjusted for mortality, adjusted OR, 95% CI Dependent functional health status prior to surgery: 2.92, 95% CI 1.48 to 5.75 Pre-operative dialysis use: 11.74, 95% CI 2.34 to 58.91	VP/KP only **Any complications, % (n/N)**[***] 30 days 5.8% (140/2433) Adjusted analysis for any complications, adjusted OR, 95% CI Dependent functional health prior to surgery: 1.78, 95% CI 1.15 to 2.76 Pre-operative sepsis/SIRS: 2.52, 95% CI 1.06 to 2.48 Disseminated cancer: 1.94, 95% CI 1.08 to 3.5 Wound infection pre-operatively: 3.47, 95% CI 1.54 to 7.80

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
				Disseminated cancer: 7.09, 95% CI 3.49 to 14.38 Chronic steroid use: 3.59, 95% CI 1.90 to 6.76 Inpatient admission status: 4.95, 95% CI 2.39 to 10.25	Impatient admission status: 3.22, 95% CI 2.18 to 4.77 **Most common AEs, % (n/N)** 30 days Urinary tract infection 2.1% (51/2433) Adjusted analysis for urinary tract infection, adjusted OR, 95% CI CVA/Stroke: 20.37, 95% CI 1.72 to 241.21 Inpatient admission status: 2.36, 95% CI 1.34 to 4.17 Respiratory complications 1.7% (42/2433) Adjusted analysis for respiratory complications, adjusted OR, 95% CI Dependent function health status prior to surgery: 2.28, 95% CI 1.13 to 4.59 COPD: 2.65, 95% 1.36 to 5.15 Chronic steroid use: 2.38, 95% CI 1.17 to 4.81 Inpatient admission status: 5.86, 95% CI 2.53 to 13.58 Vertebroplasty: 3.28, 95% CI 1.56 to 6.88 30-day re-admissions, % (n/N) 30 days 10.6% (258/2433)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Adjusted analysis for re-admission, adjusted OR, 95% CI Age, 76-89: 1.75, 95% CI 1.12 to 2.73 Age, 90+: 2.78, 95% CI 1.57 to 4.92 COPD: 1.77, 95% CI 1.27 to 2.48 Disseminated cancer: 2.89, 95% CI 1.87 to 4.75 Chronic steroid use: 2.21, 95% CI 1.57 to 3.10 ASA > II: 1.93, 95% CI 1.27 to 2.92 Inpatient admission status: 1.58, 95% CI 1.20 to 2.08 Vertebroplasty: 1.64, 95% CI 1.11 to 2.43 30-day reoperations, % (n/N) 30 days 3.6% (88/2433) Adjusted analysis for re-operation, adjusted OR, 95% CI Age, 90+: 3.27, 95% CI 1.31 to 8.13 Pre-operative Sepsis/SIRS: 2.54, 95% CI 1.04 to 6.21 Disseminated cancer: 2.38, 95% CI 1.11 to 5.09
Kim, 2022	1 month	NR	NR	VP/KP only	VP/KP only

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Database: ACS-NSQIP				**Mortality, % (n/N)** 1 month 2.1% (40/1935) Multivariate logistic regression for mortality, adjusted OR, 95% CI ASA classification (1 and 2 vs. 4): 16.6, 95% CI 1.96 to 140.96 Creatine ≥1.3 mg/dl: 3.49, 95% CI 1.13 to 10.82 MAC or local anesthesia: 3.05, 95% CI 1.20 to 7.76	**Overall complications, % (n/N)** 1 month 8.6% (166/1932) **Minor complications, % (n/N)** 1 month 2.7% (53/1932) Multivariate logistic regression for minor complications, adjusted OR, 95% CI History of chronic obstructive pulmonary disease: 2.60, 95% CI 1.06 to 6.37 **Major complications, % (n/N)** 1 month 4.9% (95/1932) Multivariate logistic regression for major complications, adjusted OR, 95% CI Albumin: 2.39, 95% CI 1.31 to 4.83 White blood cell count: 1.12, 95% CI 1.04 to 1.20 **Most common AEs, % (n/N)** 1 month Urinary tract infection: 1.6% (30/1932) Pneumonia: 0.9% (18/1932)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Reoperation, % (n/N) 1 month 3.2% (61/1932) **Major complications, % (n/N)** 1 month Pulmonary embolism: 0.7% (13/1932)
Sun, 2023 Retrospective case series	NR	NR	VP/KP only **Pulmonary cement embolism, % (n/N)** Median 412 days 17.2% (64/373)	NR	NR
Wang, 2014 Retrospective case series	NR	VP/KP only **New vertebral fractures, % (n/N)** Timing NR 12.6% (45/358) **New symptomatic vertebral fractures, % (n/N)** Timing NR 7.3% (26/358) Multiple logistic regression for new vertebral fractures, adjusted OR, 95% CI Age: 1.06, 95% CI 1.01 to 1.12 BMD: 2.70, 95% CI 1.14 to 6.39	VP/KP only **Cement leakage, % (n/N)** Timing NR 0% (0/358)	NR	VP/KP only **Major complication, % (n/N)** Timing NR 0% (0/358) **Re-operation due to new symptomatic fractures, % (n/N)** Timing NR 7.3% (26/358)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
		Intravertebral clefts: 0.22, 95% CI 0.08 to 0.63			
Zhang, 2020 Retrospective case series	NR	NR	VP/KP only **Cement leakage, % (n/N)** Timing NR 32.5% (87/268)[†††] Multivariate logistic regression, adjusted OR, 95% CI Delayed surgery: 2.74, 95% CI 1.35 to 5.59 Preoperative compression ratio: 0.13, 95% CI 0.02 to 0.84 Upper endplate disruption: 2.74, 95% CI 1.14 to 6.56	NR	NR
Sacroplasty					
Beall, 2022 Database: The Sacroplasty Registry	6 months	NR	SP only **Any leakage, % (n/N)** 6 months 17.7% (18/102) **Symptomatic leakage, % (n/N)** 6 months 1.0% (1/102)	SP only **Mortality, % (n/N)** 6 months 0% (0/102)	SP only **Readmission (rolling total), % (n/N)** 1 month: 11.8% (6/51) 1 to 3 months: 1.9% (1/52) 3 to 6 months: 2.0% (1/49)
Heo, 2017 Retrospective case series	Mean 15.86 (5.69) months	NR	SP only **Non-symptomatic leakage, % (n/N)**	NR	SP only **Major AEs, % (n/N)** 0% (0/68)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
			2.9% (2/68)		
Kortman, 2013 Retrospective case series	Minimum 1 year	SP only **Progressive fracture dislocation, % (n/N)** 0.5% (1/204) **New symptomatic fractures, % (n/N)** 1.5% (3/204)	SP only **Symptomatic leakage, % (n/N)** 0.5% (1/204)	SP only **Procedure-related deaths, % (n/N)** 0% (0/204)	SP only No infections, pulmonary emboli or hemorrhages reported **Reoperation, % (n/N)** 1.5% (3/204)

ACS-NSQIP = American College of Surgeons National Surgical Quality Improvement Program; CI = confidence interval; F/U = follow-up; KP = kyphoplasty; NR = not reported; OR = odds ratio; SP = sacroplasty; VCF = vertebral compression fracture; VP = vertebroplasty.
* Only analyzed amongst patients with single level VCFs
† All cement leakage was asymptomatic.
‡ No deaths were related to the procedure
§ Two of these patients died during their hospital stay. One from a mitral valve endocarditis, another from cardiac insufficiency.
** All re-operations due to new symptomatic fractures
†† Calculated using the totals for cement leakage in the different fracture groups.
‡‡ An additional 35 patients had asymptomatic bone cement displacement.
§§ Calculated by combining adjacent and non-adjacent fracture groups.
*** Authors report a list of possible AEs, and analyses adjusted for numerous other variables.
††† Authors report 96 patients and 32.5%. Because they report 32.5% in text multiple times, and only report 96 patients once in a table, we back-calculated to receive the numerator.

Appendix Table K1. Efficacy Results of Case Series of Malignancies

Author (year)	F/U	Function	Pain	Quality of Life	Other
Vertebroplasty					
Moulin, 2020 Retrospective case series	1 month (all non-mortality outcomes)	NR	VP only **NRS (0-10), mean (SD)** Baseline (n=50) 5.0 (1.8) 1 month (n=50) 1.7 (1.4)	NR	VP only **Mean opioid consumption, mean mg/d (SD)** Baseline (n=50) 76 (42) 1 month (n=50)

Author (year)	F/U	Function	Pain	Quality of Life	Other
	Mean 401 days (mortality)		**Marked Improvement in NRS pain (decrease 50%), % (n/N)** 78% (31/40) **Moderate improvement in NRS (decrease 30% to 50%), % (n/N)** 2% (1/40) **Unchanged NRS, % (n/N)** 20% (8/40) **Increased pain on NRS (increase 30%), % (n/N)** 0% (0/40) "painful group" subgroup **NRS (0-10), mean (SD)** Baseline (n=NR) 6.0 (1.2) 1 month (n=NR) 2.1 (1.3)		45 (37) "painful group" subgroup **Mean opioid consumption, mean mg/d (SD)** Baseline (n=NR) 79 (35) 1 month (n=NR) 34 (30)
Rocha Romero, 2020 Retrospective case series	Up to two years	NR	VP only **NRS (0-10), mean (SD)** Baseline (n=44) 5.16 (NR) 1 month (n=44) 1.07 (NR) 3 months (n=44) 1.48 (NR) 5 months (n=44)	VP only **KPS (0-100), mean (SD)** Baseline (n=44) 78.6 (NR) 1 month (n=44) 78.0 (NR) 3 months (n=44) 76.7 (NR) 5 months (n=44)	VP only **Morphine Equivalent Daily Dose, mean (SD)** Baseline (n=44) 33.4 (NR) 1 month (n=44) 24.0 (NR) 3 months (n=44) 29.4 (NR)

Author (year)	F/U	Function	Pain	Quality of Life	Other
			1.77 (NR) 12 months (n=44) 1.77 (NR) 15 months (n=44) 1.45 (NR) 18 months (n=44) 1.76 (NR) 24 months (n=44) 1.68 (NR)	75.8 (NR) 12 months (n=44) 77.2 (NR) 15 months (n=44) 75.6 (NR) 18 months (n=44) 74.2 (NR) 24 months (n=44) 77.9 (NR)	5 months (n=44) 29.4 (NR) 12 months (n=44) 28.2 (NR) 15 months (n=44) 28.2 (NR) 18 months (n=44) 32.0 (NR) 24 months (n=44) 21.0 (NR)
Kyphoplasty					
Garcia-Maroto, 2015 Retrospective case series	Mean 11 months	NR	KP only **VAS (0-10), mean (SD)** Timing NR Baseline (n=75) 7.49 (1.19) Follow-up (n=NR) 3.21 (0.95)	KP only **KPS (0-100), mean (SD)** Timing NR Baseline (n=75) 60.2 (10) Follow-up (n=NR) 80.7 (12.1)	KP only **Major opioid use, % (n/N)** Baseline 53.3% (40/75) Follow-up 12% (9/75) **Minor opioid use, % (n/N)** Baseline 26.6 (20/75) Follow-up NR **NSAID use, % (n/N)** Baseline 20% (15/75) Follow-up 42% (32/75)
Wu, 2023 Retrospective case series	1 year	KP only ODI (0-100), mean (SD) Baseline (n=92) 70.9 (7.1) 3 days (n=92)	KP only VAS (0-10), mean (SD) Baseline (n=92) 6.3 (2.0) 3 days (n=92)	KP only SF-36 total (0-100), mean (SD) Baseline (n=92) 89.7 (16.1)	NR

Author (year)	F/U	Function	Pain	Quality of Life	Other
		31.4 (4.7) 1 month (n=92) 31.2 (3.5) 3 months (n=92) 31.2 (3.5) 1 year (n=92) 30.4 (3.2)	3.3 (1.5) 1 month (n=92) 2.3 (1.1) 3 months (n=92) 2.8 (1.2) 1 year (n=92) 3.4 (1.1)	1 year (n=92) 99.5 (19.7)	

F/U = follow-up, NR = not reported; NRS = numerical rating scale; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SP = sacroplasty.

Appendix Table K2. Safety Results of Case Series of Malignancies

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Vertebroplasty					
Cui, 2022 Retrospective case series	NR	NR	VP only **Any leakage, % (n/N)** Timing NR 34.9% (185/530 levels)	NR	NR
Moulin, 2020 Retrospective case series	1 month (all non-mortality outcomes) Mean 401 days (mortality)	VP only **New Fracture, % (n/N)** 1 month 10.0% (5/50)	NR	VP only **Mortality, % (n/N)** Total: 34.0% (17/50) Within 100 days: 10.0% (5/50)	VP only No major complications reported **Reoperation, % (n/N)** 6.0% (3/50) **Other Skeletal-Related Event, % (n/N)** 4.0% (2/50)

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
					Cement pulmonary embolism, % (n/N) 2.0% (1/50) **Other AE, % (n/N)** 12.0% (6/50)
Rocha Romero, 2020 Retrospective case series	1 year	VP only **New Fracture, % (n/N)** 1 year 30% (13/44)*	NR	NR	NR
Kyphoplasty					
Garcia-Maroto, 2015 Retrospective case series	Mean 11 months	KP only 12 months: 14.7% (11/75)	KP only **Any leakage, % (n/N)** 5.7% (7/122)	KP only **Mortality, % (n/N)** 3 months: 1.3% (1/75) 9 months: 9.3% (7/75)	NR
Molloy, 2016 Retrospective case series	3 months	NR	KP only **Any leakage, % (n/N)** 3 months 27.8% (44/158)	KP only **Mortality, % (n/N)** 3 months: 0% (0/158)	KP only No wound infections, chest infections, urine infections, myocardial infarctions, DVTs or Pes reported
Wu, 2022 Retrospective case series	1 year	NR	KP only **Any leakage, % (n/N)** 12.1% (26/215 levels) **Symptomatic leakage, % (n/N)** 0% (0/215 levels)	KP only **Mortality, % (n/N)** 18.8% (22/117)	KP only **Major AEs, % (n/N)** 0% (0/117) **Any AE, % (n/N)** 74.3% (87/117)
Wu, 2023	1 year	NR	KP only	NR	KP only

Author (year)	F/U	New fractures	Cement Leakage	Mortality	Other
Retrospective case series			**Any leakage, % (n/N)** 13.0% (12/92)		**Major AEs, % (n/N)** 0% (0/92) **Minor AEs, % (n/N)** 3.3% (3/92)
Zou, 2010 Retrospective case series	1 year	NR	KP only **Not clinically significant leakage, % (n/N)** 4.7% (2/43 levels)	NR	KP only No neurological, embolic, or cardiovascular complications were observed at final follow-up
Mixed Vertebroplasty and Kyphoplasty					
Burton, 2011 Retrospective case series	2 months	VP/KP only **Any new fracture, % (n/N)** 2 months 24.6% (100/407)† **Adjacent fracture, % (n/N)** 2 months 17.6% (72/408)	VP/KP only **Any leakage: 93.4%, % (n/N)** 2 months 127/136)‡	NR	VP/KP only **Reoperation, % (n/N)** 19.2% (78/407) **Other procedural AE, % (n/N)** 5.1% (7/136)‡ Authors report that the exact incidence of complications is unknown

AE = adverse events; F/U = follow-up; KP = kyphoplasty; NR = not reported; VP = vertebroplasty.
* n's back-calculated.
† 259 total new fractures in 100 patients.
‡ Total number of procedures with recorded AEs

APPENDIX L. Inclusion and Exclusion Criteria of Included Trials

Appendix Table L1. Inclusion and Exclusion Criteria for Vertebroplasty Trials in Patients with Osteoporotic Vertebral Compression Fractures

Author (year)	Inclusion	Exclusion
Vertebroplasty versus Sham		
Carli, 2023	• Aged ≥50 years • Focal back pain (score ≥5 on VAS) at the level of the VCF for ≥3 months at the time of spinal radiography • One or more VCFs on spinal radiograph with vertebral height loss of ≥15% between the fifth thoracic vertebra and fifth lumbar vertebra • Diminished bone density (T-score less than -1) on dual-energy x-ray absorptiometry scan • Bone edema of the fractured vertebra at MRI	• Severe cardiopulmonary condition • Untreatable coagulopathy • Suspected underlying disease
Clark, 2016; Diamond, 2020*	• Aged ≥60 years • Back pain <6 weeks • NRS score ≥7 • MRI confirming one or two recent fractures	• Inability to provide informed consent • Chronic back pain requiring opiate use • Substantial fracture retropulsion • Acute infection • Spinal malignancy • Neurological complications • >2 VCFs
Firanescu, 2018; Firanescu, 2019	• Aged ≥50 years • One to three VCFs • T5-L5 focal back pain at the level of fracture for ≤6 weeks • ≥5 on VAS • Diminished bone density (T-score -1 or less) on a dual energy x-ray absorptiometry scan • ≥15% loss of vertebral height • Bone oedema on MRI • Due to issues with recruitment six months after study initiation, also included patients with pain up to nine weeks	• Severe cardiopulmonary morbidity • Untreatable coagulopathy • Systemic or local spine infection • Suspected malignancy • Neurological symptoms • Inability to undergo MRI
Hansen, 2019	• Osteoporotic VCF from T5-L5 • ≥70 on VAS • <8 weeks of back pain	• History of malignancy • Age <50 years • Known allergy toward PVP components • Dementia as determined on the MMSE

	• MRI-STIR sequence showing edema using a Phillips Achieva 1.5 Tesla scanner	• Osteoporotic fractures of the long bones • Unable to consent
Kallmes, 2009; Comstock, 2013	• Age ≥50 years • Diagnosis of one to three painful osteoporotic VCFs between levels T4 and L5 • Inadequate pain relief with standard medical therapy • Current ≥3 on VAS pain • Fractures ≤1 year old • For fractures of uncertain age, additional requirement was marrow edema on MRI or increased vertebral-body uptake on bone scanning	• Evidence or suspicion of neoplasm in the target vertebral body • Substantial retropulsion of bony fragments • Concomitant hip fracture • Active infection • Uncorrectable bleeding diatheses • Surgery within previous 60 days • Lack of access to a telephone • Inability to communicate in English • Dementia
Buchbinder, 2009	• Recruited from the practices of general practitioners and specialists and from hospital inpatient and emergency departments. • Presence of back pain for ≤12 months • Presence of one or two recent vertebral fractures, defined as vertebral collapse of grade ≥1 according to grading system of Genant. • Edema, a fracture line, or both within the vertebral body on MRI	• Presence of >2 recent vertebral fractures • Spinal cancer • Neurological complications • Osteoporotic vertebral collapse >90% • Fracture through or destruction of the posterior wall • Retropulsed body fragment or bony fragments impinging on the spinal cord • Medical conditions that would make the patient ineligible for emergency decompressive surgery if needed • Previous vertebroplasty • Inability to give informed consent • Likelihood of noncompliance with follow-up
Vertebroplasty versus Usual care		
Blasco, 2012	• Acute, painful osteoporotic vertebral fractures from T4-L5 • Clinical onset <12 months • Confirmed by spine radiograph and by presence of edema on MRI or activity on bone scan • VAS score ≥4	• Untreatable coagulopathy • Active local or systemic infection • Current malignancy • Vertebral canal occupation by a fragment of the vertebral body or non-osteoporotic vertebral fracture • Active associated disorders (fibromyalgia or spondyloarthropathies) • Other disorders (dementia) that may interfere with correct assessment of quality of life and pain

