Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Home Affairs
Federal Office of Public Health FOPH
Health and Accident Insurance Directorate
Section Health Technology Assessment

Health Technology Assessment (HTA)

HTA Report

Title	Vertebroplasty or Kyphoplasty in Painful Osteoporotic Vertebral Compression Fractures Unresponsive to Non-Surgical Treatment
Author/Affiliation	JonHenry Jacobsen, Royal Australasian College of Surgeons Alvin Atlas, Royal Australasian College of Surgeons Magdalena Moshi, Royal Australasian College of Surgeons Elise Rochet, Royal Australasian College of Surgeons Joanna Duncan, Royal Australasian College of Surgeons Ning Ma, Royal Australasian College of Surgeons Ross McLeod, eSYS Development Thomas Vreugdenburg, Royal Australasian College of Surgeons
Technology	Percutaneous Vertebroplasty and Percutaneous Balloon Kyphoplasty
Date	30 November 2020
Keywords	Vertebroplasty, kyphoplasty, vertebral fracture, osteoporosis

Bundesamt für Gesundheit
Sektion Health Technology Assessment
Schwarzenburgstrasse 157
CH-3003 Bern
Schweiz
Tel.: +41 58 462 92 30
E-mail: hta@bag.admin.ch

Conflicts of Interest

The authors have no financial, academic, personal or any other conflicts of interest to declare in relation to this project.

Executive Summary

This report evaluates the clinical effectiveness, safety, costs, and cost-utility associated with percutaneous vertebroplasty (PVP) and percutaneous balloon kyphoplasty (PBK) in patients with painful osteoporotic vertebral compression fractures (OVCF). In addition, legal, social, ethical and organisational issues associated with PVP and PBK are explored.

Clinical Evaluation

Percutaneous Vertebroplasty

The safety and clinical effectiveness of PVP was informed by 12 randomised controlled trials (RCTs), 2 non-RCTs, 2 database analyses and 15 single-arm trials. The included RCTs were of high to moderate quality, and the non-RCTs and single-arm trials were of moderate to low quality.

Compared to CT, PVP led to significant reductions in pain (mean difference [MD] -1.52; 95% confidence interval [CI] -2.86, -0.17; p = 0.03), Oswestry disability index (ODI) (MD -16.27; 95% CI -23.53, -9.01; p < 0.0001) and Roland-Morris Disability Questionnaire (RDQ) (MD -2.03; 95% CI -3.06, -1.01; p = 0.0002) at 1 month. These results remained statistically different at 12 months, but were heterogenous and did not surpass the lower bounds of minimum clinically important differences (MCIDs) (noting, the applicability of the MCIDs are uncertain). Sub-group analysis was performed on fracture age to investigate heterogeneity. Fractures younger than 8 weeks (acute) reported statistical and clinically important reductions in pain at 1 month. By 12 months, only the statistical effect persisted. Similarly, for fractures older than 8 weeks, there were statistical differences in pain at 1 and 12 months; however, they did not surpass identified MCIDs.

Compared to sham, PVP statistically reduced pain at 1 month (MD -0.76; 95% CI -1.21, -0.31; p = 0.0009) and 12 months (MD -0.88; 95% CI -1.47, -0.29; p = 0.003). The effects were heterogenous, and unlikely to translate to clinically relevant differences. Results for the remaining outcomes were inconsistent. Fractures younger than 8 weeks reported statistical differences in pain (up to 12 months) and EQ-5D (up to 6 months) favouring PVP, but the clinical importance of these differences was uncertain. Fractures older than 8 weeks reported statistical differences in pain but not EQ-5D at 1 month and 12 months. The effects for pain did not surpass identified MCIDs and did not persist at later timepoints.

The sham and CT arms were pooled for the analysis of safety. Overall, there was no statistical difference in mortality, adverse events or new fractures across the RCTs and non-RCTs. Analyses of the US Medicare database reported the relative incidences of mortality and most adverse events

(bedsores, cardiac complications, infection and pneumonia) were significantly lower at 30 days, 5 years and 10 years following PVP compared to CT (noting the absolute event rate was not reported).

Percutaneous Balloon Kyphoplasty

The safety and clinical effectiveness of PBK was informed by 4 RCTs, 4 non-RCTs, 2 databases and 6 single-arm trials. The included RCTs were of high to moderate quality, and the non-RCTs and single-arm trials were of moderate to low quality.

Compared to CT, PBK demonstrated a statistical and clinically meaningful reduction in pain at 1 week (MD -3.63; 95% CI -5.59, -1.68; $p < 0.001$). By 12 months the effect did not surpass MCID thresholds (MD -1.27; 95% CI -2.04, -0.51; $p < 0.01$). There were insufficient trials to perform sub-group analysis based on fracture age.

Mortality and adverse events were similar between PBK and CT in the RCTs and non-RCTs. Analyses of the US Medicare database reported the relative incidences of mortality, hospital readmission and adverse events were significantly lower at 30 days and 10 years following PBK compared to CT.

No studies evaluated PBK compared to sham.

Costs and Cost-Effectiveness

A decision analytic model was created to evaluate the cost-effectiveness of PVP and PBK vs CT, with probabilistic and univariate sensitivity analyses used to evaluate uncertainty and the impact of key assumptions. When considering trials that enrolled both acute (younger than 8 weeks) and sub-acute (older than 8 weeks) fractures there were no QoL improvements following PVP. The intervention was not cost-effective in this broad population given the comparator is lower cost than the intervention.

Trials evaluating acute fractures reported significant improvements in EQ-5D at some time points following PVP, and consequently an economic analysis was undertaken for this sub-population. The model determined the incremental cost-effectiveness ratio (ICER) for PVP (vs CT) to be CHF19,669 per quality-adjusted life year (QALY) using the baseline adjusted results of the VERTOS II trial at 12 months. PBK was cost-effective compared to CT with an ICER of CHF18,405 per QALY at 1 year, noting, however, that the model was informed by 1 trial in patients who had fractures for 3 months or less.

The probabilistic sensitivity analyses determined an 85% probability that PVP is superior (i.e. cost-effective) compared to CT at a willingness-to-pay threshold of CHF100,000/QALY at 12 months using results of the adjusted baseline analysis from VERTOS II. PBK was found to have an 87% probability of being superior to CT at a willingness-to-pay threshold of CHF100,000/QALY using results of the

FREE trial. Univariate sensitivity analyses indicated that the cost-effectiveness of PVP and PBK was most influenced by assumed costs for CT.

A budget impact analysis using three substitution scenarios (100%, 75% and 50% of patients substituting from PVP and PBK to CT) was undertaken to determine the financial impact of delisting PVP and PBK. If 100% of patients substituted from PVP to CT, there is a net saving of CHF6.5 million in 2020. Likewise, if 100% of patients substituted from PBK to CT, there is a net saving of CHF3.8 million in 2020. If both procedures would be delisted, there would be a collective net saving of CHF10.3 million in 2020, increasing to CHF13.5 million by 2024.

Legal, Social, Ethical and Organisational Issues

Disinvesting from PVP and PBK may impact the utilisation of healthcare resources, as the procedures had a shorter length of stay, and patients were more likely to be discharged home, compared to CT. The lack of consistent differences between the sham and intervention arms makes it unclear whether PVP exhibited true clinical effectiveness or whether the effects were attributable to a placebo- or confounding-effect.

Conclusion

PVP and PBK appeared to have a beneficial effect on pain in the short-term compared to CT and sham, and acute fractures (less than 8 weeks old) appeared to be more responsive to these procedures; however, the differences generally do not persist over time. PVP and PBK reported comparable safety compared to CT and sham from RCTs and non-RCTs; however, results from larger database analyses indicated PVP and PBK reduced mortality and adverse event rates at 30 days and 10 years post-intervention.

PVP was not cost-effective when using estimates derived from studies that included both acute and sub-acute fractures, as there were no QoL improvements. If results of the VERTOS II trial were considered, which included patients with acute fractures only, PVP was cost-effective compared to CT at 12 months. Similarly, PBK was cost-effective compared to CT, noting the EQ-5D estimates were informed by only 1 trial. Delisting PVP and PBK would result in a net cost saving for the payer.

Zusammenfassung

In diesem Bericht werden die klinische Wirksamkeit, Sicherheit, Kosten sowie das Kosten-Nutzen-Verhältnis der perkutanen Vertebroplastie (PVP) und der perkutanen Ballon-Kyphoplastie (PBK) bei Patienten mit schmerzhaften osteoporotischen Wirbelkompressionsfrakturen (OVCF) evaluiert. Zudem wird auf rechtliche, soziale, ethische und organisatorische Probleme im Zusammenhang mit der PVP und der PBK eingegangen.

Klinische Beurteilung

Perkutane Vertebroplastie

Die Sicherheit und klinische Wirksamkeit der PVP wurde in 12 randomisierten kontrollierten Studien (RKS), 2 Nicht-RKS, 2 Datenbankanalysen und 15 einarmigen Studien untersucht. Die eingeschlossenen RKS waren von hoher bis mittlerer Qualität, und die Nicht-RKS sowie die einarmigen Studien waren von mittlerer bis geringer Qualität.

Im Vergleich zur konservativen Therapie (KT) führte die PVP nach 1 Monat zu signifikanten Reduktionen der Schmerzen (mittlere Differenz [MD] -1,52; 95 %-Konfidenzintervall [KI] -2,86, -0,17; p = 0,03), im Oswestry Disability Index (ODI) (MD -16,27; 95 %-KI -2353, -9,01; p < 0,0001) sowie im Roland-Morris Disability Questionnaire (RDQ) (MD -2,03; 95 %-KI -3,06, -1,01; p = 0,0002). Die statistischen Unterschiede zwischen diesen Ergebnissen bestanden auch nach 12 Monaten. Die Ergebnisse waren jedoch heterogen und die unteren Grenzen der minimalen klinisch bedeutsamen Unterschiede (MCID) wurden nicht überschritten (wobei anzumerken ist, dass die Anwendbarkeit der MCID mit Unsicherheiten behaftet ist). Zur Untersuchung der Heterogenität wurde eine Subgruppenanalyse hinsichtlich des Frakturalters durchgeführt. Bei Frakturen, die jünger als 8 Wochen waren (akut), wurde nach 1 Monat eine statistisch und klinisch bedeutsame Reduktion der Schmerzen festgestellt. Nach 12 Monaten war lediglich der statistische Effekt vorhanden. Ebenfalls wurden bei Frakturen, die älter als 8 Wochen waren, nach 1 Monat und nach 12 Monaten statistische Unterschiede hinsichtlich der Schmerzen festgestellt. Diese überschritten jedoch nicht die identifizierten MCID.

Im Vergleich zum Scheineingriff führte die PVP zu einer statistischen Reduktion der Schmerzen nach 1 Monat (MD -0,76; 95 %-KI -1,21, -0,31; p = 0,0009) und nach 12 Monaten (MD -0,88; 95 %-KI -1,47, -0,29; p = 0,003). Die Effekte waren heterogen und es ist unwahrscheinlich, dass sie zu klinisch relevanten Unterschieden führen. Die Ergebnisse hinsichtlich der übrigen Endpunkte waren inkonsistent. Bei Frakturen, die jünger als 8 Wochen waren, wurden statistische Unterschiede bei Schmerzen (bis zu 12 Monate) und EQ-5D (bis zu 6 Monate) zugunsten der PVP festgestellt. Die klinische Bedeutung dieser Unterschiede war jedoch unklar. Für Frakturen, die älter als 8 Wochen

waren, wurden nach 1 Monat und 12 Monaten statistische Unterschiede bei den Schmerzen berichtet, jedoch nicht beim EQ-5D. Die Effekte hinsichtlich der Schmerzen überstiegen nicht die identifizierten MCID und wurden zu späteren Zeitpunkten nicht mehr festgestellt.

Die Scheineingriff- und KT-Arme wurden für die Sicherheitsanalyse gepoolt. Insgesamt gab es in den RKS und den Nicht-RKS keinen statistischen Unterschied betreffend Mortalität, unerwünschte Ereignisse oder neue Frakturen. Analysen der US-amerikanischen Medicare-Datenbank haben aufgezeigt, dass die relativen Inzidenzen der Mortalität und der meisten unerwünschten Ereignisse (Dekubitus, kardiale Komplikationen, Infektionen und Pneumonie) 30 Tage, 5 Jahre und 10 Jahre nach der PVP im Vergleich zur KT signifikant niedriger waren (wobei die absolute Ereignisrate nicht angegeben wurde).

Perkutane Ballon-Kyphoplastie

Die Sicherheit und klinische Wirksamkeit der PBK wurde in 4 RKS, 4 Nicht-RKS, 2 Datenbankanalysen und 6 einarmigen Studien untersucht. Die eingeschlossenen RKS waren von hoher bis mittlerer Qualität, und die Nicht-RKS sowie die einarmigen Studien waren von mittlerer bis geringer Qualität.

Im Vergleich zur KT führte die PBK nach 1 Woche zu einer statistisch und klinisch bedeutsamen Reduktion der Schmerzen (MD -3,63; 95 %-KI -5,59, -1,68; $p < 0,001$). Nach 12 Monaten überschritt der Effekt die MCID-Schwellenwerte nicht (MD -1,27; 95 %-KI -2,04, -0,51; $p < 0,01$). Eine Subgruppenanalyse hinsichtlich des Frakturalters konnte aufgrund nicht genügender Anzahl an Studien nicht durchgeführt werden.

Für die PBK und die KT wurden in den RKS und Nicht-RKS ähnliche Daten betreffend Mortalität und unerwünschte Ereignisse berichtet. Analysen der US-amerikanischen Medicare-Datenbank haben aufgezeigt, dass die relativen Inzidenzen der Mortalität, der Re-Hospitalisierung und der unerwünschten Ereignisse nach 30 Tagen und 10 Jahren nach der PBK im Vergleich zur KT signifikant niedriger war.

Eine Beurteilung der PBK im Vergleich zum Scheineingriff wurde in keiner der Studien vorgenommen.

Kosten und Kosteneffektivität

Ein entscheidungsanalytisches Modell wurde erstellt, um eine Evaluation der Kosteneffektivität der PVP und der PBK im Vergleich zur KT durchzuführen, wobei zur Beurteilung der Unsicherheit und der Auswirkungen der wichtigsten Annahmen probabilistische und univariate Sensitivitätsanalysen verwendet wurden. Bei der Analyse von Studien, in die sowohl akute (jünger als 8 Wochen) als auch subakute (älter als 8 Wochen) Frakturen eingeschlossen wurden, wurden nach der PVP keine

Verbesserungen der Lebensqualität festgestellt. Die Intervention war in dieser breiten Population nicht kosteneffektiv, da sich der Komparator im Vergleich zur Intervention als kostengünstiger erwies.

In Studien, in denen akute Frakturen auswertet wurden, konnten zu einigen Zeitpunkten nach der PVP signifikante Verbesserungen des EQ-5D festgestellt werden. Folglich wurde eine ökonomische Analyse für diese Subpopulation durchgeführt. Anhand des Modells wurde unter Verwendung der Baseline-angepassten Ergebnisse der VERTOS II-Studie das inkrementelle Kosten-Effektivitäts-Verhältnis (ICER) für die PVP (gegenüber der KT) von CHF 19'669 pro qualitätsangepasstes Lebensjahr (QALY) nach 12 Monaten festgestellt. Mit einem ICER von CHF 18'405 pro QALY nach 1 Jahr war die PBK im Vergleich zur KT kosteneffektiv. Zu beachten ist jedoch, dass das Modell lediglich eine Studie berücksichtigte, die sich mit Patienten befasste, deren Frakturen 3 Monate alt oder jünger waren.

In probabilistischen Sensitivitätsanalysen, in denen Baseline-angepasste Ergebnisse von VERTOS II zur Anwendung kamen, wurde eine Wahrscheinlichkeit von 85 Prozent für die Überlegenheit (d. h. Kosteneffektivität) der PVP gegenüber der KT nach 12 Monaten bei einer Zahlungsbereitschaftsschwelle von CHF 100’000/QALY festgestellt. Anhand der Ergebnisse der FREE-Studie wurde festgestellt, dass die PBK mit einer Wahrscheinlichkeit von 87 Prozent bei einer Zahlungsbereitschaftsschwelle von CHF 100’000/QALY gegenüber der KT überlegen ist. Univariate Sensitivitätsanalysen haben aufgezeigt, dass die Kosteneffektivität der PVP und der PBK von den angenommenen Kosten der KT am stärksten beeinflusst wurde.

Eine Budget-Impact-Analyse wurde unter Verwendung von drei Substitutionsszenarien (100 %, 75 % und 50 % der Patienten, die von der PVP und der PBK auf die KT umgestellt wurden) durchgeführt, um die finanziellen Auswirkungen der Streichung der PVP und der PBK aus der Liste zu bestimmen. Eine Umstellung von 100 Prozent der Patienten von der PVP auf die KT entspricht einer Nettoeinsparung von CHF 6,5 Millionen für das Jahr 2020. Ebenfalls hätte eine Umstellung von 100 Prozent der Patienten von der PBK auf die KT zu einer Nettoeinsparung von CHF 3,8 Millionen für das Jahr 2020 geführt. Eine Streichung beider Verfahren aus der Liste ergäbe eine Nettoeinsparung von CHF 10,3 Millionen im Jahr 2020, die bis 2024 auf CHF 13,5 Millionen ansteigen würde.

Rechtliche, soziale, ethische und organisatorische Probleme

Die Streichung der PVP und der PBK aus der Liste kann sich auf die Inanspruchnahme von Gesundheitsressourcen auswirken, da die Verfahren mit einer kürzeren Verweildauer einhergehen und die Patienten im Vergleich zur KT eher nach Hause entlassen werden konnten. Aufgrund des Fehlens konsistenter Unterschiede zwischen dem Scheineingriffs- und dem Interventionsarm besteht

Unklarheit darüber, ob die PVP eine echte klinische Wirksamkeit aufwies oder die Effekte auf einen Placebo- oder Confounding-Effekt zurückzuführen waren.

Fazit

Im Vergleich zur KT und zum Scheineingriff schienen die PVP und die PBK kurzfristig eine positive Auswirkung auf die Schmerzen zu haben. Zudem schienen akute Frakturen (weniger als 8 Wochen alt) besser auf diese Verfahren anzusprechen. Im Verlauf der Zeit hatten die Unterschiede im Allgemeinen jedoch keinen Bestand. Im Vergleich zur KT und zum Scheineingriff haben die PVP und die PBK in den RKS und den Nicht-RKS eine vergleichbare Sicherheit aufgezeigt. In grösseren Datenbankanalysen ergaben sich jedoch Hinweise dafür, dass die PVP und die PBK 30 Tage und 10 Jahre nach der Intervention eine Reduktion der Mortalitätsraten sowie der Raten unerwünschter Ereignisse zur Folge hatten.

Die PVP war bei Verwendung von Schätzungen, die aus Studien, die sowohl akute als auch subakute Frakturen einschlossen, abgeleitet wurden, nicht kosteneffektiv, da keine Verbesserungen der Lebensqualität festgestellt wurden. Bei der Verwendung von Ergebnissen aus der Studie VERTOS II, bei der nur Patienten mit akuten Frakturen berücksichtigt wurden, war die PVP im Vergleich zur KT nach 12 Monaten kosteneffektiv. Die PBK war im Vergleich zur KT ebenfalls kosteneffektiv, wobei zu beachten ist, dass die EQ-5D-Schätzungen auf lediglich einer Studie beruhten. Die Streichung der PVP und der PBK aus der Liste würde zu einer Nettokosteneinsparung für den Kostenträger führen.

Synthèse

Le présent rapport évalue l'efficacité clinique, la sécurité, les coûts et le rapport coût/efficacité associés à la vertébroplastie percutanée (VP) et à la cyphoplastie percutanée à ballonnets (CPB) chez les patients atteints de fractures vertébrales ostéoporotiques par compression (FVOC) s'accompagnant de douleurs. De plus, il explore les questions juridiques, sociales, éthiques et organisationnelles associées à la VP et à la CPB.

Évaluation clinique

Vertébroplastie percutanée

La sécurité et l'efficacité clinique de la VP ont été déterminées par 12 essais contrôlés randomisés (RCT pour *randomised controlled trials*), 2 essais non randomisés, 2 analyses de bases de données et 15 essais à bras unique. Les RCT considérés étaient de qualité élevée à modérée et les essais non randomisés et à bras unique étaient de qualité modérée à faible.

Par rapport au traitement conservateur (TC), la VP entraînait une réduction significative de la douleur (différence moyenne [DM] -1,52 ; intervalle de confiance [IC] à 95 % -2,86, -0,17 ; p = 0,03) et des scores aux questionnaires d'Oswestry (ODI) (DM -16,27 ; 95 % IC -23,53, -9,01 ; p < 0,0001) et de Roland-Morris (RDQ) (DM -2,03 ; IC à 95 % -3,06, -1,01 ; p = 0,0002) à 1 mois. Ces résultats restaient statistiquement différents à 12 mois ; cependant, ils étaient hétérogènes et ne dépassaient pas les limites inférieures des différences minimales cliniquement importantes (MCID). Notons qu'il n'est pas certain que les MCID soient applicables ici. Une analyse de sous-groupes a été effectuée en fonction de l'âge de la fracture, afin de rechercher l'origine de l'hétérogénéité des résultats. Les fractures de moins de 8 semaines (aiguës) étaient associées à des réductions statistiquement et cliniquement importantes de la douleur à 1 mois. À 12 mois, seul l'effet statistique persistait. De même, pour les fractures de plus de 8 semaines, il y avait des différences statistiques dans la douleur à 1 et à 12 mois ; toutefois, ces différences ne dépassaient pas les MCID identifiées.

Par rapport au placebo, la VP induisait une diminution statistique de la douleur à 1 mois (DM -0,76 ; IC à 95 % -1,21, -0,31 ; p = 0,0009) et à 12 mois (DM -0,88 ; IC à 95 % -1,47, -0,29 ; p = 0,003). Les effets étaient hétérogènes et peu susceptibles de se traduire par des différences cliniquement pertinentes. Les résultats sur les autres critères examinés étaient inégaux. Les fractures de moins de 8 semaines étaient associées à des différences statistiques dans la douleur (jusqu'à 12 mois) et le score EQ-5D (jusqu'à 6 mois) en faveur de la VP, mais l'importance clinique de ces différences était incertaine. Les fractures de plus de 8 semaines étaient associées à des différences statistiques dans la douleur, mais pas dans le score EQ-5D, à 1 et à 12 mois. Concernant la douleur, les effets ne dépassaient pas les MCID identifiées et ne persistaient pas dans le temps.

Le bras placebo et le bras TC ont été regroupés pour l'analyse de la sécurité. Dans l'ensemble, il n'y avait pas de différence statistique en termes de mortalité, d'événements indésirables ou de nouvelles fractures dans les RCT et les essais non randomisés. Des analyses de la base de données américaine *Medicare* ont révélé que les incidences relatives de la mortalité et de la plupart des événements indésirables (escarres, complications cardiaques, infections et pneumonies) étaient significativement plus faibles 30 jours, 5 ans et 10 ans après une VP, comparativement au TC. Remarquons que le taux absolu d'événements n'était pas indiqué.

Cyphoplastie percutanée à ballonnets

La sécurité et l'efficacité clinique de la CPB ont été déterminées par 4 RCT, 2 essais non randomisés, 2 analyses de bases de données et 6 essais à bras unique. Les RCT considérés étaient de qualité élevée à modérée et les essais non randomisés et à bras unique étaient de qualité modérée à faible.

Par rapport au TC, la CPB entraînait une diminution statistique et cliniquement significative de la douleur à 1 semaine (DM -3,63 ; IC à 95 % -5,59, -1,68 ; $p < 0,001$). À 12 mois, l'effet ne dépassait pas les seuils de MCID (DM -1,27 ; IC à 95 % -2,04, -0,51 ; $p < 0,01$). Il n'y avait pas suffisamment d'essais pour effectuer une analyse de sous-groupes en fonction de l'âge de la fracture.

Dans les RCT et les essais non randomisés, la mortalité et les événements indésirables étaient similaires entre la CPB et le TC. Des analyses de la base de données américaine *Medicare* ont révélé que l'incidence relative de la mortalité, de la réhospitalisation et des événements indésirables était significativement plus faible à 30 jours ainsi que 10 ans après une CPB comparativement au TC.

Aucune étude n'a évalué la CPB par rapport à une intervention placebo.

Coûts et rapport coût-efficacité

Un modèle analytique de décision a été créé pour évaluer le rapport coût-efficacité de la VP et de la CPB par rapport au TC. L'incertitude et l'impact des hypothèses principales ont été évalués au moyen d'analyses de sensibilité univariées et probabilistes. Si l'on se base sur les essais portant sur des fractures à la fois aiguës (moins de 8 semaines) et subaiguës (plus de 8 semaines), aucune amélioration de la qualité de vie n'a été constatée après une VP. Dans cette vaste population, le rapport coût-efficacité de l'intervention était défavorable, étant donné le coût supérieur de l'intervention par rapport au traitement reçu par le groupe témoin.

Les essais portant sur des fractures aiguës ont décrit des améliorations significatives dans le score EQ-5D à certains points dans le temps après la VP. Par conséquent, une analyse économique a été réalisée pour cette sous-population. En se basant sur les résultats après ajustement de l'essai VERTOS II à 12 mois, le modèle a déterminé que le rapport coût-efficacité différentiel (ICER) pour la VP (par rapport au TC) était de 19 669 CHF par année de vie pondérée par la qualité (QALY). La

CPB présentait un rapport coût-efficacité favorablecomparativement au TC, avec un ICER de 18 405 CHF par QALY à 1 an. Notons toutefois que le modèle se fondait sur un seul essai clinique dans lequel les patients souffraient de fractures depuis 3 mois ou moins.

D'après les analyses de sensibilité probabilistes, la VP avait 85 % de probabilité d'être supérieure (c.-à-d. d'un meilleur rapport coût-efficacité) au TC avec un seuil de consentement à payer s'élevant à 100 000 CHF par QALY à 12 mois, selon les résultats de l'analyse après ajustement de VERTOS II. En se basant sur les résultats de l'essai FREE, la CPB avait 87 % de probabilité d'être supérieure au TC avec un seuil de consentement à payer de 100 000 CHF par QALY. Les analyses de sensibilité univariées indiquaient que le rapport coût-efficacité de la VP et de la CPB était surtout influencé par les coûts présumés pour le TC.

Une analyse d'incidence budgétaire fondée sur trois scénarios de substitution (en faisant passer 100 %, 75 % et 50 % des patients de la VP et de la CPB au TC) a été réalisée afin de déterminer l'impact financier de l'exclusion de la VP et de la CPB. Si 100 % des patients étaient passés de la VP au TC, il y aurait eu une économie nette de 6,5 millions de CHF en 2020. De même, si 100 % des patients étaient passés de la CPB au TC, il y aurait eu une économie nette de 3,8 millions de CHF en 2020. Si les deux procédures avaient été exclues, il y aurait eu une économie nette collective de 10,3 millions de CHF en 2020, qui atteindrait 13,5 millions de CHF d'ici 2024.

Questions juridiques, sociales, éthiques et organisationnelles

Un désinvestissement en matière de VP et de CPB pourrait avoir un impact sur l'utilisation des ressources de santé, car ces procédures sont associées à une durée d'hospitalisation plus courte et à un retour plus fréquent des patients à la maison, comparativement au TC. Le manque de cohérence dans les différences entre le bras placebo et le bras intervention ne permet pas de déterminer avec certitude si la VP a une véritable efficacité clinique ou si les effets sont attribuables à un effet placebo ou à un biais de confusion.

Conclusion

Il semble que la VP et la CPB aient un effet bénéfique sur la douleur à court terme par rapport au TC et à une intervention placebo. Les fractures aiguës (moins de 8 semaines) semblent plus sensibles à ces procédures ; toutefois, les différences ne persistent généralement pas dans le temps. La VP et la CPB présentaient une sécurité comparable à celle du TC et du placebo dans les RCT et les essais non randomisés. Cependant, d'après les résultats d'analyses de bases de données plus importantes, la VP et la CPB réduisaient les taux de mortalité et d'événements indésirables 30 jours et 10 ans après l'intervention.

En se basant sur les estimations tirées d'études incluant à la fois des fractures aiguës et subaiguës, la VP avait un rapport coût-efficacité défavorable car il n'y avait pas d'amélioration de la qualité de vie. En revanche, en prenant en compte les résultats de l'essai VERTOS II, qui incluait uniquement des patients souffrant de fractures aiguës, la VP avait un rapport coût-efficacité favorable comparée au TC à 12 mois. De même, la CPB avait un rapport coût-efficacité favorable par rapport au TC. Il convient de noter que les estimations des scores EQ-5D provenaient d'un seul essai. L'exclusion de la VP et de la CPB entraînerait une réduction nette des coûts pour le payeur.

Executive Summary

Nel presente rapporto vengono valutati l'efficacia clinica, la sicurezza, i costi e il rapporto costo-utilità della vertebroplastica percutanea (PVP) e della cifoplastica percutanea con palloncino (PBK) in pazienti con fratture osteoporotiche vertebrali da compressione (OVCF) dolenti. Sono inoltre investigati alcuni aspetti legali, sociali, etici e organizzativi legati alla PVP e alla PBK.

Valutazione clinica

Vertebroplastica percutanea

Le informazioni riguardo alla sicurezza e all'efficacia clinica della PVP sono tratte da 12 trial controllati randomizzati (RCT), 2 trial controllati non randomizzati (non-RCT), 2 analisi di database e 15 trial a braccio singolo. Gli RCT inclusi erano di qualità da alta a moderata, mentre i non-RCT e i trial a braccio singolo erano di qualità da moderata a bassa.

In confronto al trattamento clinico (CT), la PVP induceva riduzioni significative del dolore (differenza media [MD] -1.52; 95 % intervallo di confidenza [CI] -2.86, -0.17; p = 0.03), indice di disabilità Oswestry (ODI) (MD -16.27; 95 % CI -23.53, -9.01; p < 0.0001) e il questionario Roland-Morris sulla disabilità (RDQ) (MD -2.03; 95 % CI -3.06, -1.01; p = 0.0002) a 1 mese. Questi risultati rimanevano statisticamente diversi a 12 mesi, ma erano eterogenei e non superavano i limiti inferiori delle differenze minime clinicamente importanti (MCID) (da notare tuttavia che l'applicabilità delle MCID era incerta). Per analizzare l'eterogeneità si è effettuata un'analisi di sottogruppo sull'età della frattura. Nelle fratture risalenti a meno di 8 settimane (acute) erano riportate riduzioni statisticamente e clinicamente importanti del dolore a 1 mese. A 12 mesi persisteva unicamente l'effetto statistico. Differenze statistiche del dolore a 1 e 12 mesi erano analogamente riportate nelle fratture risalenti a più di 8 settimane, ma senza superare le MCID identificate.

In confronto al trattamento fittizio (sham), la PVP riduceva statisticamente il dolore a 1 mese (MD -0.76; 95 % CI -1.21, -0.31; p = 0.0009) e a 12 mesi (MD -0.88; 95 % CI -1.47, -0.29; p = 0.003). Gli effetti erano eterogenei, ma difficilmente traducibili in differenze clinicamente rilevanti. I valori relativi ai restanti esiti risultavano inoltre incoerenti. Nelle fratture risalenti a meno di 8 settimane erano riportate differenze statistiche a livello di dolore (fino a 12 mesi) e di EQ-5D (fino a 6 mesi) a favore della PVP, ma l'importanza clinica di queste differenze era incerta. Le fratture risalenti a più di 8 settimane riportavano differenze statistiche a livello di dolore ma non di EQ-5D a 1 mese e 12 mesi. Gli effetti relativi al dolore non superavano le MCID identificate e non persistevano in punti temporali successivi.

Per l'analisi della sicurezza sono stati congiunti i bracci sham e CT. Nell'insieme non si è rilevata alcuna differenza statistica a livello di mortalità, eventi avversi o nuove fratture né negli RCT né nei

non-RCT. Da analisi del database di US Medicare risultava che mortalità ed eventi avversi gravi (piaghe da decubito, complicazioni cardiache, infezioni e polmoniti) avevano incidenze relative significativamente più basse a 30 giorni, 5 anni e 10 anni dopo la PVP rispetto al CT (da notare tuttavia che non era riportato il tasso assoluto di eventi).

Cifoplastica percutanea con palloncino

Le informazioni sulla sicurezza e l'efficacia clinica sono tratte da 4 RCT, 4 non-RCT, 2 analisi di database e 6 trial a braccio singolo. Gli RCT inclusi erano di qualità da alta a moderata, mentre i non-RCT e i trial a braccio singolo erano di qualità da moderata a bassa.

In confronto al CT, la PBK mostrava una riduzione statisticamente e clinicamente significativa del dolore a 1 settimana (MD -3.63; 95% CI -5.59, -1.68; $p < 0.001$). A 12 mesi l'effetto non superava le soglie delle MCID (MD -1.27; 95 % CI -2.04, -0.51; $p < 0.01$). I trial non erano tuttavia sufficienti a condurre analisi di sottogruppo in funzione dell'età della frattura.

Mortalità e eventi avversi di PBK e CT erano simili sia negli RCT che nei non-RCT. Da analisi del database di US Medicare risultava che mortalità, riammissione in ospedale ed eventi avversi avevano incidenze relative significativamente più basse a 30 giorni e a 10 anni dopo la PBK rispetto al CT.

Nessuno studio valutava la PBK rispetto allo sham.

Costi e rapporto costo-efficacia

Per valutare l'efficacia in rapporto al costo della PVP e della PBK rispetto al CT si è creato un modello analitico decisionale utilizzando analisi di sensibilità probabilistica e univariata per valutare l'incertezza e l'impatto delle ipotesi chiave. Considerando i trial che arruolavano sia fratture acute (risalenti a meno di 8 settimane) che sub-acute (risalenti a più di 8 settimane), non si riscontravano miglioramenti in termini di qualità di vita dopo la PVP. L'intervento risultava quindi inefficace in rapporto al costo in questa popolazione allargata, essendo i costi del comparatore più bassi rispetto a quelli dell'intervento.

I trial che valutavano fratture acute riportavano invece nel EQ-5D miglioramenti significativi in alcuni punti temporali dopo la PVP, per cui si è proceduto a un'analisi economica per questa sotto-popolazione. Il modello ha determinato il rapporto costo-efficacia incrementale (ICER) della PVP (rispetto al CT) in 19 669 franchi per anno QALY (anni di vita corretti in funzione della qualità) usando i risultati dell'analisi di baseline corretta del trial VERTOS II a 12 mesi. La PBK risultava efficace in rapporto al costo rispetto al CT con un ICER di 18 405 franchi per QALY a 1 anno, considerando tuttavia che il modello attingeva informazioni da 1 trial su pazienti con fratture risalenti a 3 mesi o meno.

Le analisi di sensibilità probabilistica hanno determinato una probabilità dell'85 % che la PVP fosse superiore (ovvero efficace in termini di costo) in confronto al CT a una soglia di disponibilità a pagare di 100 000 franchi/QALY a 12 mesi utilizzando i risultati dell'analisi di baseline corretta di VERTOS II. La PBK mostrava invece una probabilità dell'87 % di essere superiore al CT a una soglia di disponibilità a pagare di 100 000 franchi/QALY utilizzando i risultati del trial gratuito. Le analisi di sensibilità univariate indicavano che l'efficacia in rapporto al costo della PVB e della PBK era influenzata in massima parte dai presunti costi del CT.

Per determinare le ripercussioni finanziarie dovute alla cancellazione di interventi di PVP e PBK si è proceduto a un'analisi di impatto sul budget usando tre scenari di sostituzione (100 %, 75 % e 50 % di pazienti che sostituiscono una PVP e una PBK con il CT). Se il 100 % dei pazienti avesse sostituito una PVP con un CT, per il 2020 sarebbe stato realizzato un risparmio netto di 6,5 milioni di franchi. Analogamente, se il 100 % dei pazienti avesse sostituito una PBK con il CT per il 2020 sarebbe stato realizzato un risparmio netto di 3,8 milioni di franchi. Se entrambe le procedure fossero state cancellate, per il 2020 sarebbe stato realizzato un risparmio netto complessivo di 10,3 milioni di franchi che sarebbe aumentato a 13,5 milioni di franchi di qui al 2024.

Aspetti legali, sociali, etici e organizzativi

Disinvestire da interventi di PVP e PBK potrebbe ripercuotersi sull'utilizzo di risorse sanitarie nella misura in cui tali procedure comportavano soggiorni in ospedale più brevi e i pazienti avevano maggiori probabilità di essere dimessi prima rispetto al CT. L'assenza di differenze coerenti tra i bracci sham e d'intervento rende tuttavia difficile capire se la PVP presentasse una vera efficacia clinica o se gli effetti fossero attribuibili a fattori placebo o che potessero indurre a confusione.

Conclusioni

Sia la PVP che la PBK hanno mostrato di avere un effetto benefico sul dolore nel breve termine in confronto al CT e allo sham e le fratture acute (risalenti a meno di 8 settimane) si sono dimostrate più reattive a queste procedure. In generale, tuttavia, tali differenze non perdurano nel tempo. Sebbene la PVP e la PBK mostrassero una sicurezza comparabile rispetto al CT e sham, sia negli RCT che nei non-RCT, i risultati di analisi di database di maggior dimensioni indicavano tuttavia che le prime riducevano la mortalità e i tassi di eventi avversi a 30 giorni e a 10 anni dopo l'intervento.

Utilizzando stime derivate da studi che includevano fratture sia acute che sub-acute, la PVP risultava inefficace in rapporto al costo in quanto non dava miglioramenti a livello di QoL. Considerando i risultati del trial VERTOS II, che includeva unicamente pazienti con fratture acute, la PVP risultava invece efficace in rapporto al costo rispetto al CT a 12 mesi. Pur tenendo conto che le stime EQ-5D

traevano informazioni da un solo trial, risultava efficace rispetto al costo anche la PBK. Cancellare interventi di PVP e di PBK si tradurrebbe dunque in un risparmio di costi netto per il pagante.

Table of Contents

List of Figures

List of Tables

Abbreviations and acronyms

ADL	Activities of daily living
CHUV	Centre Hospitalier Universitaire Vaudois
CIRSE	Cardiovascular and Interventional Radiological Society of Europe
EUnetHTA	European Network for Health Technology Assessment
CHF	Swiss franc
CT	Conservative treatment
CUA	Cost utility analysis
EQ-5D	EuroQol 5 dimension questionnaire
FOPH	Swiss Federal Office of Public Health
GRADE	Grading of recommendations, assessment, development and evaluations
HR-QoL	Health-related quality of life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
MCID	Minimum clinically important difference
MRI	Magnetic resonance imaging
MMSE	Mini-Mental state examination
MSAC	Medical Services Advisory Committee
NA	Not applicable
NICE	National Institute of Health and Care Excellence
Non-RCT	Non-Randomised controlled trial
NRS	Numerical rating scale
NSAIDs	Non-Steroidal anti-inflammatory drugs
NSM	Non-Surgical management
ODI	Oswestry disability index
OVCF	Osteoporotic vertebral compression fractures

PBK	Percutaneous balloon kyphoplasty
PICO	Patients, intervention, comparator, outcome
PMMA	Polymethyl methacrylate
PVP	Percutaneous vertebroplasty
QALY	Quality-Adjusted life year
QUALEFFO	Quality of life questionnaire of the European Foundation for Osteoporosis
RCT	Randomised controlled trial
RDQ	Roland-Morris disability questionnaire
SF-36	Short form 36 questionnaire
SOF-ADL	Study of osteoporotic fractures–activities of daily living
STIR	Short-TI inversion recovery
TCM	Traditional Chinese medicine
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
WHO	World Health Organization

Objective of the HTA report

The Federal Office of Public Health (FOPH) is reviewing the public reimbursement of percutaneous vertebroplasty (PVP) and percutaneous balloon kyphoplasty (PBK) for the treatment of osteoporotic painful vertebral compression fractures.

The process to evaluate health technologies involves multiple phases: 1) the pre-scoping phase, 2) the scoping phase, and 3) the health technology assessment (HTA) phase. This document represents the outcome of the HTA phase.

The objective of an HTA is to generate a focused assessment of various aspects of a health technology. The analytical methods applied to assess the value of using a health technology are described. The analytical process is comparative, systematic and transparent, and involves multiple stakeholders. The domains covered in an HTA report include clinical effectiveness and safety; costs, cost-effectiveness and budget impact; and legal, social, ethical and organisational issues. The purpose is to inform health policy and decision-making, thereby promoting an efficient, sustainable, equitable and high-quality health system.

1 Policy Question and Context

PVP and PBK are used to treat vertebral fractures. These procedures are indicated for a range of fracture types, most commonly for the treatment of painful osteoporotic vertebral compression fractures (OVCFs). In 2012–2013, 2,894 PVP procedures and 1,278 PBK procedures were conducted in patients with OVCFs in Switzerland. Considerable regional variation exists in age- and sex-standardised procedure rates, ranging from 1.0 to 10.1 per 10,000 persons across hospital service areas.[1] This regional variability is not wholly explained by population demographics or socioeconomic factors and may represent differences in clinician preferences.[1]

There is discord in the scientific literature, clinical practice guidelines and international reimbursement policies regarding the utilisation of PVP and PBK for patients with OVCFs:

- In the literature, randomised-controlled trials (RCTs) evaluating PVP for OVCFs report conflicting results regarding pain, disability and quality of life outcomes.[2-8]
- Clinical practice guidelines have differing recommendations, with the American Academy of Orthopaedic Surgeons recommending against PVP for OVCFs (2010) and the American College of Radiology, the American Society of Neuroradiology, the American Society of Spine Radiology, the Society of Interventional Radiology, and the Society of Neurointerventional Surgery endorsing vertebral augmentation procedures for OVCFs (June 2019).[9] [10] The Cardiovascular and Interventional Radiological Society of Europe (CIRSE) recommend vertebral augmentation for OVCFs (2017).[11]
- The National Institute of Health and Care Excellence (NICE) in the United Kingdom (UK) recommends the use of PVP and PBK for patients with severe ongoing pain attributable to the fracture who have failed conservative treatments.[12] In contrast, Australia removed PVP and PBK from the private reimbursement list in 2011, following an HTA evaluation concluding there was insufficient evidence of clinical benefit and unacceptable cost-effectiveness.[13] However, the procedure was relisted for reimbursement in 2020 for acute OVCFs.[14]
- In Switzerland, PVP is reimbursed without restrictions and PBK is reimbursed for patients with acute symptomatic fractures within 8 weeks of onset, that are unresponsive to conservative treatment, with more than 15º of kyphosis and/or more than one-third loss of vertebral body height.[15]

To address concerns regarding the high degree of regional variability, and the ongoing debate regarding the relative clinical and cost-effectiveness of PVP and PBK, an HTA evaluation has been commissioned to inform a policy decision on the reimbursement of these procedures.

2 Medical Background

2.1 Health Condition

Osteoporosis is a common musculoskeletal disorder and a leading cause of fractures among the elderly.[16 17] The disorder is characterised by an imbalance in bone remodelling.[18 19] In brief, remodelling requires the sequential activation of 2 bone cells: osteoclasts and osteoblasts.[20 21] Osteoclasts dissolve existing bone via acidification and enzymatic digestion, leading to bone resorption. Osteoblasts are subsequently activated and secrete matrix proteins, which promotes bone formation.[21 22] Under normal conditions, the rate of bone resorption is balanced with the rate of deposition, thus maintaining bone strength and integrity. As we age, the rate of resorption slightly outpaces deposition. During osteoporosis however, the balance shifts towards excessive resorption, resulting in a loss of bone mass, increased bone porosity and deterioration of the bone's microarchitecture.[20 21 23 24] Resulting changes promote bone fragility and increase the risk of fracture from minimally traumatic events (fragility fractures) such as falls or lifting small objects.[25]

There is no single cause of osteoporosis. Rather, the disorder is polygenic and influenced by multiple extrinsic factors including lifestyle, medication use, nutrition and sex hormones; and intrinsic factors such as age, ethnicity, family history and sex.[26] These factors alter the pathophysiology of osteoporosis and give rise to the different types of the disorder—primary and secondary osteoporosis.[27 28]

Primary osteoporosis includes idiopathic, type I and type II osteoporosis. Idiopathic osteoporosis generally occurs in children and young adults without any identifiable cause.[29] Type I osteoporosis occurs in post-menopausal females and is characterised by increased bone resorption and calcium excretion due to oestrogen deficiency.[22] Type II osteoporosis occurs in males and females aged 70 years and older and is characterised by cellular senescence.[22] The ability of bone marrow to synthesise osteoblast precursors is reduced, thereby leading to increase bone resorption.[20] Secondary osteoporosis is generally caused by underlying medical conditions, for example, inflammatory conditions, hypogonadism and endocrinopathies; or medications including, corticosteroids, antiepileptics, selective serotonin reuptake inhibitors and chemotherapy.[30]

Osteoporosis is asymptomatic and individuals are often unaware they have the disorder until a fracture occurs, noting many fractures may also be asymptomatic.[31] Owing to the systemic nature of osteoporosis, fractures can occur in any bone, however, hips, forearms and vertebrae are the most commonly affected areas.[24] Of particular relevance to this HTA report, are osteoporotic vertebral compression fractures (OVCFs). Compression fractures result from an inability of the vertebral bodies to support an applied load/force.[32] The three types of compression fractures differ in their presentation: wedge fractures are characterised by the compression of the anterior segment and preservation of the

posterior segment of the vertebral body; biconcave fractures are characterised by preservation of the anterior and posterior segments but collapse of the middle segment; and in burst fractures the entire vertebrae breaks, often resulting in a bone fragment protruding into the spinal canal (retropulsion).[33] Irrespective of the type, OVCFs may be symptomatic or asymptomatic[31] and generally occur in the thoracolumbar and mid-thoracic regions with comparatively few fractures in the sacral and cervical regions.[34] OVCFs cause substantial morbidity and increase the risk of mortality partly due to compromised pulmonary function.[35-37] OVCFs also increase the risk of developing subsequent fractures causing further pain and disability. It is therefore imperative to identify and appropriately treat these fractures.[36]

OVCFs are difficult to diagnose clinically (without radiographs) because they are generally asymptomatic.[38] If symptoms are present, they are typically non-descript and applicable to many disease processes. Consequently, diagnosis relies heavily upon radiography, where a vertebral compression fracture is indicated if vertebral body height reduces by 20%.[39] Other indications of compression fractures include lack of parallelism between the vertebral endplates, the presence of endplate deformities, and an altered appearance compared to adjacent vertebrae.[40] In addition, MRI is often used to determine age and activity (healing status) of the fracture, while noting that osteoporotic fractures often have a heterogeneous appearance due to reduced bone marrow and increased fat content.[41] Generally, T1-weighted, T2-weighted and Short-TI Inversion Recovery (STIR) image sequences are utilised. New and non-healed fractures appear hypointense on T1-weighted images, hyperintense on STIR sequences, and hyperintense or have a heterogeneous intensity on T2-weighted scans owing to the presence of oedema.[11] A linear black signal on MRI may also indicate non-union (fracture fails to heal).[42] MRI can also aid in discerning between benign and malignant vertebral compression fractures.[43]

The main symptom of OVCF is acute and chronic pain.[33 39] Acute pain arrives quickly, is severe, and related to soft tissue damage and strain. Acute pain is thought to arise from fracture mobility[2] and the strain of back muscles compensating for changes in the spinal colomn.[44] Pain persisting beyond the normal healing time (2 to 3 months) is termed chronic pain and is caused by aberrant repair processes.[45] Without effective treatment, acute and chronic pain impair mobility, reduce quality of life and increase the risk of death.[46-48] These effects are more pronounced in older adults who often already have complex medical conditions and are socially isolated.[32 49]

Kyphosis—abnormal curvature of the spine[50]—is another common symptom of OVCF. Kyphosis generally results from multiple adjacent wedge fractures. As the anterior sides of the vertebrae fuse, the spine curves forward.[33] Kyphotic deformity decreases the space between ribs, and the space between the rib cage and pelvis.[44 51] It is associated with a number of complications, including impaired digestion,

difficulty breathing and loss of mobility, leading to reduced quality of life.[44 51-53] In the context of this HTA, kyphosis is considered a surrogate outcome and therefore, is not the focus of this report.

2.2 Incidence and Prevalence of Osteoporosis Vertebral Compression Fractures

Osteoporosis disproportionately affects older adults. As life expectancy and the aging population continue to grow, the incidence and burden of osteoporosis is anticipated to increase. For example, in 2010 approximately 27.6 million adults aged 50 to 84 years met the World Health Organization (WHO) diagnostic criteria for osteoporosis (bone mineral density T-score $\leq$ 2.5 standard deviations) across the European Union. This is projected to increase to 33.9 million adults by 2025, with females and adults older than age 70 accounting for approximately 79% and 65% of new cases, respectively. Likewise, the annual number of osteoporotic fractures is anticipated to increase from 3.5 million in 2010 to 4.5 million in 2025, with two-thirds of fractures occurring in females. There were 516,266 new OVCFs in 2010, which is anticipated to rise by 24% in 2025. With respect to burden of disease, osteoporotic fractures accounted for 2 million disability adjusted life years (DALY) lost and 1.2 million quality-adjusted life years (QALY) lost, and was implicated in 43,000 deaths in 2010. This is forecast to increase by 2025.[24]

In Switzerland, 3 million adults age 50 or older were at risk of developing osteoporosis in 2010. Of these, 458,547 met the WHO diagnostic criteria for osteoporosis, with the disease disproportionally affecting females (80% of all cases). In 2010, there was an estimated 74,000 osteoporotic fractures, approximately 15% of which (11,000 cases) were OVCFs. The highest incidence of OVCFs was reported in females (67%) and those age 80 and above (51% of all OVCFs).[54] Of patients with OVCFs, 29.1% of males and 22.3% of females were hospitalised. The highest incidence of hospitalisation was among males age 85 and above.[55] There were 208 deaths causally related to OVCFs in 2010.[54] Like in the EU, the number of patients with osteoporosis and OVCF is forecast to increase in Switzerland.[54]

3 Technology

3.1 Technology Description

PVP and PBK are minimally invasive vertebral augmentation procedures used to treat vertebral fractures.[56] Both procedures involve application of cement to the fractured vertebral body in an attempt to stabilise the spine. These procedures are indicated for individuals with painful OVCFs refractory to conservative treatment (CT), painful vertebrae due to osteolysis or tumour invasion, and vertebral fracture due to osteonecrosis.[11] The procedures are particularly useful for older adults who are often poor surgical candidates or unable to receive braces or casts. Further, the procedures lower the risk of developing adverse events associated with prolonged periods of bed rest or certain medications (non-steroidal anti-inflammatory drugs [NSAIDs] and opioids).[56] However, there are risks associated with the PVP and PBK including, cement leakage and adjacent vertebral fracture.[57 58]

3.1.1 PVP

PVP is the injection of cement, most often polymethyl methacrylate (PMMA), into a fractured vertebral body. The aim of the procedure is to relieve pain and strengthen the bone to prevent further fractures.[4 12] PVP patients are given analgesic medication and a local anaesthetic, with or without conscious sedation. The procedure often utilises a bipedicular approach whereby 2 needles are used, 1 either side of the pedicle, to inject cement at the same vertebral level to provide more even distribution. Unipedicular approaches are also used according to the type of fracture. In general, PVP is indicated for relatively simple compression fractures and in Switzerland it is reimbursed without limitations.

3.1.2 PBK

PBK is a variant of PVP involving the insertion of balloon-like devices called tamps into the vertebral body.[12] The balloon tamp is inserted by vertebral paracentesis with a needle cannula under image guidance (fluoroscopy), and the injection device connected. A wider needle cannula (usually 8-gauge) is needed to allow for the balloon tamp to be inserted. There are at least 2 versions of the PBK procedure: (i) the balloon is inflated with bone cement (usually PMMA) until the normal height of the vertebral body is restored;[59] (ii) the balloon is inflated with fluid then removed and cement is injected into the cavity created.[60] PBK aims to reduce pain and restore fractured vertebrae to the normal vertebral height.[12] The procedure is predominately performed under general anaesthetic and requires patients to remain in hospital overnight.[61 62]

Due to limited Switzerland-specific evidence regarding PBK, the FOPH implemented mandatory nationwide reporting of each PBK procedure performed. To support government decision-making, the SwissSpine registry was created in March 2005 to assess real-world safety and effectiveness of PBK.[63]

PBK is currently reimbursed in Switzerland for OVCFs only for patients with fresh thoracolumbar fractures (less than 8 weeks old) associated with pain visual analogue scale (VAS) ≥ 5 and significant deformation, such as thoracic kyphosis >15° or lumbar kyphosis >10°.[15]

3.1.3 Conduct of the Procedures

PVP or PBK are treatment options for patients with severe, ongoing pain after a recent vertebral fracture, where the level of fracture is confirmed by physical examination and imaging, and for whom medical pain management is ineffective.[12] Choice of technique is dependent on the fracture type and location, bone quality, and the patient's activity level.[64] Having a wider spectrum of indication, PVP is used to treat simple compression fractures. Where there is kyphotic deformity, especially in the thoracolumbar junction, PBK may be the preferred option.[65 66] The procedures are predominately performed on an inpatient basis. For example, in 2018, approximately 94% of PVP procedures (n = 2,542) and all PBK procedures were inpatient procedures as inferred by DRG and TARMED data. Further, an interventional radiologist usually performs PVP, while a qualified spinal surgeon generally performs PBK.[15 67] Because these procedures are performed under fluoroscopic guidance a hospital must have high-quality imaging equipment available.[68] Necessary materials include radiopaque bone cement and a delivery system.[69] Patients must recline in a supine position for 1 to 2 hours post-procedure while the cement hardens. A short-term prescription for analgesics may be given for immediate procedure-site pain.[69]

Compared to PVP, PBK additionally includes the insertion of a balloon tamp during PBK (either deflated or left in place),[59 60] has reportedly longer operating times,[57] requires a more expensive delivery system (additional US$3,000 for PBK), and more often requires patients to stay overnight. Consequently, PBK is more costly than PVP (according to USA data).[61] However, PBK can also be conducted as a day surgery procedure under neuroleptic IV sedation and may then be as quick as PVP (clinical reviewer, personal communication).

Cement flow during these procedures cannot be completely controlled. Cement leakage and adjacent vertebral fracture are common complications that can be identified by CT and MRI scans of the treated vertebrae.[57 70] These complications may be symptomatic or asymptomatic, which is an important distinction for assessment of safety.[71] Adjacent fractures may result in further pain and disability, while leakage may lead to complications if cement enters the spinal canal, lungs or veins.[72] New fractures, especially in adjacent vertebrae, are commonly recorded in RCTs.[57 71 73] New OVCFs either remain asymptomatic or require treatment by PVP or PBK.

Needle insertion for both procedures may cause potential adverse reactions including bleeding, systemic infection and damage to neural structures.[12]

3.1.4 Contraindications

Contraindications for PVP and PBK generally overlap and can be delineated into absolute and relative contraindications. Absolute contraindications include asymptomatic vertebral fractures, coagulopathies, localised or systemic infections (e.g. osteomyelitis or discitis), tumour infiltration of the spinal canal, allergy to bone cement or anaesthetic agents, and patient improvement with CT. For PBK, burst fractures are an additional absolute contraindication.[11 56]

Relative contraindications to PVP and PBK include radicular pain, loss of vertebral body height by 70%, spinal canal stenosis, and patients with a high tumour burden.[11 56]

There is current debate about whether PVP or PBK should be performed prophylactically on patients with osteoporosis.

3.1.5 Incidence of PVP and PBK in Switzerland

In Switzerland in 2018, 2,714 PVP and 1,501 PBK procedures were conducted, as inferred by Federal diagnosis-related group (DRG) and TARMED data. Approximately 94% of PVP procedures (n = 2,542) and all PBK procedures were performed on an inpatient basis. It is important to note that DRG and TARMED data describe the number of PVP and PBK procedures performed in Switzerland, they do not specify the indication for surgery, how many procedures were performed per patient, or whether patients utilised both PVP and PBK. Therefore, it is unclear how many of the procedures were specific to patients with OVCF. However, in Switzerland in 2012–2013 only 9% of PBK and PVP procedures were performed for indications other than OVCF.[1] It was subsequently estimated that approximately 2,894 PVP and 1,278 PBK procedures were performed for OVCF.[1] Swiss hospital data from 2015 reported that 2,073 PVP and 1,052 PBK procedures were performed in individuals age 17 and older, although this data was not specific to OVCF.

As mentioned, considerable regional variation exists in age- and sex-standardised procedure rates, ranging from 1.0 to 10.1 per 10,000 persons across hospital service areas.[1] Those in the greater Bern area, Uri and Schwyz, had the highest PVP/PBK age- and sex-standardised procedure rates (6.9–10.1 per 10,000 persons). The lowest PVP/PBK procedure rates (1.0–2.0 per 10,000 persons) were found in Zurich, Jura, Basel, Glarus, Geneva and the Western Valais.[1] Two-thirds of this variation cannot be explained by demographic or socio-economic factors, and is unlikely to be driven by regional variation in patient need or preference. Therefore, most of the observed variation is likely unwarranted and due to physician preference.[1] The most recent estimates from 2017 also reported a large variation in the incidence of PVP (range 0–4.3 per 10,000) and PBK (range 1.0–10.8 per 10,000) among 20 hospital regions in Switzerland.[74 75]

3.2 Alternative Technologies

The alternative treatment for this population is CT—a comprehensive, multifaceted approach which includes analgesics (with or without opiates), bed rest, back braces, physiotherapy and lifestyle changes. CT is recommended as first-line treatment for patients with OVCF.[11 76]

Bed rest is an immobilisation strategy designed to limit fracture movement, promote healing, and prevent secondary fractures. Optimal rest duration has not been defined, with studies utilising both short- (days) and long-term (weeks) periods.[77] Bed rest can result in adverse effects, often more pronounced in older patients. For example, bone density and muscular strength are reduced and the risk of developing bedsores and deep vein thrombosis (DVT) is increased.[78] After bed rest, patients should engage in rehabilitation exercises under the supervision of a physiotherapist. Physiotherapy begins with education on how to avoid pain during daily activities. Exercise directed by a physiotherapist can reduce pain, build strength and prevent future fractures in OVCF patients.[79]

Braces are used to support muscular deconditioning, promote appropriate posture, and provide comfort and neuromuscular re-education for OVCF patients. Bracing after fracture can be an important treatment, and in some cases, braces may provide enough support to facilitate healing. Each brace is individually tailored for comfort and function. As pain declines the brace should be worn less frequently before ceasing altogether.[80]

Medications most commonly used to treat OVCF-related pain include NSAIDs, paracetamol, opioids, lidocaine patches and muscle relaxants.[39 81] Opioids can often relieve OVCF pain, however the side effects can be serious, including constipation, nausea and cognitive impairment.[82] Patients with OVCF pain may respond to NSAIDs if the pain relates to inflammation in the soft tissue. Problematic side effects of NSAIDs are stomach ulcers, nausea, and gastritis.[83] If the patient's ability to perform daily functions improves, medications should be gradually reduced to avoid significant morbidity.[80]

If CT fails to provide significant improvement in pain or function, approaches such as nerve blocks and neuromodulation may be indicated.[60 84] These approaches can also serve as alternatives to PVP/PBK for the management of spinal pain in patients contraindicated for traditional surgical procedures. Nerve blocking involves the use of an anaesthetic agent to disrupt the transmission of pain signals along nerve fibres. Common nerve blocks include epidural block and spinal anaesthesia.[85 86] Neuromodulation utilises electrical spinal stimulation to inhibit pain pathways. A subcutaneously implanted pulse generator creates an electrical field around the spinal column and dorsal pathways, which interrupts the pain pathways.[87 88]

3.3 Concomitant Treatments

Surgical and non-surgical approaches to managing OVCF should be used in combination with osteoporosis medication.[89] Medical treatment for primary osteoporosis includes adequate intake of calcium and vitamin D, followed by pharmacological treatments or hormone replacement therapy.[24] The choice of pharmacological treatment is influenced by several factors, including whether the patient has primary or secondary osteoporosis. In general, pharmacological treatments should be used as concomitant therapy in patients age 70 or older with minimal trauma fracture/s, low bone density, and who are on prolonged, high dose corticosteroid treatment. Common pharmacological treatments for reducing bone loss include raloxifene, strontium ranelate and teriparatide medications. Bisphosphonate medicines may be used for the prevention of osteoporotic fractures, although their use is controversial with reports of prolonged bisphosphonate therapy leading to atypical subtrochanteric fractures and jaw osteonecrosis.[90] Denosumab is recommended for prevention of fractures in post-menopausal women, although its use is also under investigation for safety reasons. Hormone replacement therapy can be given to women at any stage of menopause and aims to preserve and increase bone mineral density.[91]

Physicians should also review any medicines or environmental factors that may contribute to falls in the elderly patient.[89]

3.4 Treatment Pathway

The Centre Hospitalier Universitaire Vaudois (CHUV)[92] and the CIRSE clinical practice guidelines[11] are summarised below (***Figure 1***). The CIRSE and CHUV guidelines do not mention nerve block or neuromodulation. These techniques have been included in the treatment pathway, noting their applicability is uncertain. Clinical practice guidelines from the American Academy of Orthopaedic Surgeons and the American College of Radiology are used to supplement CHUV and CIRSE in domains with limited information (conservative management).[76 93]

Management of OVCF is multidisciplinary involving general practitioners, radiologists, endocrinologists, surgeons, allied health professionals (such as physiotherapists) and the patient.[11] Treatment decisions should reflect the patient's needs and consider age, severity of osteoporosis, presence of comorbidities and suitability for surgical procedures. The team should also conduct a detailed examination to confirm that the vertebral fracture is the likely source of pain and rule out other potential causes. This includes both physical and radiographic examinations (generally MRI, see ***Section 2.1*** for further information).

CT is the first-line treatment for individuals with OVCF.[11] These likely include bed rest, bracing, physiotherapy, analgesics (paracetamol, NSAIDs and opioids) and osteoporosis medication (bisphosphonates, calcium and vitamin D).[11 76 93] Patients contraindicated to certain medications or refractory to CT, as indicated by continuing pain and disability 3 to 4 weeks post-treatment, may be

considered for nerve blocks and neuromodulation.[76] If the patient is unsuitable for, or has failed these techniques, PVP and PBK is indicated, depending on the type of fracture, the degree of deformity and the presentation of symptoms. The CIRSE guidelines note that PVP can also be considered within days of a painful OVCF if the patient is at high risk of developing deep vein thrombosis, pneumonia or bedsores (pressure ulcer).[11] The CHUV guidelines suggest PVP and PBK should be performed within 2 to 6 weeks following the fracture. The delay provides sufficient time to determine whether patients are refractory to CT.[92]

PVP is indicated primarily for relatively simple compression fractures and in Switzerland is reimbursed without limitations. In contrast, PBK is reimbursed in Switzerland only for patients with fresh thoracolumbar fractures (less than 8 weeks duration) associated with pain (VAS ≥ 5) and significant deformation such as thoracic kyphosis >15° or lumbar kyphosis >10°.

An important consideration when selecting PVP or PBK is the duration and activity of the fracture, however, at present there is little consensus regarding what constitutes an acute fracture.[2 5 94-96] Trials evaluating PVP or PBK for OVCF have defined “acute” as fractures present for less than 2, 6 or 9 weeks.[2 5 95 96] As previously mentioned, the reimbursement of PBK in Switzerland is restricted to fractures of 8 weeks or less. However, exceptions can be made for fractures older than 8 weeks with signs of activity. “Active” (non-healed) fractures are generally indicated by the presence of an oedema in the fractured vertebral body.[41]

3.5 Regulatory Status/Provider

In Switzerland, medical device regulations are generally consistent with *European Union* directives, as the country participates in the European single market.[97] Numerous PVP and PBK kits have received regulatory approval (CE mark) and are therefore available for sale within Europe[98 99], for example, *Medtronic*[99 100], *Merit Medical*[101 102], and *Joimax*.[98 103]

In Switzerland, both PVP and PBK to treat OVCF are reimbursed though mandatory health insurance. PVP is reimbursed without limitation, whereas PBK is only reimbursed for patients with OVCF resulting in vertebral deformity that requires correction and for whom conservative management has not offered sufficient pain relief.[15] PVP and PBK can be performed by board-certified interventional radiologists or neurosurgeons.

Other European countries also fund PVP and PBK to treat OVCF. These countries include, but are not limited to, the UK and Ireland,[104 105] which fund the procedures through their respective national health services.[105] In Ireland, PVP and PBK are reimbursed through the funding code for vertebroplasty.[104]

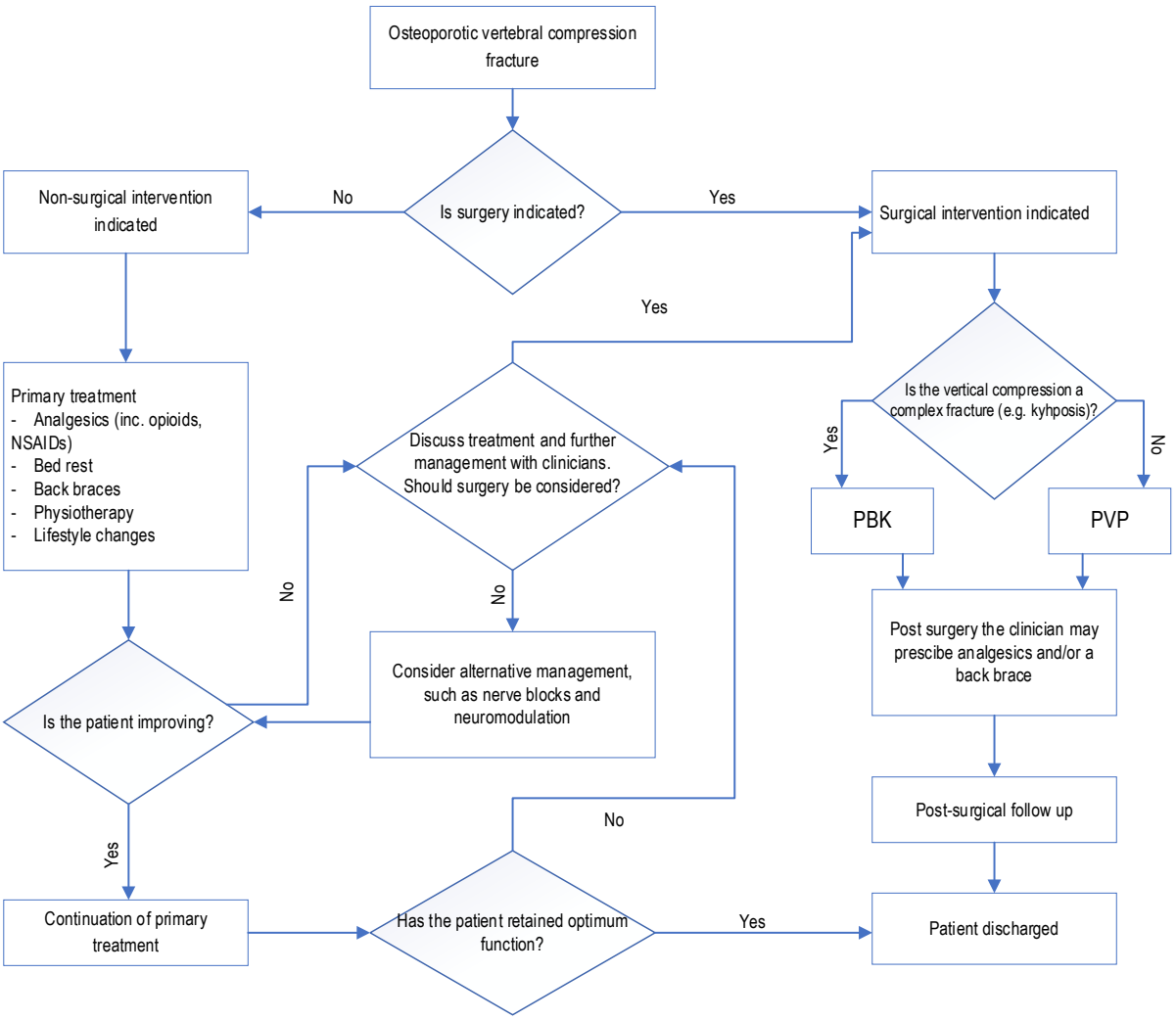


Figure 1 Treatment management pathway for patients with OVCF

Notes

The figure was adapted from the American Academy of Orthopaedic Surgeons, the American College of Radiology, CIRSE and CHUV guidelines.[11 76 92 93]

4 Population, Intervention, Comparator, Outcome (PICO)

4.1 Patients

4.1.1 PVP

The eligible patient population is defined as patients with OVCF who are non-responsive to CT. As PVP is reimbursed without restriction in Switzerland, no limitations will be placed with respect to the severity of pain, duration of fracture or degree of kyphosis.[15] PVP for other types of fracture, such as non-osteoporotic trauma or malignancy, is not the focus of this report and is thus excluded from this HTA.[15]

4.1.2 PBK

The patient population for the assessment of PBK has been defined according to the Verordnung des EDI über Leistungen in der obligatorischen Krankenpflegeversicherung and reflects the reimbursement criteria.[15] PBK is currently reimbursed for patients with thoracolumbar fractures less than 8 weeks old that are unresponsive to analgesics, are painful (VAS ≥ 5), and causing deformation (i.e. thoracic kyphosis >15°, lumbar kyphosis >10°, and/or vertebral body height reduction of more than one-third compared to adjacent bodies).[15 67] Older fractures are eligible for reimbursement in patients with OVCF if the preceding conditions have been met and the fracture is considered active on MRI and still causing pain.[15 67] PBK is not indicated for an osteoporotic patient with normal signs on MRI.[15 67] Only osteoporotic fractures, not those arising from non-osteoporotic trauma or spinal tumours, are relevant to this investigation.[15 67]

4.2 Intervention

The procedures under investigation are PVP and PBK conducted under fluoroscopic guidance. Procedural details are described in ***Section 3***.

Procedural variations that may impact clinical outcomes include the training background of the interventional radiologist or surgeon involved, the cement type (e.g. PMMA or calcium phosphate), and the choice of unipedicular or bipedicular approaches for insertion of cement into the vertebrae. These variables were proposed as relevant sub-groups in the HTA report but will not be evaluated due to the limited literature addressing the type of interventionalist, the widespread use of PMMA, and the lack of delineation between uni- and bipedicular approaches within the same trials.

Concomitant procedures whereby another intervention is conducted along with the vertebral augmentation (i.e. PVP with pedicle screws, PBK with expandable devices) confound the effect of PVP/PBK and are not relevant to the present investigation. Vertebral augmentation will only be

investigated in cases where the fracture has already occurred.[106] Vertebral augmentation given as a prophylactic treatment will not be considered.

4.3 Comparator

Sham controls provide the best evidence for the relative safety and effectiveness of PVP and PBK. Sham procedures simulate PVP and PBK procedures without injection of cement into fractured vertebrae. Patients receive the same anaesthetic, the same needles are inserted near the fractured vertebrae, and the cement is prepared within the operating room so the patient can smell the mixture.

CT is the main unblinded comparator for both PVP and PBK. Patients in sham trials often also receive CT including oral analgesics (with or without opiates), bed rest, back braces, physiotherapy and lifestyle changes. European guidelines recommend that patients undergo CT for at least 3 weeks before undergoing PVP or kyphoplasty procedures.[11]

4.4 Outcomes

4.4.1 Clinical effectiveness

The primary aim of PVP and PBK is to relieve debilitating pain associated with OVCF that limits the individual's quality of life and ability to function. In this context, the critical clinical effectiveness outcomes include pain, physical function, and quality of life. RCTs comparing PVP or PBK to sham or CT will provide the most robust evidence. Lower levels of evidence will not be included for these outcomes where adequate RCT data is available. Outcomes will be delineated into three timepoints: short-term (postoperative up to 1 month), intermediate (3–11 months), and long-term (≥ 12 months). Pain relief associated with PVP and PBK may be instantaneous, therefore no limitations were placed on the minimum follow-up duration for included studies. Durability of the treatment effect will be evaluated in trials with long-term follow-up (i.e. 12–24 months).

For each effectiveness outcome, clinically relevant measures will be considered to provide optimal indications of patient improvement. In addition to the outcomes listed below, fracture deformity outcomes were considered but not included in the PICO criteria (e.g. kyphotic height loss and wedge angle). These outcomes were not included because they are surrogate measures for patient relevant outcomes (e.g. mobility, quality of life). The direct effects of vertebral height loss and kyphosis will be captured through direct, patient-reported outcomes.