Yang, 2016	• VCF after acute minor or mild trauma, with ≥5 on VAS of back pain • Low signal on T1-weighted and high signal on T2-weighted in MRI • Level of fracture of T5 or lower • Independent living sans use of wheelchair prior to trauma • Decreased bone mineral density (T score less than -1	• Chronic back pain prior to trauma • Suspected underlying malignant disease • Spine infection retropulsion of bony fragments • Spinal cord compression syndrome • Concomitant hip fracture • Severe cardiopulmonary comorbidity • Major coagulopathy
Leali, 2016	• Acute pain from severe spinal fracture • VCF height of the visible loss of vertebral body in radiography and standard • Evidence of osteoporosis to bone densitometry • Bone marrow edema of the affected VCF visible on MRI of the spine • Presence of evidence of an acute fracture imaging RM	• Pathological fracture due to myeloma/metastasis • Retropulsion mass of bone fragments in the spinal canal • Unstable cardiopulmonary conditions • Incurable Coagulopathy • Systemic infection in progress • Local infection of the spine • Radicular syndrome or spinal cord compression
Chen 2014	• Consecutive patients with chronic osteoporotic VCFs on MRI (low signal on T1-weighted and high signal on T2-weighted) • Persistent back pain for ≥3 months	NR
Farrokhi, 2011	• VCF with 10% to 70% loss of vertebral body height on x-ray of the spine • Severe back pain related to VCF that was refractory to analgesic medication for >4 weeks and <1 year • Focal tenderness on physical examination related to the level of vertebral fracture • Bone attenuation (T-score less than -2.5) on bone densitometry • Vacuum phenomenon or bone marrow edema of the vertebral fracture on MRI • Unresponsiveness to the medical therapy before entering the trial	• Uncorrected coagulopathy • Local or systemic infection • Secondary osteoporosis • Inability to give informed consent • Impaired cardiopulmonary function • Dementia • Posterior wall defect of the vertebral body on CT studies • Painless VCF • Spinal cancer • Traumatic fracture • Neurological complications
Klazen, 2010	• VCF on x-ray of the spine (minimal 15% loss of height) • Level of VCF T5 or lower • Back pain ≤6 weeks at time of x-ray • Aged ≥50 years • Bone edema on MRI • Focal tenderness on VCF level	• Severe cardio-pulmonary condition • Untreatable coagulopathy • Systemic or local infection of the spine • Indication of alternative underlying disease • Radicular and/or myelum compression syndrome

	• Decreased bone density T-scores less than -1	
Rousing, 2009	• Intractable pain because of either acute (<2 weeks) or subacute (between two and eight weeks) osteoporotic VCFs preventing the patient from taking care of oneself • Sufficient cognitive function to complete the study	• Aged <65 years • Uncorrected therapeutic anticoagulation • Senile dementia • Impaired cognitive functions or other cerebral disease • Infection in the spine or the overlying skin • Malignant disease • Bone metabolic disease • Fracture of tubular bone • Allergy to radiopaque agents
Voormolen, 2007	• VCF with height loss of vertebral body (minimal 15%) on x-ray of the spine • Invalidating back pain related to the VCF refractive to medical therapy for >6 weeks and <6 months • Focal tenderness on physical examination related to the level of the VCF • Bone attenuation T-scores less than -2.0 • Bone marrow edema of the affected VCF on MRI scan of the spine	• Poor cardiopulmonary condition • Untreatable coagulopathy • Ongoing systemic infection or local infection of the spine • Radicular and/or myelum compression syndrome • Indication of other underlying disease than osteoporosis • No informed consent
Yi, 2014[†]	• First time symptomatic osteoporotic VCFs • Serious low back pain • High signal in T2 MRI image • Diagnosed by severe osteoporosis	NR
Vertebroplasty versus Nerve Block		
Tan, 2023	• Aged ≥70 years admitted to hospital • Ambulatory prior to injury • <6 weeks from date of injury • ≥7 or more on 0-11 pain scale while standing • MRI confirmed oedema at the site of the vertebral fracture • Ability to adhere to study procedures and complete follow-up	• Debilitating chronic back pain not relieved despite opiate use • Substantial fracture retropulsion; acute infection, spinal malignancy • Three or more acute vertebral fractures • Bed bound prior fracture • Receiving palliative care • Lack of capacity and no consultee • Spinal deformity which contraindicates VP

Wang, 2016	• Severe pain caused by acute (fracture occurred within 2 weeks) or subacute (fracture occurred within two to eight weeks) VCFs.	• Age <55 years • Posterior margin body or spinal cord damage • Long-term use of anti-coagulant drugs • Senile dementia • Cognitive damage or other cerebral disease • Malignant disease • Spinal infection or skin disease • Metabolic bone disease • Tubular bone fractures or allergy to radiopaque agents
Vertebroplasty versus Kyphoplasty		
Wang, 2015	• Recent lumbar or thoracic VCFs (proven by radiographs and MRI) • Unsatisfactory pain relief (VAS ≥5) after ≥4 weeks of conventional therapy • Confirmed diagnosis of osteoporosis or osteopenia (proven by DEXA)	• Burst fractures • Infection • Radicular syndrome • Primary bone tumors • Spinal metastases
Liu, 2010; Liu 2015	• Confirmed osteoporotic VCF at the thoraco-lumbar junction	NR
Griffoni, 2020	• 55 years or older • Pain for ≥4 weeks attributable to one or more osteoporotic VCF below the fourth thoracic vertebra • No pain control by brace and analgesics • Still evidence of bone marrow edema of the fractured vertebra • Shared decision between physician and patient based also on the patient's tolerance of the brace	• Oncological treatment at the affected level • Infection • Stenosis of the vertebral canal at the affected level • Coagulation alterations • Allergy to iodinated contrast medium • Pregnancy
Evans, 2016	• Patients 50 years of age or older • Pain that had occurred in the previous 12 months attributable to one or more compression fractures of the vertebrae in the areas T4-L5 confirmed with a physical examination and imaging • Fractures detected on plain radiography • Pain from compression fractures of ≤5 out of 10 • Candidates for minimally invasive surgery • Able to successfully complete a battery of health questionnaires	• Neurological deficits related to the compression fracture • Other contraindications to vertebral augmentation • No history of surgery within 60 days • No history of open back surgery • No concomitant hip fracture, rib fracture, or sacral insufficiency fracture • No malignant tumor deposit (multiple myeloma), tumor mass, or tumor extension into the epidural space at the level of the fracture to be treated

	• Available and willing to participate in follow-up	
Endres, 2012	• Osteoporosis proven on DXA scan • Fresh painful single-level osteoporosis with sintering fractures in the middle and lower thoracic spine and lumbar spines • Conservative therapies (analgesics according to the WHO scheme, physiotherapy, physical therapy, orthotics adjustment) proven ineffective for ≥4 weeks	• No painful vertebral deformation • Considerable degenerative damage • Vertebral deformation (e.g., vertebra plana) • Tumor and metastases • Local or systemic infection • Untreated clotting disorder
Dohm, 2014	• Patients with One to three acute, painful VCF from T5 to L5 due to osteoporosis • Correlative clinical findings with edema on MRI, uptake on radionuclide bone scans, or acute vertebral height loss within 6 months by CT, MRI, or x-ray	• Back pain not attributable to VCF • More than three acute fractures • VCFs >6 months old • Fractures due to cancer or high-energy trauma • Required procedures other than balloon kyphoplasty or vertebroplasty for fracture stabilization • Contraindications such as irreversible coagulopathy or known allergies to bone cement or contrast • Evidence of local or systemic infection
Vogl, 2013	• Up to 3 painful VCFs between T4 and L5 that did not respond to ≥6 weeks of conservative treatment, or admitted to the hospital for acute back pain and treated soon after admission. • Required to have a confirmed diagnosis of osteoporosis or osteopenia.	• Burst fractures • Planar collapse • Vertebral bodies having inadequate space between endplates for cavity creation • Infection • Current cancer treatment at the affected level • Multiple myeloma • Prior cement placement at the affected level
Yi, 2014[†]	• First time symptomatic osteoporotic VCFs • Serious low back pain • High signal in T2 MRI image • Diagnosed by severe osteoporosis	NR
Wang, 2018	• Patients received bilateral resection of ovarian cancer • VCFs were diagnosed by x-ray examination • Bone density examination showed T≤2.5 • Patients signed written informed consent	• Patients with severe osteomalacia or osteoporosis • Patients with coagulation disorders

Wang, 2023	• Meet WHO diagnostic criteria for osteoporosis • Imaging examination shows that the posterior vertebral wall is intact and they can tolerate PVP or KP surgery	• Pathological fracture for other reasons • Fracture time >3 weeks • Severe conditions

CT = computed tomography; DXA = Dual X-ray Absorptiometry; KP = kyphoplasty; MRI = magnetic resonance imaging; NR = not reported; NRS = numerical rating scale; PVP = percutaneous vertebroplasty; VAS = visual analogue scale; VCF = vertebral compression fracture; VP = vertebroplasty.
* Diamond, 2020 was a subgroup analysis of patients with fractures less than 6 weeks old.
† Yi, 2014 includes three arms: vertebroplasty, kyphoplasty, and usual care. Inclusion criteria included under all relevant groupings for completeness.

Appendix Table L2. Inclusion and Exclusion Criteria for Kyphoplasty Trials in Patients with Osteoporotic Vertebral Compression Fractures

Author (year)	Inclusion	Exclusion
Kyphoplasty versus Usual Care		
Li, 2017	• ≥65 years of age • Course of the disease lasting 2 hours to 2 weeks • Patients having imaging features consistent with clinical manifestations and confirmed as having thoraco-lumbar compression fractures • Voluntarily provided written informed consent	• Patients with VCFs or lumbo-dorso pain that could not be rules out the causes of other potential lesions • Patients with malignancy, severe cardiopulmonary disease, administration of long-term steroids or systemic infection • Patients with coagulopathy which could not be cured
Yi, 2014*	• First time symptomatic osteoporotic VCFs • Serious low back pain • High signal in T2 MRI image • Diagnosed by severe osteoporosis	NR
Liu, 2019	• Elderly patients with multiple osteoporotic VCFs admitted to the hospital between January 2016 and 2017 • Underwent x-ray and computed tomography examination • Satisfied the diagnostic standards of multiple osteoporotic VCFs	NR
Wardlaw, 2009; Boonen, 2011; Van Meirhaeghe, 2013	• One to three VCFs from T5 through L5 • At least one fracture needed to have oedema assessed by MRI. At least one fracture had to show a 15% loss of height or more. Single fractures were to meet both criteria • Fractures due to osteopenia arising from primary or secondary osteoporosis, multiple myeloma, or osteolytic metastatic tumors were included†	• Younger than 21 years • Chronic fractures (estimated fracture age more than 3 months) • Pedicle fractures • Previous vertebroplasty • Neurological deficit • Radicular pain • Spinal cord compression • Canal narrowing

	• Patients with up to three contiguous or non-contiguous fractures at any level could be included if these additional fractures also had MRI signal changes, progressive height loss, or pseudoarthrosis • Back pain score ≥4 on 0-10 scale	• Taking uninterruptible anticoagulation therapy • Had allergies to kyphoplasty materials or contraindications to MRI • Dementia or were unable to walk before fracture (walking aids allowed) • VCFs were from primary bone tumors, osteoblastic metastases, or high energy trauma
Kyphoplasty versus Other Surgical Procedures		
Werner, 2013	• One or more osteoporotic VCF of the thoracic, thoracolumbar, or lumbar spine • Fresh fractures as demonstrated on MRI with use of transverse short tau inverted recovery • Marked pain	• Pregnancy • High-energy trauma • Poly-trauma • Previous major spine surgery within one year prior to admission • Bone metastases • Additional posterior spinal instrumentation

KP = kyphoplasty; MRI = magnetic resonance imaging; VCFs = vertebral compression fractures.
* Yi, 2014 includes three arms: vertebroplasty, kyphoplasty, and usual care. Inclusion criteria included under all relevant groupings for completeness.
† Nearly all patients had fractures due to osteoporosis.

Appendix Table L3. Inclusion and Exclusion Criteria for Kyphoplasty Trials in Patients with Malignancies (Berenson, 2011)

Author (year)	Inclusion	Exclusion
KP vs. Usual care		
Berenson, 2011	• Aged ≥21 years • Cancer and one to three painful VCFs (T5-L5) clinically diagnoses in conjunction with either plain radiographs or MRI • Pain numeric rating score ≥4 • Roland-Morris Disability Questionnaire score ≥10	• Osteoblastic tumors • Primary bone tumors (e.g., osteosarcoma • Plasmacytoma at the index VCF • Enrolled in a concurrent phase 1 investigational anticancer treatment study • Substantial clinical morbidities (aside from VCF and cancer) • VCF morphology deemed unsuitable for kyphoplasty by the treating physician (e.g., vertebra plana, comminuted fractures, fractures that did not have cortical integrity or that had posterior wall involvement, or those with epidural involvement and a tumor noted) • Needed additional surgical treatment for the index fracture • Needed treatment with high-dose steroids • Intravenous pain medication • Nerve blocks to control chronic back pain unrelated to index VCFs

KP = kyphoplasty; MRI = magnetic resonance imaging; VCFs = vertebral compression fractures.

APPENDIX M. FDA Approved Devices

Appendix Table M1: Summary of newly approved FDA devices since the 2016 signal update

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
Balloon Kyphoplasty			
GS Medical Co., Ltd.	Tracker Plus Kyphoplasty System	K211797 (10/28/2021)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes percutaneous vertebral augmentation. This system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
GS Medical Co., Ltd.	Tracker Kyphoplasty System	K192335 (12/4/2019)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes use during percutaneous vertebral augmentation. This system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Jiangsu Changmei Medtech Co., Ltd.	Kyphoplasty Balloon Catheter	K223709 (8/16/2023)	Intended to be used for the reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine during balloon kyphoplasty (for use with cleared spinal PMMA bone cements).
OK MediNet Korea Co., Ltd.	Kyphoplasty Balloon System	K221142 (4/21/2023)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes use during percutaneous vertebral augmentation. This system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Joline GmbH & Co.	Joline Kyphoplasty System Allevo	K192449 (5/27/2020)	Intended to be used for the reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine during balloon kyphoplasty (for use with cleared spinal PMMA bone cements).
Stryker Corporation	Stryker iVAS Elite Inflatable Vertebral Augmentation System (Stryker iVAS Elite Balloon Catheter)	K181752 (12/21/2018)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine. This includes use during percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements and Cortoss Bone Augmentation Material indicated for use during percutaneous vertebral augmentation procedures, such as kyphoplasty.
Hanchang Co. Ltd.	SpineKure Kyphoplasty System	K172871 (5/29/2018)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes use during percutaneous vertebral augmentation. This system is to be used with cleared

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
			spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
G-21 s.r.l.	Modified Winch Kyphoplasty (15 and 20 mm) 11 Gauge Balloon Catheters	K172214 (8/23/2017)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes use during percutaneous vertebral augmentation. This system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Pan Medical Ltd.	13G InterV Kyphoplasty Catheter (Micro) and 11G InterV Kyphoplasty Catheter (Mini-Flex)	K162453 (11/1/2016)	Intended to be used for reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine during balloon kyphoplasty (for use with cleared spinal PMMA bone cements).
Pan Medical Ltd	InterV Kyphoplasty Catheter (Balloon Length: 10, 15 and 20mm) InterV Kyphoplasty Catheter (Mini) (Balloon Length: 10, 15 and 20mm)	K150322 (3/6/2015)	Intended to be used for reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine during balloon kyphoplasty. (for use with cleared PMMA bone cements).
Imedicom Co. Ltd.	Medinaut Kyphoplasty System	K153296 (7/29/2016)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Carefusion	AVAflex Vertebral Balloon System	K151125 (11/24/2015)	Intended for the reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine for kyphoplasty (for use with CareFusion Radiopaque Bone Cement†).
Osseon LLC	Osseoflex SB Straight Balloon 10g/4ml Osseoflex SB Straight Balloon 10g/2ml	K150607 (4/9/2015)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
BM Korea Co. Ltd.	GUARDIAN-SG Inflatable Bone Expander System	K143006 (1/16/2015)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
			includes percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Zavation LLC	ZVPLASTY	K141419 (9/12/2014)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine, tibia, radius, and calcaneus. This includes percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Medtronic, Inc.	Kyphon Express II Inflatable Bone Tamps	K123771 (12/21/2012)	Intended to be used as a conventional bone tamp for the reduction of fractures and/or creation of a void in cancellous bone in the spine (including use during balloon kyphoplasty with a PMMA-based bone cement that is cleared for use in kyphoplasty procedures), hand, tibia, radius, and calcaneus.
Medtronic, Inc.	Kyphon Xpander II Inflatable Bone Tamp	K101864 (10/14/2010)	Intended to be used as a conventional bone tamp for the reduction of fractures and/or creation of a void in cancellous bone in the spine (including use during balloon kyphoplasty with a PMMA-based bone cement that is cleared for use in kyphoplasty procedures), hand, tibia, radius, and calcaneus.
Dragon Crown Medical Co., Ltd.	DCM Kyphoplasty System	K162283 (2/6/2017)	Intended to be used for the reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine. This includes use during percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Shanghai Kinetic Medical Co., Ltd.	KMC Kyphoplasty System	K113742 (9/17/2012)	Intended to be used for the reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine. This includes use during percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Soteira, Inc.	Shield Kyphoplasty System	K093477 (12/8/2011)	Intended to provide control of cement flow during injection of PMMA bone cement that has been cleared for use in vertebral augmentation for the treatment of acute, persistently painful (after a minimum of 6 weeks of conservative care), stable, anterior column osteoporotic compression fractures (wedge or concave) of the vertebrae at levels T4-L5 in the adult spine.
Mechanical Vertebral Augmentation			
EBI, L.P.	EBI Vertebroplasty Systems	K060148 (3/13/2006)	Indicated to deliver bone cement legally cleared for use in the spine for the treatment of compression fractures of a vertebral body.

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
Cook, Inc.	Vertefix Vertebroplasty Procedure Set	K042691 (11/8/2005)	Indicated for the fixation of vertebral compression fractures during a vertebroplasty procedure. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Benvenue Medical Inc.	Kiva VCF treatment system	K141141 (8/14/2014)	Indicated for use in the reduction and treatment of spinal fractures in the thoracic and/or lumbar spine from T6-L5. It is intended to be used in combination with the Benvenue Augmentation Cement Kit†.
Arthrocare Corporation	Parallax Contour Vertebral Augmentation	K100479 (9/21/2010)	Indicated for use during kyphoplasty or vertebral augmentation procedures to create a void in the vertebral body and fill the void with Parallax Acrylic Resin† (bone cement).
Arthrocare Corporation	Arthrocare Parallax Contour Vertebral Augmentation Device	K110183 (2/16/2011)	Indicated for use during kyphoplasty or vertebral augmentation procedures to create a void in the vertebral body and fill the void with Parallax Acrylic Resin† (bone cement). The painful pathological vertebral compression fractures may result from osteoporosis, benign or malignant lesions such as metastatic cancers and myeloma.
Neuro Therm Inc.	Parallax Balloon Inflatable Bone Tamp-10 Mm Balloon Parallax Balloon Inflatable Bone Tamp-15 Mm Balloon Parallax Balloon	K122503 (6/25/2013)	Intended to be used as a conventional bone tamp for the reduction and fixation of fractures and/or creation of a void in cancellous bone in the spine. This includes use during percutaneous vertebral augmentation. The Parallax Balloon Inflatable Bone Tamp is to be used with cleared spinal PMMA bone cement indicated for use during percutaneous vertebral augmentation, such as kyphoplasty.
Vexim SA	SpineJack Expansion Kit	K181262 (8/30/2018)	Indicated for use in the reduction of painful osteoporotic vertebral compression fractures. It is intended to be used in combination with Stryker Vertaplex and Vertaplex HV bone cement.
Hyprevention SAS	V-Strut Vertebral Implant	K191709 (3/5/2020)	Indicated for use in the treatment of vertebral fractures in the thoracic and lumbar spine from T9 to L5. It is intended to be used in combination with Teknimed F20 bone cement.
Depuy Synthes Spine	Synthes Synflate Vertevral Balloon System	K130146 (5/20/2023)	Intended to be used for the reduction of fractures and/or creation of a void in cancellous bone in the spine. This includes use during percutaneous vertebral augmentation. The system is to be used with cleared spinal PMMA bone cements indicated for use during percutaneous vertebral augmentation procedures, such as kyphoplasty.
DFINE, Inc.	StabiliT Vertebral Augmentation System	K090986 (12/30/2009)	Intended for percutaneous delivery of StabiliT ERx Bone Cement in vertebroplasty or kyphoplasty procedures in the treatment of pathological fractures of the vertebrae. Painful vertebral compression fractures may result from osteoporosis,