Critical

Pain is the primary OVCF symptom impacting quality of life. Pain related to spinal fracture is most often reported using visual analogue scale (VAS) and numerical rating scale (NRS) measured on a per-patient basis and presented as a mean difference across included patients. The clinically relevant differences

in patient pain generally range from 1.5 to 4 on a 10-point scale, with variability likely reflective of the underlying pathology.[23 107-111] The scores typically reflect an individual patient and it is unclear whether they are reflective of group differences. For further pain-related minimum clinically important differences (MCIDs) refer to ***Section 17.5*** (***Appendix E***).

Physical function can be impacted by both pain and kyphosis caused by OVCF. Function can be measured using a variety of scales, including the Roland-Morris Disability Questionnaire (RDQ) and the Oswestry Disability Index (ODI). Clinically relevant differences in RDQ range from 2 to 3 [23 108 109 112 113] and from 4 to 15 for ODI.[108 109 114] For additional information relating to MCIDs refer to ***Section 17.5*** (***Appendix E***).

Measuring physical function with objective personal instruments such as pedometers, smart watches, smart phones and wearable fitness trackers is gaining popularity in clinical studies as a complement to subjective data collection via self-administered questionnaires and VAS. If available, this form of data would be an acceptable measure of physical function in the assessment.

Quality of life has been measured using both generic scales (e.g. Short Form-36 [SF-36], EuroQol 5 dimension questionnaire [EQ-5D]) and disease-specific scales (e.g. quality of life questionnaire of the European Foundation for Osteoporosis [QUALEFFO]) in trials evaluating PVP and PBK. Functional measures of quality of life include discharge home, ability to execute activities of daily living, and independent living or admission to nursing home accommodation. Clinically relevant differences in EQ-5D range from 0.17 to 0.24 [23 111 115] and in SF-36 from 1.2 to 3.0.[114 116] No MCIDs were identified for QUALEFFO. For further quality of life-related MCIDs refer to ***Section 17.5*** (***Appendix E***).

Important

Concomitant analgesia usage, specifically long-term opioid use, is a surrogate outcome used to measure the effectiveness of an intervention at relieving pain.

4.4.2 Safety

While both procedures are low risk, PVP and PBK do carry safety concerns related to cement leakage. All study designs (i.e. RCTs, non-randomised trials [non-RCTs] and single-arm trials) were considered relevant when identifying safety issues related to PVP and PBK. However, only prospectively designed studies were included due to the limitations associated with retrospective collection of safety data. Large databases and registry trials were exempt from this exclusion criterion given the general under-reporting of safety outcomes in prospective trials.

Critical

Serious adverse events (e.g. cement leakage, infection) and **all-cause mortality** are critical safety outcomes associated with the use of PVP and PBK. In this context, a serious adverse event is characterised as an event that is life-threatening, requires hospitalisation, is disabling or permanently damaging, requires intervention or causes death, or any other event deemed serious by the study investigators.[117]

It has been hypothesised that internal fixation therapies such as PVP and PBK may increase the likelihood of **new symptomatic adjacent vertebral fracture** in patients with osteoarthritis.[118] Adjacent vertebral fracture may be symptomatic or asymptomatic (i.e. only appearing on radiographic evidence). This review is primarily concerned with symptomatic adjacent fracture.

Important

Exposure to radiation (patient and physician), **adverse events** and **radiographic evidence of fracture** are important safety outcomes.

4.4.3 Comparative cost-effectiveness

As warranted by the clinical investigation, an economic evaluation comparing the cost-effectiveness of PVP or PBK to CT was performed. To ensure the applicability of the economic evaluation, the evaluation will be conducted using Swiss cost information (e.g. TARMED, DRGs, Spezialitätenliste). Model and parameter uncertainties will be investigated using both probabilistic and deterministic sensitivity analyses. The impact of any significant uncertainties will be interpreted in the Swiss context. A cost-utility analysis (CUA) is the most likely modelling approach, which will evaluate the cost in Swiss francs (CHF) per utility gained (via quality-adjusted life year [QALY]) for PVP or PBK and CT.

4.4.4 Budgetary impact

The budgetary impact of removing PVP and PBK was evaluated. The 5-year projected impact of withdrawing PVP and PBK from the reimbursement list was calculated in terms of the net cost differences. Uncertainties in the estimated budgetary impact were investigated by sensitivity analyses.

4.5 Deviations from the Scoping Report

Deviations from the PICO criteria defined in the scoping report are as follows:

- For the population, the impact of procedural variations including the training background, cement type and approach will not be investigated via sub-group analysis.
- For outcomes, timed up-and-go and radiographic evidence of new fractures are included. Procedure-related mortality was amended to all-cause mortality, noting a narrative description of procedure-related mortality will be provided where evidence is available.
- For the assessment of effectiveness and safety, the type of eligible study was broadened to include database and registry trials.
- The PICO was changed from non-surgical management to conservative treatment owing to the lack of studies evaluating other forms of non-surgical treatment for OVCF.

4.6 PICO-Boxes

Table 1 PICO criteria for PVP

P:	Patients with painful OVCF that does not respond to CT *Exclusions: fractures arising from non-osteoporotic trauma or spinal tumours*
I:	PVP *Exclusions: concomitant treatments including pedicle screw fixation, prophylactic augmentation*
C:	Conservative treatments (optimal medical therapy, physiotherapy, bracing) or sham procedure
O:	**Clinical effectiveness:** • Pain (NRS, VAS) • Physical function (ODI, RDQ) • Quality of life (EQ-5D, SF-36, QUALEFFO) • Analgesia usage • Proportion of patients able to return to independent living compared to proportion requiring assisted accommodation (i.e. nursing homes) **Safety:** • Mortality • Serious adverse events • Any adverse events • New symptomatic and radiographic vertebral fractures • Cement leakage • Patient/physician exposure to radiation
S:	**Clinical effectiveness:** • RCTs • In the absence of randomised trials, other comparative study designs will be considered *Exclusions: narrative reviews, letters to the editor, case reports, single-arm studies* **Safety:** • RCTs • Non-RCT and cohort trials with at least 10 patients in each treatment • Registry/databases and prospective single-arm trials with at least 50 patients *Exclusions: narrative reviews, letters to the editor, case reports*

Abbreviations
CT = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **NRS** = numerical rating scale, **ODI** = Oswestry disability index, **OVCF** = osteoporotic vertebral compression fracture, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RCT** = randomised controlled trial, **RDQ** = Roland-Morris disability questionnaire, **SF-36** = short form 36 questionnaire, **VAS** = visual analogue scale.

Table 2 PICO criteria for PBK

P:	1) Patients with painful OVCF less than 8 weeks old that does not respond to CT, with pain (VAS ≥ 5) and vertebral deformity (thoracic kyphosis >15°, lumbar kyphosis >10°, and/or vertebral body height reduction of more than one-third compared to adjacent body) 2) Patients with fractures older than 8 weeks that meet the preceding criteria and are considered active on MRI (i.e. bone oedema) and still cause pain *Exclusions: fractures arising from non-osteoporotic trauma or spinal tumours*
I:	PBK *Exclusions: concomitant treatments including pedicle screw fixation, prophylactic augmentation, kyphoplasty with other expandable devices including Sky bone expander, stents etc*
C:	Conservative treatments (optimal medical therapy, physiotherapy, bracing) or sham procedure
O:	**Clinical effectiveness:** • Pain (NRS, VAS) • Physical function (ODI, RDQ) • Quality of life (EQ-5D, SF-36, QUALEFFO) • Analgesia usage • Proportion of patients able to return to independent living vs assisted accommodation **Safety:** • Mortality • Serious adverse events • Any adverse events • New symptomatic and radiographic vertebral fractures • Cement leakage • Patient/physician exposure to radiation
S:	**Clinical effectiveness:** • RCTs • In the absence of randomised trials, other comparative study designs will be considered *Exclusions: narrative reviews, letters to the editor, case reports, single-arm studies* **Safety:** • RCTs • Non-RCT and cohort trials with at least 10 patients in each treatment • Registry/databases and prospective single-arm trials with at least 50 patients *Exclusions: narrative reviews, letters to the editor, case reports*

Abbreviations
EQ-5D = EuroQol 5 dimension questionnaire, **NRS** = numerical rating scale, **ODI** = Oswestry disability index, **OVCF** = osteoporotic vertebral compression fracture, **PBK** = percutaneous balloon kyphoplasty, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RCT** = randomised controlled trial, **RDQ** = Roland-Morris disability questionnaire, **SF-36** = short form 36 questionnaire, **VAS** = visual analogue scale.

5 HTA Key Questions

For evaluation of technologies, the following key questions covering the central HTA domains—as designated by the European Network for Health Technology Assessment (EUnetHTA) Core Model® (clinical effectiveness, safety, costs, cost-effectiveness, budget impact, legal, social, ethical and organisational aspects)—are addressed:

1. Are PVP and PBK clinically effective compared to CT or sham procedure?
2. Are PVP and PBK safe compared to CT or sham procedure?
3. What are the costs of PVP and PBK?
4. Is PVP and PBK cost-effective compared to CT or sham procedure?
5. What is the budget impact of PVP and PBK?
6. Are there legal, social or ethical issues related to PVP and PBK?
7. Are there organisational issues related to PVP and PBK?

5.1 Additional Questions

Key sub-questions of relevance to PVP and PBK have been informed by the EUnetHTA Core Model® (Version 3.0)[119] and are outlined in ***Section 17.2*** (***Appendix B***). The sub-questions were used to frame the responses to the key questions for each assessment domain (i.e. effectiveness, safety, cost-effectiveness, ethical, patient/social, legal, organisational).

6 Methodology Literature Search

6.1 Databases and Search Strategy

A systematic literature search was conducted on 8 biomedical databases (PubMed, Embase, the Cochrane Library, CINAHL, York Centre for Reviews and Dissemination, CEA Registry, Econlit and Ethmed) from inception to 4 April 2019. An updated search was performed to identify additional studies published between completion of the scoping report and commencement of the HTA. The updated search was run from 4 April 2019 to 13 December 2019. In addition, ongoing or unpublished clinical trials were searched from the following databases: ClinicalTrals.gov, Cochrane Central Register of Controlled Trials, EU Clinical Trials Registry, WHO International Clinical Trials Registry Platform, Current Controlled Trials MetaRegister, and Australian New Zealand Clinical Trials Registry.

Search terms comprised a combination of keywords and medical subject headings (MeSH) relating to PVP, PBK and OVCF. The full search strategy for each database is reported in ***Appendix A: Source of Literature (databases and websites)***. No search filters were applied. All languages were screened by title and abstract. Selection of studies was limited to English, French, German and Italian language studies. Relevant studies in additional languages were identified to estimate the likelihood of language bias in the search results.

To capture known adverse events associated with CT such as NSAIDs, opioids and paracetamol, an additional non-systematic search was performed to identify contemporary meta-analyses. The search was conducted in PubMed and included keywords such as osteoporosis, elderly and paracetamol, and those relating to specific opioids (e.g. morphine, codeine) and NSAIDs (e.g. celecoxib, diclofenac).

6.2 Other Sources

Grey literature databases were also searched and are listed in ***Appendix A: Source of Literature (databases and websites)***.

6.3 Study Selection

Study selection was conducted by 2 authors. Both authors independently reviewed all records by title and abstract, and then full text. Title and abstract selection were conducted using Rayyan software (Qatar Computing Research Institute).[120] Studies were included if they met the PICO criteria outlined in ***Sections 4.1–4.4*** and ***Section 4.6***. Differences in study selections were settled via consensus at each stage of the selection process.

Non-RCTs, cohort and single-arm trials study selection

To ensure the comparative incidence of mortality, adverse events and new fractures was captured, non-RCTs and cohort trials were eligible for inclusion. Single-arm trials and any single arm from a comparative trial reporting cement leak were also included (as comparative rates are not applicable to this outcome). The systematic searches identified a large number of non-RCTs, cohort and single-arm trials, and a comprehensive assessment of all identified studies was not possible in the given timeframe. To expedite the process and to focus on more applicable and informative trials, additional exclusion criteria were applied when reviewing full text. Studies were excluded based on the following:

- studies outside Europe or North America
- sample size less than 50 for single-arm trials and less than 20 (10 per arm) for non-RCTs
- used other vertebral augmentation methods, novel cement types or procedural variations
- retrospective trials

Database analyses/registry trials

Database analyses/registry trials were deemed eligible for inclusion because they may provide long-term safety information and are therefore important when considering the extended assessment of harms. The Medicare & Medicaid Services Medicare Provider Analysis and Review File was the most widely utilised database. Publications were selected for inclusion if they reported mortality, delineated between the types of adverse events, and included a CT comparator arm.

Studies utilising data from the Swiss spine registry were included in the single-arm trials. No other registries were identified.

Meta-analyses study selection

To address potential long-term harms associated with CT, existing meta-analyses evaluating mortality and adverse events rates of NSAIDs, opioids and paracetamol were included. Meta-analyses were screened against the PICO criteria; however, none were identified in the population of interest. Consequently, the inclusion criteria were expanded to identify studies that resembled the population of interest. The following criteria (listed in order of priority) were used to select relevant meta-analyses:

1. analyses in patients with osteoporotic vertebral fractures, vertebral fractures, or osteoporosis
2. analyses in older adults (age 70 and above) or individuals with spine pathologies or osteoarthritis
3. follow-up duration greater than 12 months
4. studies from European countries

After screening the identified studies, 2 meta-analyses most closely reflecting these criteria were selected.

7 Clinical Effectiveness and Safety

7.1 Summary Statement Clinical Effectiveness and Safety

PVP vs CT

The evidence base comparing PVP to CT was comprised of 8 RCTs of moderate-quality. At 1 month, there were statistically significant and clinically meaningful differences between PVP and CT for pain (VAS) and function-related outcomes (ODI and RDQ), favouring PVP. These differences persisted to 12 months, but did not surpass the lower bounds of identified MCIDs, so were not considered clinically relevant. Furthermore, the results were subject to moderate-to-considerable levels of heterogeneity at most timepoints. There were limited statistical differences for analgesic use and quality of life outcomes (QUALLEFO and EQ-5D) at 1 and 12 months.

To investigate potential causes of heterogeneity, sub-group analysis compared fractures of different ages. At 1 month, fractures younger than 8 weeks (acute) reported greater reductions in pain (VAS) compared to fractures older than 8 weeks. The effects did not persist to 12 months. The remaining outcomes were not compared owing to different methods of analysis.

CT and sham cohorts were pooled for the assessment of safety as both groups effectively received CT. Collectively, 12 RCTs, 2 non-RCTs, 2 database analyses and 15 single-arm studies assessed safety. There were no statistically significant differences in mortality, adverse events and new fractures between PVP and CT. Cement leaks occurred in 55.0% of treated vertebrae and were mostly asymptomatic. The findings of the non-RCTs and single-arm trials echoed results from the RCTs, however, the databases reported significantly lower rates of mortality and adverse events—such as cardiac complications, DVT, infection and pneumonia—following PVP. By contrast, the incidence of pulmonary embolism was significantly higher following PVP. It is important to note that the database analyses reported relative effects and did not report the absolute number of patients affected. ***Table 63*** shows the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) summary of findings table for PVP vs CT.

PVP vs Sham

The evidence base comparing PVP to sham consisted of 4 high-quality RCTs. In general, the effects were inconsistent, were subject to significant heterogeneity, and were not clinically significant if the upper bounds of MCIDs were used. At 1 month, pain (VAS/NRS) statistically differed between PVP and sham arms, but not analgesic use, QUALEFFO or RDQ. By 12 months however, pain and QUALEFFO were both statistically different. The results for EQ-5D were not pooled because the correlation coefficient could not be calculated (both studies had opposite results to one another).

Sub-group analysis was performed on fracture age in an attempt to explain heterogeneity. At 1 month, acute fractures reported slightly greater reductions in pain (VAS) and EQ-5D compared to older fractures (greater than 8 weeks). The effects persisted for EQ-5D but not pain at longer timepoints (6 and 12 months). The remaining outcomes could not be compared owing to different methods of analysis. For the GRADE summary of findings table for PVP vs sham refer to ***Table 64***.

PBK vs CT

The evidence base comparing PBK to CT consisted of a small number of low-to-moderate quality trials (4 RCTs, 4 non-RCTs, 2 databases analyses and 6 single-arm trials). From 1 day to 1 week, there were statistical and clinically significant differences between PBK and CT cohorts with respect to pain. The statistical effect remained at 12 months, however the results were unlikely to translate to clinically important differences. All timepoints were subject to considerable heterogeneity and inconsistency. There were statistically significant differences between PBK and CT cohorts with respect to ODI, RDQ and EQ-5D, but the results were informed by only 1 study and clinical relevance was unclear on the basis of identified MCIDs.

There were no differences between PBK and CT for mortality, adverse events or new fractures across the RCTs and non-RCTs. However, analyses of US Medicare databases suggested PBK reduced relative mortality, hospital readmission and adverse event rates compared to CT (absolute rates not reported). For the GRADE summary of findings table for PBK vs CT refer to ***Table 65***.

7.2 Methods

7.2.1 Appraisal

Two independent researchers conducted the quality appraisal, with differences settled via consensus. RCTs were appraised for risk of bias using the Cochrane risk-of-bias tool for randomised trials (version 2.0[121]), non-RCTs were appraised using the ROBINS-I tool[122], single-arm trials were appraised using the Institute of Health Economics (IHE) case series tool[123] and meta-analyses were appraised using the AMSTAR tool.[124] The overall quality of the evidence per outcome was assessed using the GRADE approach.[125] One researcher appraised the outcomes using GRADE, which was checked by an independent researcher.

7.2.2 Meta-Analysis of Dichotomous Outcomes

For dichotomous outcomes with at least 2 RCTs, a meta-analysis was performed using Review Manager Version 5.3.[126] Dichotomous outcomes were analysed using the Mantel-Haenszel statistical method with random-effects models. The results of the analyses were reported as risk ratios (RR) with 95% confidence intervals (CI). Random-effects models were used to account for variation in fracture severity and other population-based factors, and differences in the conduct of the interventions across the included studies. The interpretation of RRs is in accordance with the Cochrane Handbook (version 6.0).[126] A RR of 1 indicated that the estimated effects were the same for the intervention (PVP or PBK) and comparator (CT or sham). A RR greater than 1 indicated an increased probability of the event occurring in the intervention group relative to the comparator group. A RR less than 1 indicated a reduced probability of the event occurring in the intervention group relative to the comparator group.

For outcomes with less than 2 RCTs, or where it was inappropriate to pool trials, the results were described narratively.

7.2.3 Meta-Analysis of Continuous Outcomes

For continuous outcomes with at least 2 RCTs, a longitudinal meta-analysis was performed. Meta-analysis of longitudinal studies combines effect sizes measured at pre-determined timepoints and accounts for the intrinsic within-study and between-study correlations. This method likely provides a more robust method of analysis than considering each timepoint as an independent event relative to other timepoints. The approach of performing separate univariate meta-analyses at individual timepoints ignores the dependence between longitudinal effect sizes which can result in imprecise parameter estimates.[127] The meta-analysis was conducted in R utilising the metafor package with two-stage analysis, multivariate function for longitudinal data (rma.rv). This exploration accounts for both within- and between-study correlations. The longitudinal meta-analyses took a first order autoregression covariance structure to smooth out missing data due to unreported timepoints in some studies. Within-

study covariance was calculated for each study using a method adapted from Horváth (2009).[128] A point estimate (mean difference) and 95% CI were calculated for each timepoint for each outcome of interest.

For pain outcomes, studies reporting NRS and VAS were pooled, as Bahreini (2015)[129] and Gajasinghe (2010)[130] suggest the scales are generally equivalent and are highly correlative. For studies reporting both measures, the most frequently reported measure was included in the meta-analysis. For safety outcomes, all timepoints were pooled.

7.2.4 Sub-Group Analysis

Outcomes with 2 or more studies underwent further analysis based on the following fracture sub-groups: acute (less than 8 weeks duration) and sub-acute (greater than 8 weeks duration). The sub-groups were analysed using longitudinal meta-analyses as previously described. At least 2 studies per sub-group were required to perform meta-analyses.

For sub-groups with only 1 study, the mean and standard deviation were converted to mean difference and 95% CI for consistency. To determine whether the intervention and comparator group statistically differed in this sub-group, the statistics provided in the respective study were used. For studies reporting the overall effect but not individual timepoints, the mean difference and 95% CI were used to infer significance. Sub-groups analysed using this method are not comparable to sub-groups analysed using longitudinal meta-analysis. For outcomes reporting both methods, the impact of fracture duration cannot be determined.

7.2.5 Heterogeneity

The results of the meta-analysis were presented using forest plots, for a visual representation of variability in the reported effect sizes across studies. Heterogeneity and inconsistency were assessed statistically using the Chi^2 test ($p < 0.10$ representing significant heterogeneity) and the I^2 statistic for the meta-analysis of dichotomous outcomes, and Tau^2 and I^2 for continuous outcomes. The thresholds for low, moderate, substantial and considerable heterogeneity followed those proposed in the Cochrane handbook (I^2 = 0–40% might not be important; 30–60% moderate; 50–90% substantial; 75–100% considerable heterogeneity).[131] The importance of the I^2 result was dependent on the size and direction of the measured effect, and the strength of evidence for heterogeneity (i.e. Chi^2 and Tau^2).

7.2.6 Extended Assessment of Harms

The extended assessment of harms aimed to identify adverse events which may have been missed from the RCTs owing to insufficient power or limited follow-up duration. The assessment encompassed database analyses with long-term follow-up to ascertain the comparative harm of PVP and PBK. In addition, existing meta-analyses (or pooled analyses) were utilised to determine harms of specific CTs (NSAIDs, opioids and paracetamol).

PVP and PBK

Long-term adverse events associated with PVP and PBK were assessed by considering non-RCT studies, single-arm trials and database/registry analyses identified in the systematic search. Studies were selected based on the PICO and inclusion criteria outlined in ***Section 6.3***. However, studies evaluating trauma- and cancer-related fractures were eligible for inclusion provided they did not comprise more than 15% of the assessed fracture population (OVCFs accounted for ≥ 85% of fractures). Results from non-RCT studies and single-arm trials were narratively described or pooled where appropriate. Database analyses were narratively described.

CT

A targeted non-systematic search was performed to identify contemporary meta-analyses of CTs. Studies were selected based on the criteria outlined in ***Section 6.3***. The results were narratively summarised and tabulated to provide a naïve comparison. This approach had several limitations including the applicability of assessed populations and the lack of statistical comparisons, however, performing a network meta-analysis was beyond the scope of this HTA.

7.2.7 Assessment of Publication Bias

Clinical trial registries were searched for unpublished studies as a means of narratively describing the risk of publication bias. There were no outcomes with the minimum number of studies required to perform funnel plot asymmetry analysis.

7.2.8 Missing Values

Missing standard deviations (SDs) were obtained from available standard errors (SE) and CI using the following formula:

$$SD = SE \times \sqrt{N}$$

$$SD = \sqrt{N} * (\text{upper limit} - \text{lower limit}) / 3.92^{*} \quad (^{*}95\%\ CI)$$

For studies only reporting outcomes graphically, Webplot digitizer was used to generate numerical values.

EQ-5D, QUALEFFO and RDQ scores listed in Klazen (2010)[132] were obtained from Buchbinder (2018)[94] because the study did not report measures of variability. It was noted that Buchbinder (2018) obtained the results from the study authors.

7.2.9 Efficacy and Effectiveness

The delineation between efficacy and effectiveness trials was not considered for this HTA.

Statistical interpretation of studies using an active comparator differs from that of placebo trials. A lack of statistically significant difference between treatment groups could indicate that 2 interventions are equally effective, equally ineffective, or that there is no difference between the 2 groups.

7.2.10 Safety

For safety-related outcomes, the number of patients experiencing an event was reported, unless otherwise stated. CT and sham patients were pooled for the assessment of safety because patients undergoing the sham procedure also received CT. Therefore, the risk profile in these populations should be relatively similar.

When defining severe adverse events, the definition within the study was used. (Retrospectively applying the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human use guidelines, for example, is likely inappropriate given the general under-reporting of adverse events and frequent lack of detail.[133]) The lack of standardisation of adverse events may limit the conclusions of the safety sections, however, as the true effect may be under- or over-estimated.

To ascertain the comparative harm of PVP and PBK, the assessment of safety encompassed RCTs, non-RCTs and database analyses. Single-arm trials were also included for the assessment of cement leak, as comparative rates were not applicable.

7.3 PRISMA Flow Diagram

Results of the systematic literature searches are presented in ***Figure 2***. Database searches and pearling of relevant studies yielded a total of 9,409 results. (The results from each database are listed in ***Section 17.1,*** (***Appendix A***).) After removal of duplicates, 6,716 citations were reviewed by title and abstract, and of these, 896 were reviewed by full text. A total of 56 publications evaluating PVP (k = 37) , PBK (k = 15) or both (k = 4) were identified (noting the corresponding arms from trials evaluating both interventions were reported when discussing the number of trials for PVP and PBK).

A total of 12 RCTs, 2 non-RCTs, 2 database-analyses trials and 15 single-arm studies met the inclusion criteria for the assessment of PVP, noting an individual trial may have been reported across multiple publications. (The single-arm studies consisted of 13 single-arm trials and 2 comparative trials with a PVP and PBK arm). Four unique RCTs, 4 non-RCTs, 2 database analyses and 6 single-arm trials met the inclusion criteria for the assessment of PBK, with several trials reported across multiple publications. The database analyses and 2 studies with single arms that evaluated both PVP and PBK are included in these totals. The database analyses included both PVP and PBK arms. A list of all excluded trials is not provided, however notable excluded trials are listed in ***Section 17.10*** (***Appendix J***).

English, French, German and Italian articles were eligible for inclusion in this report. Articles written in other languages were not included in the scoping report but were screened by title and abstract.

PRISMA diagrams were not provided for ethical, legal, social and organisational issues as the searches were conducted in both a systematic and non-systematic manner.

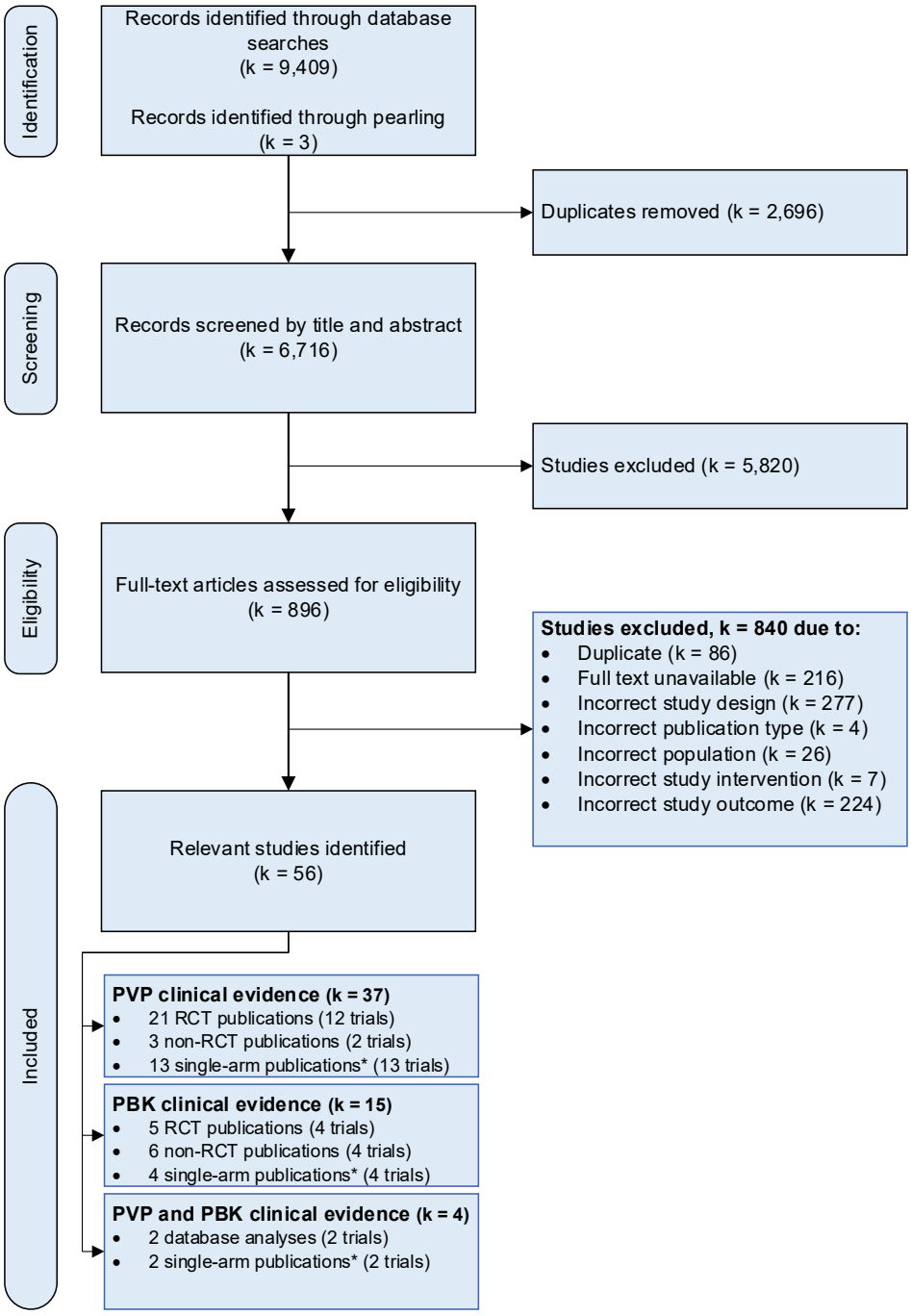


Figure 2 PRISMA flow chart for study inclusion

Notes
Several RCTs and non-RCTs were reported across multiple publications.
When discussing the number of trials for each intervention, only the number of unique trials were discussed, not the total number of publications.
* = single-arm publication includes single-arm trials and comparative trials with a PVP or PBK arm.

7.4 Study Characteristics

In total, 12 unique RCTs evaluating the clinical effectiveness and safety of PVP compared to CT or sham were included. In addition, 2 non-RCTs, 15 single-arm studies and 2 database analyses assessing the safety of PVP were included. The single-arm studies consisted of 13 single-arm trials and 2 comparative trials with a PVP and PBK arm. The evidence evaluating PBK to CT was relatively smaller, comprising 4 RCTs, 4 non-RCTs, 6 single-arm studies and 2 database analyses. Several trials were reported across multiple publications and the trials reporting both PVP and PBK were included in the aforementioned totals. In addition, 6 meta-analyses evaluating specific CTs (NSAIDs, opioids and paracetamol) were included for the extended assessment of harms. The following sections highlighted the characteristics of each intervention separated according to study design.

7.4.1 PVP

RCTs

Overall, 21 studies were included in the assessment of safety (k = 18) and clinical effectiveness (k = 17) (***Table 3***). Of the included studies, 12 were original studies and 9 were extension studies. Given that the extension studies contain all or part of the original trial population, they will not be discussed below to prevent double counting of the evidence base.

The included RCTs consisted of single- (k = 4) and multi-centre (k = 8) trials conducted in Europe (k = 7), Australia (k = 3), United States of America (USA) (k = 2), China (k = 2) and Iran (k = 1). Studies performed in Iran and China were included owing to the limited RCT evidence base evaluating PVP (noting the applicability of trials—see ***Section 7.6***). No study was fully conducted in Switzerland. One international multicentre trial had a centre in Freiburg, Switzerland, but the number of patients treated at this institution was not reported. Five trials were fully or partially conducted in central or western European countries, including Denmark, France, Italy and the Netherlands.

Patients were recruited from primary care centres (general practitioner or specialist clinic) and hospitals. To be eligible, patients required a confirmed osteoporotic fracture as indicated by a reduction in vertebral body height and the presence of oedema on MRI, focal tenderness at the level of the fracture, and pain refractory to medical therapy. Seven studies required minimum VAS or NRS scores ranging from 3 to 7 out of 10. Five trials did not report minimum pain requirements. The minimum duration of pain varied across the included trials, ranging from 6 weeks to 1 year. Patients were excluded if they had cardiopulmonary comorbidities, coagulopathy, systemic or local spine infection, and if the fracture was caused by cancer, trauma or secondary osteoporosis.

The median sample size was 120 patients, ranging from 34 to 400. Patients were mostly older females (age 70 and above) with few comorbidities and in a significant amount of pain, as inferred by baseline

VAS scores. The duration of pain prior to enrolment ranged from a mean of 5.5 days to 7 months. The number of baseline fractures varied from 1 to 3 across the included trials. Leali (2016) studied PVP in post-menopausal women only.

PVP was performed as an inpatient or outpatient procedure by a radiologist, neurosurgeon or orthopaedic surgeon. Patients received local anaesthetic, conscious sedation or general anaesthetic and were placed into the prone position for the procedure. Under fluoroscopic guidance, an 11- or 13-gauge needle was placed using a uni- or bi-lateral approach. Bone cement (PMMA) was injected until it reached the posterior aspect of the vertebral body or leaked into extraosseous structures or veins. Patients generally received analgesia as needed following the procedure.

The comparators were either CT or sham. CT included bed rest, bracing, physiotherapy and analgesia (mostly NSAIDs and opioids). Sham interventions simulated the procedure, although cement was not injected into the damaged vertebrae. Patients received local anaesthetic and verbal and physical cues associated with the procedure. PMMA was often mixed in close proximity to the patient to ensure the smell and sounds of the procedure were copied. All sham trials were double-blind trials, with the patient and outcome assessor unaware of which intervention the patient received. Trials evaluating CT were inherently open-label as the interventions were surgical or medicinal in nature. All patients received osteoporosis medication (bisphosphonates, Vitamin D and calcium supplements) throughout the duration of the trial.

The median follow-up time for safety and clinical effectiveness outcomes was 12 months, ranging from 6 to 36 months. The critical effectiveness outcome of pain was the most frequently reported outcome (k = 11). Fewer studies evaluated the remaining outcomes, for example, quality of life (k = 9) or adverse events (k = 9). For further information, refer to ***Table 3***.