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
			benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Bone Cements			
Kyphon, Inc.	Kyphx HV-R Bone Cement	K041584 (07/07/2004)	Indicated for the treatment of pathological fractures of the vertebral body due to osteoporosis, cancer, or benign lesions using a balloon kyphoplasty procedure. Cancer includes multiple myeloma and metastatic lesions, including those arising from breast or lung cancer, or lymphoma. Benign lesions include hemangioma and giant cell tumor.
Globus Medical, inc.	Concord Plus Radiopaque Bone Cement	K162618 (01/23/2017)	Indicated for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Globus Medical, inc.	Concord Radiopaque Bone Cement	K042168 (01/14/2005)	Indicated for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Cardinal Health	Radiopaque Bone Cement	K043518 (05/11/2005)	Indicated for the fixation of pathological fractures of the vertebral body.
DePuy Spine, Inc.	Vertebroplastic Radiopaque Bone Cement	K043406 (07/15/2005)	Indicated for the treatment, using vertebroplasty or kyphoplasty procedures of pathological fractures of the vertebral body caused by osteoporosis, benign lesions (hemangioma), or malignant lesions (metastatic cancers, myeloma).
Teknimed SA	Cohesion Bone Cement	K103816 (02/04/2011)	Used for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures of the vertebral body may result from osteoporosis, benign lesions (hemangioma), or malignant lesions (metastatic cancers, myeloma).
Teknimed SA	Vertecem	K090435 (12/21/2009)	Used for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures of the vertebral body may result from osteoporosis, benign lesions (hemangioma), or malignant lesions (metastatic cancers, myeloma).
Teknimed SA	Opacity + Bone Cement	K080873 (08/28/2008)	Used for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures of the vertebral body may result from osteoporosis, benign lesions (hemangioma), or malignant lesions (metastatic cancers, myeloma).
Teknimed SA	Spine-Fix Biommetic Bone Cement	K043593 (03/17/2006)	Used for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
			fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Disc-O Tech Medical Technologies, LTD	Confidence Ex High Viscosity Bone Cement	K062424 (09/14/2006)	Indicated for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancer, myeloma).
Disc-O Tech Medical Technologies, LTD	Confidence High Viscosity Bone Cement	K060300 (06/21/2006)	Indicated for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty procedures. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancer, myeloma).
DFINE, Inc.	Stabili ERX Bone Cement	K090986 (12/30/2009)	intended for the treatment of pathological fractures of the vertebrae using a vertebroplasty or kyphoplasty procedure. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
DFINE, Inc.	Space CPSXL Bone Cement	K061531 (08/30/2006)	Indicated for the treatment of pathological fractures of the vertebrae using a vertebroplasty or kyphoplasty procedure. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Biomet Manufacturing Corporation	Cobalt V Radiopaque Vertebroplasty Bone Cement	K070015 (11/30/2007)	Indicated for the fixation of pathological fractures of the vertebral body due to osteoporosis, benign lesions and malignant lesions using a vertebroplasty or kyphoplasty procedure.
Stryker Corp.	Vertaplex HV High Viscosity Radiopaque Bone Cement	K192818 (03/31/2020)	intended to restore the integrity of the spinal column even in the absence of fusion for a limited time period in patients with advanced stage tumors involving the thoracic and lumbar spine in whom life expectancy is of insufficient duration to permit achievement of fusion. Vertaplex HV High Viscosity Radiopaque Bone Cement and the ES2 Augmentable Spinal System are for use together at spinal levels where the structural integrity of the spine is not severely compromised.
Stryker Corporation	Verteplex High Viscosity (HV) Radiopaque Bone Cement	K150582 (06/12/2015)	Indicated for the fixation of pathological fractures of the vertebral body using vertebroplasty or kyphoplasty. It is also indicated for the fixation of pathological fractures of the sacral vertebral body or ala using sacral vertebroplasty or sacroplasty. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Stryker Corporation	Stryker Vertaplex Radiopaque Bone Cement	K072118 (12/07/2007)	Indicated for the fixation of pathological fractures of the vertebral body using vertebroplasty of kyphoplasty procedures. Painful vertebral compression

Manufacturer	Procedure/Device	510(k) no. (Date cleared)	Indication*
			fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).
Medtronic, Inc.	Kyphon Xpede Bone Cement (update)	K163032 (02/27/2017) K151227 (11/16/2015) K102397 (02/28/2011)	Indicated for the treatment of pathological fractures of the vertebral body due to osteoporosis, cancer, or benign lesions using a cementoplasty (i.e. kyphoplasty or vertebroplasty) procedure. Cancer includes multiple myeloma and metastatic lesions, including those arising from breast or lung cancer, or lymphoma. Benign lesions include hemangioma and giant cell tumor. Pathological fracture may include a symptomatic vertebral body microfracture (as documented by appropriate imaging and/or presence of a lytic lesion) without obvious loss of vertebral body height.
Medtronic, Inc.	Kyphon (R) HV-R Bone Cement	K180700 (05/18/2018) K160983 (08/24/2016) K093828 (08/12/2010)	Indicated for the treatment of pathological fractures of the vertebral body due to osteoporosis, cancer, or benign lesions using a cementoplasty (i.e. kyphoplasty or vertebroplasty) procedure. Cancer includes multiple myeloma and metastatic lesions, including those arising from breast or lung cancer, or lymphoma. Benign lesions include hemangioma and giant cell tumor. Pathological fracture may include a symptomatic vertebral body microfracture (as documented by appropriate imaging and/or presence of a lytic lesion) without obvious loss of vertebral body height.
Tecres S.p.A	Kyphon VuE Bone Cement	K220131 (04/18/2022)	Indicated for the treatment of pathological fractures of the vertebral body due to osteoporosis, cancer, or benign lesions using a cementoplasty (i.e. kyphoplasty or vertebroplasty) procedure. Cancer includes multiple myeloma and metastatic lesions, including those arising from breast or lung cancer, or lymphoma. Benign lesions include hemangioma and giant cell tumor. Pathological fracture may include a symptomatic vertebral body microfracture (as documented by appropriate imaging and/or presence of a lytic lesion) without obvious loss of vertebral body height.
Orthovita, Inc.	Cortoss Bone Augmentation Material	K080108 (06/05/2009)	Indicated for the fixation of pathological fractures of the vertebral body using vertebral augmentation. Painful vertebral compression fractures may result from osteoporosis, benign lesions (hemangioma), and malignant lesions (metastatic cancers, myeloma).

PMMA = polymethylmethacrylate.
* Information on indications abstracted directly from 510(k) Premarket Notification summaries in the US FDA medical device databases. No device summaries included contraindications.
† Does not appear in FDA 510(k) database.

APPENDIX N. Information for Economic Studies

Appendix Table N1: Summary of Selected Systematic Reviews of Economic Studies

SR, Search dates	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Pron, 2022[10]; Pron, 2023[9] * Up to May 2021 Medline, Embase, CINAHL, EconLit, Cochrane, DARE	VP vs. usual care, KP vs. usual care, VP vs. KP Osteoporotic VCFs	Cost effectiveness EQ-5D-based QALYs, SF-6D-based QALYs	10† Cost-effectiveness reports	Yes (BMJ Checklist)	No	**ICERs for VP vs. Usual Care:** A CUA analysis in one UK trial with a 1-year time horizon resulted in an ICER per QALY of USD $33,395. In three other trials, ICERs per QALY ranged from USD $39,774 (inpatient) to USD $12,293 (outpatient) in the US, a cost saving of USD -$3,273 in the UK, and USD $16,221 in England and Wales. A CEA analysis reported ICERs ranging from USD $2,452 to USD $13,543 per life-year gained across age-gender subgroups with a 3-year time horizon. **ICERs for KP vs. Usual Care:** CUA analyses across three trials reported ICERs of USD $39,122 for a 3-year time horizon in Japan, USD $134,043 for a 2-year time horizon in Sweden, and USD $17,745 for a lifetime time horizon in the UK. In three other trials, ICERs per QALY ranged from USD $10,922 (outpatient) to USD $43,455 (inpatient) in the US, USD $3,954 in the UK, and USD $32,442 in England and Wales. A CEA analysis reported ICERs per life-year gained ranging from USD $1,863 to USD $6,687 across age-gender subgroups with a 3-year time horizon. **ICERs for VP vs. KP:** CUA analyses across two trials indicated that the cost-effectiveness relationship was variable and highly dependent on modeling scenarios and sensitivity analyses.

SR, Search dates	Interventions Condition	Primary Outcomes	Evidence Base	Risk of Bias Assessed (Tool)	Quantitative Synthesis	Primary Conclusions
Borgström, 2015[1] Up to March 2013 Embase, PubMed, EconLit, NHS EED	KP vs. non-surgical management	Cost effectiveness EQ-5D-based QALYs	4 Cost-effectiveness reports 1 HTA	None	No	**ICERs for KP vs. Non-Surgical Management:** One trial reported an ICER of €19,706, another reported €10,900 per QALY, a third reported €3,337 per QALY, and a fourth reported €92,154 per QALY. **ICERs for VP vs. Non-Surgical Management:** One trial reported an ICER of €22,685 per QALY.

BMJ = British Medical Journal; CEA = cost effectiveness analysis; CUA = cost utility analysis; EQ-5D = EuroQol 5-Dimensions; HTA = Health Technology Assessment; ICER = incremental cost-effectiveness ratio; KP = kyphoplasty; QALY = quality adjusted life-year; SF-6D = Short Form-6 Dimensions; SR = systematic review; USD = United States Dollar; VCF = vertebral compression fracture; VP = vertebroplasty.

* Includes three of the same reports as Pron, 2023, and is therefore summarized here.

† One study (Takura, 2017) was not truly comparative, and therefore not included in the present HTA.

APPENDIX O. Economic Study Tables

Appendix Table O1: U.S. Based Full Economic Studies

Study Characteristics	Edidin 2012	Hopkins 2020
Population	Medicare population (≥65 years) with newly diagnosed OVCF; demographics/patient characteristics other than age, sex not reported **Total Population: N=1,007,070** **Exclusion criteria (14.7% w/ N=148,092):** <65 years old, VCF diagnosis in prior 12 months, patients enrolled in health maintenance organizations, and patients not enrolled in both Part A (hospital insurance) and Part B (medical insurance) of Medicare. **Study cohort: N=858,978** (i.e., VCF patients from Medicare database: 2005–2008). Charlson comorbidity index (Dartmouth-Manitoba version) calculated. VCF patients stratified in 2 patient groups: ▪ "operated" patients in treatment group (21.3% w/ n=182,946 patients): Kyphoplasty (13.9% w/ n=119,253 patients) and Vertebroplasty (7.4% w/ 63,693 patients) ▪ "non-operated" patients in control group (78.7% w/ 676,032 patients): no surgical procedure In cost analysis: ▪ 5,670 kyphoplasty patients and 3,539 vertebroplasty patients	Between 2014 and 2016, **75,524 patients diagnosed with VCF (source: CMS).** **Matched Group 1 (KP inpatients vs CMM): n=2,071 x2** **Matched Group 2 (KP outpatients vs CMM): n=3,708 x 2** **Matched Group 3 (VP inpatients vs CMM): n=720 x 2** **Matched Group 1 (VP outpatients vs CMM): n=1,042 x 2** **Exclusion criteria:** no diagnosis of cancer, continued Medicare enrollment (w/ allowed 30d gap), no history of KP/VP procedures in 6 months baseline. **Age:** 65+ **Female (%):** female patients % ranges from 78.4% (KP outpatients) to 82.3% (both KP and VP inpatients). **Charlston Score Group (%):** 42.4% for KP inpatients having a score of 2+, 56.6% for KP outpatients having a score of 0, 46.8% for VP inpatients having a score of 2+, 55.9% for VP outpatients having a score of 2+.

Study Characteristics	Edidin 2012	Hopkins 2020
	▪ 57,809 non-operated patients	**Diagnosis of Osteoporosis (%):** ranges from 56.0% (VP outpatients) to 70.8% (KP inpatients).
Intervention(s)	Vertebroplasty, kyphoplasty	Vertebroplasty, kyphoplasty
Comparator(s)	Nonsurgical management	Conservative medical management
Country	US	US
Funding	Medtronic	Medtronic
Study design	Cost-effectiveness: Cost per life-year gained;	CUA
Perspective	Payer (Medicare)	Payer (Medicare)
Time horizon	NR (appears to be 3 years (based on cost source); survival modeling to 48 months	Lifetime (2 year-simulation)
Analytic model	Based on survival analysis; models adjusted for age, gender, race, census region, public support for Medicare premium, patient health status, fracture type (pathologic, traumatic), setting (inpatient, outpatient), per capita income, year of diagnosis	Markov Model
Effectiveness outcome	Life-year gained	QALY
Effectiveness outcome components	Survival analysis/life expectancy Weibull survival model (adjusted for comorbidities)	Euroqol 5-dimension (EQ-5D) (from FREE-2 trial) and recalculation of patient utility values using US-specific preference weights from published literature.
Source for effectiveness data	Life years gained LYG -calculated directly from the survival characteristics of patient cohorts in the Medicare populations, longitudinal administrative claims data from CMS.	Medicare claims payments, w/ propensity-score matching performed for KP and VP vs CMM.
Costing year	2010 USD 3-year analysis (2005-2008) adjusted to 2010 (November) US dollars	2016 USD
Currency	USD	USD
Discounting	3% per annum (costs and outcomes)	3% per annum (costs and outcomes)

Study Characteristics	Edidin 2012	Hopkins 2020
Components of cost data	Cumulative Medicare payment; Inpatient and outpatient claim data (i.e., cumulative Medicare payment for each patient for up to 3 years following VCF diagnosis), physician/carrier, skilled nursing facility, home health agency, hospice and durable medical equipment claims files.	Inpatients and outpatient claim data, home health, skilled nursing facility, hospice, inpatient rehabilitation, and readmission.
Cost sources	Median cumulative Medicare payment for each patient was identified for up to 3 years following VCF diagnosis;	CMS
Sensitivity analysis	Discount rate 0% and 5% for sensitivity by age	One way sensitivity analysis with HR = 1(no mortality benefit), probabilistic sensitivity analysis
QHES	53/100	82/100
Results:		
Cost	**Cumulative median costs for:** ▪ Kyphoplasty ranged from USD $15,410 to $26,410 in Year 1; from USD $18,890 to $30,470 in Year 2; from USD $20,530 to $32,790 in Year 3 (across all age-gender groups). ▪ Vertebroplasty ranged from USD $11,520 to $25,080 in Year 1; from USD $15,510 to $32,390 in Year 2; from $18,190 to $36,770 in Year 3 (across all age-gender groups).	**KP; inpatient (n=2,071)** Discounted Costs: USD $58,986 QALY: 2.08 **KP; outpatient (n=3,708)** Discounted Costs: USD $32,972 QALY: 3.88 **VP; inpatient (n=720)** Discounted Costs: USD $61,342 QALY: 2.23 **VP; outpatient (n=1,042)** Discounted Costs: USD $32,301 QALY: 3.71
Cost comparator(s)	▪ Non-operated patients ranged from USD $4,840 to $9,960 in Year 1; from USD $6,900 to $12,850 in Year 2; from USD $7,950 to $14,290 in Year 3 (across all age-gender groups).	**CMM; inpatient (n=2,071)** Discounted Costs: USD $32,324 QALY: 1.47 **CMM; outpatient (n=3,708)** Discounted Costs: USD $24,234 QALY: 3.08 **CMM; inpatient (n=720)** Discounted Costs: USD $31,005 QALY: 1.47

Study Characteristics	Edidin 2012	Hopkins 2020
		CMM; outpatient (n=1,042) Discounted Costs: USD $23,789 QALY: 3.02
ICER: Cost per life-year = ratio between the discounted incremental cost and the discounted years of life gained.	Kyphoplasty vs. Non-surgical **At 3% discount rate:** Cost/LYG ranges from USD $1,863 to $3,751 for female patients across all age groups. Cost/LYG ranges from USD $2,318 to $6,687 for male patients across all age groups. Vertebroplasty vs. Non-surgical **At 3% discount rate:** Cost/LYG ranges from USD $2,452 to $6,603 for female patients across all age groups. Cost/LYG ranges from USD $6,621 to $10,602 for male patients across all age groups. Vertebroplasty vs kyphoplasty **At 3% discount rate:** Cost/LYG ranges from minus USD $284 to $2,399 for female patients across all age groups. Cost/LYG ranges from minus USD $4,878 to $2,763 for male patients across all age groups.	ICER for KP inpatient vs CMM: USD $43,455/QALY gained. ICER for KP outpatient vs CMM: USD $10,922/QALY gained. ICER for VP inpatient vs CMM: USD $39,774/QALY gained. ICER for VP outpatient vs CMM: USD $12,293/QALY gained.
One-way SA	Kyphoplasty vs. Non-surgical: **At 0% discount rate:** Cost/LYG ranges from USD $1,425 to $3,340 for female patients across all age groups. Cost/LYG ranges from USD $1,935 to $6,167 for male patients across all age groups. **At 5% discount rate:** Cost/LYG ranges from USD $2,219 to $4,046 for female patients across all age groups. Cost/LYG ranges from USD $2,608 to $7,050 for male patients across all age groups. Vertebroplasty vs. Non-surgical **At 0% discount rate:**	**Mortality benefit "turned off" (HR = 1):** ICER for KP inpatient vs CMM: $283,579/QALY gained. ICER for KP outpatient vs CMM: USD $55,485/QALY gained. ICER for VP inpatient vs CMM: USD $314,958/QALY gained. ICER for VP outpatient vs CMM: USD $53,077/QALY gained. Results confirmed mortality was key driver of model results. **Varying KP and VP utility weights within 95% CI limits:** ICER ranges for KP inpatient: USD $37,152 to

Study Characteristics	Edidin 2012	Hopkins 2020
	Cost/LYG ranges from USD $2,243 to $6,311 for female patients across all age groups. Cost/LYG ranges from USD $7,796 to $56,435 for male patients across all age groups. **At 5% discount rate:** Cost/LYG ranges from USD $2,599 to $6,802 for female patients across all age groups. Cost/LYG ranges from USD $6,748 to $13,651 for male patients across all age groups. Vertebroplasty vs kyphoplasty **At 0% discount rate:** Cost/LYG ranges from minus USD $287 to $2,179 for female patients across all age groups. Cost/LYG ranges from minus USD $4,511 to $2,555 for male patients across all age groups. **At 5% discount rate:** Cost/LYG ranges from minus USD $279 to $2,555 for female patients across all age groups. Cost/LYG ranges from minus USD $5,136 to $2,905 for male patients across all age groups.	$53,321. ICER ranges for KP outpatient: USD $8,698 to $15,022. ICER ranges for inpatient VP: USD $34,673 to 47,357. ICER ranges for VP outpatient: USD $9,599 to 17,566. **Varying age-specific risks of subsequent fractures within 95% CI limits:** ICER ranges for KP inpatient: USD $42,375 to $45,400. ICER ranges for KP outpatient: USD $10,004 to $11,900. ICER ranges for inpatient VP: USD $38,516 to $41,494. ICER ranges for inpatient KP: USD $10,712 to $13,915.
Other SA	NR	Probabilistic analysis: 80% probability of C/E at $50K WTP threshold for KP across all model simulations (models appear to assume mortality benefit).
Author's Conclusion	• Vertebral compression fracture treatments are cost effective in the Medicare population when compared with nonsurgical management, while among patients for whom surgical treatment was indicated, kyphoplasty was found to be cost effective, and perhaps even cost saving, compared with vertebroplasty • Kyphoplasty group had the longest median life expectancies, followed by vertebroplasty patients, and then by non-operated patients	• VP and KP are more expensive vs. CMM short term but may be C/E among patients eligible for surgery at US WTP threshold. • Both KP and VP C/E vs. CMM at a US WTP threshold of $50,000/QALY in 80% and 100% of 500 model simulations, respectively.
Limitations	• Limited sensitivity analyses; no evaluation of assumptions or drivers of cost-effectiveness.	• Only 2-year of follow-up data and extrapolation while modeling over lifetime

Study Characteristics	Edidin 2012	Hopkins 2020
	• Administrative data (possibility of misclassification unclear) and proportions of patients with comorbidities that might contribute to mortality by treatment group. Causal association cannot be inferred. • No demographic, patient characteristics or comorbidity data by treatment group beyond age and sex were reported; cannot assess balance between groups or adjustment for important prognostic factors • Unclear relationship between survival observed in Medicare data and use of vertebroplasty or kyphoplasty; accuracy of survival data unclear	horizon). • Administrative data (possibility of selection bias, confounding, misclassification unclear) and proportions of patients with comorbidities that might contribute to mortality by treatment group. Causal association cannot be inferred.