Table 3 PVP: characteristics of included RCTs assessing clinical effectiveness and safety

Author; Year; Country; Trial name	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
PVP vs CT				
Blasco 2012[8] Spain NR	OVCF from T4–L5, clinical onset <12 months, pain measured as VAS ≥4, confirmed by X-ray and presence of oedema on MRI or activity on bone scan n = 125	RCT, open-label Single-centre Recruited from primary care centres, specialists from hospital inpatient, outpatient & emergency departments 12 months	***PVP*** Bilateral, transpedicular, PMMA cement, in C-arm or in a biplane angiography suite ***CT*** Analgesics & rescue therapy	***Clinical Effectiveness*** Analgesic use Pain (VAS) Quality of life (QUALEFFO) ***Safety*** Cement leakage New vertebral fracture Mortality
Chen 2014[134] China NR	OVCF confirmed with MRI, persistent back pain for > 3 months n = 96	RCT, open-label Single-centre Recruited from department of orthopaedics 12 months	***PVP*** Transpedicular, PMMA, fluoroscopic guidance ***CT*** Bracing, analgesia, physiotherapy and anti-osteoporotic medication	***Clinical Effectiveness*** Analgesic use Function (ODI, RDQ) Pain (VAS) ***Safety*** New fractures
Farrokhi 2011[135] Iran NR	OVCF with 10-70% vertebral height loss, severe back pain refractory to analgesics for ≥4 weeks to 1 year, focal tenderness on clinical exam related to fracture level, bone attenuation, bone oedema or vacuum phenomenon on MRI, unresponsive to medical therapy n = 82	RCT, single-blinded Single-centre Recruited from outpatient centres 36 months	***PVP*** Unilateral, PMMA cement, fluoroscopic guidance ***CT*** Optimal medical management i.e. mix of paracetamol, codeine, ibuprofen, calcium, vitamin D, alendronate and calcitonin	***Clinical Effectiveness*** Functional (ODI) Pain (VAS) ***Safety*** Cement leakage New vertebral fracture Mortality

Author; Year; Country; Trial name	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Klazen 2010a[132] Klazen 2010b[136] Venmans 2010[137] Venmans 2011[138] Netherlands VERTOS II	Painful (VAS≥5) thoracolumbar OVCF, minimum 15% vertebral height loss, back pain for 6 weeks or less, bone oedema on MRI, focal tenderness on physical examination, decreased bone density (T scores ≤–1) n = 202	RCT, open-label Multicentre (n = 5) Recruited from radiology departments 12 months	***PVP*** Transpedicular, bilateral, PMMA cement, continuous fluoroscopic monitoring for cement extravasation ***CT*** Pain medication—analgesics in ascending order: paracetamol, tramadol, tramadol and paracetamol, morphine. Osteoporosis medication	***Clinical Effectiveness*** Analgesic usage Function (RDQ) Pain (VAS) Quality of life (QUALEFFO, EQ-5D) ***Safety*** Adverse events Cement leakage New vertebral fracture Mortality
Leali 2016[139] Italy, France, Switzerland NR	Post-menopausal women, 1 thoracolumbar OVCF (primary or secondary osteoporosis), acute pain from severe fracture (not defined), bone oedema present on MRI n = 400	RCT Multicentre (n = 4) 6 months	***PVP*** Transpedicular, PMMA cement, fluoroscopic monitoring, osteoporosis medication and pain medication ***CT*** Pain medication, osteoporosis medication, physiotherapy or bracing	***Clinical Effectiveness*** None ***Safety*** Adverse events Mortality
Rousing 2009[96] Rousing 2010[140] Denmark NR	OVCF with intractable pain less than 8 weeks, MRI confirmed VCF n = 49	RCT, open-label Single-centre 12 months	***PVP*** PMMA cement, fluoroscopic monitoring for cement extravasation ***CT*** Brace treatment, pain medication, general mobilising physiotherapy	***Clinical Effectiveness*** Function (TUG) Pain (VAS) Quality of life (SF-36, EQ-5D) ***Safety*** Adverse events Mortality New vertebral fracture

Author; Year; Country; Trial name	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Voormolen 2007[141] Netherlands VERTOS I	OVCF with minimum 15% height loss on spine X-ray, debilitating back pain relating to the fracture with 6 weeks to 6 months duration refractory to medical therapy, focal tenderness related to level of fracture on exam, bone attenuation T-scores less than -2.0, bone marrow oedema at fracture on spine MRI, patient age ≥50 years n = 34	RCT, open-label Multicentre (n = 3) 12 months	***PVP*** Transpedicular, PMMA cement, under fluoroscopic guidance ***CT*** Optimal pain medication i.e. paracetamol, NSAIDs or opiate derivatives	***Clinical Effectiveness*** Analgesic use Function (RDQ) Pain (VAS) Quality of life (QUALEFFO) ***Safety*** Adverse event
Yang 2016[6] China, USA NR	OVCF from acute mild/minor trauma, back pain (VAS ≥5), low signal on T1-weighted and high signal on T2-weighted MRI, fracture level T5 or lower, living independently without need for wheelchair prior to trauma, decreased BMD (T score ≥-1) n = 107	RCT Multicentre (n = 4) Recruited from emergency room or outpatient clinics 12 months	***PVP*** Transpedicular, PMMA cement, under fluoroscopic guidance ***CT*** Bed rest, bracing, physiotherapy & NSAIDs. Tramadol and morphine if needed	***Clinical Effectiveness*** Pain (VAS) Quality of life (ODI, QUALEFFO) ***Safety*** Adverse events Cement leakage New vertebral fractures

Author; Year; Country; Trial name	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
PVP vs Sham				
Buchbinder 2009[7] Kroon 2014[142] Staples 2015[143] Australia NR	Back pain <12 months, 1-2 recent vertebral fracture (collapse grade 1 or higher), MRI-confirmed acute VCF (oedema or fracture line) n = 78	RCT, double-blinded Multicentre (n = 4) Recruited from general practitioners, specialists at hospital inpatient and emergency departments 24 months	***PVP*** PMMA cement, unipedicular, biplane imaging or image intensifier screen rotated to monitor progress ***Sham*** Sham procedure—subcutaneous lidocaine injection with needle advancement and tapping, mimicking PVP procedures	***Clinical Effectiveness*** Analgesic use Function (RDQ) Pain (NRS/VAS) Quality of life (QUALEFFO, EQ-5D) ***Safety*** Any adverse events Mortality New vertebral fracture
Clark 2016[2] Australia VAPOUR	Osteoporotic patients, 1 or 2 VCF < 6 weeks, pain NRS > 7, MRI confirmed VCF n = 120	RCT, double-blinded Multicentre (n = 4) Recruited from practitioners, specialists at hospital inpatient and emergency departments 6 months	***PVP*** PMMA cement, unipedicular or bipedicular, fluoroscopic guidance ***Sham*** Sham procedure—blunt needle advancement and tapping, mimicking PVP procedure	***Clinical Effectiveness*** Analgesic use Function (RDQ) Pain (NRS, VAS) Quality of life (QUALEFFO, SF-36, EQ-5D) ***Safety*** Any adverse events Cement leakage Mortality New vertebral fracture ***Other*** Length of stay
Firanescu 2011[144] Firanescu 2018[95] Firanescu 2019[3] Netherlands VERTOS IV	1-3 painful (VAS ≥5) thoracolumbar OVCF of up to 6 weeks duration [a], diminished bone density (T score -1 or less), ≥15% loss of vertebral height, bone oedema on MRI n = 180	RCT, double-blinded Multicentre (n = 4) Recruited from outpatient clinics 12 months	***PVP*** Transpedicular, bilateral, PMMA cement, postoperative CT for cement extravasation ***Sham*** Sham vertebroplasty procedure without cement injection	***Clinical Effectiveness*** Analgesic usage Function (RDQ) Pain (VAS) Quality of life (QUALEFFO) ***Safety*** Any adverse events New vertebral fracture Mortality

Author; Year; Country; Trial name	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Kallmes 2009[4] Comstock 2013[145] USA, UK, Australia INVEST	1-3 OVCFs from T4–L5, VCF <12 months, age >50 years, refractory to medical therapy, pain score at least 3/10 n = 131	RCT, double-blinded Multicentre (n = 11) Recruited from outpatient clinics 12 months	***PVP*** PMMA cement, in fluoroscopy suite, under conscious sedation, unilateral ***Sham*** Sham procedure, needle insertion, no cement injection	***Clinical Effectiveness*** Analgesic use Function (SOF-ADL, RDQ) Pain (NRS/VAS) Quality of life (EQ-5D, SF-36) ***Safety*** Adverse events Mortality

Abbreviations

BMD = bone mineral density, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **MRI** = magnetic resonance imaging, **NR** = not reported, **NRS** = numerical rating scale, **NSAIDs** = nonsteroidal anti-inflammatory drugs, **ODI** = Oswestry disability index, **OVCF** = osteoporotic vertebral compression fracture, **PMMA** = polymethyl methacrylate, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RCT** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **SF-36** = short form 36 questionnaire, **SOF-ADL** = study of osteoporotic fractures–activities of daily living, **TUG** = timed up-and-go, **UK** = United Kingdom, **USA** = United States of America, **VAS** = visual analogue scale, **VCF** = vertebral compression fracture.

Notes

a = after 6 months the authors broadened the inclusion to patients with fractures up to 9 weeks old, due to recruitment difficulties.

Non-RCTs and database analyses

Overall, there were 2 database analyses[146 147] and three non-RCTs[148-150] included in the assessment of safety (***Table 4***). Of the non-RCTs, 2 were original studies[148 150] and 1 was an extension study.[149] To prevent double-counting of the evidence base, only the original trials were reported.

The non-RCTs were single-centre studies conducted in Australia and Romania. The database analyses obtained information from United States Centres for Medicare & Medicaid Services Medicare Provider Analysis and Review File database. The databases analyse data sets from 2006[146] and 2005–2014.[147] There were no studies performed in Switzerland, consequently the applicability of the evidence base is discussed in ***Section 7.6***.

Patients were recruited from hospital emergency or inpatient departments in Diamond (2003).[150] Andrei (2017) identified patients from a prospective registry who had undergone PVP or CT at a University Hospital from 2009 to 2012.[148] PVP and CT patients were matched 1:1 based on age, sex, and the level and type of fracture. Both trials enrolled patients with painful OVCF confirmed by imaging. Patients were excluded if the fractures were non-osteoporotic in nature, if there was retropulsion of bony fragments into the spinal canal, if there was coagulopathy, or if there were neurological deficits related to fracture or cognitive impairments that prevented accurately assessing pain or quality of life measures. Andrei (2017) further restricted eligibility based on fracture age (2 months or less). The database analyses

identified relevant patients using primary diagnosis and treatment codes (international classification of diseases codes) corresponding to vertebral compression fracture, PVP, PBK and CT. Patients younger than 65 with renal disease or malignant neoplasms were excluded, in an attempt to restrict analysis to fractures associated with osteoporosis. It was unclear how many patients with non-osteoporotic fractures were included in the analyses.

Sample size of the non-RCTs ranged from 66 to 126, and from 68,752 to 2,077,944 in the database analyses. Patients in the non-RCTs were typically female, with mean age 66 to 76 years, and in a moderate amount of pain as inferred by baseline VAS scores (2/5 and 6/10). Approximately 40% of patients enrolled in Diamond (2003) had secondary osteoporosis due to hyperparathyroidism and 65% were vitamin D deficient.[150] Further demographic information was not provided in Andrei (2017).[148] As with the non-RCTs, the patients included in the databases were predominately older females (age 75 and above) with comorbidities (68% had Charlson comorbidity index score greater than 1).

PVP was performed by a radiologist via a transpedicular approach under local anaesthesia as an inpatient (53%) or outpatient (47%) procedure in Diamond (2003).[150] Andrei (2017) and the database analyses did not provide PVP procedural information. One database reported that the average length of stay was 5.7 ± 4.8 days following PVP, suggesting the procedure was performed on an inpatient basis.

Patients who refused PVP were assigned to the CT arm. CT included analgesia, hot packs, gentle mobilisation and osteoporosis medication. No information was provided regarding CT in the database analyses.

Length of follow-up was 12 months in the non-RCTs and ranged from 30 days to 10 years in the database analyses. Adverse events were the most frequently reported outcomes across the evidence base. For further information refer to ***Table 4***.

Table 4 PVP: characteristics of included non-RCTs assessing safety

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Non-RCTs				
Andrei 2017[148] Romania	Osteoporotic vertebral fracture n = 66	Prospective, blinded Single-centre 12 months	***PVP*** PMMA, fluoroscopic guidance ***CT*** NR	***Safety*** Adverse event
Diamond 2003[150] Diamond 2006[149] Australia	Severe vertebral fracture pain 1–6 weeks, unresponsive to non-opiate analgesia, evidence of osteoporosis n = 126	Prospective Single-centre 24 months	***PVP*** PMMA, fluoroscopic guidance, transpedicular approach ***CT*** Paracetamol, opiates, COX inhibitors, hot packs, gentle mobilisation	***Safety*** Any and severe adverse events Cement leakage Mortality New fractures
Database analyses				
Chen 2013[146] USA	Age > 65 years did not have end-stage renal disease or malignant neoplasm n = 68,752	USA Medicare & Medicaid database NR 30 days–6 months	***PVP*** ***PBK*** ***Nonsurgical management***	***Safety*** Adverse events Mortality Readmissions ***Other*** Length of stay Discharge to home Additional vertebral procedures
Ong 2018[147] USA	Diagnosed VCF, hospital record extending 12 month before VCF, age > 65 years n = 2,077,944	USA Medicare claims database NR 1–10 years	***PVP*** ***PBK*** ***Nonsurgical management***	***Safety*** Adverse events Mortality Readmissions ***Other*** Length of stay Discharge to home

Abbreviations

CT = conservative treatment, **n** = number of patients, **PBK** = percutaneous balloon kyphoplasty, **PMMA** = polymethyl methacrylate **PVP** = percutaneous vertebroplasty, **USA** = United States of America, **VCF** = vertebral compression fracture.

Single-arm trials

Fifteen single-arm studies (13 single-arm trials[151-163] and 2 comparative trials with single-arms[164 165]) evaluating PVP were included (***Table 5***). All studies were prospective and most were conducted in a single centre (11 studies). Duration of follow-up ranged from 6 months to more than 5 years. All studies included more than 50 patients and 11 included more than 100 patients.

All studies required patients to have at least 1 painful vertebral fracture, however the method for assessing this differed between studies. The majority of studies confirmed the fracture with both radiographical evidence (i.e. MRI, CT scan, loss of vertebral height, x-ray and/or kyphosis) and clinical evidence (presence of pain, usually measured using VAS). Bae (2012) listed a minimum VAS threshold to enter the study (≥ 5 out of 10).[152] Some studies also included neurological examinations and other health assessments as part of the eligibility criteria (e.g. SF-36 health survey, EQ-5D or activity of daily living scale).

Fourteen studies focused on osteoporotic fracture only,[152-165] whereas Al-Ali (2009) also included trauma and cancer vertebral fractures, noting osteoporosis accounted for the majority of fractures (n = 357/404, 88%).[151]

All patients were required to have failed CT for a certain amount of time to be eligible for PVP. This requirement varied between studies, from at least 1 month,[152 154] 6 weeks,[158 162 163] or 2 months.[156 159]

In all studies, most patients were female (range 59% to 96%) and the mean age of patients ranged from 68 to 94 years. There was considerable variability in the duration of the fracture before surgery, ranging from a few days to several months. Baseline pain was usually measured with VAS and for most studies mean baseline pain was in the “severe” category (> 7.5 out of 10).[166]

Vertebroplasty was conducted using PMMA of a variety of brands, characteristics (i.e. low or medium viscosity), and additives (i.e. opacifying agents). The unilateral transpedicular approach was the most commonly used approach.

Co-interventions were generally poorly reported, however, when this information was included, these consisted mainly of osteoporotic prophylaxis (calcium, vitamin D and bisphosphonates), anticoagulation medication and pain relief. Parathyroid hormone (PTH) or teriparatide therapy was used for patients with more than three vertebral compression fractures.

Table 5 PVP: characteristics of included single-arm trials assessing cement leak

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention
Al-Ali (2009)[151] USA	Patients with painful vertebral fractures who failed CT n = 357	Prospective case series Single-centre Follow-up: up to 1 year	**PVP** PMMA, fluoroscopic guidance, unipedicular (95%) or bipedicular approach
Bae (2012)[152] USA	Patients with painful vertebral fractures or radiographic evidence who failed CT (4-52 weeks) n = 256	PMMA vs Cortoss cement Multicentre Follow up: 24 months	**PVP** PMMA and Cortoss, fluoroscopic guidance, transpedicular or extrapedicular approach
DePalma (2011)[153] USA	Patients with incapacitating pain due to vertebral fractures who failed CT n = 123	Prospective case series Single-centre Follow-up: 24 months	**PVP** PMMA, fluoroscopic guidance, approach NR
Dohm (2014)[164] USA	Patients with acute painful vertebral fractures with clinical evidence who failed CT n = 404	PVP vs PBK Multicentre Follow up: 24 months	**PVP** PMMA, fluoroscopic guidance, bilateral or unilateral approach also included PBK arm
Fenoglio (2008)[154] Italy	Osteoporotic patients with painful vertebral fractures who failed CT (at least 1 month) n = 52	Prospective case series Single-centre Median follow up 20.4 months (range 6-24 months)	**PVP** PMMA, CT guidance, unipedicular approach
Kotwica (2011) [155] Poland	Patients with single osteoporotic vertebral fractures who failed CT n = 200	Prospective case series Single-centre Follow-up: minimum 12 months, 2 years for 80 patients	**PVP** PMMA, guidance NR, unilateral transpedicular approach
Masala (2012)[157] Italy	Patients with symptomatic osteoporotic vertebral collapse from low-energy trauma who failed CT n = 80	Prospective case series Single-centre Follow-up: up to 1 year	**PVP** PMMA, fluoroscopic guidance, transpedicular (left unilateral) approach

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention
Masala (2009)[156] Italy	Patients with painful vertebral fractures who failed CT (at least 2 months) n = 308	Prospective case series Single-centre Follow-up: up to 3 years	**PVP** PMMA, fluoroscopic guidance, transpedicular or intercostovertebral approach
Nieuwenhuijse (2012)[159] Netherlands	Patients with painful osteoporotic vertebral fractures who failed CT (at least 2 months) n = 115	Prospective case series Single-centre Follow-up: up to 1 year	**PVP** PMMA, fluoroscopic guidance, transpedicular approach, unipedicular in 131 fractures (60.6%) and bipedicular in 85 fractures (39.4%).
Niuewenhuijse (2010)[158] Netherlands	Patients with painful osteoporotic vertebral fractures who failed CT (at least 6 weeks) n = 64 (low viscosity cement: 30, medium viscosity cement: 34)	Low vs medium viscosity PMMA Single-centre Follow-up: 1 year	**PVP** PMMA, guidance NR, uni- or bipedicular approach
Pitton (2008)[160] Germany	Patients with painful osteoporotic vertebral fractures who failed CT n = 191	Prospective case series Single-centre Follow-up: mean 19.7 months	**PVP** PMMA, computed tomography fluoroscopic guidance, transpedicular, intercostotransverse or dorsolateral approach
Santiago (2010)[165] Spain	Patients with non-traumatic or low-energy fractures diagnosed with primary osteoporosis who failed CT n = 60	PVP vs PBK Single-centre Follow-up: up to 1 year	**PVP** PMMA, extrapedicular (9 patients) or bilateral transpedicular (21) approach, also included PBK arm
Saracen (2014)[161] Poland	Patients with multiple osteoporotic vertebral fractures n = 160	Prospective case series Single-centre Follow-up: at least 24 months	PVP PMMA, fluoroscopic guidance, unilateral transpedicular approach
Voormolen (2006a)[163] Netherlands	Patients with painful osteoporotic vertebral fractures who failed CT (at least 6 weeks) n = 77	Prospective case series Single-centre Follow-up: minimum 6 months	PVP PMMA, fluoroscopic guidance, approach NR

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention
Voormolen (2006b)[162] Netherlands	Patients with painful vertebral fractures who failed CT (at least 6 weeks) n = 112	Prospective case series Single-centre Follow-up: mean 10.4 months	PVP PMMA, fluoroscopic guidance, bilateral transpedicular approach, bipedicular injection

Abbreviations

CT = conservative treatment, **n** = number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **PMMA** = polymethyl methacrylate, **PVP** = percutaneous vertebroplasty, **USA** = United States of America.

7.4.2 PBK

RCTs

There were 5 studies comparing PBK to CT[42 59 60 167 168], of which 4 were original studies and 1 was an extension study (***Table 6***). Given that the extension study contained all or part of the original trial population, it was not discussed below. No studies were identified comparing PBK to sham.

Three of the included studies were single-centre trials in China. One European multicentre trial included centres in Austria, Belgium, France, Germany, Italy, Sweden, the Netherlands and the UK. Studies from non-European settings were included owing to the limited RCT evidence base evaluating PBK. The applicability of these populations is addressed in ***Section 7.6***.

The RCTs provided limited inclusion and exclusion criteria. In general, patients were included if they had back pain attributable to an osteoporotic vertebral fracture and were age 60 or older. Wardlaw (2009) noted vertebral fractures required the presence of oedema and a loss of vertebral body height by 15%.[168] This study also included fractures attributable to primary or secondary osteoporosis, multiple myeloma and osteolytic metastatic tumours. The remaining studies limited the inclusion criteria to primary osteoporosis. Exclusion criteria generally encompassed non-osteoporotic fractures, cardiopulmonary comorbidity or coagulopathy, and systematic infection.

The median sample size was 98 patients, ranging from 41 to 300. Patients were generally older adults (age 70 and above), female and in a significant amount of pain (mean baseline VAS 7–9). The majority of patients had 1 fracture located at the thoracolumbar junction. Wardlaw (2009) noted that approximately 70% of patients had used non-pharmacological therapies prior to enrolment, however less than half were using osteoporosis medication.[168]

PBK was performed by a radiologist or surgeon. Patients received general or local anaesthesia and were placed in the prone position for the procedure. Under fluoroscopic guidance, a needle was inserted—using a bilateral or transpedicular approach—into the vertebral body to create a working channel. A balloon was advanced through the working channel and gradually inflated to create a cavity. The balloon was inflated until the kyphosis angle was adequately reduced and then the cavity was filled with PMMA cement.

The comparator, CT, included bed rest, bracing, physiotherapy, analgesic (NSAIDs, opioids and paracetamol) and osteoporosis medication. Trials evaluating CT were inherently open-label owing to the different nature of the interventions.

The median follow-up time for safety and clinical effectiveness outcomes was 12 months, ranging from 6 to 24 months. One study did not report the length of follow-up and consequently, it was only included for the assessment of safety. Pain was the most frequently studied effectiveness outcome (k = 3), with

2 studies measuring function or quality of life outcomes. Three studies assessed safety outcomes, of which adverse events was the most commonly reported (k = 3). For further information refer to ***Table 6***.

Table 6 PBK: characteristics of included RCTs assessing clinical effectiveness and safety

Author; Year; Country; Trial name	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Jin 2018[42] China NR	Single level thoracolumbar OVCF in patients ≥60 years, local pain and injured vertebra on clinical exam, linear black signal on MRI n = 41	RCT, open-label Single-centre 12 months	***PBK*** PMMA cement, transpedicular, unilateral, fluoroscopic guidance, balloon deflated and removed ***CT*** Analgesics and osteoporosis treatment	***Clinical Effectiveness*** Pain (VAS) Quality of life (SF-36) ***Safety*** None
Li 2017[60] China NR	OVCF patient age ≥ 65 years, duration 2 hours to 2 weeks, fracture confirmed with x-ray, computed tomography or MRI scan n = 80	RCT, open-label Single-centre 6 months	***PBK*** PMMA cement under constant fluoroscopic guidance, balloon deflated and removed ***CT*** Physiotherapy and bed rest	***Clinical Effectiveness*** Pain (VAS) Function (ODI) ***Safety*** Any adverse event
Liu 2019[59] China NR	Multiple OVCF confirmed with x-ray and computed tomography scans n = 116	RCT, open-label Single-centre NR [a]	***PBK*** Cement type NR, fluoroscopic guidance, balloon deflated and removed ***CT*** Analgesics, physiotherapy, fixation and bed rest	***Clinical Effectiveness*** Not included ***Safety*** Any adverse event Cement leak
Wardlaw 2009[168] Van Meirhaeghe 2013[167] Austria, Belgium, France, Germany, Italy, Sweden, Netherlands, UK FREE trial	>1 acute T5–L5 VCF, bone marrow signal changes on MRI, decreased vertebral height compared with adjacent vertebrae, pain score at least 4/10 n = 300	RCT, open-label Multicentre (n = 21) 24 months	***PBK*** PMMA cement, fluoroscopic guidance ***CT*** Analgesics, bed rest, bracing, physiotherapy, rehabilitation programs and walking aids, calcium and vitamin D	***Clinical Effectiveness*** Pain (VAS) Function (RDQ) Quality of Life (SF-36, EQ-5D) ***Safety*** Any and severe adverse event Cement leak Mortality New vertebral fracture

Abbreviations

CT = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **MRI** = magnetic resonance imaging, **NR** = not reported, **OVCF** = osteoporotic vertebral compression fracture, **ODI** = Oswestry disability index, **PBK** = percutaneous balloon kyphoplasty, **RDQ** = Roland-Morris disability questionnaire, **SF-36** = short form 36 questionnaire, **UK** = United Kingdom, **VAS** = visual analogue scale.

Notes

a = the length of follow-up was not reported and therefore the study was omitted from the effectiveness analysis.

Non-RCTs and database analyses

Overall, 6 non-RCTs[38 169-174] and 2 database analyses[146 147] were included for the assessment of clinical effectiveness and safety (***Table 7***). The 6 non-RCTs consisted of 4 original trials and 2 extension studies. The extension studies were not discussed further to prevent double counting of the evidence base. The study characteristics of the database analyses were addressed in ***Section 7.4.1*** and were not discussed further here.

The included studies were single- (k = 3) and multi-centre (k = 1) trials conducted at university and public hospitals in Germany, Italy and Slovenia. No studies were performed in Switzerland, consequently the applicability of the evidence base is discussed in ***Section 7.6***. When reported, patients were recruited from hospital inpatient and outpatient clinics. Patients were included if they had an osteoporotic vertebral fracture with pain localised to the fractured vertebrae and PBK was technically feasible. Movrin (2010) further restricted inclusion to those patients who had failed medical therapy, had a kyphotic deformity greater than 30˚, progressive loss of vertebral height and VAS scores greater than 5.[174] Edit-Koch (2011) noted minimum pain requirements (VAS > 5) but did not mention kyphotic deformity or vertebral height loss.[169] One study specified that fractures had to be younger than 6 weeks[174] and 1 study specified fractures must be younger than 3 months.[169] Kasperk (2005) reported that fractures had to be older than 12 months.[173]

The median sample size was 84 patients, ranging from 50 to 124 patients. The patient population was predominately older females (age 70 and above). The number of fractures per patient varied. For example, Movrin (2010) noted 92% of patients had 1 vertebral fracture,[174] whereas 73% of patients in Kasperk (2005) had three or more fractures.[173] Two studies did not report number of fractures per patient. The thoracolumbar region was the most common fracture site across all studies.

Kasperk (2005) noted PBK was performed as an inpatient procedure.[173] The remaining studies did not specify procedure location. Patients received local or general anaesthesia and were placed in the prone position. The procedure was performed by radiologists or surgeons using a bilateral approach in three studies. Cannulae were inserted into the vertebral body with cavities subsequently inflated using balloon tamps. After inflation, the cavity was filled with PMMA cement (Movrin 2010[174] and Giannotti 2012[170]), or PMMA or calcium phosphate cement (Kasperk 2005[173]). Edit-Koch (2011) did not provide specific procedural information, rather a general overview of the procedure.[169] Patients who refused PBK were assigned to CT, which included osteoporosis medication, analgesics and physiotherapy. CT was not defined in 2 studies.[169 170]

The length of follow-up ranged from 12 to 36 months. New fractures and cement leaks were the most commonly reported safety outcome (k = 3). Pain was the most commonly reported effectiveness outcome (k = 2). For further information, refer to ***Table 7***.

Table 7 PBK: characteristics of included non-RCTs assessing clinical effectiveness and safety

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Non-RCTs				
Eidt-Koch 2011[169] Germany	Age > 50 years, painful osteoporotic fracture (T5–L5), fracture < 3 months, VAS > 5 points n = 124	Prospective Multicentre 12 months	***PBK*** PMMA, fluoroscopic guidance ***CT*** NR	***Effectiveness*** EQ-5D RDQ ***Safety*** Mortality
Giannotti 2012[170] Italy	Osteoporotic vertebral fracture n = 50	Prospective Single-centre 12 months	***PBK*** PMMA, biplanar imaging, fluoroscopic guidance bipedicular approach ***CT*** NR	***Safety*** Cement leakage New fractures
Kasperk 2005[173] Grafe 2005[171] Kasperk 2010[172] Germany	Painful osteoporotic fracture > 12 months, chronic back pain > 1 year n = 60	Prospective Single-centre 36 months	***PBK*** PMMA or calcium phosphate cement ***CT*** Analgesic medication, physiotherapy	***Effectiveness*** Pain (VAS) ***Safety*** Adverse events Cement leakage New and adjacent fractures, Mortality
Movrin 2010[174] Slovenia	Painful vertebral fracture < 6 weeks, kyphotic deformity > 30º, VAS > 5 points, able to tolerate general anaesthesia n = 107	Prospective Single-centre 12 months	***PBK*** PMMA, fluoroscopic guidance transpedicular approach ***CT*** Bed rest, analgesic medication	***Effectiveness*** Pain (VAS) ***Safety*** New and adjacent fracture Cement leakage

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention; Comparator	Relevant outcomes
Database analyses				
Chen 2013[146] USA	Diagnosed vertebral fracture, age > 65 years, no end-stage renal disease or malignant neoplasm n = 68,752	USA Medicare & Medicaid database 30 days–3 years	***PVP*** ***PBK*** ***CT***	***Safety*** Mortality, complications, readmissions ***Other*** Length of stay, discharge to home, additional vertebral procedures
Ong 2017[147] USA	Diagnosed vertebral fracture, hospital record extending 12 month before VCF, age > 65 years n = 2,077,944	USA Medicare claims database 1–10 years	***PVP*** ***PBK*** ***CT***	***Safety*** Mortality, readmissions, complications ***Other*** Length of stay, discharge to home

Abbreviations
CT = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **PBK** = percutaneous balloon kyphoplasty, **PMMA** = polymethyl methacrylate, **PVP** = percutaneous vertebroplasty, **RCT** = randomised controlled trial, **RDQ** = Roland-Morris disability questionnaire, **USA** = United States of America, **VAS** = visual analogue scale, **VCF** = vertebral compression fracture.

Single-arm trials

Four single-arm trials[63 175-177] and 2 comparative trials with single arms evaluating PBK were included (***Table 8***).[164 165] The studies were mainly prospective, single- or multi-centre trials with long-term follow-up ranging from 6 months to 2 years. All the studies included more than 50 patients; 3 studies included more than 400 patients.

All trials required patients to have at least 1 painful vertebral fracture. The method for assessing this differed between studies. The majority of studies confirmed the fracture with both radiographic evidence (i.e. MRI, CT scan, loss of vertebral height, X-ray and/or kyphosis) and clinical evidence (presence of pain, usually measured using VAS). Two studies specified a minimum VAS threshold to enter the study ranging from ≥ 5 out of 10[177] to ≥ 6 out of 10.[176] Some studies also included neurological examination and other health assessments as part of the eligibility criteria (e.g. SF-36 health survey, EQ-5D or activity of daily living scale).

Four studies focused on osteoporotic fracture only,[164 165 175 177] whereas Hubschle (2014) also included trauma and cancer vertebral fractures, noting osteoporosis accounted for the majority of fractures (84%, n = 522/625).[63] Prokop (2012) did not include any criteria, only noting that their study was in osteoporotic patients.[176]

In all studies, most patients were female (range 70–80%) and the mean age of patients ranged from 69 to 76 years. There was considerable variability in the duration of the fracture before surgery, ranging from a few days to several months. Baseline pain was usually measured via VAS, and for most studies mean baseline pain was in the "severe" category (> 7.5 out of 10).[166]

Kyphoplasty was mostly conducted using PMMA and a bilateral transpedicular approach was the most common. Co-interventions were generally poorly reported, however, when this information was available, these consisted mainly of osteoporotic medication (calcium, vitamin D and bisphosphonates).

Table 8 PBK: characteristics of included single-arm trials assessing cement leak

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention
Dohm 2014[164] United States	Patients with acute painful vertebral fractures [1-3] with clinical evidence who failed CT n = 404	PVP vs PBK Multicentre 24 months	***PVP and PBK*** PMMA, fluoroscopic guidance, approach: bilateral or unilateral
Hillmeier 2004[175] Germany	Patients with painful osteoporotic vertebral fractures n = 102	Prospective comparative study (PMMA vs calcium phosphate cement) Multicentre 6–12 months	***PBK*** PMMA (138 fractures), calcium phosphate cement (54 cases), guidance NR, dorsal and transpedicular bilateral approach
Hubschle 2014[63] Switzerland	Patients with osteoporosis, trauma and cancer diagnoses n = 625	Retrospective case series Multicentre 12 months	***PKB*** PMMA, guidance NR, approach NR
Prokop 2012[176] Germany	NR n = 564	Case series Single-centre NR	***PBK*** PMMA, fluoroscopic guidance, approach NR
Robinson 2008[177] United States	Patients with painful vertebral fractures who failed CT (12 weeks) n = 102	Prospective case series Single-centre 6 months	***PBK*** PMMA, guidance NR, approach NR Co-intervention: all patients received thrombosis prophylaxis (low-molecular weight heparins)

Author; Year; Country	Inclusion criteria; Sample size	Design; Setting; Follow-up	Intervention
Santiago 2010[165] Spain	Patients with diagnosed non-traumatic or low-energy fractures with primary osteoporosis who failed CT n = 60	PVP vs PBK Single-centre 1 year	***PVP and PBK*** PBK: PMMA bilateral transpedicular approach, fluoroscopic guidance for vertebrae with loss of vertebral height but no evidence of oedema

Abbreviations
CT = conservative treatment, **n** = number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **PMMA** = polymethyl methacrylate, **PVP** = percutaneous vertebroplasty.

7.4.3 CT (Extended Assessment of Harms)

Six studies were included in the extended assessment of harms,[82 83 178-181] of which 4 were meta-analyses, 1 was an umbrella review and 1 was a pooled analysis of retrospective trials. The meta-analyses implemented systematic search strategies to identify relevant trials, searching at least 2 of the following databases: CENTRAL, CINAHL, EMBASE, MEDLINE, AMED or PsychcINFO. One study identified trials via the Pfizer Corporate Clinical trials Registry.[178] Four meta-analyses included RCTs. One meta-analysis included existing reviews,[82] and another included observational studies as the authors noted RCTs were unlikely to capture adverse events due to the short follow-up duration.[180]

Two meta-analyses compared NSAIDs to other NSAIDs (traditional or cox-2 inhibitors [coxib]) or placebo.[83 178] The coxib and traditional NSAID trialists' collaboration analysis (2013) included 280 trials comparing NSAIDs to placebo (124,513 patients) and 474 trials comparing NSAIDs to another NSAIDs (229,296 patients).[83] The patients were approximately 61 years old, female (66%), Caucasian (79%) with osteoarthritis (63%). Length of follow-up was not reported. The meta-analysis by Mallen (2011) included 21 trials (9,461 patients) with a mean age of 72 years.[178] The most common NSAIDs assessed were celecoxib, naproxen, ibuprofen and diclofenac. Length of follow-up ranged from 6 to 52 weeks.

One meta-analysis and 1 umbrella review evaluated the use of opioids for chronic non-cancer pain.[82 179] The umbrella review by Els (2017) included 16 reviews of 14 different opioid medicines such as codeine, morphine and oxycodone.[82] The analysis compared the effect of opioids with placebo and an active comparator. The patient population variously included individuals with phantom limb pain, osteoarthritis, neuropathic pain, chronic non-cancer pain and chronic lower back pain. Further demographic information was not provided. The length of follow-up ranged from 2 weeks to 13 months. The meta-analysis by Busse (2018) included 96 RCTs (26,169 patients) and compared opioids with placebo or active comparator.[179] Mean patient age was 58 years and approximately 61% were female, with chronic, neuropathic or nociceptive pain or central sensitisation. Length of follow-up was not reported.

Two meta-analyses investigated the effects of paracetamol versus non-use or placebo.[180 181] The meta-analysis by Machado (2015) included 12 RCTs (5,366 patients) comparing paracetamol with placebo for spinal pain and osteoarthritis of the hip and knee.[181] Follow-up extended from less than 6 weeks to 20 years. The meta-analysis by Roberts (2015) included 8 trials (593,027 patients) and assessed the adverse event profile of oral paracetamol in adults >18 years compared with non-use.[180] The trials included both healthy and sick patients, with length of follow-up ranging from less than 6 weeks to more than 7 years.

Gastrointestinal adverse events were the most commonly reported outcome (k = 4) followed by adverse events (k = 3) and mortality (k = 3). Few trials reported the incidence of specific adverse events separately, most pooled events into broader categories (e.g. vascular or gastrointestinal adverse events).

Table 9 Extended assessment of harms: characteristics of the included studies assessing safety

Author; Year	Inclusion criteria; Sample size	Design; Follow-up	Intervention; Comparator	Relevant outcomes
Coxib and traditional NSAIDs trialist collaboration 2013[83]	Used NSAIDs for > 4 weeks, RCTs k = 574	Meta-analysis NR	NSAIDs (celecoxib, diclofenac, ibuprofen, naproxen and others) Placebo	***Safety*** Gastrointestinal adverse event Heart failure Mortality Vascular event
Mallen 2011[178]	Age > 65, osteoarthritis, rheumatoid arthritis, ankylosing spondylitis, using NSAIDs > 2 weeks, RCTs k = 21	Meta-analysis 6–52 weeks	NSAIDs (celecoxib, diclofenac, ibuprofen and naproxen) Placebo	***Safety*** Any adverse event Gastrointestinal adverse event
Els 2017[82]	Chronic non-cancer pain, age > 18 years, using opioids > 2 weeks, SRs k = 16	Umbrella review 2 weeks–13 months	Opioids [a] Placebo Active (non-opioid comparator) [b]	***Safety*** Any and serious adverse events Mortality Withdrawal due to adverse event

Author; Year	Inclusion criteria; Sample size	Design; Follow-up	Intervention; Comparator	Relevant outcomes
Busse 2018[179]	Chronic non-cancer pain > 3 months, RCTs k = 96	Meta-analysis NR	Opioids [c] Placebo NSAIDs	***Safety*** Constipation Dizziness Drowsiness Dry mouth Headache Nausea Pruritis
Roberts 2015[180]	Age > 18 years, consuming paracetamol (0.5–1g per 4–6 hours), observational trials k = 8	Meta-analysis and narrative summary 2–20 years	Paracetamol Non-use	Cardiovascular adverse event Gastrointestinal adverse event, Mortality
Machado 2015[181]	Non-specific spinal pain or osteoarthritis, RCTs k = 12	Meta-analysis < 6 weeks–7.2 years	Paracetamol Placebo	Abnormal liver function, Any and serious adverse events, Withdrawal due to adverse event

Abbreviations

g = grams, **k** = number of studies, **NR** = not reported, **NSAIDs** = non-steroidal anti-inflammatory drugs, **RCTs** = randomised controlled trials, **SR** = systematic reviews.