C/E = cost-effectiveness; CMM = conservative medical management; CMS = Centers for Medicare and Medicaid Services; CUA = cost utility analysis; EQ-5D = EuroQol 5-dimension; HR = hazard ratio; ICER = incremental cost-effectiveness ratio; KP = kyphoplasty; LYG = life year gained; NR = not reported; OVCF = osteoporotic vertebral compression fracture; QALY = quality adjusted life year; QHES = Quality of Health Economic Studies; SA = sensitivity analysis; USD = United States dollar; VCF = vertebral compression fracture; VP = vertebroplasty; WTP = willingness-to-pay.

Appendix Table O2: Non-U.S. Full Economic Studies: Government Reported Studies (Health Quality Ontario and UK National Health Service)

Study Characteristics	Cameron 2016 (Health Quality Ontario)	Stevenson 2014 (UK National Health Service HTA)
Population	VCF due to cancer **Population**: N=72 patients (Ontario hospital) aged 18+ with cancer (lung, breast, prostate, multiple myeloma, etc.). N=72 patients (n=36 for cancer and KP; n=9 for cancer and VP; n=27 for cancer and hybrid procedures). 90% of patients were outpatients. Mean age: 65 years old.	Osteoporotic vertebral compression fracture 70 % female patients Patient age: 70 years old **Exclusion criteria:** non-randomized studies (except for adverse events
Intervention(s)	Vertebroplasty, kyphoplasty	Vertebroplasty, kyphoplasty
Comparator(s)	NSM (NSAIDs, bed rest, radiation therapy, braces, wheelchair)	Operative placebo with local anesthesia and optimal pain management.
Country	Canada	England and Wales
Funding	Health Quality Ontario	National Institute for Health Research Health Technology Assessment Programme.
Study design	CUA*	CUA†
Perspective	Healthcare (Ontario Ministry of Health and Long-Term Care).	Healthcare
Time horizon	1 year	Lifetime
Analytic model	Markov Model	Mathematical Model/Markov model
Effectiveness outcome	QALY	QALY
Effectiveness outcome components	SF-36	EQ-5D AQoL DPQ MMSE SF-36 QUALEFFO
Source for effectiveness data	Published sources (validated by expert opinion) Utilities estimated from an industry-sponsored abstract that mapped SF-36 scores from the CAFE trial	Following Trials: Blasco 2012 Buchbinder 2009 Farrokhi 2011 FREE INVEST Liu 2010 Rousing 2009

Study Characteristics	Cameron 2016 (Health Quality Ontario)	Stevenson 2014 (UK National Health Service HTA)
		VERTOS VERTOS II
Costing year	2015	NR (SR)
Currency	CAD	GBP
Discounting	5% per annum on QALY and costs	3.5% per annum on QALY and costs (J&J model)
Components of cost data	Procedure costs, hospital costs (in-hospital stay, ED visit, day procedure costs), non-hospital costs, direct costs (nursing, diagnostic imaging, pharmacy, and laboratory), physician costs.	Consultant costs, radiology, anesthetist, and staff costs, sedation, surgical consumables, and procedure costs, hospitalization costs per day.
Cost sources	Ontario administrative sources, Ontario Schedule of Benefits for Physician Benefits, Ontario Schedule of Laboratory Fees, Ontario Drug Benefit Formulary, Ontario Case Costing Initiative database.	List prices from manufacturer
Sensitivity analysis	One-way sensitivity analysis using QALY benefit, time horizon, cancer type, mortality benefit, standardized costs. PSA using Monte Carlo simulation.	Extensive Univariate sensitivity analyses and probabilistic SA
QHES	80/100	99/100
Results:		
Cost / QALY	**KP avg procedure costs: CAD $3,695 (SD 1,432)** KP costs 1 level: CAD $2,866 (n=15 patients) KP costs 2 levels: CAD $3,164 (n=9) KP costs >2 levels: CAD $5,134 (n=12) **VP avg procedure costs: 738 CAD (SD 522 CAD)** VP costs 1 level: CAD $166 (n=1 patient) VP costs 2 levels: CAD $235 (n=1) VP costs >2 levels: CAD $891 (n=7) **KP adjusted Procedure costs, direct costs, physician fees:** CAD $8,877 **VP adjusted Procedure costs, direct costs, physician fees:** CAD $2,879 **KP Expected costs:** CAD $24,320. QALY=0.414 **VP Expected costs:** CAD $20,942. QALY=0.414	**Assuming no mortality benefit for KP, VP or OPLA:** VP procedure cost: GBP £6,118 VP QALY: 4.91 KP procedure cost: GBP £8,244 KP QALY: 4.91 **Assuming a relative risk of mortality:** VP procedure cost: GBP £6,210 VP QALY: 5.04 KP procedure cost: GBP £8,507 KP QALY: 5.27

Study Characteristics	Cameron 2016 (Health Quality Ontario)	Stevenson 2014 (UK National Health Service HTA)
Cost / QALY of comparator(s)	**NSM Expected costs:** CAD $17,073 NSM QALY: 0.197	**Assuming no mortality benefit for KP, VP or OPLA:** OPLA procedure cost: GBP £6,118 OPLA QALY: 4.83 OPM procedure cost: GBP £6,181 OPM QALY: 4.74 **Assuming a relative risk of mortality:** OPLA procedure cost: GBP £6,163 OPLA QALY: 4.89 OPM procedure cost: £6,181 OPM QALY: 4.74
ICER	**ICER for KP patient vs NSM:** **1 year:** CAD $33,471 /QALY gained. **ICER for VP patient vs NSM:** **1 year:** CAD $17,870/QALY gained.	**Assuming no mortality benefit for KP, VP or OPLA:** VP: Dominating OPLA: Dominated OPM: Dominated KP: Dominated **At WTP of £20,000 per QALY gained.** **Assuming a relative risk of mortality:** OPM: Dominated VP: GBP £312 KP: GBP £9,806
One-way SA	Variations of mortality reduction, QALY benefit, time horizon, and discount rate on ICER.	Variations of patients' characteristics (age), gender and T-score. Variations of hospitalization costs, cement price, equipment and procedure costs. Variations of discount rates, bisphosphonate usage and wane period. Variations of time of convergence (starting at 1 months, from 12 to 24 months, and from 24 to 36 months). Variations in mortality and fracture rates. Variations in EQ-5D data used from trials.
Other SA	Probabilistic SA (Monte Carlo simulation).	Probabilistic SA
Author's Conclusion	• Cost-effectiveness of KP and VP vs NSM although authors used restrictive assumptions regarding QALY.	• The cost-effectiveness ratios of the interventions were driven by the scenario chosen. If a differential mortality

Study Characteristics	Cameron 2016 (Health Quality Ontario)	Stevenson 2014 (UK National Health Service HTA)
		effect was chosen, then KP consistently had a cost-per-QALY-gained ratio below £20,000 • If a pooled beneficial effect was used, then VP consistently had a cost-per-QALY-gained ratio below £10,000. • Where no mortality effect was assumed the way that the utility was derived influenced CU • When data from the two high-quality blinded trials (Buchbinder et al. and INVEST) were used then the cost-per-QALY-gained ratios for VP and KP were often greater than £20,000, depending on the other assumptions made
Limitations	• VP one-way sensitivity results not shown in study. • Growth rate assumption in forecasting number of vertebral augmentation procedure might be limited. • Short time horizon	• No causal inference of VCF effect on mortality effect is possible.

AQoL = Assessment of Quality of Life; CAD = Canadian Dollar; CU = cost utility; CUA = cost utility analysis; DPQ = Dallas Pain Questionnaire; ED = emergency department; GBP = Great Britain Pound; HTA = health technology assessment; ICER = incremental cost-effectiveness ratio; KP = kyphoplasty; MMSE = Mini Mental State Examination; NR = not reported; NSM = non-surgical management; OPLA = operative placebo with local anesthesia; OPM = optical pain management; VP = percutaneous vertebroplasty; QALY = quality adjusted life year; QHES = Quality of Health Economic Studies; QUALEFFO = Quality of Life Questionnaire of the European Foundation for Osteoporosis; PSA = probabilistic sensitivity analysis; SA = sensitivity analysis; SF-36 = Short Form questionnaire-36 items; SR = systematic review; VCF = vertebral compression fracture; VP = vertebroplasty.

* Report also includes a systematic review and budget analysis

† Report also includes analyses of efficacy and various investigations relevant to the National Health Service.

Appendix Table O3: Non-U.S. Full Economic Studies: Non-Government Studies

Study Characteristics	Svedbom 2013 Sweden/UK	Fritzell 2011 Sweden
Population	Average age: 70 years old All female patients (OVCF)	**N=67 Swedish patients with an OVCF diagnosis.** n=63 in final analyses out of which n=32 in treatment group (i.e., KP) and n=31 in control group (i.e., UC). **Female patients:** 71% in treatment group (i.e., KP) and 78% in control group. **Mean age:** 72 (SD 10.1) in treatment vs 75 (SD 9.7) in control group. Fracture age: 1.2 months (SD 0.8) in both treatment and control groups). **Inclusion criteria: >**21 years old, severe thoracic and/or low back pain due to <3 months VCF, confirmed on MRI, 1 to 3 fractured vertebrae (Th5-L5), adjacent or separate levels. **Exclusion criteria:** relevant comorbidity, no understanding of Swedish language, previous spine surgery.
Intervention(s)	Vertebroplasty or kyphoplasty*	Kyphoplasty
Comparator(s)	Nonsurgical management	Usual Care
Country	United Kingdom	Sweden
Funding	Medtronic, Servier, Proctor and Gamble/Alliance, Eli Lilly, Merck, Sharp and Dohme, GSK/Roche, Amgen, Novartis	Medtronic
Study design	CUA	CUA
Perspective	Healthcare system perspective	Societal perspective (i.e., including both direct and indirect costs)
Time horizon	Lifetime	2 years
Analytic model	Markov simulation model	Unclear
Effectiveness outcome	QALY	QALY
Effectiveness outcome components	EQ-5D utilities and scores.	EQ-5D utilities and scores.
Source for effectiveness data	Other clinical and previous economic studies/published literature.	Other clinical and previous economic studies/published literature.
Costing year	2009	2008 (1€= SEK 9.6 and 1$=SEK 6.6). Between February 2003 and December 2005.

Study Characteristics	Svedbom 2013 Sweden/UK	Fritzell 2011 Sweden
Currency	GBP	SEK
Discounting	3.5% per annum (costs and outcomes)	NR
Components of cost data	Procedure costs (devices, consumables), radiologist, surgeon, nurse, MRI, Rx Spine, Blood test, ECG, Drug costs, fracture costs (analgesics, referrals, GP), bisphosphonate costs, hospitalization cost.	Procedure costs, hospital costs (e.g., physician costs, X-ray, CT, MRI, corset, anesthesia, rehabilitation), primary care (e.g., PT, chiropractors, pharmaceutical costs. Indirect costs such as travel, shopping, house cleaning.
Cost sources	Published literature, NHS, NHS Drug Tariff.	Patients' "Cost diary", Hospital's billing systems, Swedish National Board of Health and Welfare, interregional county hospital price lists.
Sensitivity analysis	One-way sensitivity analysis (variations of mortality (100% to 0%), QALY, risk of fracture w/ treatment, medication, cost, age, and discount rates).	One-way sensitivity analysis (variations of procedure costs, other direct costs, and QALY).
QHES	84/100	79/100
Results:		
Cost / QALY	KP Total costs: GBP £9,313 KP QALY: 5.473	KP cost per patient: SEK 160,017 (SD 151,083) equivalent to €16,668 (SD 15,735). KP cost/QALY gained: SEK 884,682 = €92,154 and $134,043.
Cost / QALY of comparator(s)	NSM Total costs: GBP £7,969 NSM QALY: 4.976 VP Total costs: GBP £7,157 VP QALY: 5.338	UC cost per patient: SEK 84,816 (SD 40,954) equivalent to €8,835 (SD 4,266).
ICER	**ICER for KP patient vs NSM:** GBP £2,706/QALY gained. **ICER for KP patient vs VP:** GBP £15,982/QALY gained.	KP cost/QALY gained: SEK 884,682 (€92,154 and $134,043).
One-way SA	Mortality (from no mortality reduction to full benefit): ICERs range from GBP £3,104 to £5,667/QALY (KP vs NSM) Increase risk of additional OVCF w/ KP: ICERs range from GBP £3,503 to £7,904/QALY (KP vs NSM) QALY: ICERs range from GBP £3,059 to £22,668 (KP vs NSM) Discount rate (0% and 7%): ICERs range from GBP £2,224 to £3,193 (KP vs NSM) Use of bisphosphonate: ICERs range from GBP £2,764 to £5,743 (KP vs NSM) Age (60- and 80-year old patients): ICERs range from GBP £2,373 to £5,905 (KP vs NSM)	Sensitivity analysis after varying costs: ICERs ranged from SEK 622,800 (€64,875 and US $94,364) to SEK 745,812 (€77,689 and US $113,002). Sensitivity analysis after varying the QALY benefit from 0.085 to 0.21 (from the FREE trial): resulted into an ICER of SEK 359,146 (€37,411 and US $54,416).

Study Characteristics	Svedbom 2013 Sweden/UK	Fritzell 2011 Sweden
	Costs (i.e., VP set at 0%, 50% and 75% of KP price): ICERs range from GBP £2,706 to £5,667 Length of stay in hospital (3 to 9 days): ICER/QALY gained dominating for KP vs NSM	
Other SA	Probabilistic SA	NR
Author's Conclusion	ICERs of KP vs VP and NSM fall below the UK WTP threshold of GBP £20,000-30,000/QALY.	Not possible to demonstrate that KP is C/E compared to UC in patients with OVCF.
Limitations	• Adverse events not considered • Female patients only. • Authors did not address risk of bias.	• Potential selection bias issue. • Short period (1 year). • Reliance on patients' "cost diaries".

C/E = cost-effective CUA = cost utility analysis; EQ-5D = EuroQol 5-dimension; GBP = Great Britain Pound; ICER = incremental cost-effectiveness ratio; KP = kyphoplasty; NHS = National Health Service; NR = not reported; NSM = non-surgical management; OVCF = osteoporotic vertebral compression fracture; VP = percutaneous vertebroplasty; QALY = quality adjusted life year; QHES = Quality of Health Economic Studies; SA = sensitivity analysis; SD = standard deviation; SEK = Swedish Krona; VCF = vertebral compression fracture; WTP = willingness-to-pay.
* Results focused on KP.

Appendix Table O3 (continued): Non-U.S. Full Economic Studies: Non-Government Studies

Study Characteristics	Takahashi 2019 Japan
Population	**Population at enrollment:** N=116 patients for KP vs N=485 patients for NSM. Average patients age: 78.3 years old for KP vs 77.7 for NSM (*p-value*=0.456). **Inclusion criteria:** presence, severity, and duration of pain. VAS pain score >=4 T scores <=-1 n=100 patients for KP met inclusion criteria (out of 116 patients). n=420 patients for NSM (out of 485 patients; 86.6% follow-up rate). **Exclusion criteria:** underlying malignant disease, dementia, neurological deficits, pathological fracture.
Intervention(s)	Surgical treatments kyphoplasty
Comparator(s)	NSM (+NSAIDs, osteoporosis treatment, post-operative rehabilitation program)
Country	Japan
Funding	No funding received from industry.
Study design	CUA
Perspective	Healthcare system perspective for ageing population.
Time horizon	Baseline + 6-month follow-up and lifetime horizon (3- and 20-year simulations).
Analytic model	Propensity Score Matching study and Markov simulation model (71 matched cases).
Effectiveness outcome	QALY
Effectiveness outcome components	Mortality reduction benefit, bisphosphonate treatment.
Source for effectiveness data	Statistics and Information Department of the Minister's Secretariat, Ministry of Japanese Health, Labor, and Welfare. Other clinical and previous economic studies/published literature.
Costing year	2018
Currency	JPY (on May 2018, 1 GBP = 147.63 JPY)
Discounting	3.5% per annum (on both costs and outcomes)
Components of cost data	Procedure costs (devices, consumables), radiologist costs (X-ray, MRI), lab costs, anesthesia costs, hospitalization cost, post-operative costs.
Cost sources	Published literature.
Sensitivity analysis	ICERs ranged from 652,181 JPY to 4,896,645 JPY (£4,418–£33,168).
QHES	77/100
Results:	

Study Characteristics	Takahashi 2019 Japan
Cost / QALY	KP Total costs: £9,006 – 1,329,629 JPY KP QALY: 0.424
Cost / QALY of comparator(s)	NSM Total costs: £6,277 – 926,642 JPY NSM QALY: 0.292
ICER	**ICER for KP patient vs NSM:** **3 years:** 4,404,158 JPY /QALY gained. **20 years:** 2,416,406 JPY /QALY gained.
One-way SA	Variations of mortality reduction, QALY benefit, bisphosphonate use (i.e., fracture prevention medication), and discount rate on ICER. **SA w/ input parameters:** ICERs ranged from 652,181 JPY to 4,896,645 JPY
Other SA	Nonparametric bootstrap resampling technique (10,000 times)
Author's Conclusion	In the absence of mortality benefit: ICER of 2,416,406 JPY (£16,368), below an accepted WTP threshold of £20,000-30,000/QALY.
Limitations	Administrative data are used • Propensity score matching may not entirely adjust for confounding or selection bias.

CUA = cost utility analysis; GBP = Great Britain Pound; ICER = incremental cost-effectiveness ratio; JPY = Japanese Yen; KP = kyphoplasty; NSAID = non-steroid anti-inflammatory drug; NSM = non-surgical management; QALY = quality adjusted life year; QHES = Quality of Health Economic Studies; RCT = randomized control trial; SA = sensitivity analysis; UC = usual care; VAS = visual analogue scale.

Additional descriptions of economic studies

Stevenson 2014: UK National Institute for Health Research (NIHR)[14]

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 sham, which they term operative placebo with local anesthesia (OPLA) and optimal pain management (OPM). They included extensive sensitivity around the following input parameters assumptions. The values in parentheses are the values in the one-way sensitivity analyses:

- Patient age: 70 years old (65; 80)
- Gender: female (male)
- T-score: -3 SD (-2.5; -3.5)
- Length of bisphosphonate use: 5 years (0 years)
- Assumed duration of a treatment-related mortality benefit: 5 years (0 years)
- Assumed duration of the relative risk of mortality following a vertebral fracture: 5 years (0 years)
- Assumed wane time associated with the relative risk of mortality following a vertebral fracture: 5 years (0 years)
- Risk of mortality in year of subsequent vertebral fracture: True (False)
- Costs associated with hospital stay: Johnson & Johnson (Medtronic; Medtronic length of stay/Johnson & Johnson costs; 0)
- Cost of PVP: low-viscosity cement GBP £800 (high-viscosity cement £1,546)
- Discount rate costs and benefits: 3.5% (0; 6)
- QALY loss associated with PVP and BKP = 0 (0.02)
- Hazard ratio on general mortality for BKP and PVP: (authors reported that academic-in-confidence information had been removed)
- Mortality effect of OPLA: half of PVP (no effect; equal to PVP)
- Regression mapping VAS to EQ-5D (using all data; excluding INVEST data)
- VAS scores convergence: 24 months (12 months)
- Cost of OPLA: equal to PVP (20%, 40%, 60% and 80% of PVP)

Using and varying these input parameters, the authors ran the following exploratory univariate analyses:

- The authors conducted exploratory univariate analyses by varying patient characteristics such as age (65 to 80 years), gender (male patients), and T-score (ranging from -2.5 to -3.5 standard deviations). Varying these parameters did not significantly affect the results. However, the net monetary benefits at a GBP £20,000 WTP threshold per QALY gained were lower for both BKP and PVP treatments when the patient's age was set at 80 years.

- They also varied hospitalization costs (reduced to £0), procedure costs (set to GBP £1,479), and cement price. When bed day costs were set to GBP £0, the net monetary benefits were negative for the BKP treatment and OPLA. Additionally, the net monetary benefits decreased when using the Medtronic length of stay in hospitals for BKP, PVP, and OPLA.
- The authors explored variations in equipment and procedure costs for OPLA (set at 20%, 40%, 60%, 80%, and 50% of the cost of PVP). The net monetary benefit for OPLA increased in each of these analyses.
- They examined the impact of varying the discount rate, bisphosphonate usage, and bisphosphonate wane period (with discount rates for future costs and benefits set to 0% and 6%, assuming no women were taking bisphosphonates, and a wane period following bisphosphonate treatment set to 0 years).
- The authors varied the assumed time of convergence (VAS convergence starting at 12 months), the trials used in the VAS to EQ-5D mapping, and the inclusion of treatment-associated adverse events.
  They also varied the EQ-5D data used from trials (using data from the FREE trial, Buchbinder, and assuming convergence between 24 and 36 months, Buchbinder with convergence between 12 and 24 months, and INVEST trial with convergence between 24 and 36 months, and INVEST trial with convergence between 12 and 24 months).

  Finally, the authors analyzed variations in mortality and fracture rates (assuming no mortality benefit, pooled mortality benefit for BKP and PVP, no mortality benefit for OPLA, mortality benefit for OPLA set equal to PVP, no increased mortality following the initial vertebral fracture, no waning of increased mortality risk after the initial fracture, and no increased risk of mortality in the year of additional vertebral fractures).

These univariate analyses did not alter the authors' overall conclusions. 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.

The results for each of the six scenarios (1 through 6) are detailed below. Additionally, the authors conducted sensitivity analyses for each scenario and reported the corresponding ICERs. Scenarios 2, 4, and 6 are further subdivided based on results from the FREE trial, Buchbinder et al., and INVEST data.

The results for each of the six scenarios (1 through 6) are detailed below. Additionally, the authors conducted sensitivity analyses for each scenario and reported the corresponding ICERs. Scenarios 2, 4, and 6 are further subdivided based on results from the FREE trial, Buchbinder et al., and INVEST data.

Differential beneficial effects on mortality assumed for BKP and PVP:

1. **Utility gain estimated via mapping of stable VAS:** the deterministic ICER results for BKP and PVP (vs OPLA) are GBP £9,802 and £312, respectively. Probabilistic ICER results for BKP and PVP (vs OPLA) are GBP £11,992 and £338, respectively. From the CEAC, the probability for the BKP treatment to be cost-effective at a GBP £20,000 WTP per QALY gained is ~80%.

a. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP and PVP (vs OPLA) peaked to GBP £10,490 and £8,184, respectively.
b. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) reached GBP £11,992 and £7,684, respectively.
c. When it was assumed that convergence of the EQ-5D scores started at 12 months and were equal at 24 months, the ICERs for BKP and PVP (vs OPLA) were at GBP £11,975 and £436, respectively.
d. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP (vs OPLA) and PVP were at GBP £11,992 and £398, respectively.
e. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPLA) were at GBP £11,033 and extendedly dominated, respectively.

2. **Utility gain estimated directly from EQ-5D in the trial.** Authors subdivided their analyses into three categories based on whether the FREE data, the Buchbinder et al. data or the INVEST data were used.
    a. **Using the FREE [15] data:** the deterministic ICER results for BKP and PVP (vs OPLA) are GBP £9,541 and £214, respectively. Probabilistic ICER results for BKP and PVP (vs OPLA) are GBP £7,616 and £302, respectively. From the CEAC, the probability for the BKP treatment to be cost-effective at a GBP £10,000 WTP per QALY gained is ~80%.
        i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP and PVP (vs OPM) peaked to £7,012 and extendedly dominated, respectively.
        ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) reached GBP £7,616 and £6,870, respectively.
        iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPLA) were at GBP £7,616 and £349, respectively.
        iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) were at GBP £7,254 and extendedly dominated, respectively.
    b. Using the Buchbinder et al. data[2]: the deterministic ICERs for BKP and PVP (vs OPLA) are GBP £9,853 and £731, respectively with a convergence between 12 and 24 months and a convergence between 24 and 36 months. Looking at the same convergence (i.e., between 12 and 24 and between 24 and 36 months), the probabilistic ICERs for BKP and PVP (vs OPLA) are GBP £10,073 and £725, respectively. From the CEAC, the probability for the BKP treatment to be cost-effective at a GBP £15,000 WTP per QALY gained is ~85%.
        i. When the hospitalization costs were set to GBP £0 per day and with a convergence between 12 and 24 months, the ICER for BKP (vs OPM) was GBP £10,196 while the ICER was at GBP £9,625 with a convergence between 24 and 36 months. The ICER for PVP was extendedly dominated for PVP (vs OPM) in both cases.
        ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP reached GBP £11,445 and was extendedly dominated, respectively.

iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPLA) were at GBP £10,072 and extendedly dominated, respectively.
iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) reached GBP £11,230 and was extendedly dominated, respectively.

c. **Using the INVEST**[6]; **data:** the deterministic ICERs for BKP and PVP (vs OPLA) are GBP £9,850 and £595, respectively with a convergence between 12 and 24 months and a convergence between 24 and 36 months. Looking at the same convergence (i.e., between 12 and 24 and between 24 and 36 months), the probabilistic ICERs for BKP and PVP (vs OPLA) are GBP £10,070 and £588, respectively. From the CEAC, the probability for the BKP treatment to be cost-effective at a GBP £15,000 WTP per QALY gained is ~85%.
    i. When the hospitalization costs were set to £0 per day and with a convergence between 12 and 24 months, the ICER for BKP (vs OPM) was GBP £9,850 while the ICER was at GBP £9,316 with a convergence between 24 and 36 months. The ICER for PVP was extendedly dominated for PVP (vs OPM) in both cases.
    ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP reached GBP £10,900 and was extendedly dominated, respectively.
    iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPLA) were at GBP £10,070 and £796, respectively.
    iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) reached GBP £10,657 and was extendedly dominated, respectively.

Identical beneficial effect on mortality assumed for BKP and PVP:

3. **Utility gain estimated via mapping of stable VAS:** the deterministic ICER results for BKP and PVP (vs OPM) are dominated and GBP £449, respectively. Probabilistic ICER results for BKP and PVP (vs OPM) are dominated and GBP £501, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~85%.
    a. When the hospitalization costs were set to £0 per day, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £5,941, respectively.
    b. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are dominated and reached GBP £5,529, respectively.
    c. When it was assumed that convergence of the EQ-5D scores started at 12 months and were equal at 24 months, the ICERs for BKP and PVP (vs OPM) were dominated and GBP £594, respectively.
    d. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP (vs OPM) and PVP were dominated and reached GBP £559, respectively.
    e. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) were dominated and GBP £7,458, respectively.

f. As above plus mortality effect of OPLA set to equal BKP and PVP, the ICERs for BKP (vs OPM) and PVP were dominated and reached GBP £31,304, respectively.

4. **Utility gain estimated directly from EQ-5D in the trial.** Authors subdivided their analyses in three categories based on whether the FREE data, the Buchbinder et al. [2]data or the INVEST data were used.
   a. **Using the FREE data:** [15] the deterministic ICER results for KP and VP (vs OPM) are dominated and GBP £342, respectively. Probabilistic ICER results for BKP and PVP (vs OPM) are dominated and GBP £336, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~80%.
      i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £4,513, respectively.
      ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are dominated and reached GBP £3,705, respectively.
      iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPM) are dominated and reached £361, respectively.
      iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £4,697, respectively.
   b. Using the Buchbinder et al. [2] **data:** the deterministic ICERs for BKP and PVP (vs OPM) are dominated and GBP £731, respectively with a convergence between 12 and 24 months and a convergence between 24 and 36 months. Looking at the same convergence (i.e., between 12 and 24 and between 24 and 36 months), the probabilistic ICERs for BKP and PVP (vs OPM) are dominated and GBP £725, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~85%.
      i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP (vs OPM) are dominated whether there is convergence between 12 and 24 moths or between 24 and 36 months. Regarding PVP (vs OPM), the ICERs reached GBP £7,065 and £6,572 when there is convergence between 12 and 24 moths and between 24 and 36 months, respectively.
      ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are dominated and reached GBP £7,997, respectively.
      iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPM) were dominated and reached GBP £852, respectively.
      iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £9,399, respectively.
      v. When combining all of these sensitivity analyses plus the mortality effect of OPLA is set to equal BKP and PVP, the ICERs for BKP and PVP (vs OPM) are both dominated.
   c. Using the INVEST[6] **data:** the deterministic ICERs for BKP and PVP (vs OPM) are dominated and GBP £662, respectively with a convergence between 12 and 24 months and a convergence between 24 and 36 months. Looking at the same convergence (i.e., between 12 and 24 and between 24 and 36 months), the probabilistic ICERs for BKP and PVP (vs OPM) are dominated and GBP

£655, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~85%.

i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP (vs OPM) are dominated whether there is convergence between 12 and 24 moths or between 24 and 36 months. Regarding PVP (vs OPM), the ICERs reached GBP £6,765 and £6,311 when there is convergence between 12 and 24 moths and between 24 and 36 months, respectively.
ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are dominated and reached GBP £7,219, respectively.
iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £756, respectively.
iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £8,342, respectively.
v. When combining all these sensitivity analyses plus the mortality effect of OPLA is set to equal BKP and PVP, the ICERs for BKP and PVP (vs OPM) are both dominated.

No effect on mortality assumed for BKP and PVP:

5. **Utility gain estimated via mapping of stable VAS:** the deterministic ICER results for BKP and PVP are dominated and dominating, respectively. The probabilistic ICER results for BKP and PVP are also dominated and dominating, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~90%.
    i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £12,757, respectively.
    ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are dominated and reached GBP £12,144, respectively.
    iii. When it was assumed that convergence of the EQ-5D scores started at 12 months and were equal at 24 months, the ICERs for BKP and PVP were dominated and dominating, respectively.
    iv. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP (vs OPM) and PVP were dominated and dominating, respectively.
    v. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) were dominated and reached GBP £31,953, respectively.
6. **Utility gain estimated directly from EQ-5D in the trial** Authors subdivided their analyses in three categories based on whether the FREE data, the Buchbinder et al. data or the INVEST data were used.
    a. **Using the FREE[15] data:** the deterministic ICER results for BKP and PVP (vs OPM) are dominated and dominating, respectively. The probabilistic ICER results for BKP and PVP (vs OPM) are also dominated and dominating, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~90%.

i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £8,885, respectively.
ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are dominated and reached GBP £6,514, respectively.
iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP are dominated and dominating, respectively.
iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) are dominated and reached GBP £9,701, respectively.

b. Using the Buchbinder et al. [2]: data: the deterministic ICERs for BKP and OPLA/PVP are dominated and dominating, respectively either with a convergence between 12 and 24 months or a convergence between 24 and 36 months. Looking at the same convergence (i.e., between 12 and 24 and between 24 and 36 months), the probabilistic ICERs for BKP and PVP are the same as the deterministic ones. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is ~90%.

i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP (vs OPM) are dominated whether there is convergence between 12 and 24 moths or between 24 and 36 months. Regarding OPLA/PVP (vs OPM), the ICERs reached GBP £33,963 and £24,336 when there is convergence between 12 and 24 moths and between 24 and 36 months, respectively.
ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP (vs OPLA) are both dominated.
iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICERs for BKP and PVP (vs OPLA) are both dominated.
iv. When combining all of these sensitivity analyses, the ICERs for BKP and PVP (vs OPM) are both dominated.
v. When combining all of these sensitivity analyses plus the mortality effect of OPLA is set to equal BKP and PVP, the ICERs for BKP and PVP are both dominated.

c. Using the INVEST[6] data: the deterministic ICERs for BKP and PVP are dominated and dominating, respectively with a convergence between 12 and 24 months and a convergence between 24 and 36 months. Looking at the same convergence (i.e., between 12 and 24 and between 24 and 36 months), the probabilistic ICERs for BKP and PVP are also dominated and dominating, respectively. From the CEAC, the probability for the PVP treatment to be cost-effective at a GBP £5,000 WTP per QALY gained is 100%.

i. When the hospitalization costs were set to GBP £0 per day, the ICERs for BKP (vs OPM) are dominated whether there is convergence between 12 and 24 moths or between 24 and 36 months. Regarding PVP (vs OPM), the ICERs reached GBP £27,577 and £20,895 when there is convergence between 12 and 24 moths and between 24 and 36 months, respectively.

ii. When the cost of OPLA procedure was set to 50% of PVP and the cost of equipment of OPLA equipment was set to 60% of PVP, the ICERs for BKP and PVP are dominated and reached GBP £67,780, respectively.
iii. When it was assumed that BKP and PVP were associated with a 0.02 QALY loss, the ICER for BKP is dominated. When it was assumed convergence between 12 and 24 months, the ICER for PVP was dominated while it was dominating in the case of a convergence between 24 and 36 months.
iv. When combining all of these sensitivity analyses, the ICER for BKP and PVP are each respectively dominated.

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

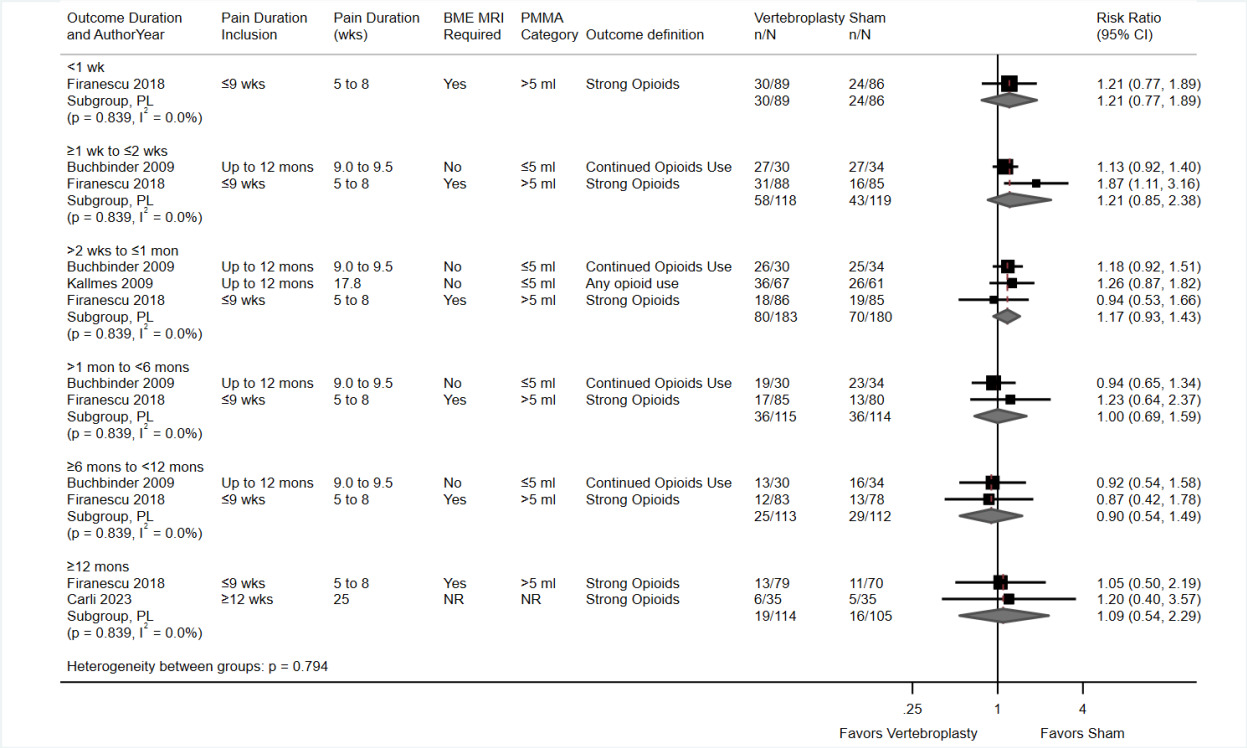
Lastly, the authors looked at the use of high-viscosity cement (vs low-viscosity cement), initial facet joint injection (prior to considering PVP or BKP) and educating patients and their effect on the ICER. They investigated whether using high-viscosity cement (vs low-viscosity cement) in all patients rather than in 15% of the patients would be cost-effective. Looking at the re-operation rate, the authors estimated that there would need to be a re-operation rate of at least 25% for using high-viscosity cement in all patients for high-viscosity cement to become a cheaper option than the alternative.

Regarding the use of an initial facet joint injection, the authors estimated that the ICER could be reduced by a third if a facet joint injection were first considered prior to a vertebral augmentation treatment (KP or VP).

Regarding adding additional costs to educate patients, the authors looked at scenario 5, scenario 6 (Buchbinder et al.) [2], and scenario 6 (INVEST) [6]; as there was no beneficial mortality effect assumed for BKP and PVP. For scenario 6[2], OPM can dominate PVP if the costs to educate patients were less than GBP £2,111 per person. For scenario 5 and scenario 6 (INVEST), the individual cost to educate patients would need to be greater than GBP £500 in scenario 5 and greater than GBP £1,800 in scenario 6 (INVEST) for PVP to have a cost/QALY gained less than the GBP £20,000 WTP threshold.

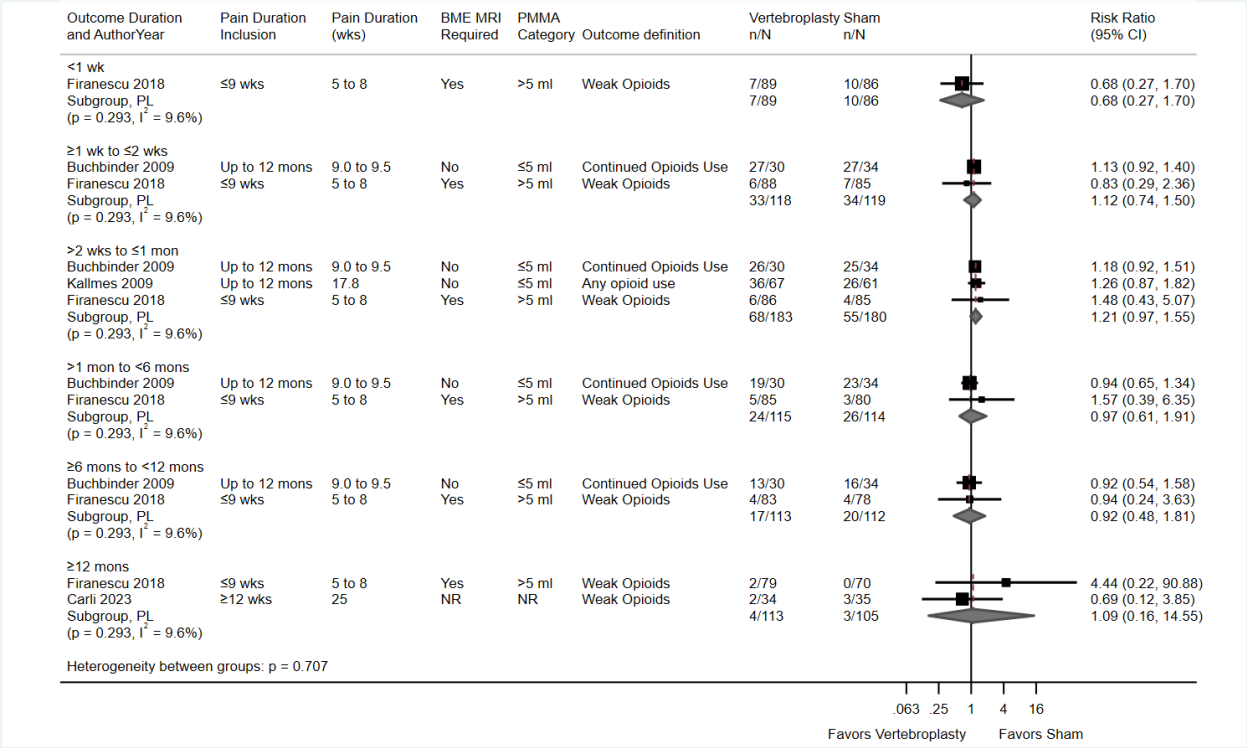
APPENDIX P. Additional Forest Plots

Appendix Figure P1. Vertebroplasty versus Sham Procedures: Strong Opioid use by Follow-up Time.