Notes

a = buprenorphine (transdermal), codeine, dextropropoxyphene, dihydrocodeine, fentanyl, hydromorphone, levorphanol, methadone, oxycodone, oxymorphone, tapentadol, tolidine, tramadol.

b = NSAIDs, tricyclic antidepressants, anticonvulsants, synthetic cannabinoids and usual care.

c = codeine, dihydrocodeine, hydrocodone, hydromorphone, meperidine, oxycodone, oxymorphone, tapentadol, tramadol.

7.5 Risk of Bias

7.5.1 PVP

RCTs

The risk of bias graphs for clinical effectiveness and safety outcomes are reported in ***Figure 3–Figure 5***. The risk of bias summaries (per study) are presented in ***Figure 6–Figure 8***. Risk of bias was assessed on a per outcome basis (clinical effectiveness and safety). Safety outcomes were further delineated to radiographic and adverse events/mortality (safety) outcomes owing to the potential impact of unblinding.

For clinical effectiveness and safety outcomes, most studies used appropriate randomisation and concealment strategies, typically consisting of block-based randomisation strategies and opaque sealed envelopes, respectively. This conclusion was reinforced by a lack of baseline differences between treatment groups. The main concern in trials comparing PVP to CT was the absence of blinding. The patient and outcome assessor were both aware of which treatment was received. This was important because knowledge of the intervention can potentially influence the reporting of subjective outcomes such as pain and quality of life measures. The lack of blinding was not a substantive issue for more objective outcomes such as new fractures, cement leak or mortality. Concerns around blinding were addressed in the sham comparison in which patient and outcome assessor were both unaware of which intervention the individual received. However, radiologists or neurosurgeons performing the procedure were inherently unblinded and it was often unclear whether they were involved with recording subjective outcomes such as pain or quality of life in sham trials.

Eight trials utilised intent-to-treat analysis.[2 4 8 95 132 135 139 141] Four trials did not report whether intent-to-treat or per-protocol analysis was used.[6 96 134 139] There were 4 predefined cross-over trials with studies typically separating clinical effectiveness results based on their original and crossed-over treatment group.[4 132 134 135] However, for safety-related outcomes, it was unclear whether the results included crossed-over patients. A further three studies noted patients in the CT (or sham) group underwent PVP during the trial period.[6 8 141] These studies did not report how patient data was subsequently analysed.[6 8 132]

There were significant baseline differences in EQ-5D in Rousing (2009) and Klazen (2010).[96 132] Klazen (2010) attempted to correct for baseline differences via regression analysis, whereas Rousing (2009) did not. Baseline imbalances were a cause of bias when estimating the effect estimate and may have led to over- or under-estimation of the true effect.

For clinical effectiveness outcomes, the reporting and analysis of outcomes was generally appropriate, with limited evidence to suggest publication bias. For safety-related outcomes, adverse events were frequently not defined and often not listed on the trials protocol.

Six trials declared no conflict of interest or receipt of grants from governmental bodies[4 6 8 96 134 135]. Four studies declared industry support.[2 7 95 132] All trials stated the sponsor had no role in the design, collection of data or preparation of manuscripts. Two trials did not report conflicts of interest.[139 141]

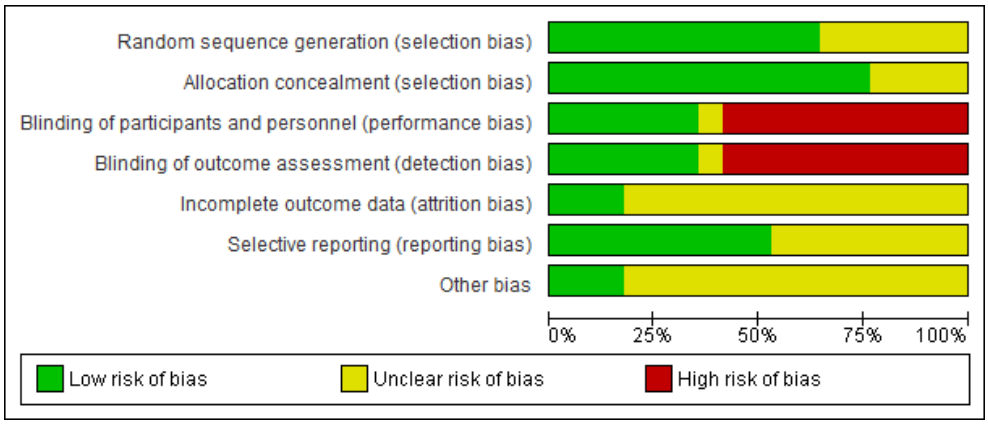


Figure 3 PVP: risk of bias graph for the RCTs assessing clinical effectiveness outcomes (17 studies)

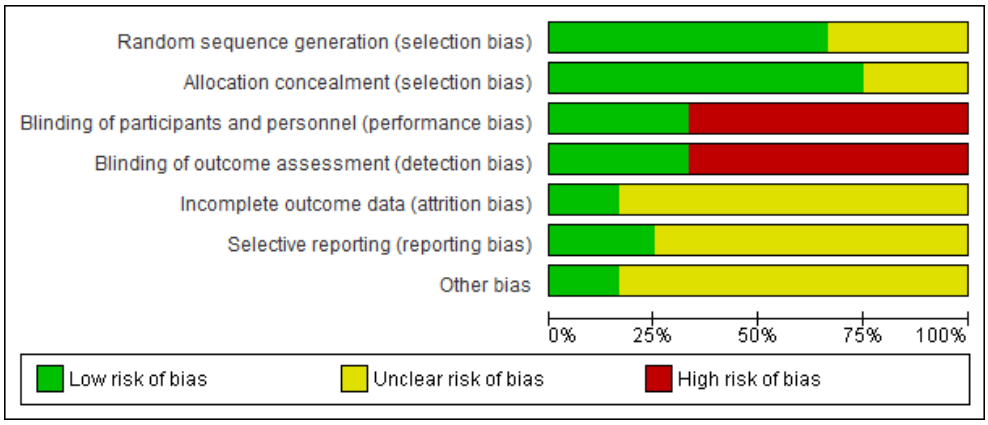


Figure 4 PVP: risk of bias graph for RCTs assessing safety outcomes (12 studies)

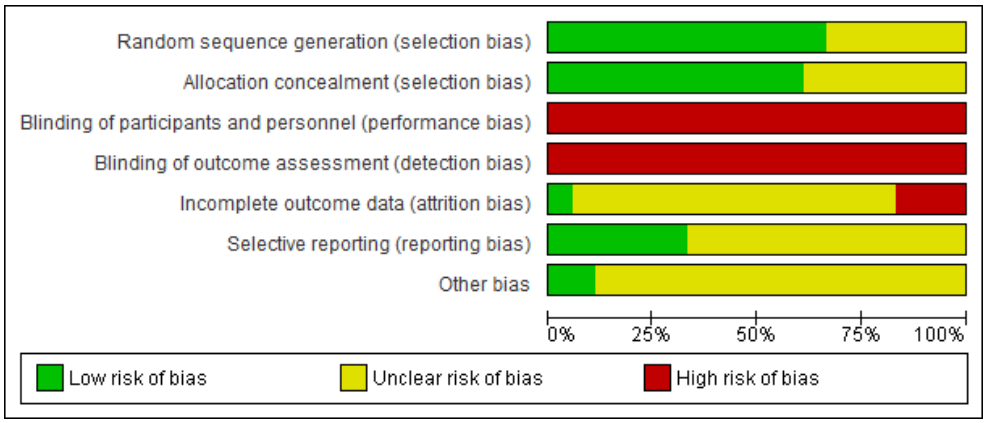


Figure 5 PVP: risk of bias graph for RCTs assessing radiographic outcomes (18 studies)

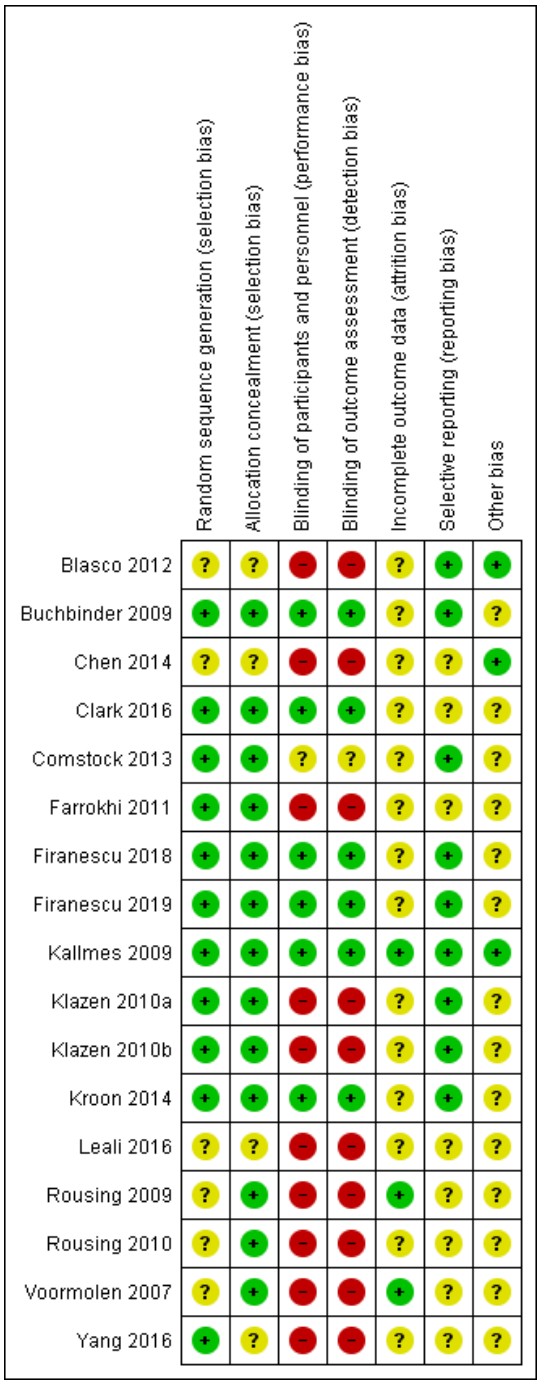


Figure 6 PVP: risk of bias summary for clinical effectiveness outcomes in the RCTs

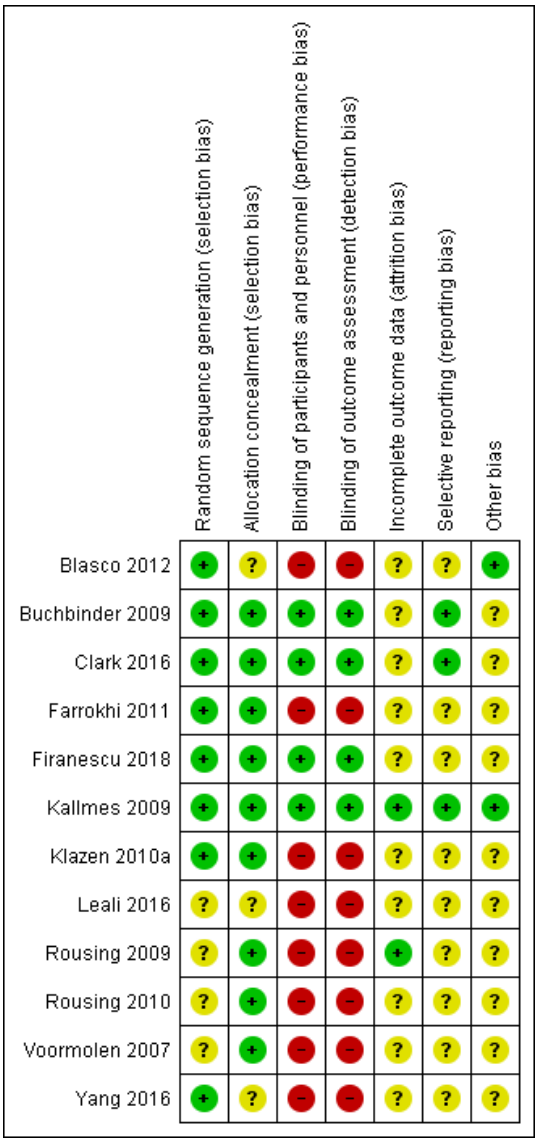


Figure 7 PVP: risk of bias summary for safety outcomes in the RCTs

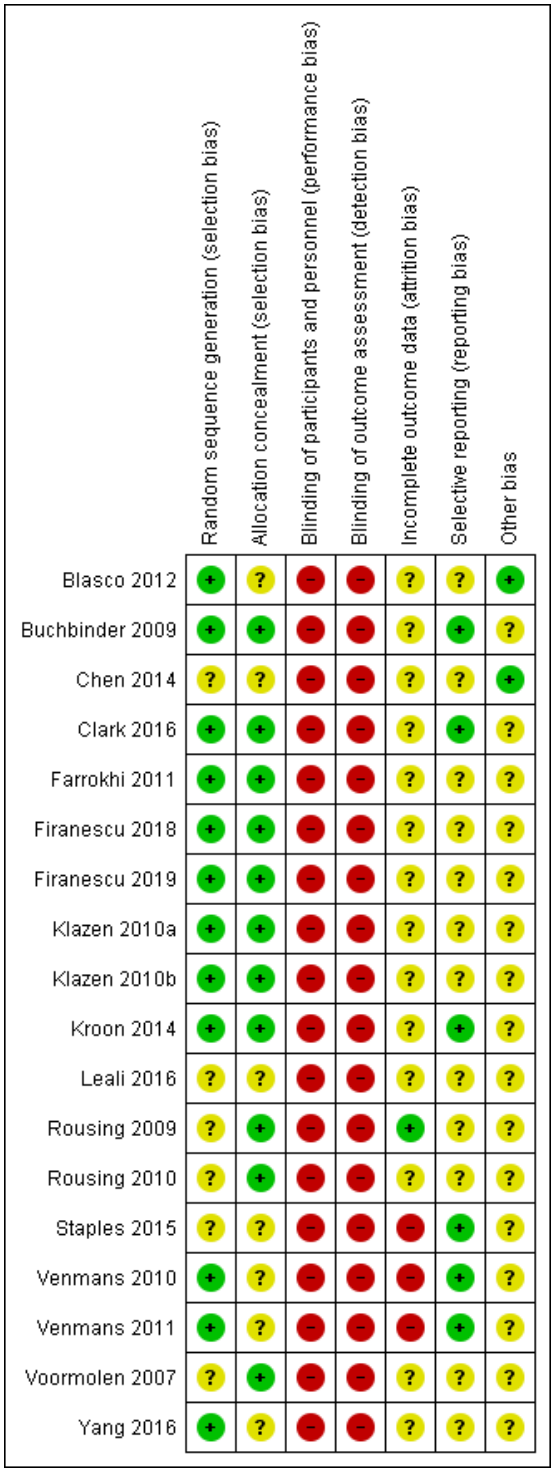


Figure 8 PVP: risk of bias summary for radiographic outcomes in the RCTs

Non-RCTs and database analyses

The risk of bias graphs are presented in ***Figure 9*** and ***Figure 10***. A summary of risk of bias (per study) is presented in ***Table 10***. Safety was further delineated into radiographic and adverse events/mortality. Many of the bias concerns were applicable to both outcomes and thus were discussed together unless otherwise stated.

The overall risk of bias was moderate-to-serious for the non-RCTs and serious for the database analyses. Aside from the under-reporting of study methods, the non-RCTs were generally well performed. For example, the potential for confounding was low, as baseline demographics and co-interventions were balanced across the 2 groups. The selection of participants may have been biased as allocation to the control group was based on refusal to undergo PVP rather than demographic factors. However, reasons for refusal of PVP were not reported and consequently, the effect on selection cannot be fully determined. Intervention and comparator information was generally lacking in both non-RCTs, although this was unlikely to significantly impact the results. The main concerns related to losses to follow-up. Data was available for 77% of participants in Diamond (2003) and 91% of participants in Andrei (2017).[148 150] Owing to the under-reporting of safety outcomes and the relatively small sample sizes, losses to follow-up disproportionally influenced the event rate. Lastly, trial assessors and patients were aware of the intervention they received. This was unlikely to result in bias for radiographic outcomes.

The main risk of bias concern in the database analyses related to patient selection (bias due to confounding). Patients were identified using ICD-9-CM codes, with codes specific to the diagnosis (vertebral fracture) and intervention (PVP or PBK). However, the codes did not provide information regarding how the fracture arose. In an attempt to limit the results to those patients with osteoporotic vertebral fractures, the studies excluded younger adults (< 65 years) and those with neoplasms. However, patients with non-osteoporotic fractures may be part of the cohort, which may have influenced the results if those patients were comparatively healthier or sicker. Furthermore, the CT cohort was poorly defined, thus it was unclear what interventions the participants received. It was unclear if any deviations were in line with usual practice. For the remaining risk of bias domains, there were no substantial bias issues.

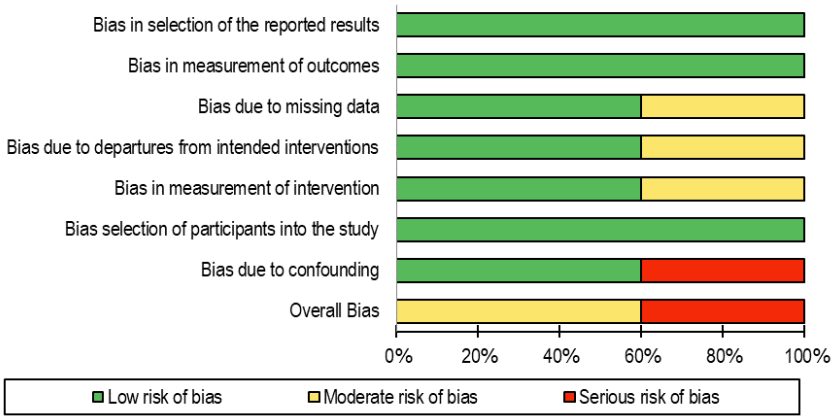


Figure 9 PVP: risk of bias graph for non-RCTs and database analyses assessing safety outcomes (3 studies)

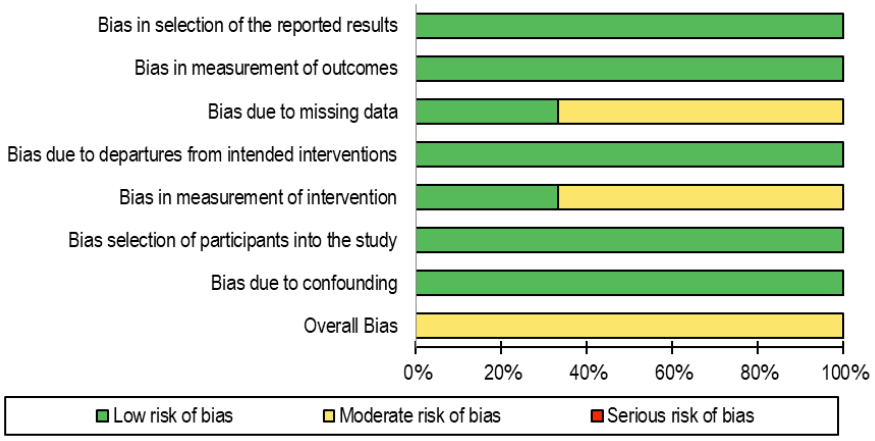


Figure 10 PVP: risk of bias graph for non-RCTs assessing radiographic outcomes (4 studies)

Table 10 PVP: risk of bias summary for safety outcomes in the non-RCTs

Author; Year	Bias due to confounding	Bias selection of participants into the study	Bias in measurement of intervention	Bias due to departures from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported results	Overall Bias
Non-RCTs (safety and radiographic)								
Andrei 2017[148]	Low	Low	Moderate	Low	Low	Low	Low	Moderate
Diamond 2003[150]	Low	Low	Moderate	Low	Moderate	Low	Low	Moderate
Diamond 2006[149]	Low	Low	Low	Low	Moderate	Low	Low	Moderate
Database analyses (safety)								
Chen 2013[146]	Serious	Low	Low	Moderate	Low	Low	Low	Serious
Ong 2018[147]	Serious	Low	Low	Moderate	Low	Low	Low	Serious

Abbreviations
Non-RCTs = non-randomised controlled trials.

Single-arm trials

Quality of the single-arm studies investigating PVP or PBK was appraised using the Institute of Health Economics (IHE) Quality Appraisal Checklist. A summary of the risk of bias is presented in ***Table 110, Appendix C*** and the corresponding risk of bias graph is presented in ***Figure 11***. Fifteen studies on PVP were included in the safety analysis. Generally, the studies were considered to be of moderate quality. All studies clearly stated their objective, used appropriate methods to measure outcomes before and after the intervention, followed patients for sufficient duration, and adequately reported adverse events. The intervention was well described by most studies and losses to follow-up were documented. Conclusions were supported by the results in all studies.

Most studies were limited by inadequate descriptions of co-interventions, and lack of blinding of outcome assessors. Patient eligibility criteria was adequately reported, however, most studies failed to report if patients entered the study at a similar point in their disease. Twelve studies explicitly stated that consecutive patients were enrolled.

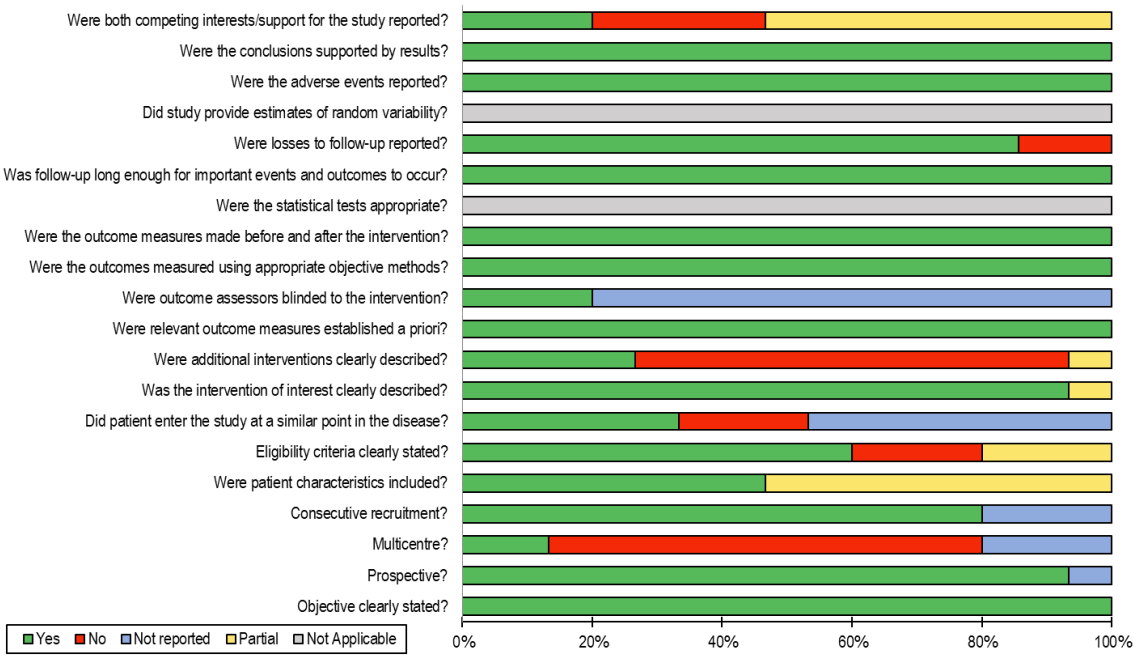


Figure 11 PVP: risk of bias graph for single-arm trials assessing safety outcomes (15 studies)

7.5.2 PBK

RCTs

The risk of bias graphs are presented in ***Figure 12–Figure 14.*** The risk of bias summaries for clinical effectiveness and safety outcomes are reported in ***Figure 15–Figure 17***. Assessment of bias was hampered by under-reporting of study methodology, which limited the ability to accurately evaluate each bias domain, an effect particularly apparent in Liu (2019).[59]

The randomisation process was adequately described in 4 studies and generally included the use of random number tables or permuted block randomisation. However, it was unclear whether treatment allocation was concealed appropriately owing to a lack of reporting across the included studies. The lack of concealment was unlikely to substantially impact the trials given that patients and radiologists were unblinded. The lack of blinding was unlikely to substantially influence objective outcomes such as adverse events and new fractures. However, knowledge of the intervention likely influenced subjective outcomes such as pain and quality of life. This was the main concern amongst PBK trials.

All studies reported substantial losses to follow-up. Owing to the limited reporting, it was unclear whether patients lost to follow-up were included in the results. Losses to follow-up were particularly important for safety-related outcomes given that most studies were already under-powered to detect differences.

Wardlaw (2009) noted that not all vertebrae were able to be read by radiologists.[168] Consequently, the incidence of new fractures was analysed in patients with images of at least 7 vertebrae (T5 to L5) at baseline and 12 months, corresponding to 81% of PBK patients and 68% of CT patients. This may have enriched or diminished the actual fracture rate. Other concerns related to the lack of published protocols, which limits the ability to accurately assess publication bias. Two trials reported the sponsor had a role in study design, data monitoring, reporting or results and paid for the statistical analysis.[167 168] Three trials declared no conflicts of interest.[42 59 60]

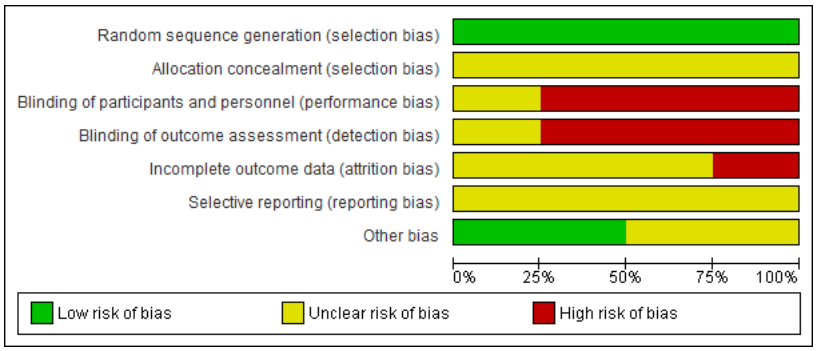


Figure 12 PBK: risk of bias graph for RCTs assessing clinical effectiveness outcomes (4 studies)

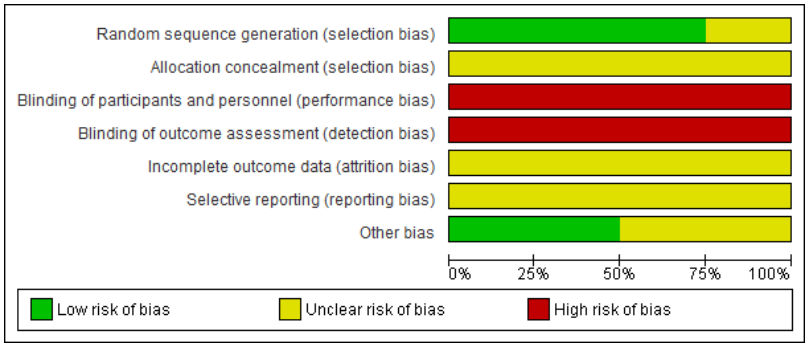


Figure 13 PBK: risk of bias graph for RCTs assessing safety outcomes (4 studies)

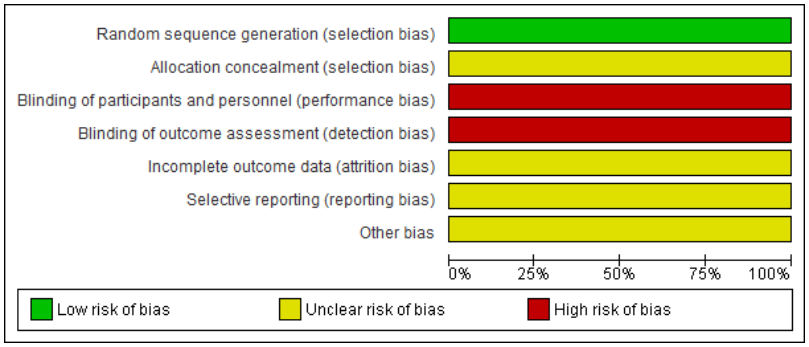


Figure 14 PBK: risk of bias graph for RCTs assessing radiographic outcomes (2 studies)

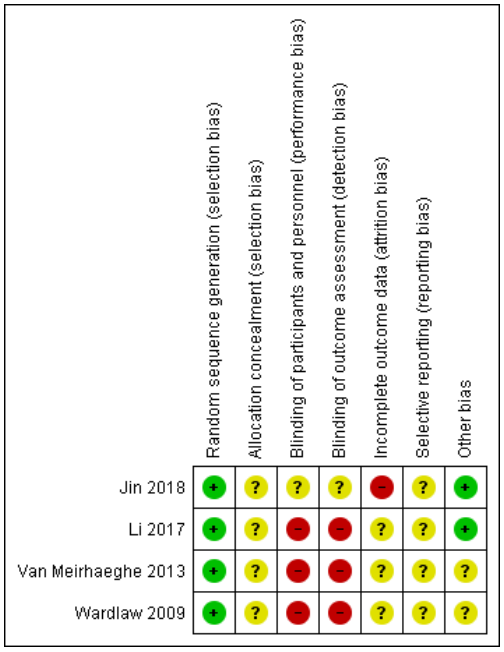


Figure 15 PBK: risk of bias summary for clinical effectiveness outcomes in the RCTs

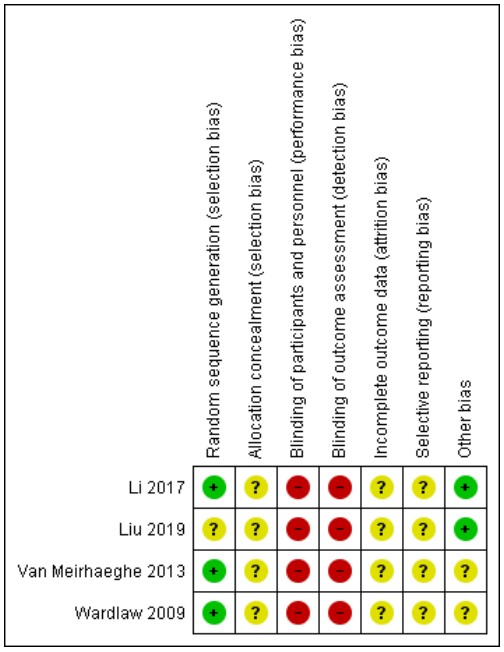


Figure 16 PBK: risk of bias summary for safety outcomes in the RCTs

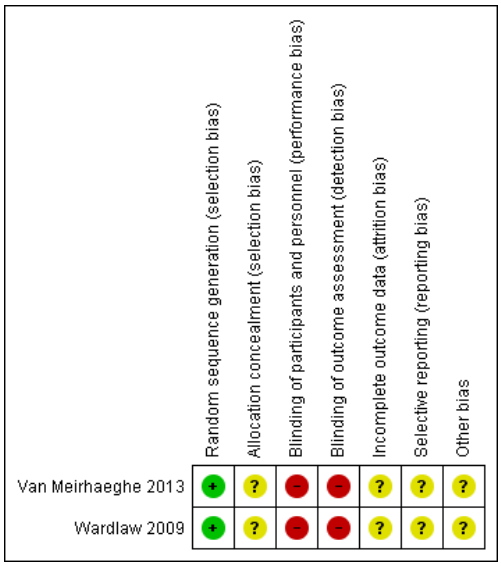


Figure 17 PBK: risk of bias summary for radiographic outcomes in the RCTs

Non-RCTs and database analyses

The risk of bias summary for clinical effectiveness, and safety outcomes is reported in ***Table 11*** and the risk of bias graphs are presented ***Figure 18***–***Figure 20***. The overall risk of bias for the non-RCTs ranged from low to moderate. The risk of bias for the database analyses was serious. Bias concerns associated with these studies was addressed in ***Section 7.5.1*** and were not discussed further.

Clinical effectiveness and safety outcomes were evaluated separately. Safety was further delineated into radiographic and adverse events/mortality. However, many of the bias concerns were applicable to all outcomes and thus will be discussed together unless otherwise stated. The non-RCTs were generally well performed, as the intervention groups were appropriately defined, there were no deviations beyond that which occurs in normal practice, and data was available for most, if not all participants. Edit-Kock (2011) failed to appropriately define the comparator group and had significant losses to follow-up.[169] Key concerns in the study by Movrin (2010) related to significant baseline differences in age, pain and kyphotic treatment angle between patients undergoing PBK and those undergoing CT.[174] The authors corrected for this when evaluating adjacent fractures but not for any other outcome. Thus, it was unclear whether the differences observed at later timepoints reflect the interventions or patient demographics.

The patient and the outcome assessor were unblinded to the intervention across all the non-RCTs. This was not a concern for objective outcomes such as new fractures. However, for subjective outcomes such as the perception of pain, knowledge of the intervention can introduce bias. Consequently, studies evaluating pain and quality of life measures were considered to have a serious risk of bias. Kasperk (2005) modified the VAS questionnaire as patients were deemed too old or fragile to answer questions regarding sex life, jogging, weight lifting and traveling.[173] It was unclear whether this modified questionnaire was administered to all patients or just those deemed too old or fragile. Giannotti (2012) provided limited methodological information consequently an accurate assessment of risk of bias could not be obtained.[170] This study was deemed to be at moderate risk of confounding, owing to limited patient demographic and co-intervention information provided throughout the study.

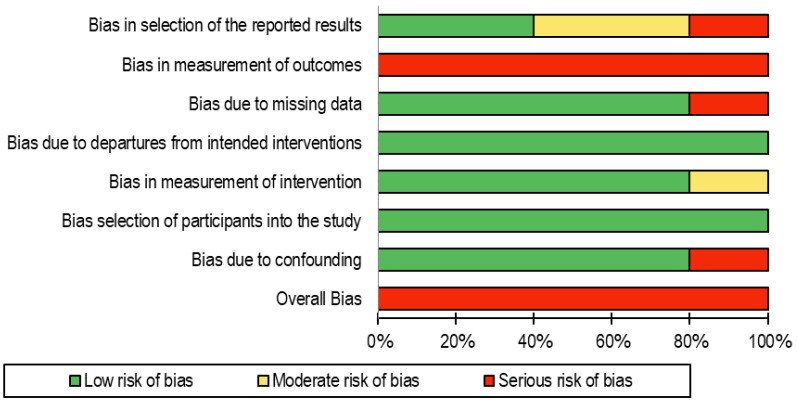


Figure 18 PBK: risk of bias graph for non-RCTs assessing clinical effectiveness outcomes (5 studies)

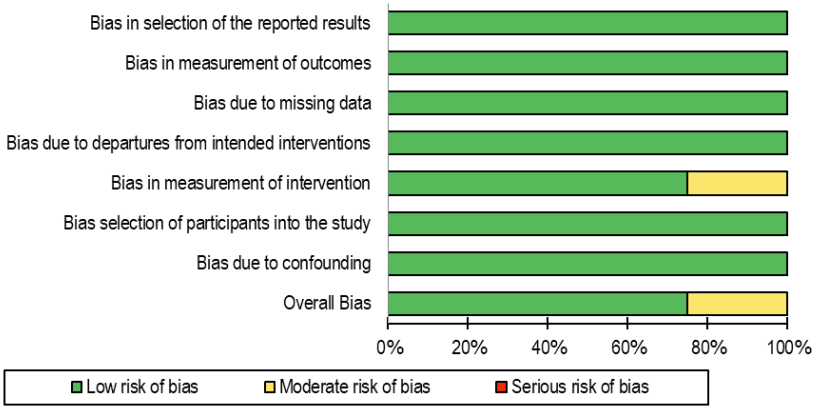


Figure 19 PBK: risk of bias graph for non-RCTs assessing safety outcomes (4 studies)

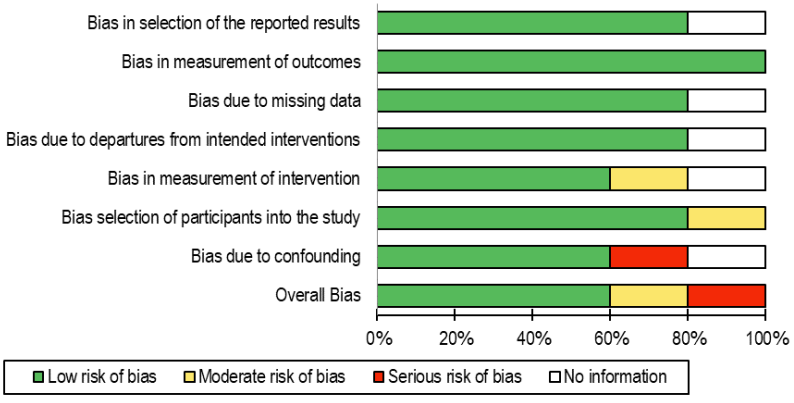


Figure 20 PBK: risk of bias graph for non-RCTs assessing radiographic outcomes (5 studies)

Table 11 PBK: risk of bias summary for clinical effectiveness, safety and radiographic outcomes in the non-RCTs

Author; Year	Bias due to confounding	Bias selection of participants into the study	Bias in measurement of intervention	Bias due to departure from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported results	Overall Bias
Clinical effectiveness								
Edit-Kock 2011[169]	Low	Low	Moderate	Low	Serious	Serious	Low	Serious
Grafe 2005[171]	Low	Low	Low	Low	Low	Serious	Low	Serious
Kasperk 2005[173]	Low	Low	Low	Low	Low	Serious	Moderate	Serious
Kasperk 2010[172]	Low	Low	Low	Low	Low	Serious	Moderate	Serious
Movrin 2010[174]	Serious	Low	Low	Low	Low	Serious	Low	Serious
Safety								
Edit-Kock 2011[169]	Low	Low	Moderate	Low	Low	Low	Low	Serious
Grafe 2005[171]	Low	Low	Low	Low	Low	Low	Low	Low
Kasperk 2005[173]	Low	Low	Low	Low	Low	Low	Low	Low
Kasperk 2010[172]	Low	Low	Low	Low	Low	Low	Low	Low
Radiographic								
Giannotti 2012[170]	NI	Moderate	NI	NI	NI	Low	NI	Moderate
Grafe 2005[171]	Low	Low	Low	Low	Low	Low	Low	Low
Kasperk 2005[173]	Low	Low	Low	Low	Low	Low	Low	Low
Kasperk 2010[172]	NI	Moderate	NI	NI	NI	Low	NI	Moderate
Movrin 2010[174]	Serious	Low	Low	Low	Low	Low	Low	Serious

Abbreviations
NI = no information.