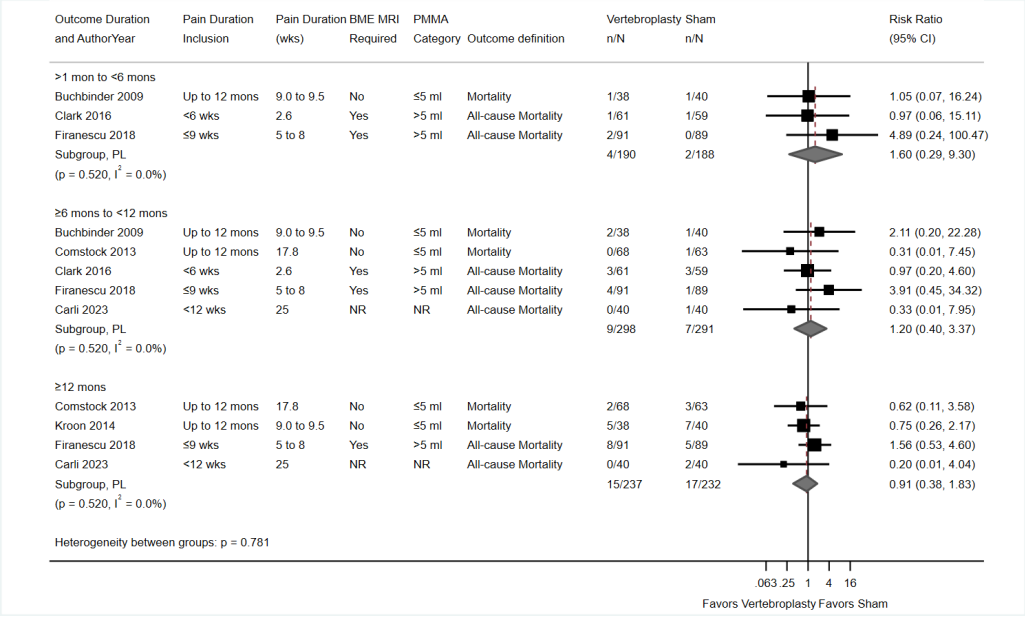
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; RR = risk ratio; Wks = weeks.

Appendix Figure P2. Vertebroplasty versus Sham Procedures: Weak Opioid use by Follow-up Time.



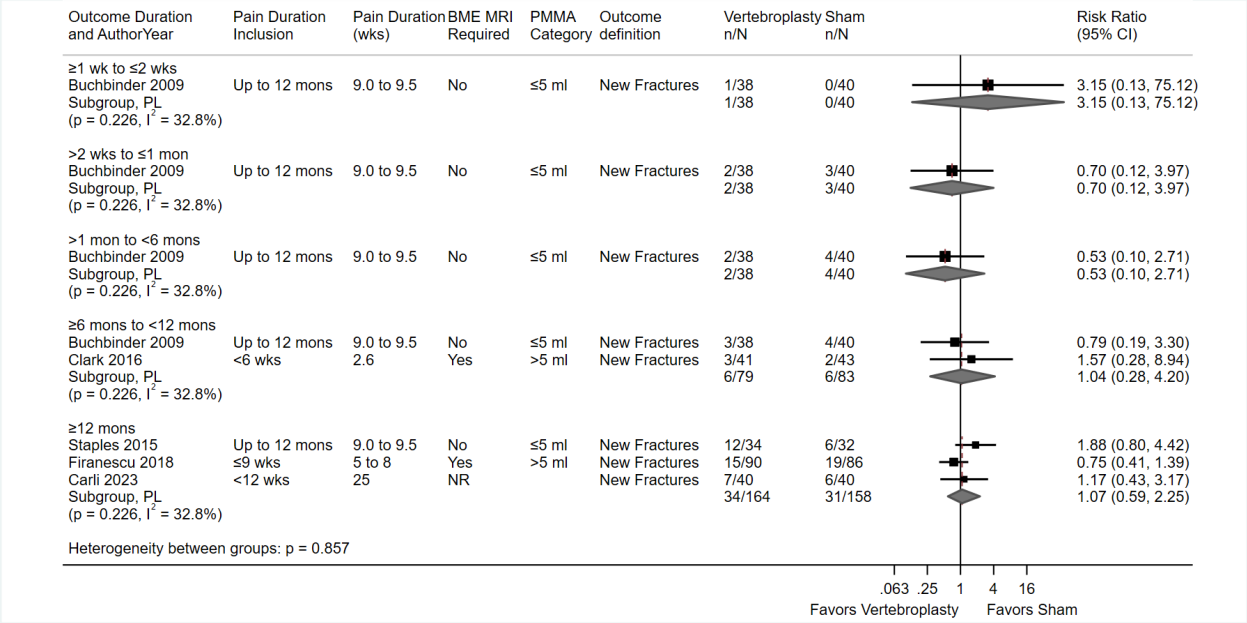
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; RR = risk ratio; Wks = weeks.

Appendix Figure P3. Vertebroplasty versus Sham Procedures: Cumulative Mortality by Follow-up Time.



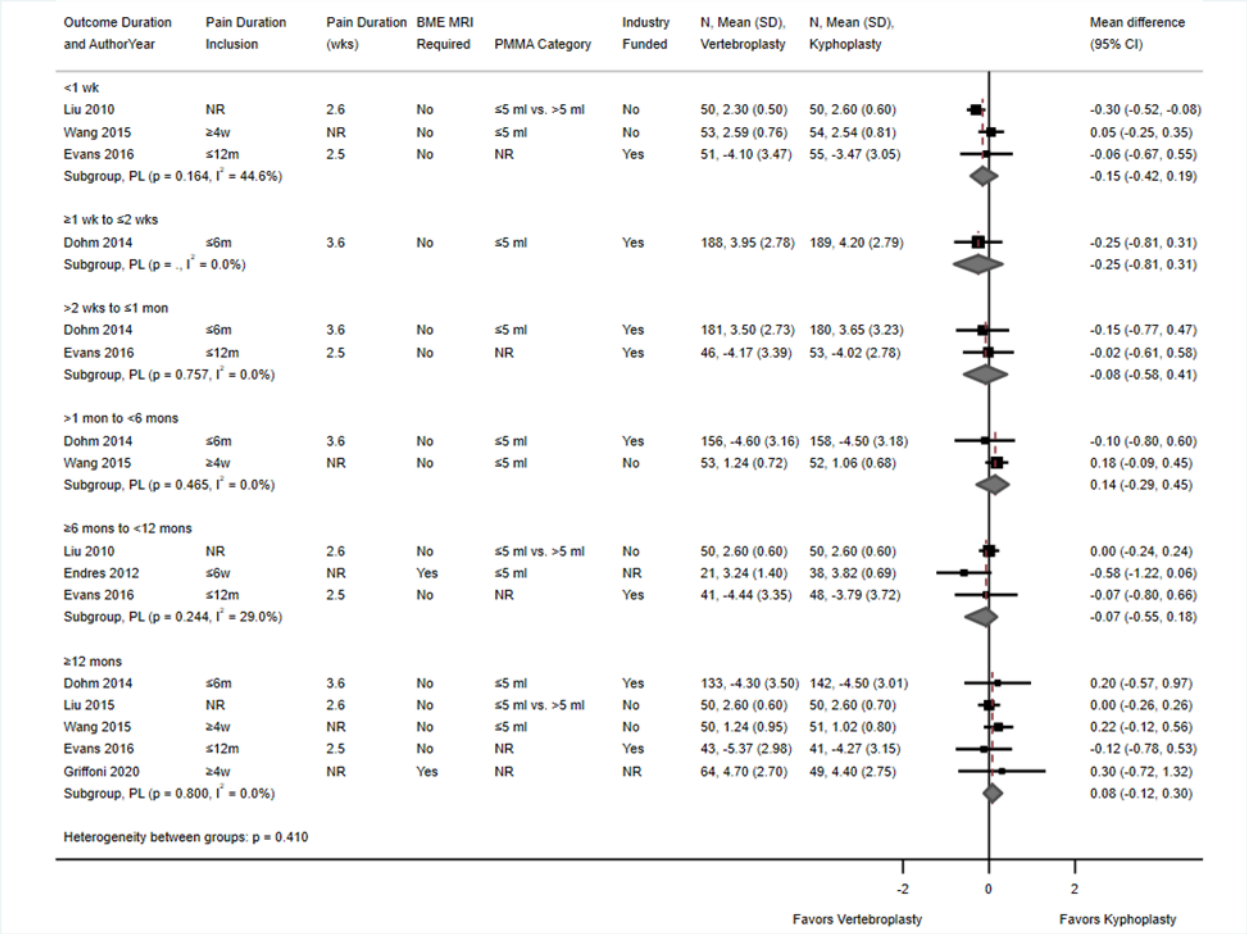
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; RR = risk ratio; Wks = weeks.

Appendix Figure P4. Vertebroplasty versus Sham Procedures: Cumulative Risk of Vertebral Fracture by Follow-up Time.



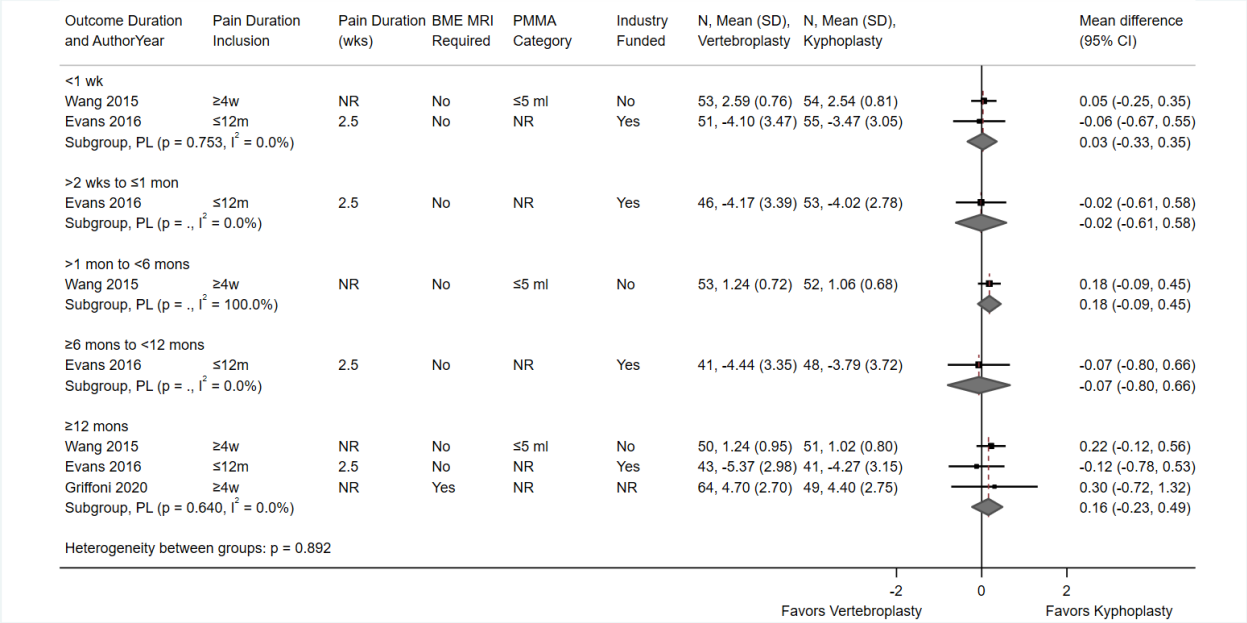
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; RR = risk ratio; Wks = weeks.

Appendix Figure P5. Vertebroplasty versus Kyphoplasty: Sensitivity Analysis of Pain Scores Excluding One Poor-quality Potential Outlier Trial (Wang, 2023).



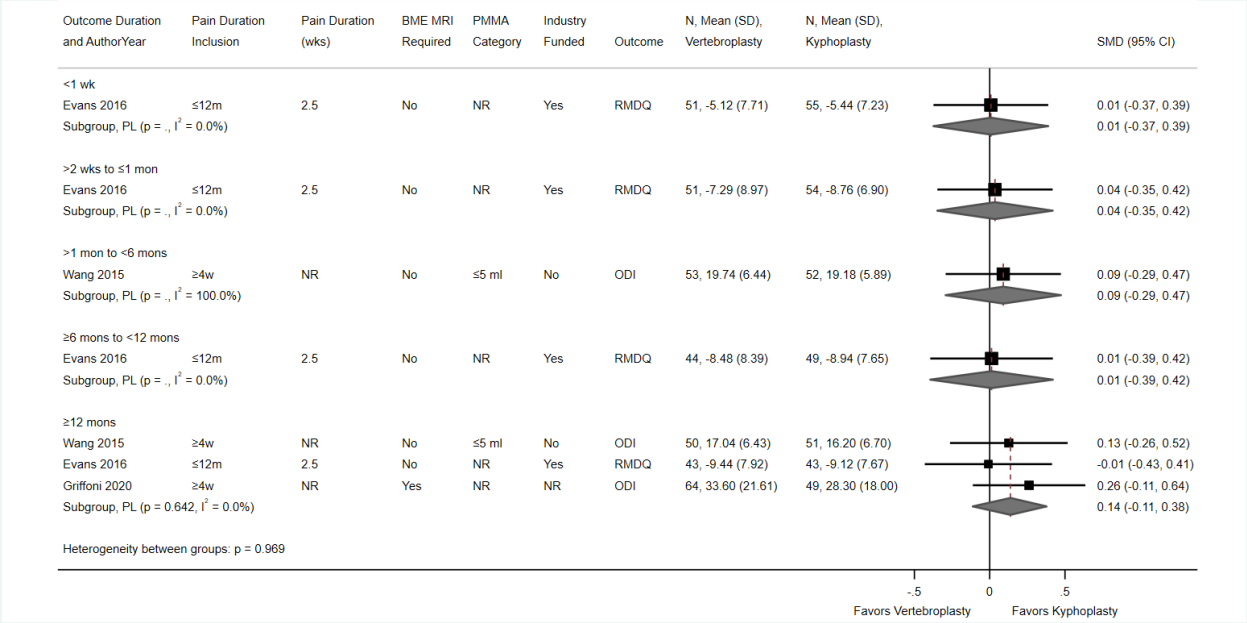
BME = bone marrow edema; CI = confidence interval; MD = mean difference; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; SD = standard deviation.

Appendix Figure P6. Vertebroplasty versus Kyphoplasty: Sensitivity Analysis of Pain Scores Excluding Poor-quality Trials (Endres 2012, Liu 2010, and Wang 2023).



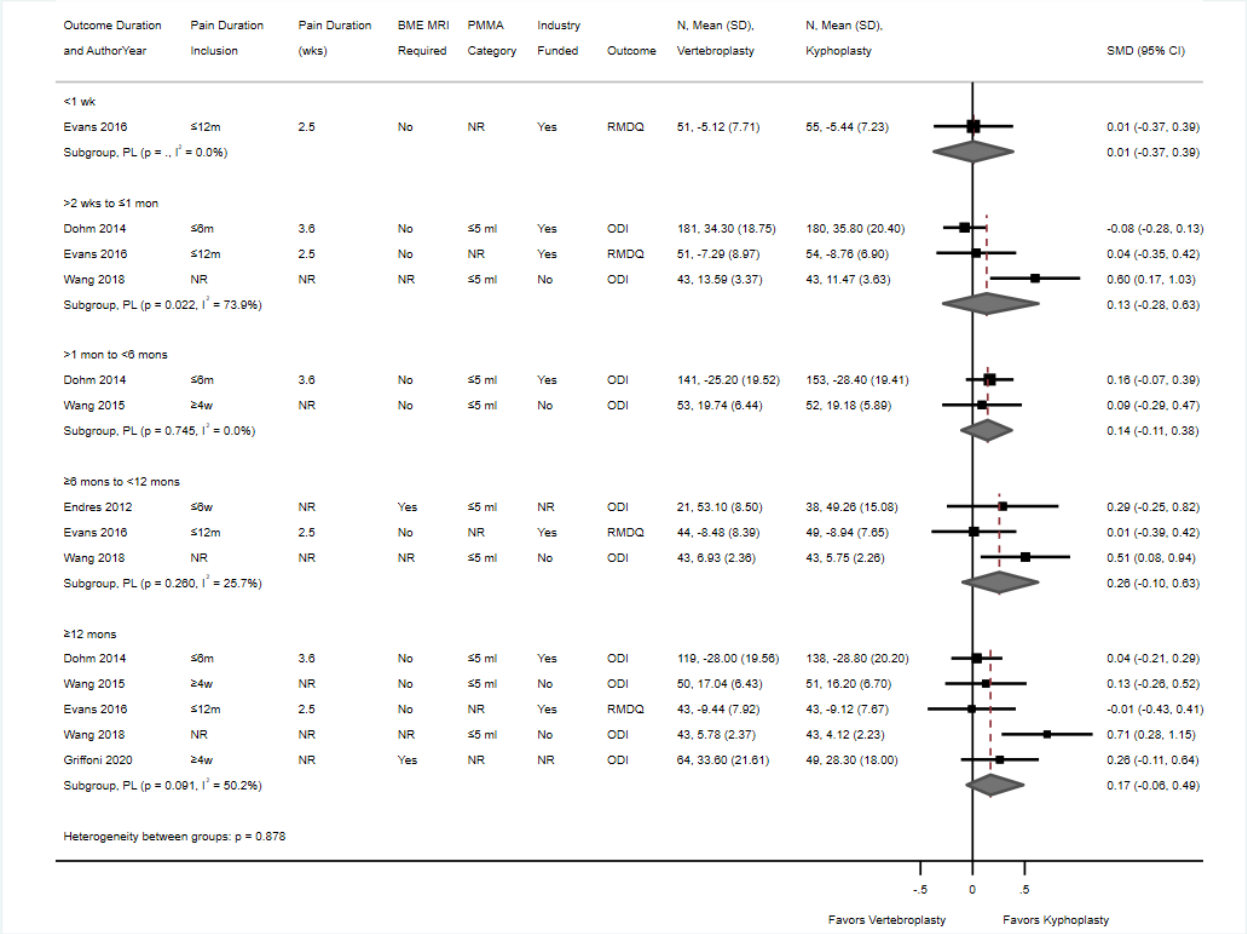
BME = bone marrow edema; CI = confidence interval; MD = mean difference; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; SD = standard deviation.

Appendix Figure P7. Vertebroplasty versus Kyphoplasty: Sensitivity Analysis of Function Scores Excluding Poor-quality Trials (Endres 2012, Liu 2010, and Wang 2023).



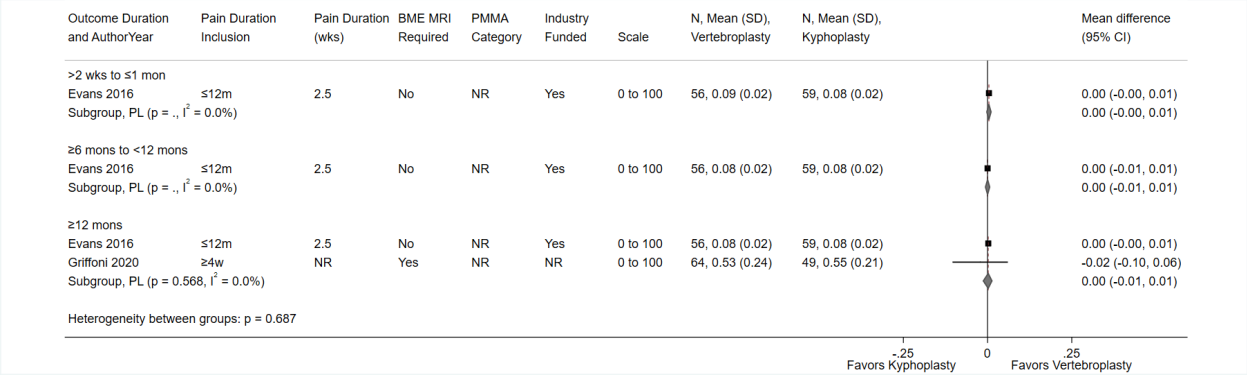
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; ODI = Oswestry Disability Index; PMMA = polymethylmethacrylate; RMDQ = Roland Morris Disability Index; SD = standard deviation; SMD = standardize mean difference.

Appendix Figure P8. Vertebroplasty versus Kyphoplasty: Sensitivity Analysis of Function Scores Excluding One Poor-quality Potential Outlier Trial (Wang, 2023).



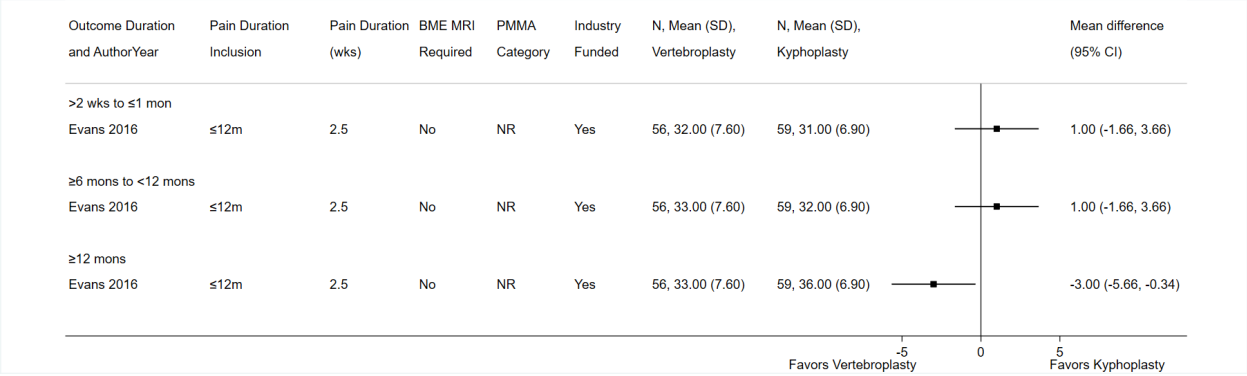
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; ODI = Oswestry Disability Index; PMMA = polymethylmethacrylate; RMDQ = Roland Morris Disability Questionnaire; SD = standard deviation; SMD = standardize mean difference.

Appendix Figure P9. Vertebroplasty versus Kyphoplasty: Sensitivity Analysis of EQ-5D QoL Scores Excluding One Poor-quality Trial (Dohm, 2024).



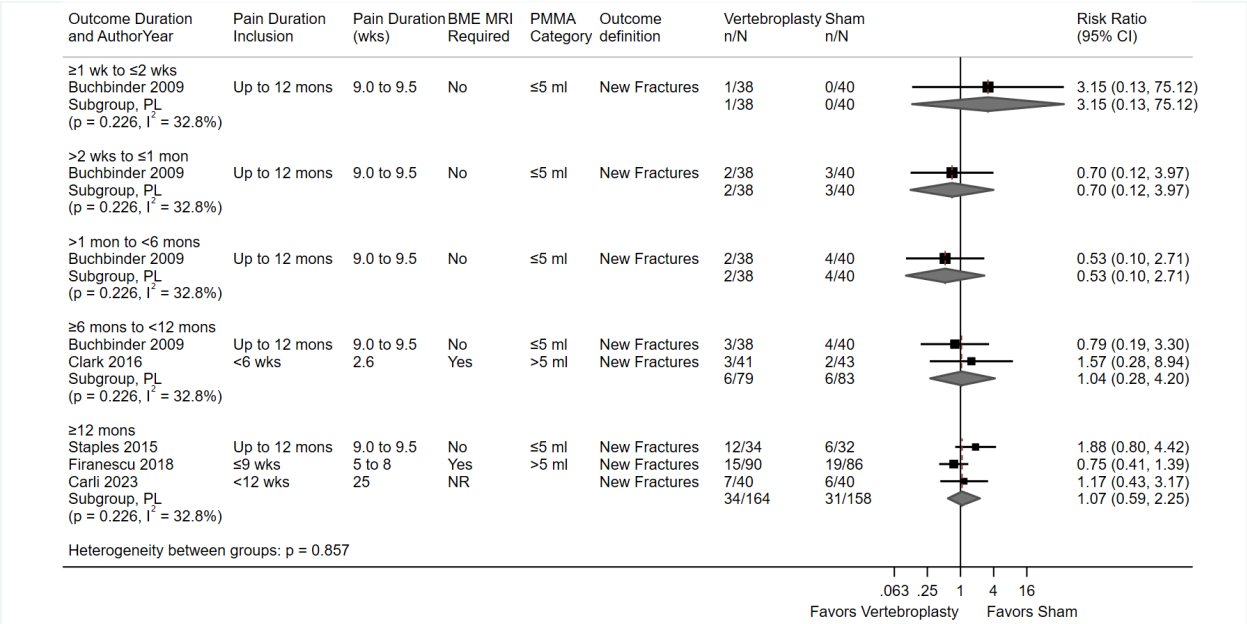
BME = bone marrow edema; CI = confidence interval; EQ-5D = EuroQol 5-Dimensions; MD = mean difference; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; SD = standard deviation.

Appendix Figure P10. Vertebroplasty versus Kyphoplasty: Sensitivity Analysis of SF-36 PCS QoL Scores Excluding One Poor-quality Trial (Dohm, 2024).



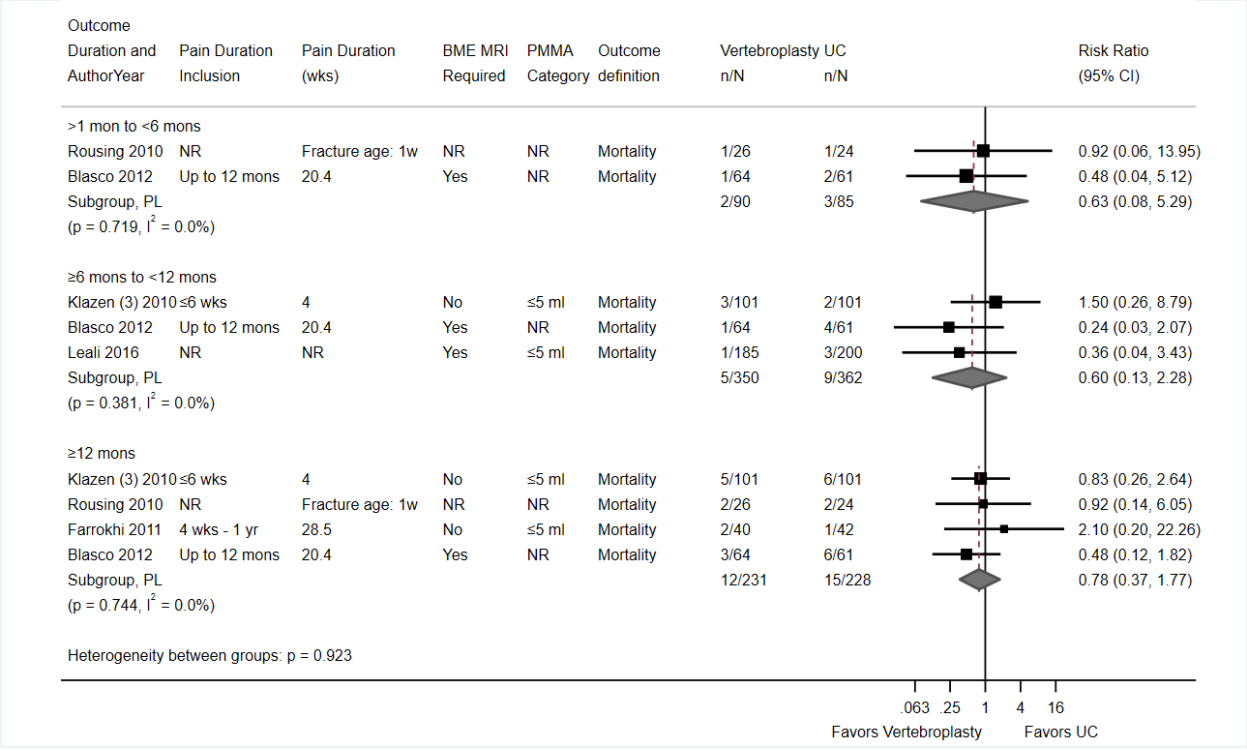
BME = bone marrow edema; CI = confidence interval; MD = mean difference; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; SD = standard deviation; SF-36 = 36-item Short Form Questionnaire.

Appendix Figure P11.VP vs. sham: analysis of new vertebral fractures stratified by timeframe



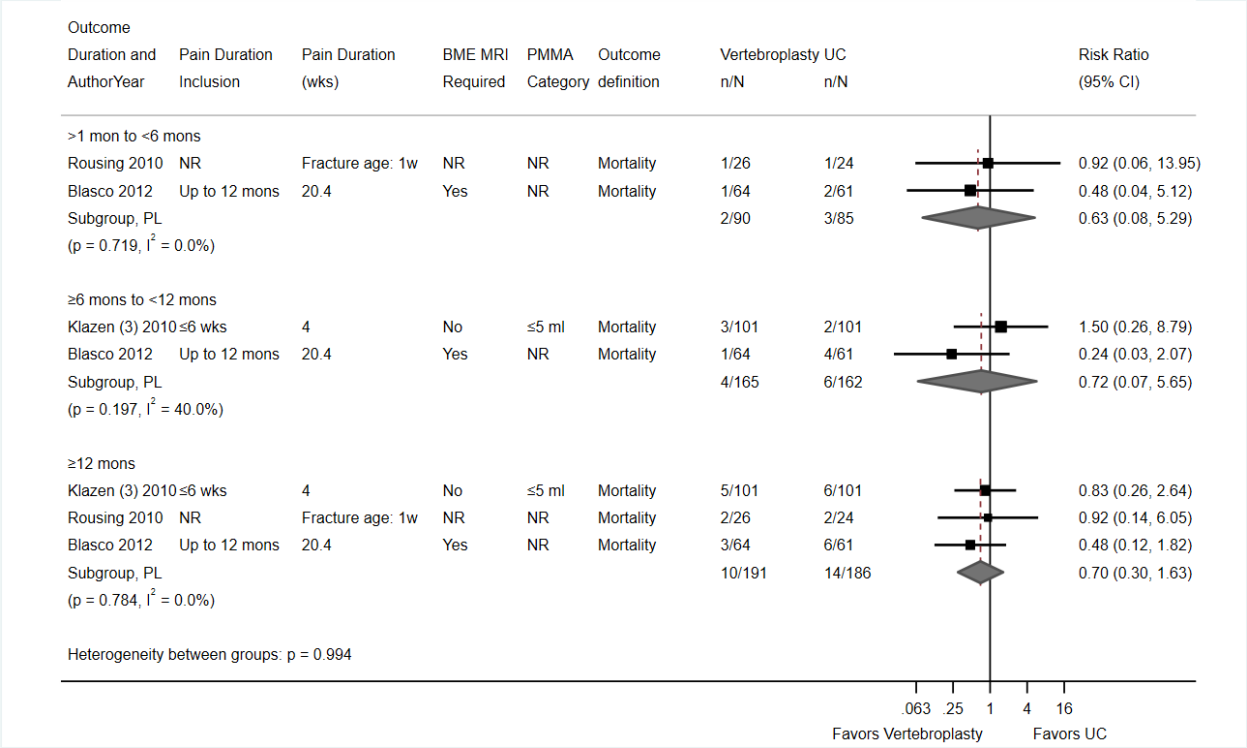
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P12. VP vs. UC: analysis of mortality stratified by timeframe



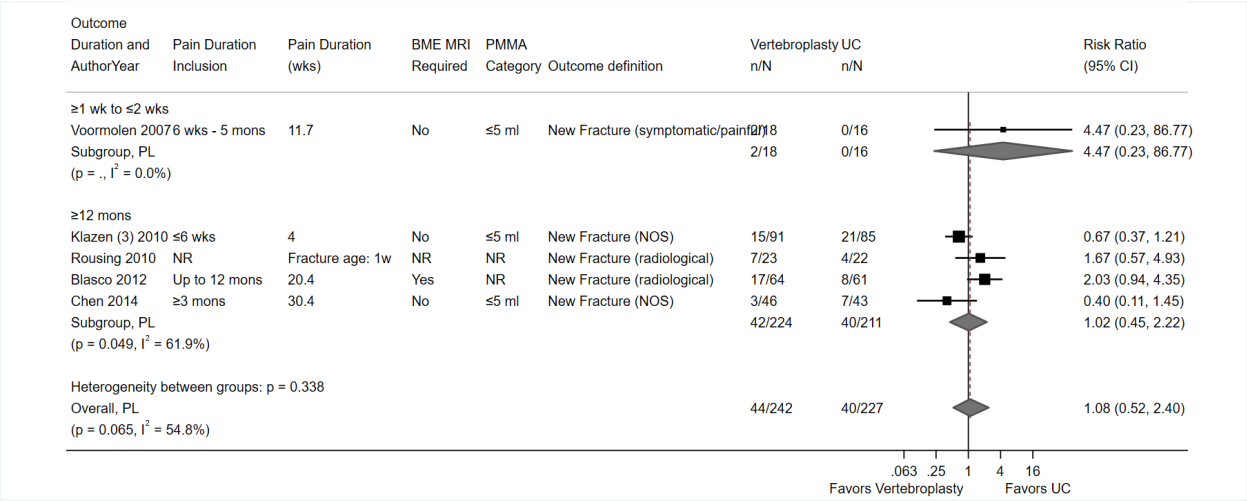
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; UC = usual care; wks = weeks.

Appendix Figure P13. VP vs. UC: sensitivity analysis of mortality excluding poor-quality trials



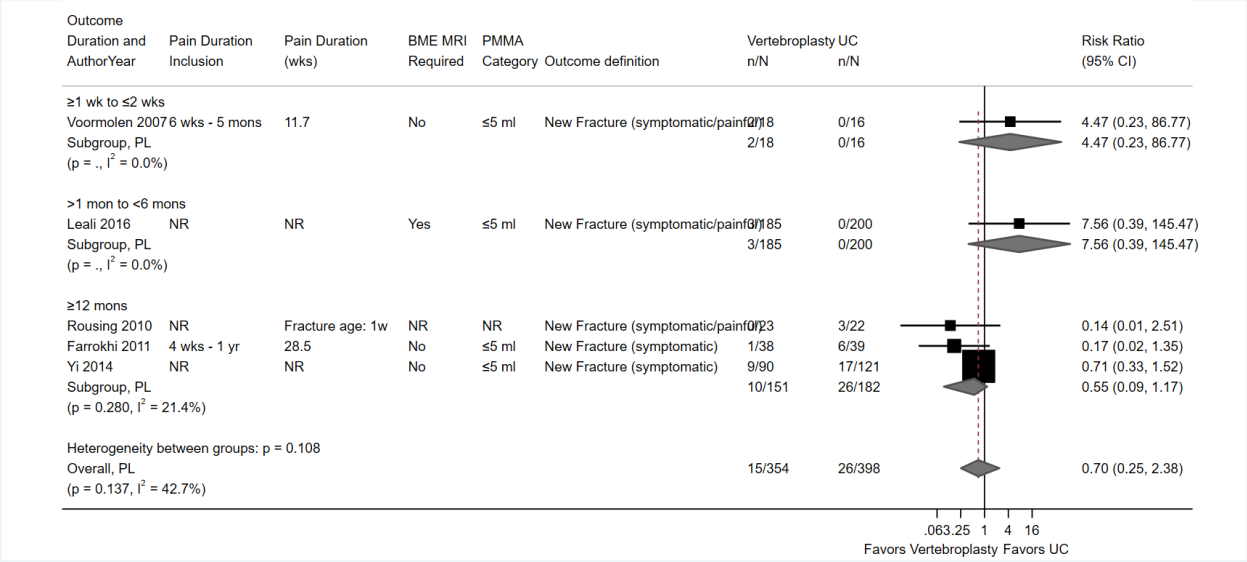
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; UC = usual care; wks = weeks.

Appendix Figure P14. VP vs. UC: sensitivity analysis of any new fracture at latest follow-up excluding poor-quality trials



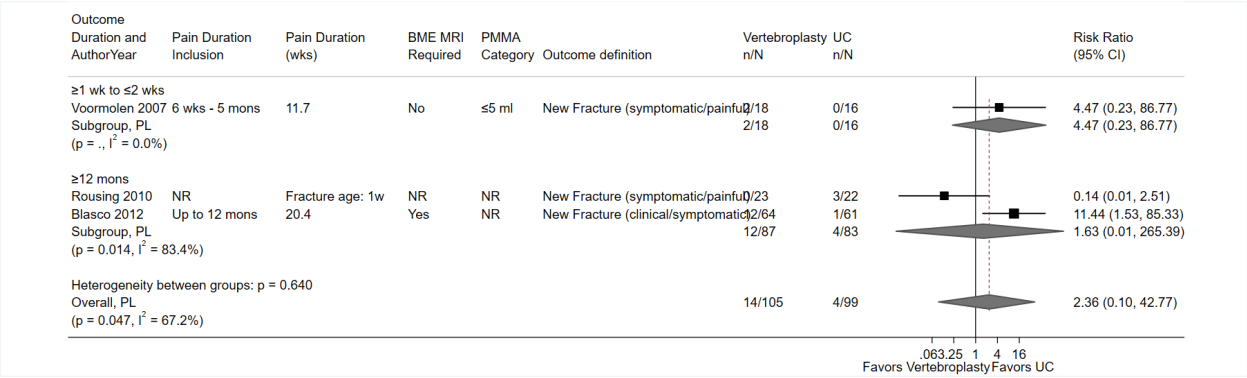
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; UC = usual care; wks = weeks.

Appendix Figure P15. VP vs. UC: sensitivity analysis of any new symptomatic fracture excluding outlier trial Blasco 2012



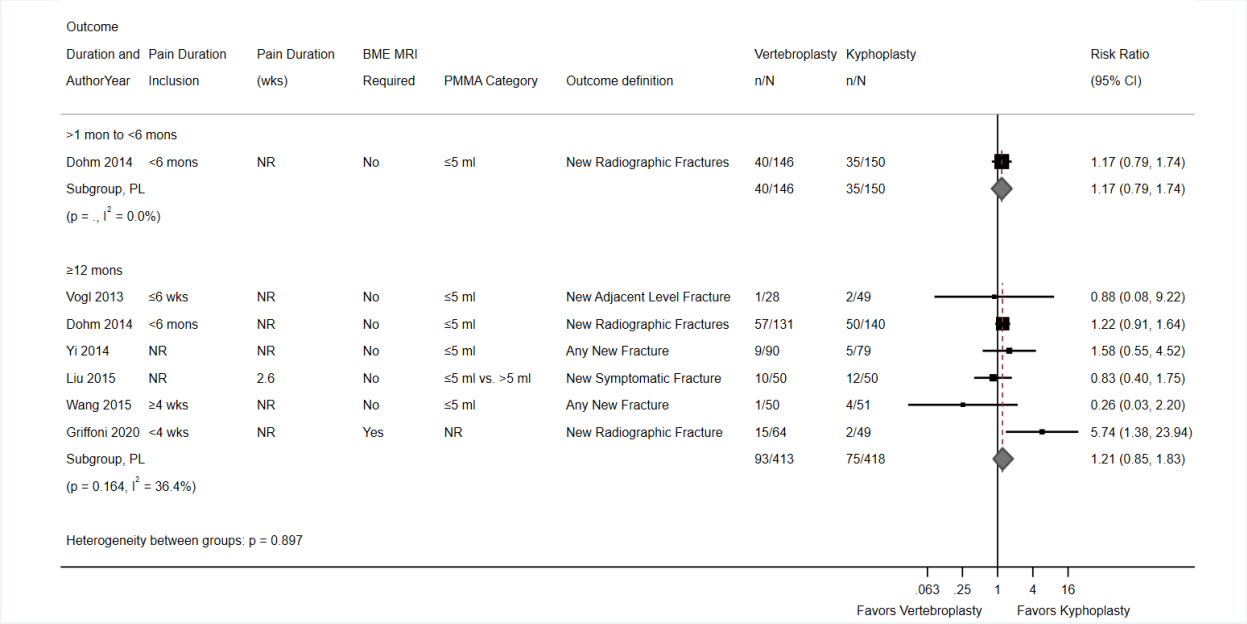
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; UC = usual care; wks = weeks.