Single-arm trials

Six studies on PBK were included in the safety analysis. The risk of bias graph is presented in ***Figure 21*** and the risk of bias summary is presented in ***Table 111***. Generally, the studies were of moderate quality. Most studies clearly stated their objective, used appropriate methods to measure outcomes before and after the intervention, followed patients for sufficient duration and adequately reported adverse events. The intervention was well described by most studies and losses to follow-up were documented. Conclusions of the studies were generally supported by the results.

Major limitations of the single-arm studies included poorly described co-interventions and that most procedures were performed in a single centre by non-blinded outcome assessors. Patient eligibility criteria was well reported in some studies, although most failed to report if patients entered the study at a similar point in their disease. Only three of the 6 studies explicitly stated that patients were consecutively enrolled.[164 165 177]

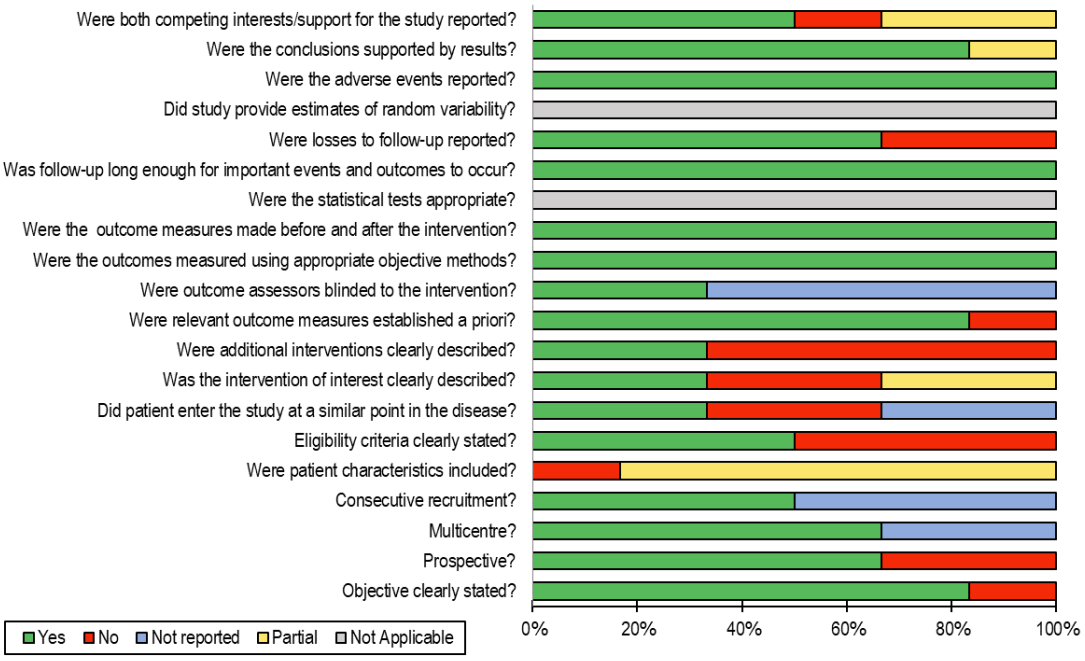


Figure 21 PBK: risk of bias graph for single-arm trials assessing safety outcomes (6 studies)

7.5.3 CT (Extended Assessment of Harms)

AMSTAR 2 was used to evaluate the studies included in the extended assessment of harms. The methodological quality of the studies ranged from low to critically low in the meta-analyses, while the quality of the retrospective pooled analysis was critically low. The risk of bias graph is presented in ***Figure 22***. The risk of bias summary is presented in ***Section 17.3 (Appendix C) Table 109***.

A PICO (or modification thereof) was reported in all 6 studies. Studies mostly reported search terms, search strategy and any deviation from the protocol. Data extraction was generally performed by 2 reviewers. Only 2 studies provided a list of excluded trials. This is likely appropriate given the volume of studies included in each of the meta-analyses.

Risk of bias assessment and quality appraisal was routinely performed in the meta-analysis with studies utilising tools such as GRADE, risk of bias 2.0 and the *Strengthening the Reporting of Observational Studies in Epidemiology* statement.[182] The approach taken in the analyses was generally appropriate with most using a conventional meta-analysis. However, heterogeneity was often not adequately explained and, while risk of bias was performed, it was infrequently discussed when interpreting results. All studies reported whether the included studies had conflicts of interest.

Of the RCTs within the meta-analyses, many did not include adequate descriptions of the randomisation procedure, allocation concealment or blinding procedures and had missing data. In spite of the limitations, the studies were unlikely to suffer from publication bias as inferred the Egger's test. However, the meta-analysis of observational/cohort studies noted channelling bias may have significantly impacted the outcomes.

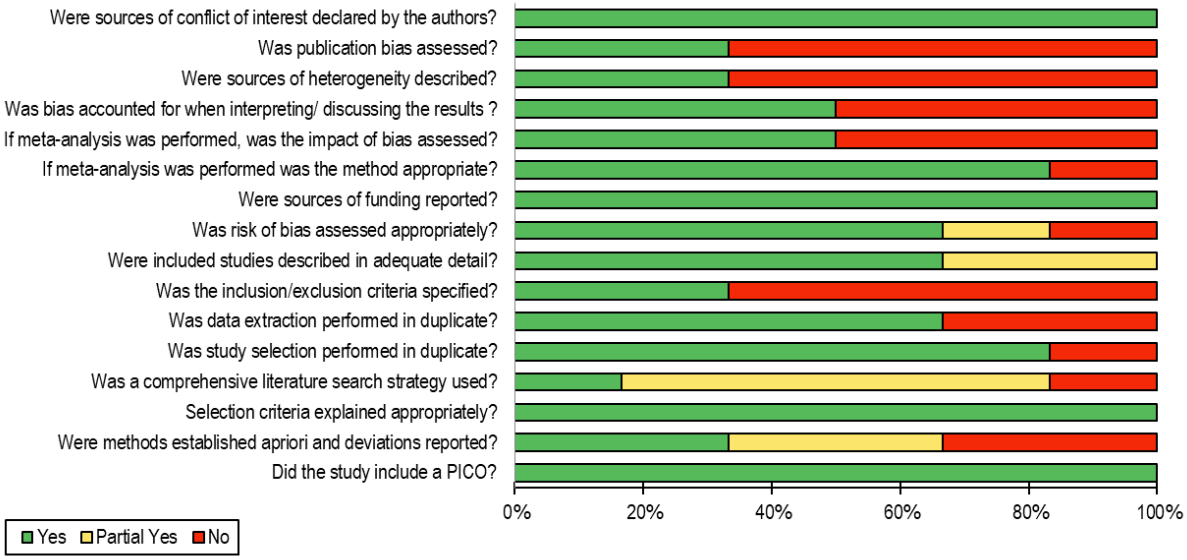


Figure 22 Extended assessment of harms: risk of bias graph for meta-analyses assessing safety outcomes (6 studies)

7.6 Applicability of Evidence Base to Switzerland

Applicability refers to the generalisability of the clinical trials to the Swiss context. It involved comparing patient demographics and clinical characteristics in the RCTs to what generally occurs in Swiss practice. An overview of the demographic and procedural characteristics associated with OVCF, PVP and PBK in Switzerland is provided in ***Table 12*** and ***Table 13***.

Table 12 Swiss demographic information and procedural characteristics associated with vertebroplasty

Parameter	Characteristics
Demographics	Osteoporotic vertebral fracture Older adults (75 and above)* Higher proportion of females[54]* More Caucasian[183 184] Few comorbidities (4.2%)[63]
Clinical Characteristics	No restrictions on use
Setting	Interventional radiologist or spinal surgeon Hospital setting with high-quality imaging equipment Mostly inpatient (96%)
Procedure	Analgesic medication and local anaesthetic, with or without conscious sedation; PMMA cement – assumed to be in accordance with CHUV[92] and CIRSE guidelines[11]
Comparator	Analgesic medication (NSAIDs, opioids and paracetamol), immobilisation (brace and bed rest) and physiotherapy

Abbreviations
CIRSE = Cardiovascular and Interventional Radiological Society of Europe, **PMMA** = polymethyl methacrylate.
Notes
* Further information on age and gender demographics provided in ***Figure 38.***

7.6.1 PVP

RCTs

Of the RCTs, 7 were conducted in Europe, with centres in Denmark, France, Italy, the Netherlands, Switzerland or the UK. Studies performed in these countries were more applicable to the Swiss context than those performed in China and Iran owing to differences in population demographics and healthcare systems.

The trial populations were largely in accordance with osteoporotic vertebral fracture patients in Switzerland. For example, the trial patients were mostly older females (age 70 and above) with few comorbidities. Swiss-specific information relating to the degree of impairment caused by OVCFs and the average age of fracture was not found, noting however, that the RCT patients were in significant

pain and had limited function as inferred by baseline VAS and RMD scores. Fracture age varied from 5.5 days to 7 months and it unclear whether patients had failed CT in several RCTs.

Treatment pathway and procedural characteristics were mostly congruent with the Swiss/European context. For example, patients were recruited from primary care centres or from outpatient and emergency departments at hospitals. Eligible patients were often refractory to medical therapy, had minimum pain scores (3–5 out of 10) and had fractures ranging from several days to months. In Switzerland, there are no pain or fracture duration requirements, therefore, the RCT population may reflect a more impaired cohort of patients. PVP was performed as an inpatient and outpatient procedure by radiologists, neurosurgeons or orthopaedic surgeons. In Switzerland, most procedures are performed on an inpatient basis with few outpatient procedures. PVP was performed under local anaesthetic, conscious sedation or general anaesthesia, with patients placed in the prone position. PMMA cement was delivered to the fracture under fluoroscopic guidance. These characteristics were consistent with the Swiss/European practice.

The conservative treatment and concomitant care administered during the RCTs also aligned with current clinical practice in Switzerland/Europe. For example, CT included analgesic medication (opioids, NSAIDs and paracetamol), bed rest, braces and physiotherapy. Patients were generally offered osteoporosis medication such as bisphosphonates, vitamin D and calcium supplements.

Non-RCTs and database analyses

The applicability of the non-RCTs and database analyses was uncertain due to the lack of methodological and patient-specific information in these trials. The non-RCTs were performed in Romania and Australia, and both database analyses utilised the US Medicare & Medicaid Services Medicare Provider Analysis and Review File database.

The inclusion criteria in the non-RCTs was consistent with that in Switzerland – no specific pain or fracture duration requirements. It was unclear whether there were specific eligibility requirements for patients undergoing PVP in the database analyses because the method of data collection (using ICD codes) omitted this information.

The demographics from the non-RCTs and database analyses were broadly consistent with the Swiss context. For example, the patients were mostly female, Caucasian, and age 65–80. Individuals had fractures for 2 months or less in Andrei (2017).[148] (Average age of fracture in Switzerland is uncertain.) However, the patients enrolled in the database analyses likely reflect an unhealthier population because the majority of patients had comorbidities (as inferred by the Charlson Comorbidity Index). Further, owing to study methodology, the patients identified in the database analyses likely included patients with non-osteoporotic fractures. This limited applicability of the databases to Swiss context.

There were several deviations from Swiss/European practice with respect to treatment management. For example, it was unclear which specialist performed PVP and whether it was performed on an inpatient and outpatient basis. Procedural information was not reported in the database analyses so it was unclear whether any deviations occurred. In the non-RCTs, the approach was likely consistent with Swiss/European practice, namely the use of local anaesthetic, a transpedicular approach, and PMMA cement administered under fluoroscopic guidance.

In the non-RCTs, patients assigned to CT received analgesia, physiotherapy and bed rest. In addition, patients in the PVP and CT groups received concomitant osteoporosis medication. This likely reflected current Swiss/European practice. It was not reported how conservatively managed patients were treated in the database analyses, therefore the applicability cannot be addressed.

Single-arm trials

The single-arm trials were generally applicable to the Swiss context. Of the 15 single-arm studies included in the review, 11 were conducted in Europe with study locations in Germany, Italy, the Netherlands, Poland and Spain. Studies conducted within Europe were more comparable to the Swiss setting than those outside this region.

Eligibility criteria were largely in agreement with Swiss demographics. Included patients were mostly older (mean age 68.5 and above, with 1 study reporting an age range of 58–91 years) females with few reported comorbidities. Patients were those diagnosed with painful osteoporotic vertebral fractures who had failed CT. Fracture duration was highly variable, ranging from 3 days to 12 months. However, there is no fracture age requirement in Switzerland for PVP.

The treatment setting was comparable with the Swiss/European setting. Patients were primarily recruited from hospitals (multidisciplinary spine centres, radiology and internal medicine departments). PVP was performed by interventional radiologists, orthopaedic surgeons or neurosurgeons. However, for most of the studies it was unclear whether the procedure was done on an inpatient or outpatient setting. PVP was performed under local anaesthesia with fluoroscopic guidance using the transpedicular, bipedicular, parapedicular or unipedicular approach. The intercostovertebral approach was used in 1 study and it is unclear how prevalent this approach is in Switzerland. PMMA was the most common cement used for PVP, noting several trials used calcium phosphate cement. These characteristics were in accordance with Swiss or European practice. Reported co-interventions included pain relief and osteoporosis medications, again aligning with Swiss and European practice.

7.6.2 PBK

Table 13 Swiss demographic information and procedural characteristics associated with kyphoplasty

Parameter	Characteristics
Demographics	Osteoporotic vertebral fracture Older adults (75 and above)* Higher proportion of females[54]* More Caucasian[183 184] Few comorbidities (4.2%)[63]
Clinical Characteristics	Fracture < 8 weeks old, unresponsive to analgesics, with pain (VAS ≥ 5) and deformation (thoracic kyphosis >15°, lumbar kyphosis >10°, and/or vertebral body height reduction of more than one-third compared to adjacent bodies)
Setting	Interventional radiologist or spinal surgeon Hospital setting with high-quality imaging equipment Inpatient only
Procedure	Analgesic medication, general anaesthesia or local anaesthesia with or without conscious sedation; PMMA cement – assumed to be in accordance with CHUV[92] and CIRSE guidelines[11]
Comparator	Analgesic medication (NSAIDs, opioids and paracetamol), immobilisation (brace and bed rest) and physiotherapy

Abbreviations
CIRSE = Cardiovascular and Interventional Radiological Society of Europe, **PMMA** = polymethyl methacrylate.
Notes
* Further information on age and gender demographics provided in ***Figure 38.***

RCTs

Of the RCTs, three were single-centre studies performed in China and 1 was a multicentre trial conducted in 8 European countries (Austria, Belgium, France, Germany, Italy, Sweden, the Netherlands and the UK). Population demographics and healthcare systems in these European countries were more consistent with Switzerland than studies based in China, which had limited applicability.

In Switzerland, PBK is reimbursed if the following conditions are met: fracture less than 8 weeks old with deformation (thoracic kyphosis >15°, lumbar kyphosis >10°, and/or vertebral body height reduction of more than one-third compared to adjacent bodies); patient unresponsive to analgesics with VAS pain score greater than 5. Wardlaw (2009) enrolled patients with a baseline VAS pain score above 4 and active fractures, as indicated by the presence of oedema on MRI and vertebral body height loss.[168] Jin (2018) enrolled patients with active, 1 week old fractures.[42] Inclusion and exclusion criteria in the remaining RCTs generally did not match the reimbursement requirements in Switzerland. However, demographics of the enrolled patients typically met several requirements in Switzerland as inferred by baseline reported VAS, Genant grades, fracture age or Cobb and kyphosis angle.

The trial populations were largely in accordance with Swiss vertebral fracture patients with respect to age and sex. For example, the trial population were predominately older (age 70 and above) females with few comorbidities. Swiss-specific information relating to the degree of impairment was unavailable. The RCT patients were incapacitated by the fractures as inferred VAS, RMD and QUALEFFO scores.

It was unclear whether the treatment pathways were consistent with the Swiss/European context, owing to under-reporting of study methodology. For example, it was not reported where or how patients were recruited, nor whether they were refractory to medical therapy. However, procedural characteristics were likely similar to Swiss practice with respect to personnel, anaesthetic and imaging requirements. For example, the procedures were performed by radiologists or surgeons (specialty not specified) who utilised fluoroscopic guidance for monitoring balloon inflation and administration of PMMA cement. Patients were placed under general or local anaesthesia. The procedure was performed as an inpatient procedure in Wardlaw (2009).[168] The remaining RCTs did not report whether the procedure was performed on an inpatient or outpatient basis.

CT and concomitant care administered to trial patients aligned with clinical practice in Switzerland/Europe. For example, CT included analgesics, bed rest, braces and physiotherapy.

Non-RCTs and database analyses

The comparative evidence base consisted of 4 non-RCTs and 2 database analyses. Applicability issues associated with the database analyses were discussed previously (see ***Section 7.6.1***).

The applicability of the non-RCTs was uncertain due to the lack of methodological and patient-specific information in the studies. The non-RCTs were conducted in Germany, Italy and Slovenia. The population demographics and healthcare systems of Germany and Italy likely reflect Switzerland more than Slovenia – thereby limiting the applicability Morvin (2010).[174]

Inclusion and exclusion criteria across the non-RCTs were generally under-reported. Movrin (2010) enrolled patients with baseline VAS pain score greater than 5, kyphotic deformity greater than 30˚ and loss of vertebrae height.[174] Edit-Koch (2011) included patients with VAS scores greater than 5 and fractures younger than 3 months.[169] For the remaining trials, inclusion criteria involved patients with back pain due to osteoporotic vertebral fractures. These criteria conflicted with current Swiss recommendations, as reported in ***Table 13***. However, demographics of the enrolled patients often met some of the requirements in Switzerland with respect to pain and degree of kyphosis. The patient demographics in the non-RCTs were broadly consistent with the Swiss context with respect to sex (mostly female) and age (70 and above).

It was unclear whether the non-RCTs were consistent with Swiss/European practice. For example, only 1 study noted that the procedure was performed on an inpatient basis.[174] Two studies noted the

procedure was performed by radiologists or surgeons.[170 173] PBK was generally performed using a bipedicular approach with balloon inflation and cement placement via fluoroscopic guidance. PMMA cement was utilised in 5 studies, with 1 study using either PMMA or calcium phosphate cement. Giannotti (2012) provided limited procedural information, thus its applicability from a practice perspective could not be determined.[170]

CT consisted of physiotherapy and pain medication, with both groups receiving osteoporotic medication (bisphosphonates, calcium and Vitamin D). These characteristics were likely consistent with the Swiss/European practice.

Single-arm trials

The applicability of the SwissSpine registry[63] is discussed separately from the remaining studies as it was directly applicable to practice in Switzerland. The main concern with the SwissSpine registry was whether osteoporosis practice differed between the evaluated period (2005–2011) and 2020. Remaining assessment domains were likely reflective of contemporary Swiss practice with respect to patient demographics, setting and procedural characteristics.

The applicability of the remaining single-arm trials was uncertain due to the lack of methodological and patient-specific information provided.

The single-arm trials were conducted in the United States (k = 2), Germany (k = 1) and Spain (k = 1). These countries were broadly congruent with population demographics and healthcare systems in Switzerland, however, it is unclear if enrolled patients were entirely reflective of Swiss practice because inclusion and exclusion criteria were not explicitly stated in three trials.[165 175 176] Of the trials reporting inclusion criteria, Robinson (2008) enrolled patients with OVCF, kyphotic deformity greater than 15° and baseline pain score greater than 5, who were refractory to medical treatment for at least 12 weeks.[177] Dohm (2014) enrolled patients with painful OVCF younger than 6 months and oedema on MRI[164] and Prokop (2012) required a minimum pain score of 6 out of 10.[176] These criteria generally matched certain aspects of the reimbursement criteria for PBK in Switzerland.

The trial populations were largely in accordance with Swiss vertebral fracture patients with respect to age and sex. For example, the trial population were predominately older (age 70 and above) females. Swiss-specific information relating to degree of impairment was unavailable, although patients were in notable pain as inferred by baseline VAS scores.

It was unclear whether the single-arm trials are consistent with Swiss/European practice. For example, no study indicated whether the procedure was performed on an inpatient basis and only three studies noted that the procedure was performed by a radiologist or surgeon.[164 165 175] PBK was performed under general anaesthesia with patients placed in the prone position in three studies.[175-177] PBK was generally

performed via a bipedicular approach with balloon inflation and cement placement performed by fluoroscopic guidance. PMMA cement was used in 4 studies.[164] [165] [176] [177] One study used PMMA or calcium phosphate cement.[175] Concomitant osteoporosis medication including calcium, vitamin D and bisphosphonates were prescribed in 1 study.[176]

7.6.3 CT (Extended Assessment of Harms)

The applicability of studies evaluating CT was uncertain. There was limited available evidence evaluating the safety of opioids, NSAIDs and paracetamol in patients with vertebral fractures or osteoporosis so the inclusion criteria were broadened. Consequently, the assessed populations generally do not match the Swiss population with respect to disease indication, age or sex. However, like patients with osteoporotic vertebral fractures, most of the patients included in the meta-analyses were using these medications to relieve pain (musculoskeletal conditions such as osteoarthritis, chronic lower back pain and rheumatoid arthritis being the most frequently reported indications for taking medication). While these conditions have a different pathogenesis to osteoporotic fractures, and medication dose and duration may differ, they all cause pain and reduce quality of life. The medications may be used at different stages of the disease and patients may be in more or less pain than those with vertebral fracture, further limiting the applicability of these studies.

It was unclear whether any of the studies were conducted wholly or partially in Switzerland, as larger meta-analyses failed to report the country in which the studies were conducted. Few trials reported detailed patient demographics. Four meta-analyses noted patient populations were predominately older (greater than 60 years), Caucasian females, which broadly matched patients with osteoporotic vertebral fractures in Switzerland.

7.7 Results: Clinical Effectiveness

7.7.1 PVP vs CT

Table 14 provides a summary of the main pooled clinical effectiveness outcomes comparing PVP to CT. The 1 month and 12 month timepoints were selected as representative timepoints for short- and long-term follow-up, respectively. By 1 month, there were statistically significant and clinically meaningful differences (indicated by dark green shading ***Table 14***) between PVP and CT for pain (VAS) and function-related outcomes (ODI and RDQ) with the results favouring PVP. There were no differences for quality of life measures (EQ-5D and QUALEFFO) or analgesic use. The statistical differences in VAS, ODI and RDQ persisted to 12 months; however, effect sizes decreased and consequently did not surpass the lower bounds of identified MCIDs. Therefore, while statistically significant, the effects did not translate to a clinically meaningful effect (indicated by light orange shading ***Table 14***). For additional information regarding each outcome, refer to the corresponding sections below.

Table 14 Summary of meta-analyses for PVP compared to CT

Outcome	Length of follow-up Mean difference (95% CI); p value					
	1 day	*1 week*	*1 month*	*3 months*	*6 months*	*12 months*
Pain (VAS) [a]	-2.99 (3.20, - 2.78) p < 0.0001	-1.88 (-2.92, -0.85) p = 0.0004	-1.52 (-2.86, -0.17) p = 0.03	-1.21 (-1.94, -0.49) p = 0.001	-1.14 (-1.77, -0.52) p = 0.02	-1.22 (-1.75, -0.69) p = 0.0003
Pain (analgesic use) [b]	NR	0.62 (0.20, 1.89) p = 0.40	0.53 (0.10, 2.69) p = 0.44	NR	0.48 (0.10, 2.42) p = 0.38	NR
Function (ODI)	-14.49 (-16.28, -12.69) p < 0.0001	-15.54 (-17.36, -13.72) p < 0.0001	-16.27 (-23.53, -9.01) p < 0.0001	-17.73 (-27.17, -8.28) p = 0.0002	-15.40 (-21.88, -8.91) p < 0.0001	-10.78 (-19.82, -1.73) p = 0.02
Function (RDQ)	-2.36 (-2.70, -2.02) p < 0.0001	-2.03 (-2.22, -1.85) p < 0.0001	-2.03 (-3.06, -1.01) p = 0.0001	-2.09 (-2.75, -1.43) p < 0.0001	-2.14 (-3.01, -1.27) p < 0.0001	-1.82 (-2.10, -1.53) p < 0.0001
Quality of life (EQ-5D)	NR	0.10 (-0.11, 0.31) p = 0.35	0.10 (-0.11, 0.31) p = 0.36	0.03 (-0.16, 0.22) p = 0.76	0.10 (-0.12, 0.32) p = 0.37	0.10 (-0.10, 0.30) p = 0.32
Quality of life (QUALEFFO)	NR	-4.89 (-9.63, -0.15) p = 0.04	-6.16 (-15.84, 3.52) p = 0.21	-6.14 (-14.34, 2.07) p = 0.14	-4.99 (-14.07, 4.09) p = 0.28	-2.78 (-9.02, 3.46) p = 0.38

Abbreviations
CI = confidence interval, **EQ-5D** = EuroQol 5 dimension questionnaire, **NA** = not applicable, **ODI** = Oswestry disability index, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RDQ** = Roland-Morris disability questionnaire, **VAS** = visual analogue scale.
Notes
For continuous outcomes, negative mean difference favours PVP, positive mean difference favours CT. For dichotomous outcomes (analgesic use only), positive risk ratio favours PVP, negative risk ratio favours CT.
a = MCIDs for pain vary from 1.5 to 4. If using the lower bounds, pain is clinically significant from 1 day to 1 month. If other estimates are used, it is unlikely to be clinically significant.

b = All types of analgesics pooled were together and the outcome of analgesic use reported as risk ratio (95% CI).
No shading = no statistically significant difference between groups (p > 0.05) or clinically meaningful effect (as inferred by MCIDs).
Light orange shading = statistically significant differences between groups (p < 0.05), however, the results do not surpass lower bounds of identified MCIDs.
Dark green shading = statistically significant difference between groups (p < 0.05) and likely to translate to a clinically meaningful effect (surpasses lower bounds of identified MCIDs).

Pain-related outcomes

PVP vs CT, pain (VAS), 1 day to 36 months

Seven studies provided evidence on pain, as measured by VAS, from 1 day to 36 months post-intervention. All 7 studies were included in the meta-analysis.[6 8 96 132 134 135 141] Overall, there was statistically significant differences between PVP and CT groups at all timepoints post-intervention (see ***Figure 23***). At 1 month the mean difference was -1.52 (95%CI -2.86, -0.17; p = 0.03) and by 12 months the difference was -1.22 (95%CI -1.75, -0.69; p = 0.0003). However, Tau^2 and I^2 statistics indicated moderate to considerable levels of heterogeneity and inconsistency for most timepoints.

Sub-groups

Sub-group analysis of vertebral fractures younger than 8 weeks determined significant differences in pain scores between PVP and CT groups at all timepoints post-intervention, as inferred by longitudinal meta-analysis.[6 96 132] At 1 month the mean difference was -2.41 (95%CI -2.61, -2.22; p < 0.0001) and by 12 months the difference was -1.26 (95%CI -1.65, -0.88; p < 0.0001). For fractures older than 8 weeks there were statistically significant differences from 1 to 12 months, but not at 1 or 2 weeks post-intervention (as inferred by a longitudinal meta-analysis).[8 135 141 146] At 1 month the mean difference was -1.12 (95%CI -1.53, -0.71; p < 0.0001) and by 12 months the difference was -1.28 (95% CI -2.24, -0.33; p = 0.01). Both sub-groups reported considerable levels of heterogeneity and inconsistency at most timepoints. For further information regarding sub-group analysis for VAS refer to ***Section 17.4*** (***Appendix D***), ***Table 114***.

The included studies utilised different methods for assessing pain. Seven studies used a 10-point VAS scale (10 representing the worst pain)[6 8 96 132 134 135 141] and 1 study used a nine-point scale (10 representing the worst pain).[135] While the studies differed slightly in scale, it was unlikely to significantly impact overall results when meta-analysed. The studies did not report the context in which the pain was felt (e.g. spontaneous or pain during activity) or who completed the VAS measurement.

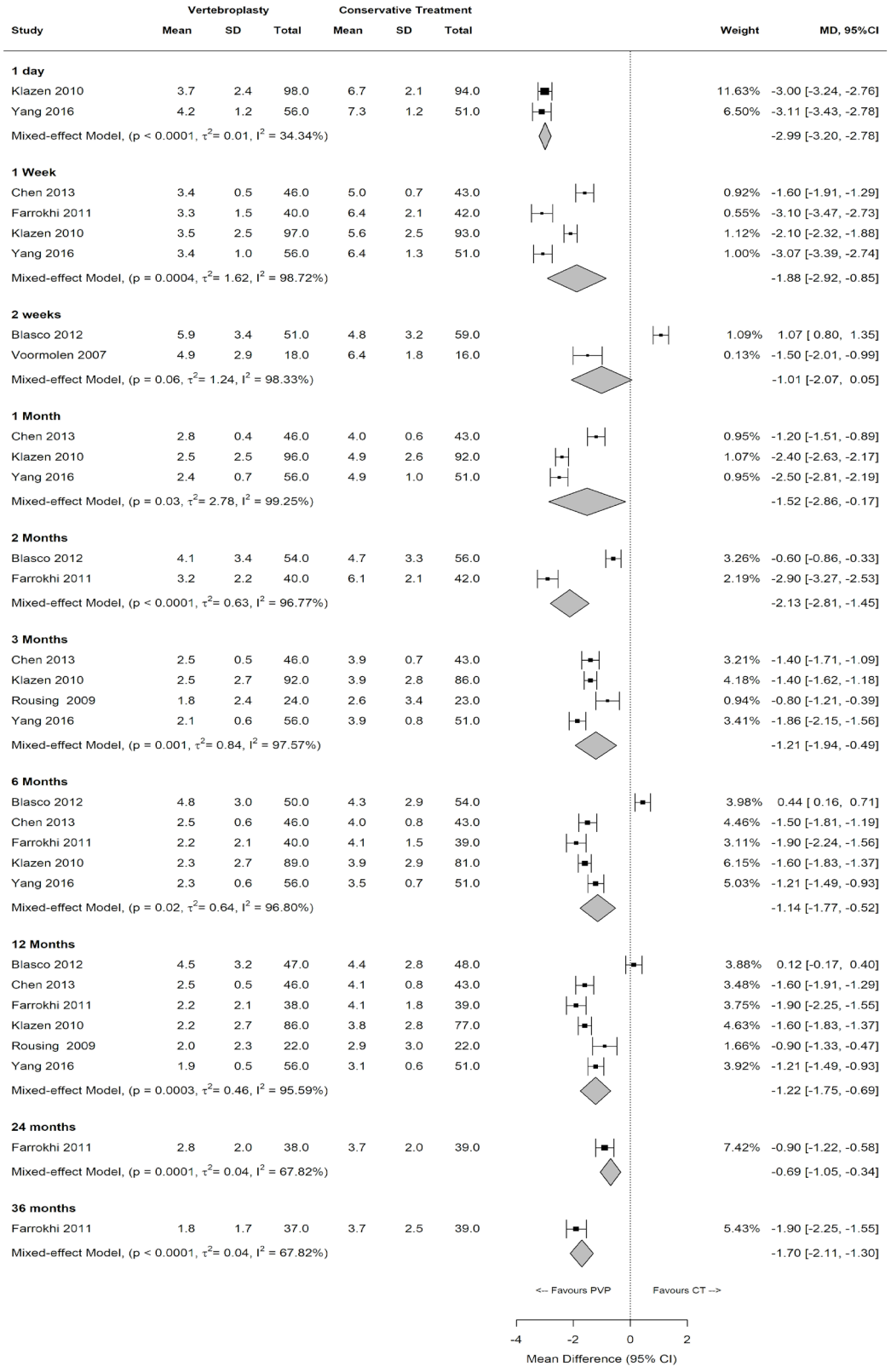


Figure 23 Mean difference in pain (VAS) for PVP compared to CT (1 day to 36 months)

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation, **VAS** = visual analogue scale.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PVP vs CT, pain (analgesic use), 1 week to 6 months

Four studies provided evidence on the number of patients using analgesics, from 1 week to 6 months post-intervention (***Table 15***).[8 132 134 141] Two studies were included in the meta-analysis (the other 2 were only described narratively). There were no statistically significant differences at 1 week (RR 0.62; 95%CI 0.20, 1.89; p = 0.40), 1 month (RR 0.53; 95%CI 0.10, 2.69; p = 0.44) or 6 months (RR 0.48; 95%CI 0.10, 2.42; p = 0.38).

Meta-analysis for the sub-groups was not performed owing to the number of available studies.

Types of analgesics were not specified in Chen (2014).[134] In Blasco (2012), analgesics included minor analgesic, minor opioid and major opioid.[8]

Table 15 PVP compared to CT: pain (analgesic use) 1 day to 6 months

Analgesic use	Number of studies	Heterogeneity	PVP n/N (%)	CT n/N (%)	Risk Ratio (RR, 95% CI)
1 week	2[8 134]	Chi^2 = 18.60 p < 0.00001 I^2 = 95%	64/110 (58.2%)	82/104 (78.9%)	0.62 (0.20, 1.89) p = 0.40
1 month	2[8 134]	Chi^2 = 18.80 p < 0.0001 I^2 = 95%	56/110 (50.9%)	71/104 (68.3%)	0.53 (0.10, 2.69) p = 0.44
6 months	2[8 134]	Chi^2 = 18.90 p < 0.0001 I^2 = 95%	54/110 (49.1%)	76/104 (73.1%)	0.48 (0.10, 2.42) p = 0.38

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **PVP** = percutaneous vertebroplasty, **RR** = risk ratio.

The 2 studies not included in the meta-analysis reported the range[141] or described the results narratively (***Table 16***).[132] Both studies concluded that there were statistically significant differences in analgesic use in the short-term between treatment groups (p value not reported in Voormolen [2007]).[132 141] Klazen (2010) reports that the differences were not significant at later timepoints (3–12 months).[132]

Table 16 PVP compared to CT: pain (analgesics use) 1 day to 2 weeks

Author; Year	Length of follow-up	PVP Mean (range) or n/N	CT Mean (range) or n/N	Mean difference (95% CI)	p value
Voormolen 2007[141]	Baseline 1 day 2 weeks	1.9 (0–3) 1.1 (0–3) 1.2 (0–3)	1.7 (0–3) 2.5 (1–3) 2.6 (2–3)	NR -1.4 (-2.1, -0.8) -1.5 (-2.3, -0.8)	NR < 0.05 < 0.05
Klazen 2010[132]	Baseline 1 day 1 week 1 month	96/101 NR NR NR	94/101 NR NR NR	NR NR NR NR	> 0.05 < 0.001 = 0.001 0.033

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **NR** = not reported, **PVP** = percutaneous vertebroplasty.