Appendix Figure P16. VP vs. UC: sensitivity analysis of any new symptomatic fracture excluding poor-quality trials



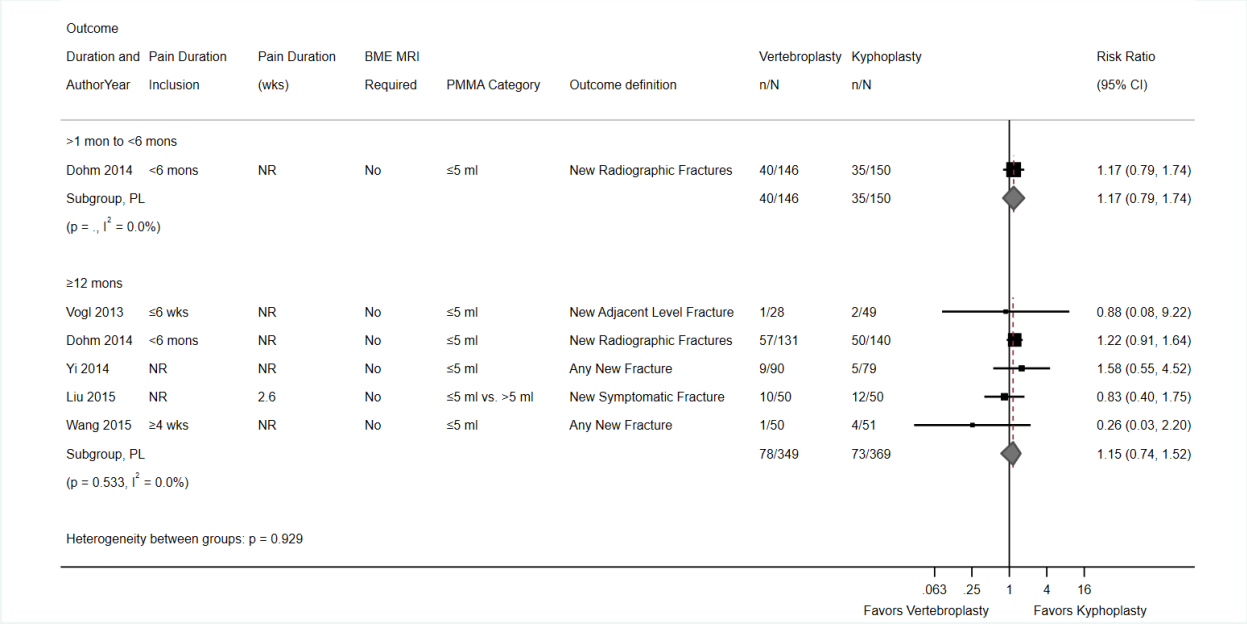
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; UC = usual care; wks = weeks.

Appendix Figure P17. VP vs. KP: analysis of new vertebral fractures stratified by timeframe



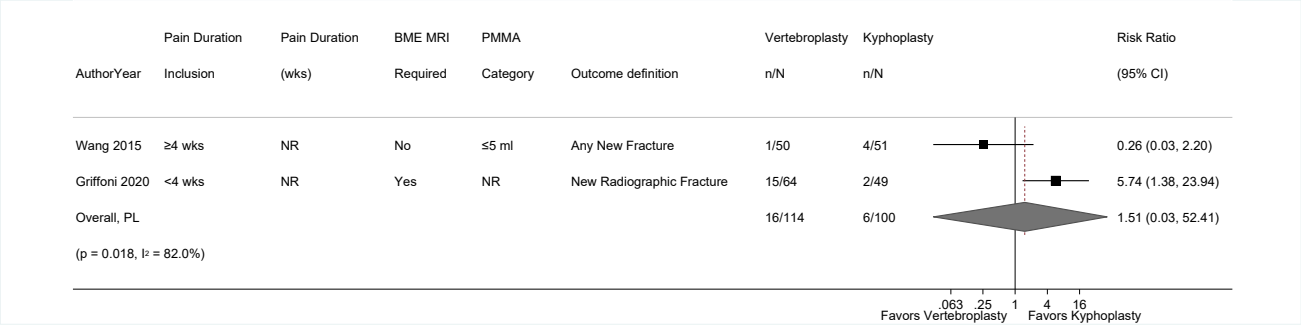
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P18. VP vs. KP: sensitivity analysis of any new vertebral fractures excluding one outlier trial (Griffoni 2020)



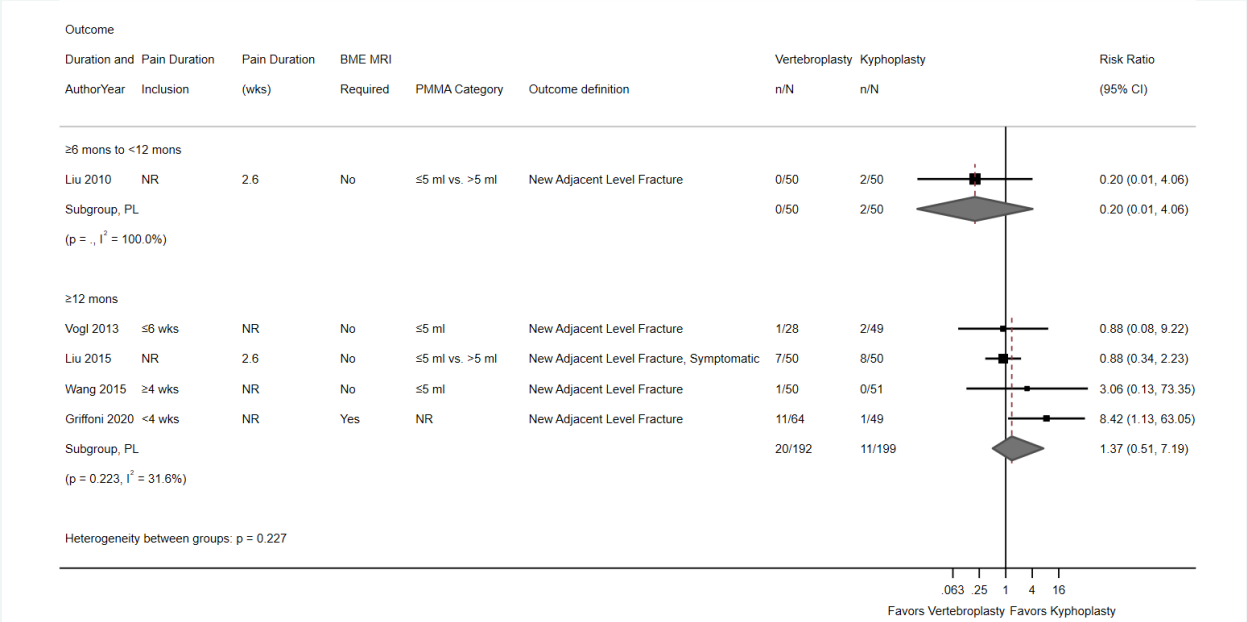
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P19.VP vs. KP: sensitivity analysis of any new vertebral fractures excluding poor-quality trials



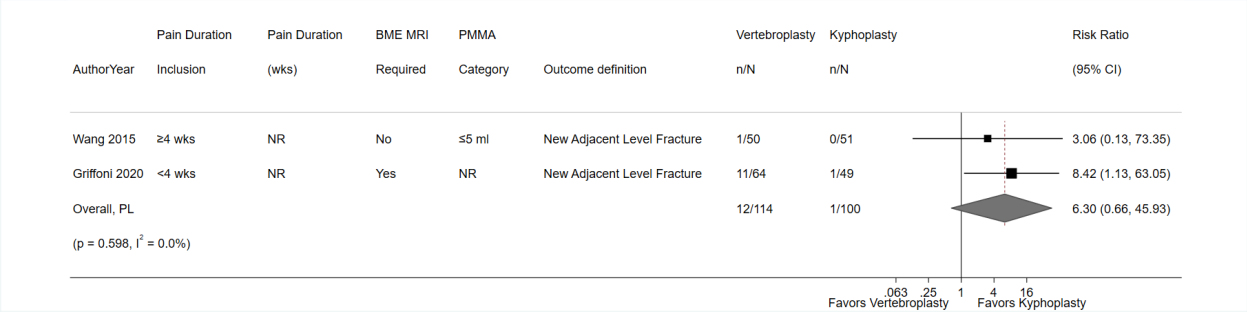
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P20. VP vs. KP: analysis of any new adjacent level vertebral fractures stratified by timeframe.



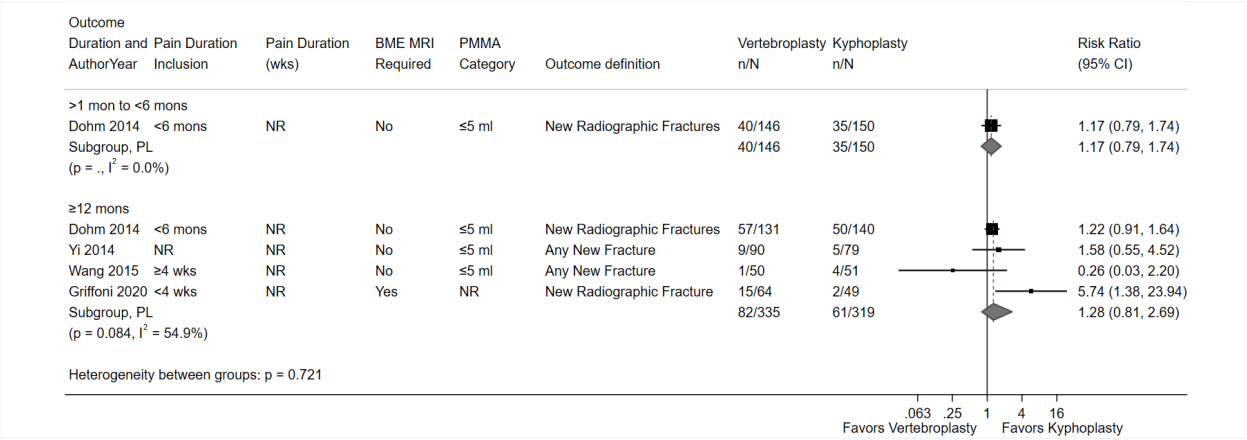
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P21. VP vs. KP: sensitivity analysis of any new adjacent level vertebral fractures excluding poor-quality trials



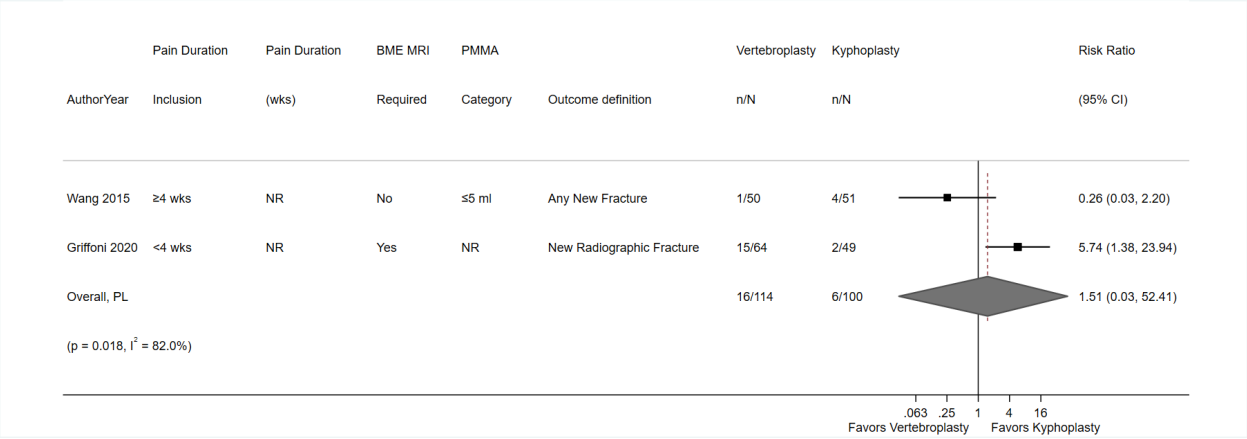
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P22. VP vs. KP: analysis of any new radiographic vertebral fractures stratified by timeframe.



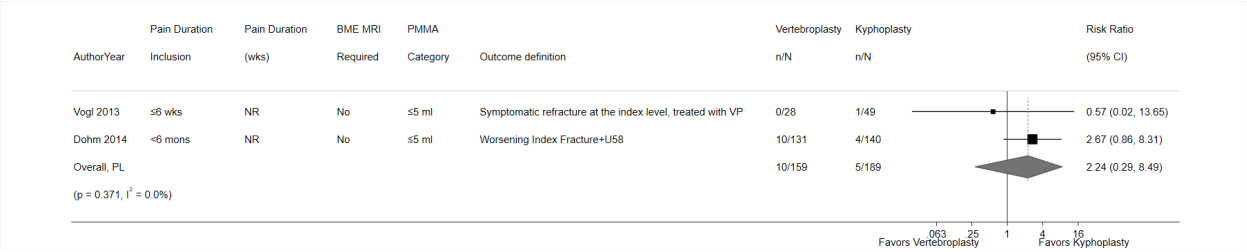
BME = bone marrow edema; CI = confidence interval; MRI = magnetic resonance imaging; mons = months; PMMA = polymethylmethacrylate; wks = weeks.

Appendix Figure P23. VP vs. KP: sensitivity analysis of any new radiographic vertebral fractures excluding poor quality trials



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

Appendix Figure P24. VP vs. KP: analysis of refracture or worsening index level fracture



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

APPENDIX Q. Summary of Mortality from Administrative Database Studies

Appendix Table Q1: 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 Funding: government and professional society	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 to 0.95)
ACS-NSQIP (Possible overlap in data)	Choo, 2018 (2012-2014) Funding: No grant or specific funding	VP: 242 (10%) KP: 2191 (90%)	30-day mortality: 2% (n=49): analyses indicate that augmentation was not an independent risk factor for mortality
	Kim, 2022 (2011-2011) Funding: no financial support received	N = 1932 VP: 197 (10%) KP: 1769 (90%)	KP vs. VP: Adj. OR 0.94 (95% CI 0.27 to 3.24); Procedure type was not a risk factor for mortality
Nationwide Inpatient Sample	Zampini, 2010 (2005) Funding: NR	N = 5766 KP: 15% Nonoperative: 84.7%	KP vs. Nonoperative: 0.3% vs. 1.6% Adj OR 0.52, p=0.003 (95% CI NR)
Mortality at Longer Follow -up (>30 days)			
Medicare	Ong, 2018* (2005-2014) Funding: Industry	VP: 117232 KP: 261756 Non-operated: 1698956†	Mortality risk overall at 10 years: 85.1% (95% CI, 84.7 to 85.5%) Propensity-adjusted results comparing groups: 19% (95% CI, 19% to 19%; p<0.001) and 7% (95% CI 7% to 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% to 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 non-op HR 0.83 (95% CI 0.82 to 0.83) VP vs. Non-op: HR 0.926 (95% CI 0.926 to 0.917) KP vs. Non-op: HR 0.81 (95% CI 0.813 to 0.806) KP vs. VP: HR 0.87 (95% CI 0.87 to 0.88)
	Edidin 2015* (2005-2009) Funding: Industry	Propensity-score matched (osteoporotic fractures) VP: 37252 KP: 36286 Non-Operated matches:	Adjusted HR at 4 years: Non-op vs. VP: HR 1.30 (95% CI 1.28 to 1.33) Non-op vs. KP: HR 1.62 (95% CI 1.60 to 1.64) KP vs. VP: HR 0.83 (95% CI 0.81 to 0.85)

		VP: Non-op n=107930 KP: Non-op n=163791	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] Funding: Funding: government and professional society	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 to 1.04); not statistically significant
Emory University Hospital	Levy, 2012 (1998 to 2007) Funding: partial NIH funding	N=250 VA (VP or KP): Non-operative (medical): No treatment:	Multivariate analyses (no treatment group reference group) VA: Adj HR 0.81 (95% CI 0.42 to 1.59) p 0.55 Non-op: Adj HR 0.83 (95% CI 0.36 to 1.89)
Private health insurance (Germany)	Lange, 2014 (2006-2010) Funding: Industry	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: Adj HR 0.58 (95% CI 0.48 to 0.70)
Taiwan National Health Insurance Research Database (NHIRD)	Lin 2017‡ (2002 to 2013) Funding: NR	Matched cohort: Early VP (≤3 months: 1773 Non-VP§:5324	Mortality incidence at 1 year: 0.46 (95% CI 0.38 to 0.56) vs. 0.63 (95% CI 0.57 to 0.70) per 100 person-months Non-VP vs. VP: HR 1.39 (95% CI 1.09 to 1.78)
	Huang 2020‡ (2003 to 2013) Funding: government	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 (95% CI 0.77 to 0.99) Open surgery vs. conservative care Adj HR 0.80 (95% CI 0.70 to 0.93)

Adj = adjusted; CI = confidence interval; HR = hazard ratio; KP = kyphoplasty; NR= not reported; OR = odds ratio; VA = vertebral augmentation VP = vertebroplasty.

* Data in Ong 2018 and Edidin 2018 overlap.

† Authors do not clearly provide n's or data for propensity matched cohort.

‡ Data in Lin 2017 and Huang 2020 overlap.

§ Defined as those that did not receive VP within 3 months of VCF.

Results across administrative data studies at longer follow-up times did not consistently show an association between lower mortality and vertebral augmentation in general versus nonoperative care. At longer follow-up times, two industry-funded studies using Medicare data with overlapping sample frames report that vertebral augmentation was associated with slightly lower mortality risk compared with nonoperative care while the third study (non-industry funded) using a 20% random sample of Medicare data with less overlap reported no association. Similarly, data from a small hospital-based study reported no difference in mortality between vertebral augmentation and no treatment or between non-operative treatment and no treatment. The two studies with overlapping samples from the Taiwan National Health Insurance Research Database reported a lower mortality with augmentation versus nonoperative care as did the study using data from private health insurance in

Germany The Medicare data base studies with overlapping data reported that KP was associated with lower mortality compared with VP.

APPENDIX R. Definitions of Magnitude of Effect

Appendix Table R1. Definitions for Magnitude of Effects, Based on Mean Between-Group Differences

Slight/Small	Moderate	Large/Substantial
Pain		
5–10 points on a 0-to 100-point VAS or the equivalent 0.5–1.0 points on a 0-to 10-point numerical rating scale or the equivalent	>10–20 points on a 0-to 100-point VAS or the equivalent >1–2 points on a 0-to 10-point numerical rating scale or the equivalent	>20 points on a 0-to 100-point VAS or the equivalent >2 points on a 0-to 10-point numerical rating scale or the equivalent
1-2 points on 0-20 scale	2-4 points on 0-20 scale	>4 points on 0-20 scale
Function		
5–10 points on the ODI	>10–20 points on the ODI	>20 points on the ODI
1–2 points on the RDQ	>2–5 points on the RDQ	>5 points on the RDQ
1-2 points on Lequesne Index	>2-5 points on the Lequesne Index	5 points on the Lequesne Index
5–10 points on the WOMAC-T	>10–20 points on the WOMAT	>20 points on the WOMAC-T
3.4-6.8 points on WOMAC PF	6.8-13.8 points on WOMAC- PF	?13.6 points on WOMAC PF
5–10 points on the KOOS	>10–20 points on the KOOS	>20 points on the KOOS
5-10 points on the IKDC	>10–20 points on the IKDC	>20 points on the IKDC
5-10 points on the Lysholm	>10–20 points on the Lysholm	>20 points on the Lysholm
Pain or function		
0.2–0.5 SMD	>0.5–0.8 SMD	>0.8 SMD
1.2 to 1.4 RR/OR	1.5 to 1.9 RR/OR	≥2.0 RR/OR

ODI = Oswestry Disability Index; RDQ = Roland Morris Disability Questionnaire; SMD = standardized mean difference; VAS = visual analogue scale. WOMAC = Western Ontario and Mc Maters Universities Osteoarthritis index with T=total, PF= physical function; IKDC=International Knee Documentation Committee knee scoring system KOOS=Knee Injury and Osteoarthritis Outcome Score

APPENDIX S. Appendix References

1. Borgström F, Beall DP, Berven S, et al. Health economic aspects of vertebral augmentation procedures. Osteoporos Int 2015;26:1239-49.
2. 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.
3. Dettori JR, Norvell DC. Discordant Systematic Reviews: Which to Believe? Global Spine J 2020;10:237-9.
4. 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.
5. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. 2011.
6. 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.
7. 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.
8. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. Ann Intern Med 1992;116:78-84.
9. 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.
10. 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.
11. Quality AfHRa. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. 2014.
12. Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. Cmaj 2020;192:E901-e6.
13. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
14. 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.
15. 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.