Function-related outcomes

PVP vs CT, function (ODI), 1 day to 36 months

Three studies provided evidence on function, as measured by ODI, from 1 day to 36 months post-intervention.[6 134 135] All three studies were included in the meta-analysis. There were statistically significant differences between PVP and CT groups at all timepoints post-intervention (see ***Figure 24***). At 1 month the mean difference was -16.27 (95%CI -23.52, -9.01; p < 0.0001) and by 12 months the difference was -10.78 (95%CI -19.82, -1.73; p = 0.02). However, Tau^2 and I^2 statistics indicated considerable levels of heterogeneity and inconsistency for most timepoints.

Sub-groups

In 1 study, vertebral fractures younger than 8 weeks showed statistically significant differences in ODI scores at all timepoints post-intervention (1 week to 12 months).[6] At 1 month the mean difference was -24.15 (95%CI -27.41, -20.89) and by 12 months the difference was -7.92 (95%CI -10.64, -5.20). For fractures older than 8 weeks there were statistically significant differences at all timepoints (1 day to 36 months), as indicated by longitudinal meta-analysis.[135 146] At 1 month the mean difference was -12.95 (95%CI -13.96, -11.93; p < 0.0001) and by 12 months the difference was -12.46 (95% CI -13.52, -11.40; p < 0.0001). However, all timepoints exhibited considerable levels of heterogeneity as inferred by the Tau^2 and I^2 statistic. For further information regarding sub-group analysis for ODI refer to ***Section 17.4*** (***Appendix D***), ***Table 115***.

All studies used the 0–100% ODI measure, whereby 0–20% represents minimal disability, 21–40% was moderate disability, 41–60% was severe disability, 61–80% was crippling back pain and 81–100% as bed-bound.[6 134 135]

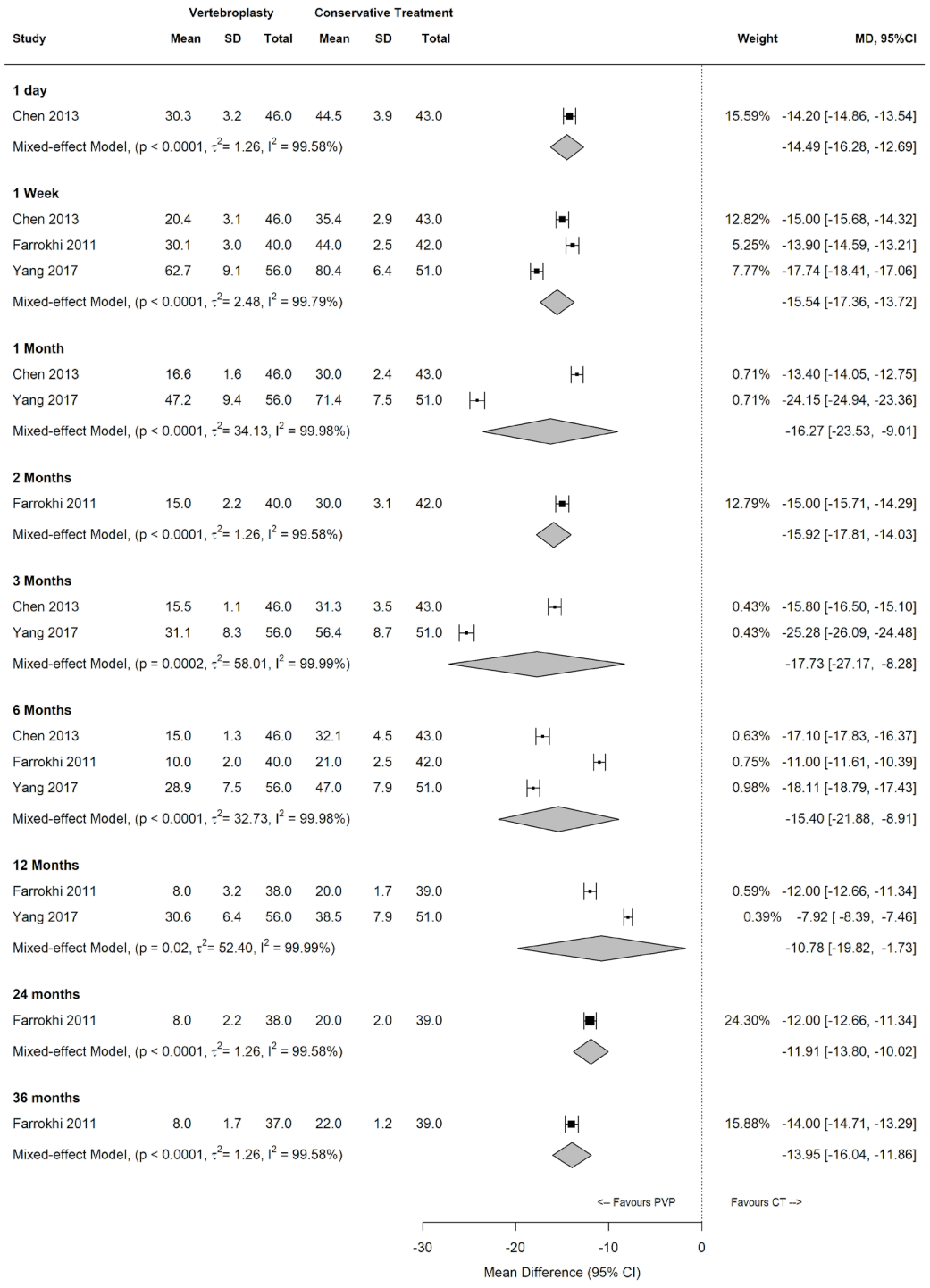


Figure 24 Forest plot indicating mean difference in function (ODI) for PVP compared to CT (1 day to 36 months)

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **MD** = mean difference, **ODI** = Oswestry disability index, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PVP vs CT, function (RDQ), 1 day to 12 months

Three studies provided evidence on function, as measured by RDQ, from 1 day to 12 months post-intervention.[132] [134] [141] Two studies were included in the meta-analysis.[132] [134] Overall, there were statistically significant differences between PVP and CT groups at all timepoints post-intervention (see ***Figure 25***). The mean difference at 1 month was -2.03 (95%CI -3.06, -1.01; p = 0.0001) and by 12 months the difference was -1.82 (95%CI -2.10, -1.53; p < 0.0001). However, the Tau^2 test and I^2 statistic indicated moderate to considerable levels of heterogeneity and inconsistency at all timepoints.

Sub-groups

Vertebral fractures younger than 8 weeks were evaluated in 1 study.[132] There were statistically significant differences in RDQ scores at 3 months post-intervention, but not at any of the remaining timepoints. At 1 month the mean difference was -1.50 (95%CI -3.22, 0.22) and by 12 months the difference was -1.90 (95%CI -4.01, 0.21). Vertebral fractures older than 8 weeks were evaluated in 1 study.[146] There were statistically significant differences in RDQ scores at all timepoints post-intervention (1 day to 6 months). At 1 month the mean difference was -2.60 (95%CI -3.26, -1.94) and at 6 months the difference was -2.60 (95%CI -2.98, -2.22). For further information regarding the sub-group analysis for RDQ refer to ***Section 17.4*** (***Appendix D***)***, Table 116.***

All studies used the 0–24-point RDQ measure, with higher scores indicating decreasing physical functioning and increasing disability.[132] [134]

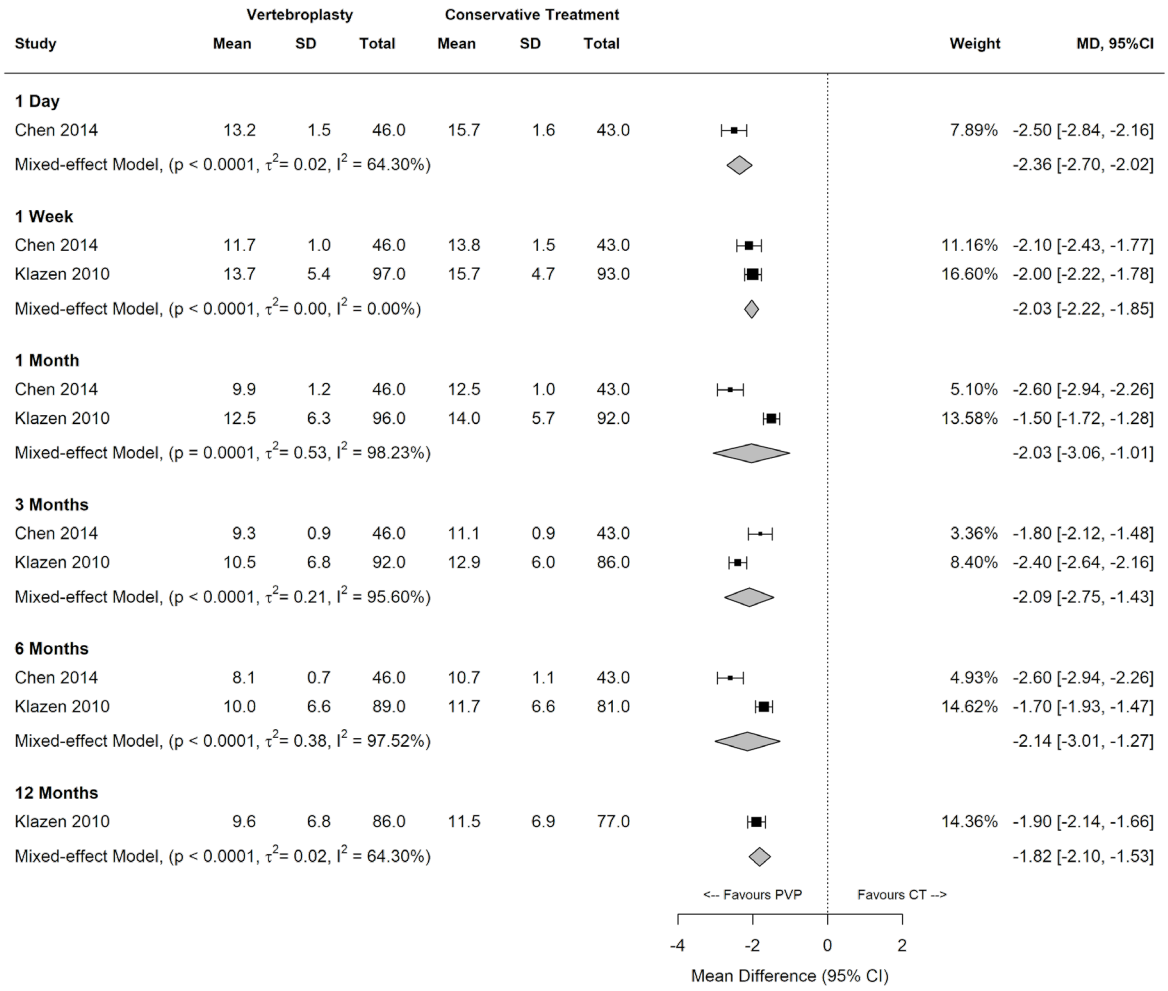


Figure 25 Forest plot indicating mean difference in function (RDQ) for PVP compared to CT (1 day to 12 months)

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **RDQ** = Roland-Morris disability questionnaire, **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

One study was not included in the meta-analysis because range was reported rather than standard deviation (***Table 17***).[141] There was a statistically significant difference between PVP and CT groups at 2 weeks post-intervention.

Table 17 PVP compared to CT: function (RDQ) 1 day to 2 weeks

Author; Year	Length of follow-up	PVP Mean (range)	CT Mean (range)	Mean difference (95% CI)	p value
Voormolen 2007[141]	Baseline 1 day 2 weeks	1.9 (0–3) 1.2 (0–3) 1.2 (0–3)	1.7 (0–3) 2.6 (2–3) 2.6 (2–3)	NR 1.4 (-2.0, -0.8) 1.4 (-2.0, -0.8)	NR < 0.05 < 0.05

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **NR** = not reported, **PVP** = percutaneous vertebroplasty, **RDQ** = Roland-Morris disability questionnaire .

PVP vs CT, function (timed up-and-go), 3 to 12 months

One study provided evidence on timed up-and-go scores at 3 and 12 months (***Table 18***).[96] There was no statistically significant difference between PVP and CT groups at either timepoint (p > 0.05).

The test involved patients rising from a chair, walking three metres, returning to, and sitting back down in the chair. A reduction in time corresponded to improved function.[96]

Table 18 PVP compared to CT: function (timed up-and-go) at 3 and 12 months

Author; Year	Length of follow-up	PVP Mean ± SD	CT Mean ± SD	p value
Rousing 2009[96]	Baseline 3 months 12 months	NR 16.0 ± 5.5 seconds 16.1 ± 7.9 seconds	NR 17.0 ± 9.7 seconds 17.3 ± 9.2 seconds	 0.75 0.67

Abbreviations
CT = conservative treatment, **NR** = not reported, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation.

Quality of life-related outcomes

PVP vs CT, quality of life (EQ-5D), 1 week to 12 months

Two studies provided evidence on quality of life, as measured by EQ-5D, from 1 week to 12 months post-intervention.[96] [132] Both studies were included in the meta-analysis. Overall, there were no statistically significant differences between PVP and CT groups at any timepoint post-intervention (see ***Figure 26***). At 1 month the mean difference was 0.10 (95%CI -0.10, 0.31; p = 0.36) and by 12 months

the mean difference was 0.10 (95%CI -0.10, 0.30; p = 0.32). The Tau^2 and I^2 statistics indicated low levels of heterogeneity and inconsistency at all timepoints.

The baseline EQ-5D score significantly differed in Rousing (2009) (p = 0.04) with patients in the PVP group reporting higher EQ-5D scores compared to the CT group.[96] Similarly, baseline EQ-5D scores differed in Klazen (2010) with patients in the PVP group reporting lower EQ-5D scores compared to the CT group (p < 0.05).[132]

Sub-group analysis was not performed because both studies enrolled participants with vertebral fractures younger than 8 weeks.

Both studies used the EQ-5D scale from 0–1, whereby 0 indicates death and 1 indicates perfect health.

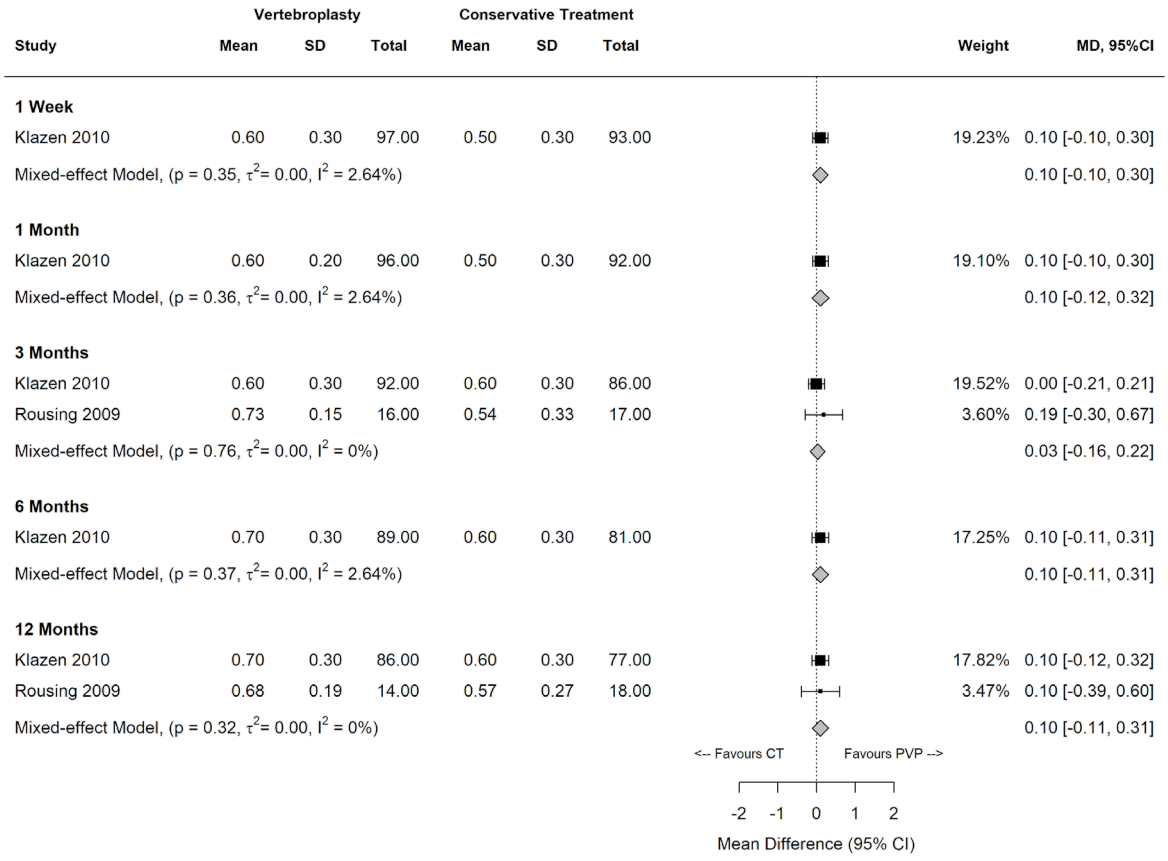


Figure 26 Forest plot indicating mean difference in quality of life (EQ-5D) for PVP compared to CT (1 week to 12 months)

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimensions questionnaire, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation.
Notes
Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PVP vs CT, quality of life (QUALEFFO), 1 week to 12 months

Four studies provided evidence on quality of life, as measured by QUALEFFO, from 1 week to 12 months post-intervention.[6 8 132 141] Three studies were included in the meta-analysis.[6 8 132] Overall, there was a statically significant difference between PVP and CT groups at 1 week (MD -4.89; 95% CI -9.63, -0.15; p = 0.04) but not at any other timepoint (2 weeks to 12 months) (see ***Figure 27***). At 1 month the mean difference was -6.16 (95% CI -15.84, 3.52; p = 0.21). By 12 months the difference was -2.78 (95% CI -9.02, 3.46; p = 0.38). The Tau^2 and I^2 statistics indicated considerable levels of heterogeneity and inconsistency at all timepoints.

Sub-groups

Sub-group analysis of vertebral fractures younger than 8 weeks found statistically significant differences in QUALEFFO scores between PVP and CT groups at 1 week and at 3 months.[6 132] Longitudinal meta-analysis found no differences at any other timepoints. At 1 month the mean difference was -9.86 (95%CI -20.97, 1.24; p = 0.08) and by 12 months the difference was -5.42 (95%CI -11.14, 0.30; p = 0.06). All timepoints exhibited considerable levels of heterogeneity and inconsistency as inferred by the Tau^2 and I^2 statistic. Vertebral fractures older than 8 weeks were evaluated in 1 study.[8] There were no statistically significant differences in QUALEFFO scores between PVP and CT groups at any timepoint post-intervention (2 weeks to 12 months). At 2 months the mean difference was 2.38 (95%CI -4.56, 9.32) and at 12 months the difference was 2.54 (95%CI -5.06, 10.14). For further information regarding the sub-groups for QUALEFFO refer to ***Section 17.4*** (***Appendix D***), ***Table 117***.

All studies used the same QUALEFFO measure (ranging from 0–100), with 0 indicating a high quality of life and 100 indicating a poor quality of life.[6 8 132 141]

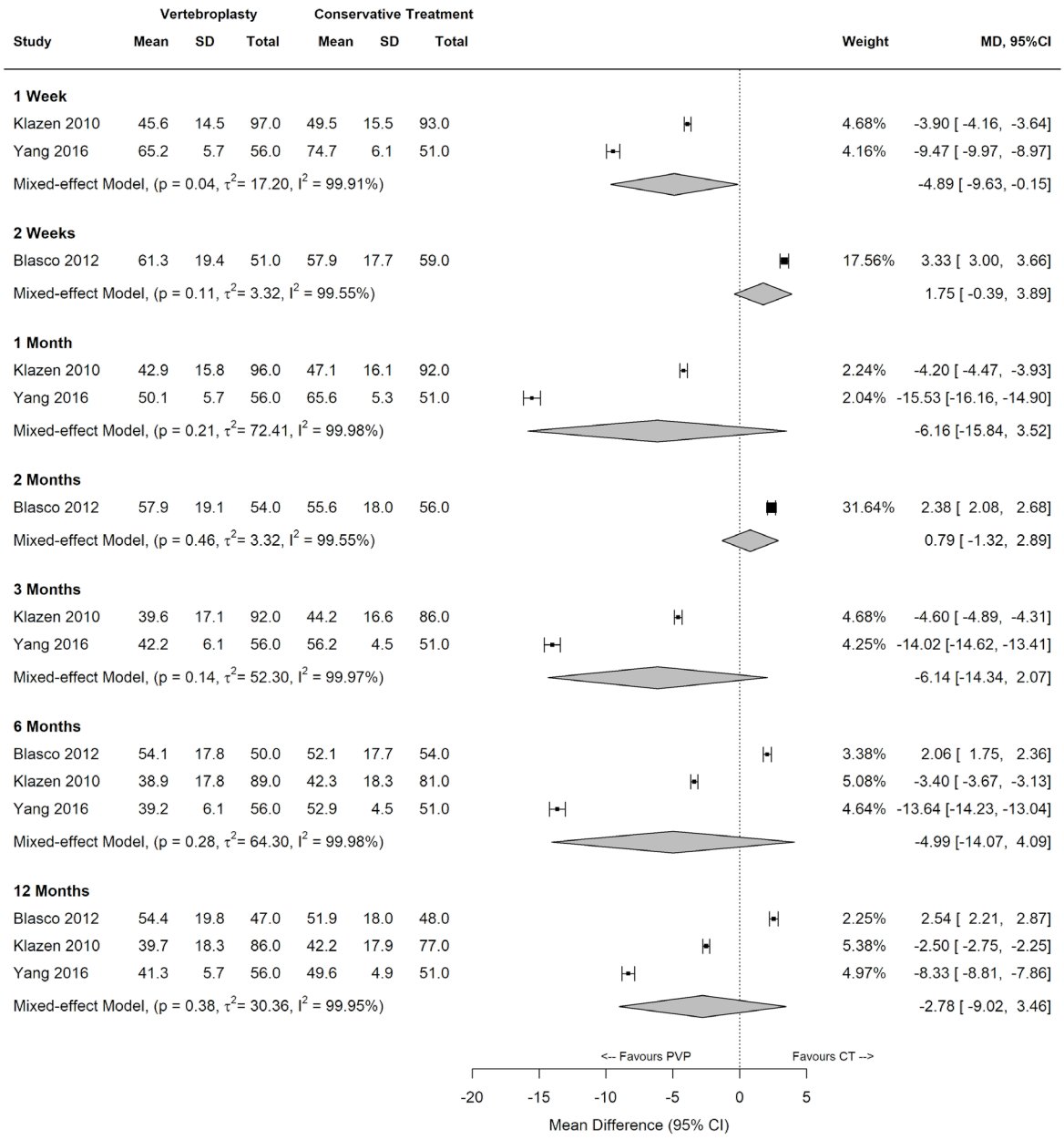


Figure 27 Forest plot indicating mean difference in quality of life (QUALEFFO) for PVP compared to CT (1 week to 12 months)

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = Quality of life questionnaire of the European Foundation for Osteoporosis, **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

One study was not included in the meta-analysis because range was reported rather than standard deviation (***Table 19***).[141] There was a statistically significant difference between the PVP and CT groups at 2 weeks post-intervention.

Table 19 PVP compared to CT: quality of life (QUALEFFO) at 2 weeks

Author; Year	Length of follow-up	PVP Mean (range)	CT Mean (range)	Mean difference (95% CI)	p value
Voormolen 2007[141]	Baseline 2 weeks	60 (37–86) 53 (28–79)	67 (38–86) 67 (40–88)	NR -14 (-24.7, -3.4)	NR < 0.05

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **NR** = not reported, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = Quality of life questionnaire of the European Foundation for Osteoporosis.

PVP vs CT, quality of life (SF-36), 3 to 12 months

One study provided evidence on SF-36 score at 3 and 12 months (***Table 20***).[96] There was no statistically significant difference between PVP and CT groups for physical or mental domains (p > 0.05).

The study used the standard 36 question survey covering 8 domains (physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional and mental health) with higher scores indicating better health and functioning.

Table 20 PVP compared to CT: quality of life (SF-36) at 3 and 12 months

Author; Year	Length of follow-up	PVP Mean (95% CI)	CT Mean (95% CI)	p value
Rousing 2009[96]	*Physical domain* Baseline 3 months 12 months	 36.7 (30.0, 43.4) 34.0 (30.1, 37.9) 32.1 (27.8, 36.3)	 33.4 (26.2, 40.7) 20.3 (24.5, 34.1) 30.5 (25.2, 35.7)	 NR 0.12 0.63
Rousing 2009[96]	*Mental domain* Baseline 3 months 12 months	 49.7 (43.6, 55.8) 48.9 (43.8, 54.0) 46.2 (39.2, 53.2)	 49.6 (41.9, 57.3) 46.2 (39.2, 53.2) 49.0 (43.9, 54.1)	 NR 0.51 0.93

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **NR** = not reported, **PVP** = percutaneous vertebroplasty, **SF-36** = short form 36 questionnaire.

PVP vs CT, Function (SOF-ADL)

No study reported this outcome.

Healthcare resource utilisation

PVP vs CT, length of hospitalisation

One study reported length of hospitalisation following PVP and CT. The mean length of hospital stay was significantly lower following PVP (9.6 days; 95% CI 8.0, 11.2 days) compared to CT (11.7 days; 95% CI 9.1, 14.3 days; p = 0.01).[96]

7.7.2 PVP vs Sham

Table 21 provides a summary of the main pooled clinical effectiveness outcomes comparing PVP to sham. The 1 month and 12 months were selected as representative timepoints for short- and long-term follow-up, respectively. By 1 month, there were statistically significant differences between PVP and sham cohorts for pain (VAS) (light orange shading ***Table 21***). Effect sizes did not surpass identified MCIDs. There were no differences in analgesic use, RDQ and QUALEFFO. At 12 months, statistical but not clinically meaningful differences were observed for pain (VAS) and QUALEFFO but not RDQ or analgesic use. Results for EQ-5D could not be pooled owing to different methods of reporting results (final scores and change from baseline), with studies demonstrating opposing results. For additional information regarding each outcome, refer to the corresponding sections below.

Table 21 Summary of meta-analyses for PVP compared to sham

Outcome	Length of follow-up Mean difference (95% CI); p value					
	1 day	*1 week*	*1 month*	*3 months*	*6 months*	*12 months*
Pain (NRS/VAS)	0.34 (0.09, 0.59) p = 0.01	-1.52 (-4.19, 1.15) p = 0.26	-0.76 (-1.21, -0.31) p = 0.0009	-0.67 (-1.12, -0.22) p = 0.004	-0.69 (-1.14, -0.23) p = 0.003	-0.88 (-1.47, -0.29) p = 0.003
Pain (analgesic use) [a]	1.07 (0.77, 1.50) p = 0.68	1.03 (0.88, 1.20) p = 0.72	0.99 (0.79, 1.24) p = 0.94	0.88 (0.67, 1.17) p = 0.39	0.85 (0.68, 1.07) p = 0.16	1.19 (0.88, 1.62) p = 0.27
Function (RDQ)	0.50 (-0.16, 1.15) p = 0.14	-0.74 (-3.86, 2.39) p = 0.64	-0.28 (-1.70, 1.15) p = 0.70	-1.07 (-1.58, -0.57) p < 0.0001	-0.80 (-2.19, 0.59) p = 0.26	-0.76 (-1.85, 0.32) p = 0.17
Function (ODI)	NR	NR	NR	NR	NR	NR
Quality of life (EQ-5D)	NA	NA	NA	NA	NA	NA
Quality of life (QUALEFFO)	NR	-3.38 (-11.81, 5.05) p = 0.43	-1.39 (-3.24, 0.47) p = 0.14	-0.75 (-1.49, -0.01) p = 0.05	-2.01 (-6.48, 2.46) p = 0.38	-2.15 (-4.08, -0.22) p = 0.03

Abbreviations
CI = confidence interval, **EQ-5D** = EuroQol 5 dimension questionnaire, **NA** = not applicable, **NR** = not reported, **NRS** = numerical rating scale, **ODI** = Oswestry disability index, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RDQ** = Roland-Morris disability questionnaire, **VAS** = visual analogue scale.
Notes
For continuous outcomes, a negative mean difference favours PVP, positive mean difference favours CT. For dichotomous outcomes (analgesic use only), positive risk ratio favours PVP, negative risk ratio favours CT.
a = all types of analgesics were pooled together and the outcome of analgesic use reported as risk ratio (95% CI).
No shading = no statistically significant difference (p > 0.05) or clinically meaningful effect (as inferred by MCIDs) between groups.
Light orange shading = statistically significant differences between groups (p < 0.05), however, the results do not surpass lower bounds of identified MCIDs.

Pain-related outcomes

PVP vs sham, pain (NRS or VAS), 1 day to 24 months

Four studies provided evidence on pain, as measured by NRS or VAS, from 1 day to 24 months post-intervention.[2 4 7 95] All 4 studies were included in the meta-analysis. Overall, there were statistically significant differences between PVP and sham groups at 1 day and 1, 3, 6, 12 and 24 months, but not at 3 days or 1 to 2 weeks post-intervention (see ***Figure 28***). At 1 month the mean difference was -0.76 (95%CI -1.21, -0.31; p = 0.0009) and by 12 months the difference was -0.88 (95%CI -1.47, -0.29; p = 0.003). However, the Tau^2 and I^2 statistics indicated considerable levels of heterogeneity and inconsistency at most timepoints.

Sub-groups

Sub-group analysis of vertebral fractures younger than 8 weeks determined that there were statistically significant differences in pain scores at 1 day, 3 days, 2 weeks, 6 months and 12 months.[2 95] The longitudinal meta-analysis indicated there were no differences at the remaining timepoints. At 1 month the mean difference was -0.90 (95%CI -1.88, 0.07; p = 0.07) and by 12 months the difference was -0.54 (95%CI -0.81,-0.28; p < 0.0001). However, the Tau^2 and I^2 statistics indicated moderate to considerable levels of heterogeneity and inconsistency at all timepoints. For fractures older than 8 weeks there were statistically significant differences at all timepoints post-intervention (except 2 weeks), as indicated by longitudinal meta-analysis.[4 7] At 1 month the mean difference was -0.66 (95%CI -0.91, -0.41; p < 0.0001) and by 12 months the difference was -0.74 (95%CI -1.24, -0.25; p = 0.003). The Tau^2 and I^2 statistics indicated low levels of heterogeneity from 3 days to 6 months and considerable heterogeneity at 6 and 12 months.

The included studies used different methods of assessing pain, although all studies measured pain on a 10-point scale with 0 representing no pain and 10 representing the worst pain. One study measured pain using VAS[95] and three studies used the NRS scale.[2 4 7] Further, 2 studies reported change in pain scores from baseline[2 7] and 2 studies reported the final pain scores.[4 95] For cross-over trials, only patients remaining in their original cohort were included in the analysis. There were statistically significant differences between PVP and sham groups for studies utilising the NRS scale (MD -0.74; 95% CI -1.03, -0.45; p < 0.0001), but not the VAS scale (MD -0.16; 95% CI -0.43, 0.12; p = 0.28). For further information regarding sub-group analysis for NRS and VAS refer to ***Section 17.4*** (***Appendix D***), ***Table 118***.

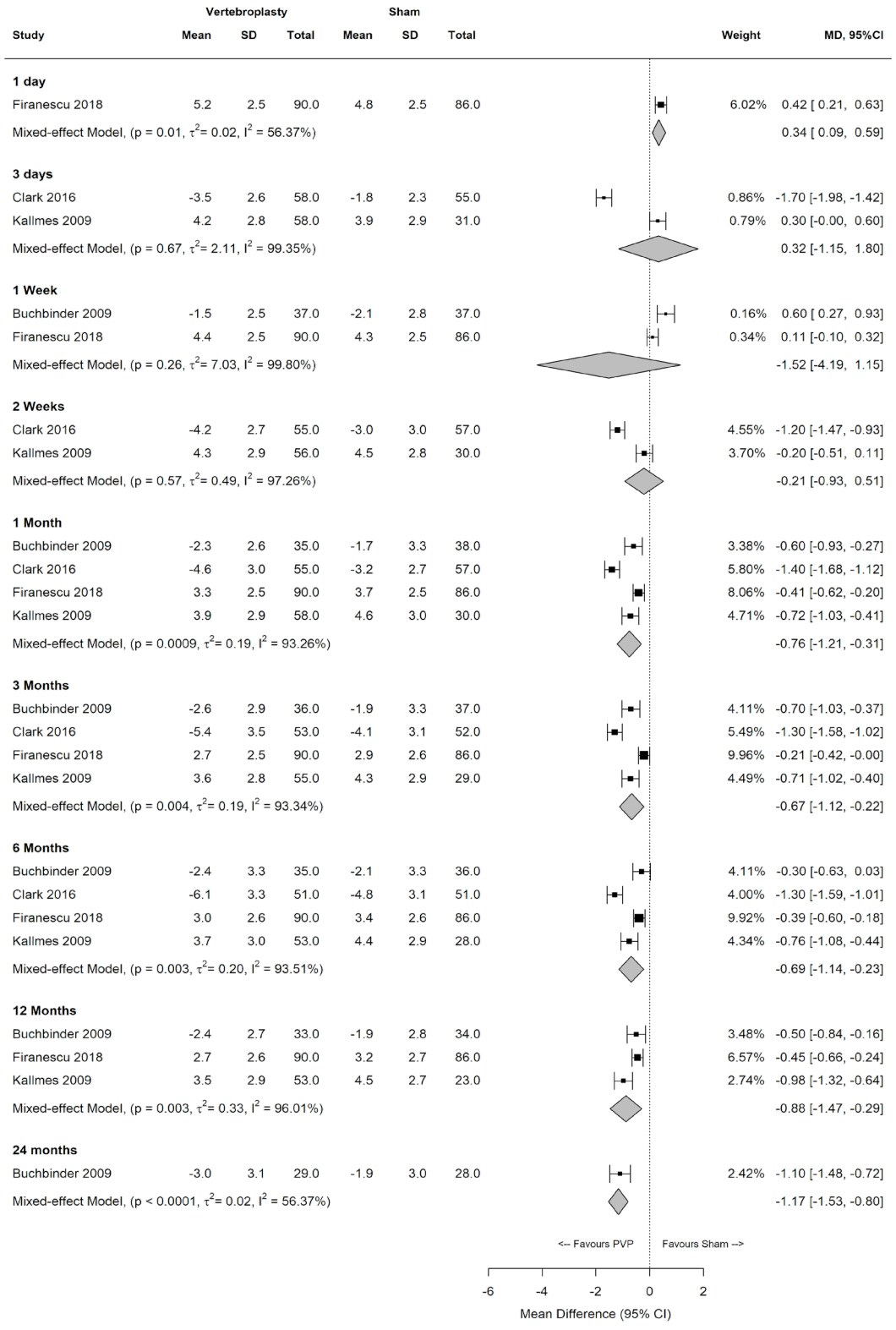


Figure 28 Forest plot indicating mean difference in pain (NRS or VAS) for PVP compared to sham (1 day to 24 months)

Abbreviations

CI = confidence interval, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PVP vs sham, pain (analgesic use), 1 day to 24 months

Four studies provided evidence on the number of patients using analgesics, from 1 day to 12 months post-intervention.[2 4 7 95] Three studies were included in the meta-analysis. There were no statistically significant differences between PVP and CT groups at any timepoint (see ***Table 22***). At 1 month the risk ratio was 0.99 (95% CI 0.79, 1.24; p = 0.94) and by 12 months the ratio was 1.19 (95% CI 0.88, 1.62; p = 0.27). The Chi^2 and I^2 statistics indicated moderate levels of heterogeneity and inconsistency at most timepoints.

Sub-groups

Sub-group analysis of vertebral fractures younger than 8 weeks determined there were no statistically significant differences at any timepoint post-intervention (1 day–12 months).[2 95] At 1 month the risk ratio was 1.09 (95% CI 0.88, 1.35; p =0.72) and at 12 months the ratio was 1.19 (95% CI 0.88, 1.62; p = 0.27). The Chi^2 and I^2 statistics indicated moderate levels of heterogeneity and inconsistency at most timepoints. Vertebral fractures older than 8 weeks were evaluated in 1 study.[4] There were no statistically significant differences in the number of patients using analgesics at 1 month (RR 1.27, 95% CI 0.88, 1.62; p = 0.19). For further information regarding the sub-groups for the number of patients utilising analgesics refer to ***Section 17.4*** (***Appendix D***), ***Table 119***.

The type of analgesic differed across the studies. Two studies reported opioid analgesics,[4 7] 1 study further differentiated between non-opioid, weak opioid and strong opioid analgesia,[95] and 1 did not specify the type of analgesic.[2] One study noted analgesic consumption in the previous 24 hours. The other studies did not report the time over which analgesics were used.

Table 22 PVP compared to sham: pain (analgesic use) 1 day to 12 months

Analgesic use	Number of studies	Heterogeneity	PVP n/N (%)	Sham n/N (%)	Risk Ratio (RR, 95% CI)
1 day	2	Chi^2 = 11.72 p <0.0006 I^2 = 91%	127/148 (85.8%)	113/143 (79%)	1.07 (0.77, 1.50) p = 0.68
1 week	2	Chi^2 = 2.51 p =0.09 I^2 = 60%	123/144 (85.4%)	116/142 (81.7%)	1.03 (0.88, 1.20) p = 0.72
1 month	3	Chi^2 = 2.60 p =0.11 I^2 = 61%	130/209 (62.2%)	128/205 (62.4%)	0.99 (0.79, 1.24) p = 0.94
3 months	2	Chi^2 = 2.60 p =0.11 I^2 = 61%	85/138 (61.6%)	91/133 (68.4%)	0.88 (0.67, 1.17) p = 0.39
6 months	2	Chi^2 = 1.30 p =0.25 I^2 = 23%	72/133 (54.1%)	84/134 (62.7%)	0.85 (0.68, 1.07) p = 0.16

Analgesic use	Number of studies	Heterogeneity	PVP n/N (%)	Sham n/N (%)	Risk Ratio (RR, 95% CI)
12 months	1	NA	44/79 (55.7%)	37/79 (46.8%)	1.19 (0.88, 1.62) p = 0.27

Abbreviations
CI = confidence interval, **n** = number of patients using analgesics, **N** = total number of patients, **NA** = not applicable, **RR** = risk ratio.

Function-related outcomes

PVP vs sham, function (RDQ), 1 day to 24 months

Three studies provided evidence on function, as measured by RDQ, from 1 day to 24 months post-intervention.[4 7 95] All three studies were included in the meta-analysis. Overall, there were statistically significant differences between PVP and sham groups at 3 months (MD -1.07; 95% CI -1.58, -0.59; p < 0.001) but not at any remaining timepoints (see ***Figure 29***). At 1 month the mean difference was -0.28 (95% CI -1.70, 1.1; p = 0.70) and by 12 months the difference was -0.76 (95% CI -1.85, 0.32; p = 0.17). However, the Tau^2 and I^2 statistics indicated considerable levels of heterogeneity and inconsistency at most timepoints.

Sub-groups

Vertebral fractures younger than 8 weeks were evaluated in 1 study.[95] There were no statistically significant differences in RDQ scores between PVP and sham groups any timepoint post-intervention (1 week to 12 months). At 1 month the mean difference was -1.12 (95%CI -2.96, 0.72). At 12 months the difference was -0.01 (95%CI -1.94, 1.92). For fractures older than 8 weeks there were statistically significant differences at 1 week and 3 months, as indicated by longitudinal meta-analysis.[4 7] There were no differences at any remaining timepoints. At 1 month the mean difference was 0.15 (95%CI -1.73, 2.03; p = 0.88) and at 12 months the difference was -1.15 (95%CI -2.84, 0.54; p = 0.18). However, all timepoints exhibited considerable heterogeneity and inconsistency as inferred by the Tau^2 and I^2 statistic. For further information regarding sub-group analysis for RDQ refer to ***Section 17.4*** (***Appendix D***), ***Table 120***.

Two studies utilised the modified 0–23-point RDQ scale[4 7] and 1 study used the 0–24 RDQ scale.[95] For both scales higher scores indicated decreasing physical functioning and increasing physical impairment. While the studies differed in the scales used, it was unlikely to significantly impact the overall results.

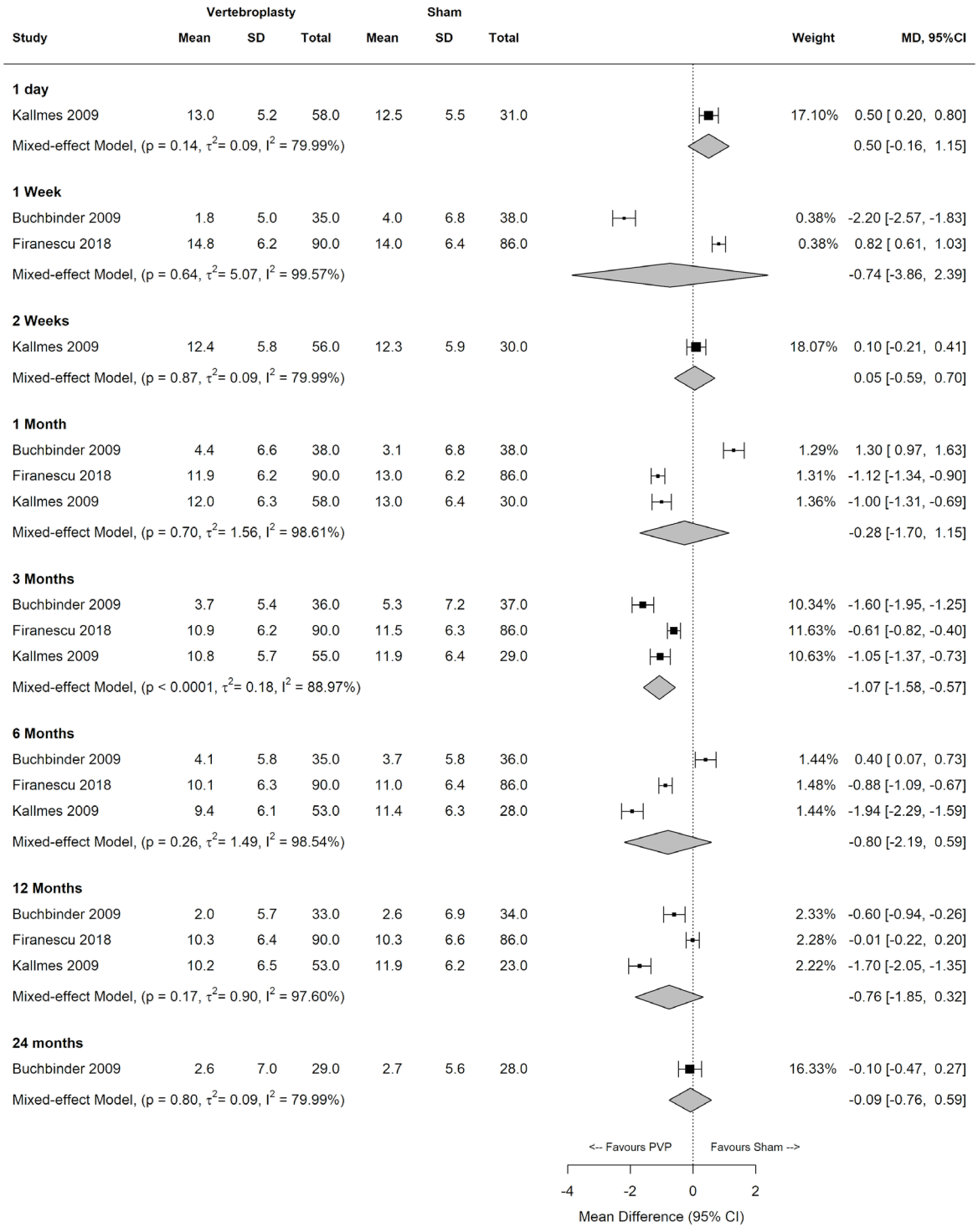


Figure 29 Forest plot indicating mean difference in RDQ for PVP compared to sham (1 day to 24 months)

Abbreviations

CI = confidence interval, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **RDQ** = Roland-Morris disability questionnaire, **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PVP vs sham, function (ODI)

No study reported this outcome.

PVP vs sham, function (timed up-and-go), 12 to 24 months

One study provided evidence on timed up-and-go scores at 12 and 24 months (***Table 23***).[7] The statistical difference was not reported by the study authors so it was unclear whether PVP and sham groups differ.

Table 23 PVP compared to CT: function (timed up-and-go) at 12 and 24 months

Author; Year	Length of follow-up	PVP Mean ± SD	Sham Mean ± SD	p value
Buchbinder 2009[7]	Baseline 12 months 24 months	20.5 ± 8.8 seconds -2.6 ± 12.2 seconds 3.5 ± 17.1 seconds	29.0 ± 15.0 seconds 4.3 ± 13.4 seconds 4.7 ± 9.7 seconds	NR NR NR

Abbreviations
NR = not reported, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation.

Quality of life-related outcomes

PVP vs sham, quality of life (EQ-5D), 1 day to 24 months

Three studies provided evidence on EQ-5D scores from 1 day to 24 months post-intervention (***Table 24***).[2 4 7] A longitudinal meta-analysis was not performed as the correlation coefficient for Buchbinder (2009) could not be calculated and Kallmes (2009) only reported EQ-5D at 1 timepoint. Therefore, the results were described narratively.

The studies reported conflicting results. One study found statistically significant differences between PVP and sham groups at 1 month and 6 months.[2] The other studies reported no statistically significant differences at any timepoint.[4 7] The studies differed in how results were presented (final scores[2 4] or change from baseline[7]) and the mean age of fractures (less[2] than or greater than 8 weeks.[4 7]

All studies used the EQ-5D scale ranging from 0 to 1, where 1 indicates perfect health and 0 indicates death.

Table 24 PVP compared to sham: quality of life (EQ-5D) questionnaire at 1 day to 12 months

Author; Year	Length of follow-up	PVP Mean ± SD	Sham Mean ± SD	p value
Clark 2016[2]	Baseline 3 days 2 weeks 1 month 3 months 6 months	0.60 ± 0.07 0.69 ± 0.11 0.69 ± 0.10 0.75 ± 0.11 0.75 ± 0.12 0.8 ± 0.11	0.59 ± 0.06 0.65 ± 0.09 0.68 ± 0.11 0.70 ± 0.11 0.71 ± 0.11 0.74 ± 0.12	NR 0.091 0.47 0.04 0.16 0.01
Buchbinder 2009[7]	1 day 1 month 3 months 6 months 12 months 24 months	0.1 ± 0.3 0.1 ± 0.3 0.2 ± 0.3 0.2 ± 0.4 0.2 ± 0.4 0.2 ± 0.4	0.1 ± 0.3 0.1 ± 0.3 0.2 ± 0.4 0.2 ± 0.4 0.2 ± 0.4 0.2 ± 0.4	NS NS NS NS NS NS
Kallmes 2009[4]	Baseline 1 month	0.57±0.18 0.70 ± 0.18	0.54±0.23 0.64 ± 0.20	NR 0.13

Abbreviations
EQ-5D = EuroQol 5 dimensions, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation, **NR** = not reported, **NS** = not significant.

PVP vs sham, quality of life (QUALEFFO), 1 day to 24 months

Three studies provided evidence on quality of life, as measured by QUALEFFO, from 1 day to 24 months post-intervention.[2 7 95] All three studies were included in the longitudinal meta-analysis. Overall, there were statistically significant differences between PVP and sham groups at 2 weeks and 3, 12 and 24 months, but not at remaining timepoints (see ***Figure 29***). At 1 month the mean difference was -1.39 (95% CI -3.24, 0.47; p = 0.14) and at 12 months the difference was -2.15 (95% CI -4.08, -0.22; p = 0.03). However, the Tau^2 and I^2 statistics indicated considerable levels of heterogeneity and inconsistency for all timepoints.

Sub-groups

Sub-group analysis of vertebral fractures younger than 8 weeks found statistically significant differences in QUALEFFO scores between PVP and sham groups at 2 weeks and at 1, 3 and 12 months.[2 95] Longitudinal meta-analysis indicated no differences at any other timepoint. At 1 month the mean difference was -2.27 (95% CI -3.71, -0.84; p = 0.002) and at 12 months the difference was -1.31 (95% CI -2.57, -0.06; p = 0.04). All timepoints exhibited considerable levels of heterogeneity and inconsistency as inferred by the Tau^2 and I^2 statistics. Vertebral fractures older than 8 weeks were evaluated in 1 study.[7] There were no statistically significant differences in QUALEFFO scores at any timepoint post-intervention (1 week to 24 months). At 1 month the mean difference was 0.40 (95% CI -4.50, 5.30) and

at 12 months the difference was -2.10 (95% CI -8.21, 4.01). For further information regarding sub-group analysis for QUALEFFO refer to ***Section 17.4*** (***Appendix D***), ***Table 121***.

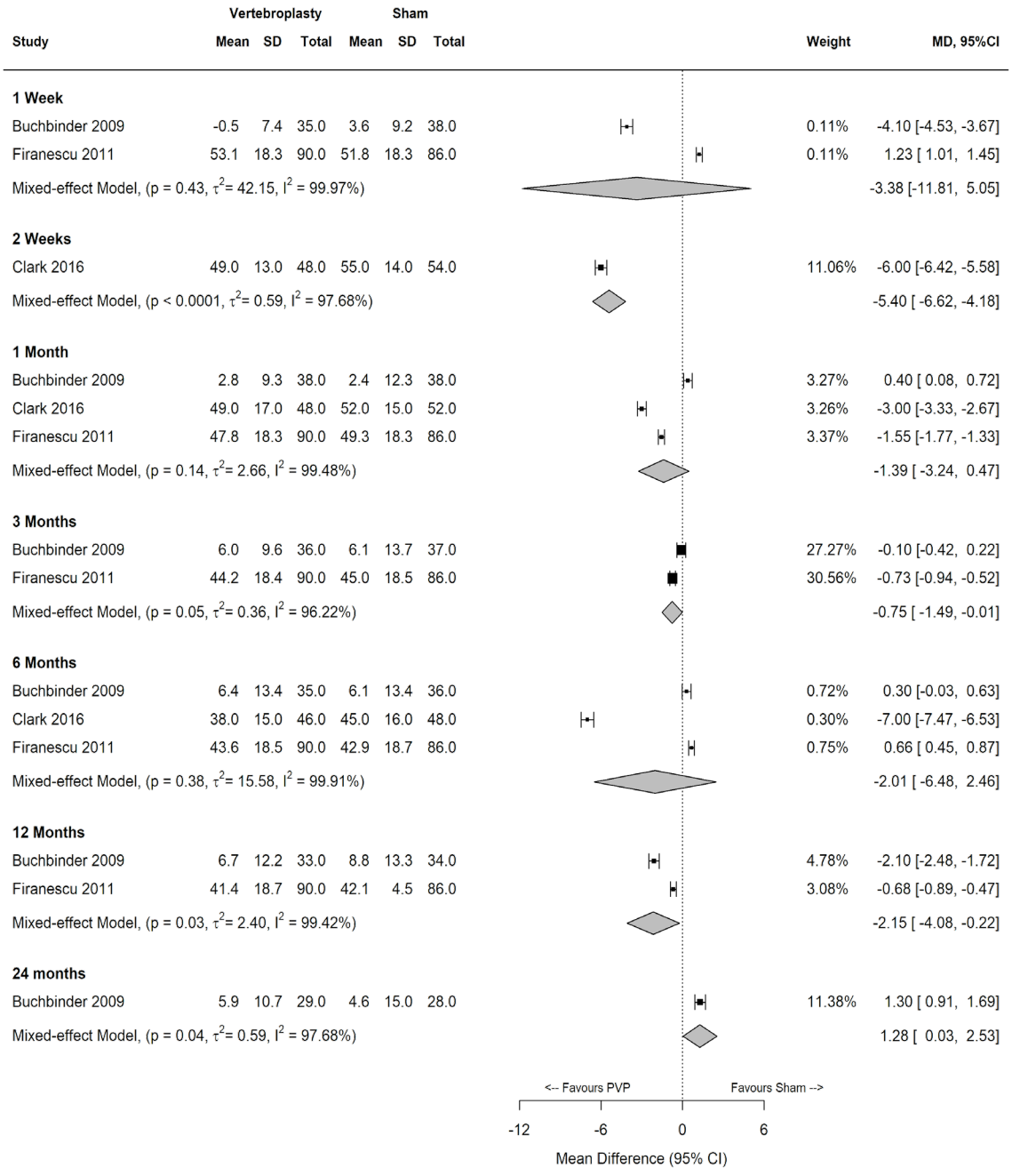


Figure 30 Forest plot indicating mean difference in QUALEFFO for PVP compared to sham (1 week to 12 months)

Abbreviations

CI = confidence interval, **MD** = mean difference, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = questionnaire of the European Foundation for Osteoporosis **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PVP vs sham, quality of life (SF-36), 1 month

One study provided evidence on the physical domain of the SF-36 score at 1 month (***Table 25***).[4] The data demonstrated no statistically significant difference between PVP and sham groups (p > 0.05).

Table 25 PVP compared to sham: quality of life (SF-36) at 1 month

Author; Year	Length of follow-up	PVP Mean ± SD	Sham Mean ± SD	p value
Kallmes 2009[4]	Baseline 1 month	25.3 ± 7.8 29.7 ± 9.6	25.3 ± 7.3 28.7 ± 8.0	NR 0.45

Abbreviations
NR = not reported, **PVP** = percutaneous vertebroplasty; **SD** = standard deviation, **SF-36** = short form 36 questionnaire.

PVP vs sham, SOF-ADL, 1 month

One study provided evidence on SOF-ADL score at 1 month (***Table 26***).[4] The data demonstrated no statistically significant difference between PVP and CT groups (p > 0.05).

Table 26 PVP compared to sham: function (SOF-ADL) at 1 month

Author; Year	Length of follow-up	PVP Mean ± SD	Sham Mean ± SD	p value
Kallmes 2009[4]	Baseline 1 month	10.0 ± 3.6 7.7 ± 3.7	10.3 ± 2.8 8.2 ± 3.6	NR 0.51

Abbreviations
NR = not reported, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation, **SOF-ADL** = study of osteoporotic fractures–activities of daily living questionnaire.

Healthcare resource utilisation

PVP vs sham, length of hospitalisation

One study reported the length of hospitalisation following PVP and sham. The median length of hospital stay was 8.5 days (IQR 4.0–13.0 days) for PVP and 14.0 days (IQR 7.0–22.0 days) for sham (p = NR). Noting only 56% and 58% of patients were treated as inpatients, respectively.[2]

7.7.3 PBK vs CT

Table 27 provides a summary of the meta-analysis for the clinical effectiveness outcome of pain. In the short-term (1 day to 1 week post-intervention), there was a clinically and statistically meaningful difference between PBK and CT for pain (VAS). The statistical difference persists for the remaining timepoints, however, the effects were unlikely to translate into a clinically meaningful difference beyond 3 months. The remaining outcomes were not pooled owing to the limited number of trials. These were described narratively in the following sections. Results from non-RCTs were also presented to supplement the lack of RCTs. The results from the non-RCTs generally echoed the RCT findings.

Table 27 Summary of the meta-analyses for PBK compared to CT

Outcome	Length of follow-up Mean difference (95% CI); p value					
	1 day	*1 week*	*1 month*	*3 months*	*6 months*	*12 months*
Pain (VAS)	-5.59 (-6.39, -4.78) p < 0.0001	-3.63 (-5.59, -1.68) p = 0.001	-0.18 (-2.15, 1.80) p = 0.86	-1.78 (-2.85, -0.71) p = 0.01	-0.48 (-1.24, 0.27) p = 0.21	-1.27 (-2.04, -0.51) p = 0.01

Abbreviations
CI = confidence interval, **VAS** = visual analogue scale.
Notes
For continuous outcomes, negative mean difference favours PVP, positive mean difference favours CT.
Light orange shading = statistically significant difference between groups (p < 0.05), however, the results do not surpass lower bounds of identified MCIDs.
Dark green shading = statistically significant difference between groups (p < 0.05) and likely to translate to a clinically meaningful effect (surpasses lower bounds of identified MCIDs).

Pain-related outcomes

PBK vs CT, pain (VAS), 1 day to 24 months, RCTs

Three studies provided evidence on pain, as measured by VAS, from 1 day to 24 months post-intervention.[42 60 168] All three studies were included in the meta-analysis. Overall, there were statistically significant differences between PBK and CT groups at 1 and 3 days, 1 week and at 3, 12 and 24 months but not at remaining timepoints (see ***Figure 31***). At 1 month the mean difference was -0.18 (95% CI -2.15, 1.80; p = 0.86) and at 12 months the difference was -1.27 (95% CI -2.04, -0.51; p < 0.01); however, the Tau^2 and I^2 statistics indicated moderate to considerable levels of heterogeneity and inconsistency at all timepoints.

The included studies utilised different methods for assessing pain. Two studies measured pain using VAS[42 60] and 1 did not specify the scale.[168] Irrespective, all studies reported the final score and measured pain on a 10-point scale with 10 representing the worst pain.

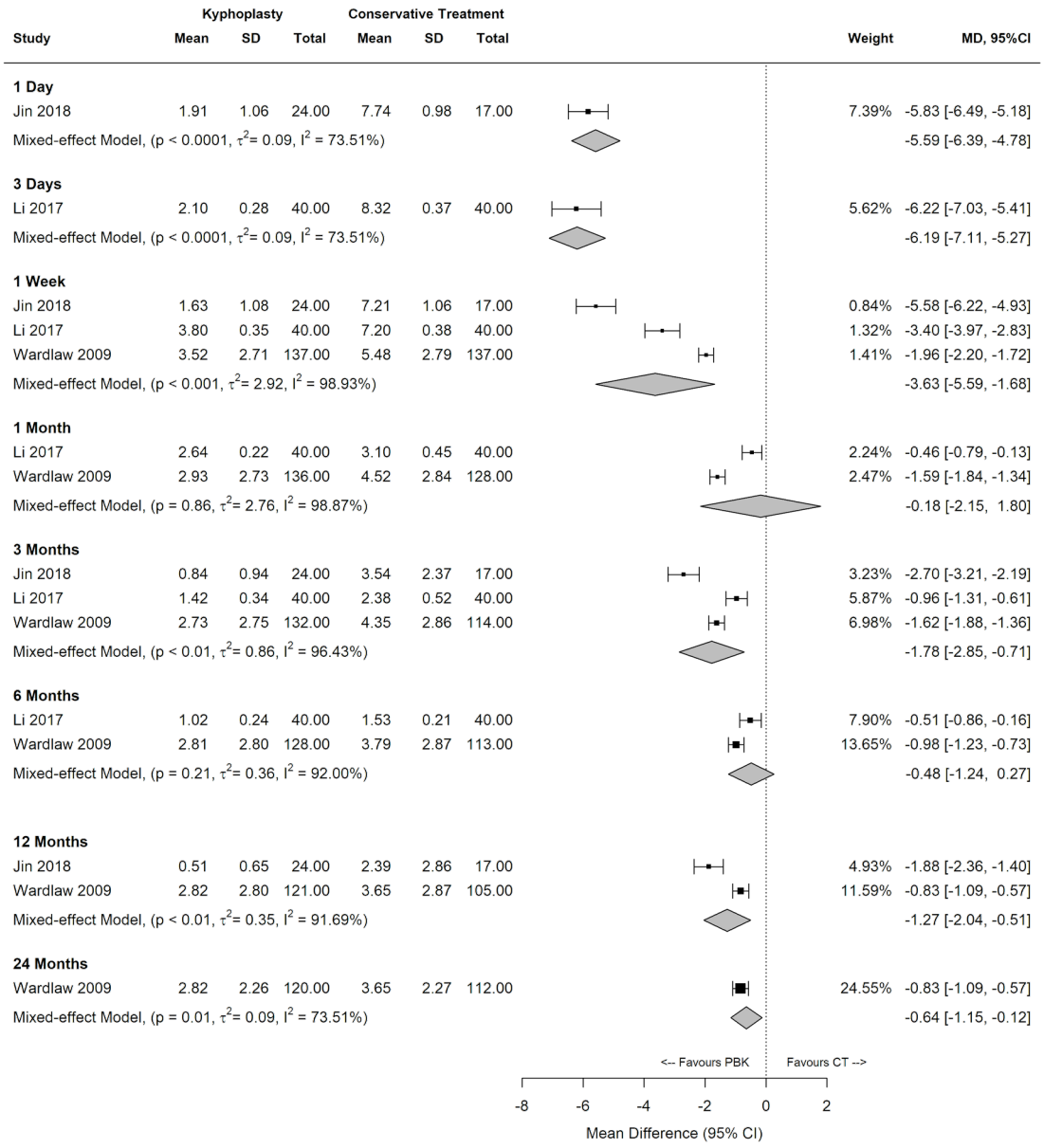


Figure 31 Forest plot indicating mean difference in pain (VAS) for PBK compared to CT (1 day to 12 months)

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **MD** = mean difference, **PBK** = percutaneous balloon kyphoplasty, **SD** = standard deviation.

Notes

Horizontal and vertical bars around the estimate (black square) depict the bounds of the confidence intervals.

PBK vs CT, pain (VAS), 3 to 12 months, non-RCTs

Two studies provided evidence on pain, as measured by VAS, from 3 to 12 months post-intervention (***Table 28***).[173 174] Overall, both studies concluded there were statistically significant differences between PBK and CT groups at 12 months (p = 0.008; p < 0.0001). Kasperk (2005) used an inverted VAS scale with a score of 0 indicating maximal pain.[173] A score of 10 corresponded to maximal pain in Movrin (2010), and baseline differences in pain score between the 2 groups confound the effect.[174]

Table 28 PBK compared to CT: pain (VAS) at 3 to 12 months (non-RCTs)

Author; Year	Length of follow-up	PBK Mean ± SD	CT Mean ± SD	p value
Kasperk 2005[173]	Baseline 3 months 6 months 12 months	26.2 ± 12.6 42.4 ± 17.9 44.2 ± 20.9 44.4 ± 19.7	33.6 ± 18.3 33.9 ± 18.4 35.6 ± 18.3 34.3 ± 19.5	NR 0.012 0.019 0.008
Movrin 2010[174]	Baseline 12 months	8.8 ± 8.1 2.0 ± 1.2	6.7 ± 7.8 3.8 ± 1.5	< 0.001 < 0.001

Abbreviations

CT = conservative treatment, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials, **SD** = standard deviation, **VAS** = visual analogue scale.

Notes

In Kasperk (2005), a lower score corresponded to more pain. In Movrin (2010), a higher score corresponded to more pain.

PBK vs CT, pain (analgesic use), 1 to 12 months, RCTs

One study provided evidence on pain as assessed by use of analgesics from 1 to 12 months (***Table 29***).[168] Overall, the number of patients taking any analgesic or a combination of analgesics (non-opioid and opioid) was smaller in the PBK group compared to the CT group. The number of patients using non-opioid and strong-opioid analgesics did not change considerably throughout the follow-up period. However, the statistical significance was not reported, thereby limiting the conclusions of the study.

Table 29 PBK compared to CT: pain (analgesic use) at 1 to 12 months (RCTs)

Author; Year	Length of follow-up	Type of analgesic	PBK n/N	CT n/N	p value
Wardlaw 2009[168]	Baseline	Any analgesic Non-opioid Combination Strong opioid	132/140 (94%) 29/140 (21%) 81/140 (58%) 22/140 (16%)	135/146 (92%) 36/146 (25%) 82/146 (56%) 17/146 (12%)	NR
Wardlaw 2009[168]	1 month	Any analgesic Non-opioid Combination Strong opioid	81/144 (71%) 28/114 (25%) 47/114 (41%) 6/114 (5%)	105/115 (91%) 31/115 (27%) 65/115 (57%) 9/115 (8%)	NR

Author; Year	Length of follow-up	Type of analgesic	PBK n/N	CT n/N	p value
Wardlaw 2009[168]	12 months	Any analgesic Non-opioid Combination Strong opioid	61/117 (52%) 28/117 (24%) 28/117 (24%) 5/117 (4%)	69/101 (68%) 32/101 (32%) 35/101 (35%) 5/101 (5%)	NR

Abbreviations
CT = conservative treatment, **n** = number of patients experiencing event, **N** = total number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.
Notes
Combination = non-opioid and opioid analgesics.

PBK vs CT, pain (analgesic use), duration not reported, non-RCTs

One study provided evidence on pain as assessed by the use of analgesics (***Table 30***).[173] Patients in the PBK group reduced opioid use whereas patients in the CT group did not, however, statistical significance and follow-up time were not reported.

Table 30 PBK compared to CT: pain (analgesic use) (non-RCTs)

Author; Year	Length of follow-up	PBK n/N	CT n/N	p value
Kasperk 2005[173]	Baseline NR	27/40 (67.0%) 22/40 (55.0%)	14/20 (70%) 13/20 (65%)	NR NR

Abbreviations
CT = conservative treatment, **n** = number of patients experiencing event, **N** = total number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

Function-related outcomes

PBK vs CT, function (ODI), 3 days to 3 months, RCTs

One study provided evidence on function, as measured by ODI, from 3 days to 3 months post-intervention (***Table 31***).[60] The data demonstrated statistically significant differences between PBK and CT groups at all timepoints post-intervention ($p < 0.05$).

Table 31 PBK compared to CT: function (ODI) at 1 week to 3 months (RCTs)

Author; Year	Length of follow-up	PBK Mean ± SD	CT Mean ± SD	p value
Li 2017[60]	Baseline 3 days 1 week 1 month 3 months	42.3 ± 6.7 20.2 ± 5.4 18.5 ± 4.3 15.1 ± 3.6 14.2 ± 4.2	41.3 ± 6.2 36.5 ± 5.1 19.7 ± 3.4 18.7 ± 5.3 18.2 ± 5.0	NS < 0.05 < 0.05 < 0.05 < 0.05

Abbreviations

CT = conservative treatment, **ODI** = Oswestry disability index, **PBK** = percutaneous balloon kyphoplasty, **NS** = not significant, **SD** = standard deviation.

PBK vs CT, function (RDQ), 1 week to 24 months, RCTs

One study provided evidence on function, as measured by RDQ, from 1 week to 24 months post-intervention (***Table 32***).[168] The reported data demonstrated statistically significant differences between PBK and CT groups from 1 month ($p < 0.0001$) to 12 months ($p < 0.001$) but not at 24 months ($p = 0.06$).

Table 32 PBK compared to CT: function (RDQ) at 1 to 24 months (RCTs)

Author; Year	Length of follow-up	PBK Mean ± SD	CT Mean ± SD	p value
Wardlaw 2009[168]	Baseline	16.9 ± 5.1	17 ± 4.3	NS
	1 week	16.9 ± 4.2	17.0 ± 4.3	NR
	1 month	10.9 ± 4.3	15.1 ± 4.3	< 0.0001
	3 months	9.2 ± 4.4	12.9 ± 4.4	< 0.0001
	6 months	8.5 ± 4.4	11.5 ± 4.5	< 0.0001
	12 months	8.6 ± 4.5	11.5 ± 4.5	< 0.001
	24 months	8.9 ± 4.5	10.3 ± 4.5	0.06

Abbreviations

CT = conservative treatment, **NR** = not reported, **NS** = not significant, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **SD** = standard deviation.

PBK vs CT, function (RDQ), 3 to 12 months, non-RCTs

One study provided evidence on function, as measured by RDQ, from 3 to 12 months post-intervention (***Table 33***).[169] The data demonstrated statistically significant differences between PBK and CT groups at all timepoints post-intervention, however, variance with each measure was not reported.

Table 33 PBK compared to CT: function (RDQ) at 3 to 12 months (non-RCTs)

Author; Year	Length of follow-up	PBK Mean ± SD	CT Mean ± SD	p value
Edit-Koch 2011[169]	Baseline	15.2 ± NR	14.4 ± NR	0.31
	3 months	10.3 ± NR	14.4 ± NR	0.004
	6 months	8.8 ± NR	14.4 ± NR	0.000
	12 months	8.9 ± NR	13.7 ± NR	0.001

Abbreviations

CT = conservative treatment, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **SD** = standard deviation

PBK vs CT, function (timed up-and-go), 1 week to 24 months, RCTs

One study provided evidence on function, as measured by timed up-and-go score, from 1 week to 24 months post-intervention (***Table 34***).[168] The data demonstrated statistically significant differences between PBK and CT groups at 3 months (p = 0.0006) to 6 months (p = 0.05) post-intervention, but not at any other timepoints.

Table 34 PBK compared to CT: function (timed up-and-go) at 1 week to 24 months (RCTs)

Author; Year	Length of follow-up	PVP Mean ± SD	CT Mean ± SD	p value
Wardlaw 2009[168]	Baseline	19.2 ± 6.6	21.6 ± 6.8	NS
	1 week	14.9 ± 6.6	18.8 ± 6.9	NR
	1 month	12.7 ± 6.7	18.7 ± 6.9	0.09
	3 months	12.7 ± 6.7	16.4 ± 6.9	0.0006
	6 months	13.5 ± 6.7	16.0 ± 6.9	0.05
	12 months	13.8 ± 6.8	16.9 ± 7.0	0.3
	24 months	13.8 ± 13.4	16.9 ± 13.5	0.1

Abbreviations
CT = conservative treatment, **NR** = not reported, **NS** = not significant, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **SD** = standard deviation.

Quality of life-related outcomes

PBK vs CT, quality of life (EQ-5D), 1 month to 24 months, RCTs

One study provided evidence on function, as measured by EQ-5D, from 1 month to 24 months post-intervention (***Table 35***).[168] The data demonstrated statistically significant differences between PBK and CT groups from 1 to 24 months. Statistical differences at 1 week were not reported in the study.

Table 35 PBK compared to CT: quality of life (EQ-5D) at 1 month to 24 months (RCTs)

Author; Year	Length of follow-up	PBK Mean ± SD	CT Mean ± SD	p value
Wardlaw 2009[168]	Baseline	0.16 ± 1.03	0.17 ± 0.99	NS
	1 month	0.54 ± 1.03	0.37 ± 1.04	< 0.0001
	3 months	0.59 ± 1.07	0.49 ± 1.04	0.002
	6 months	0.63 ± 1.03	0.50 ± 1.04	0.0009
	12 months	0.61 ± 1.03	0.51 ± 1.09	0.006
	24 months	0.61 ± 0.30	0.53 ± 0.32	0.04

Abbreviations
CT = conservative treatment, **EQ-5D** = EuroQol 5 dimensions questionnaire, **NR** = not reported, **NS** = not significant, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **SD** = standard deviation.

PBK vs CT, quality of life (EQ-5D), 3 to 12 months, non-RCTs

One study provided evidence on function, as measured by EQ-5D (VAS), from 3 to 12 months post-intervention (***Table 36***).[169] The data demonstrated statistically significant differences between PBK and CT groups at all timepoints post-intervention, however, variance with each measure was not reported.

Table 36 PBK compared to CT: quality of life (EQ-5D) at 3 to 12 months (non-RCTs)

Author; Year	Length of follow-up	PBK Mean ± SD	CT Mean ± SD	p value
Edit-Koch 2011[169]	Baseline 3 months 6 months 12 months	0.39 ± NR 0.69 ± NR 0.71 ± NR 0.72 ± NR	0.45 ± NR 0.57 ± NR 0.52 ± NR 0.60 ± NR	0.13 0.01 0.000 0.004

Abbreviations
CT = conservative treatment, **EQ-5D** = EuroQol 5 dimensions questionnaire, **NR** = not reported, **NS** = not significant, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **SD** = standard deviation.

PBK vs CT, quality of life (QUALEFFO), RCTs

No study reported this outcome.

PBK vs CT, quality of life (SF-36), 1 month to 24 months, RCTs

Two studies provided evidence on quality of life, as measured by the physical domain of the SF-36 questionnaire, from 1 month to 24 months post-intervention (***Table 37***).[42 168] Jin (2018) concluded that there were statistically significant differences between PBK and CT groups at 12 months ($p = 0.02$). Wardlaw (2009) concluded there were statistically significant differences between PBK and CT from 1 month ($p < 0.0001$) to 6 months ($p = 0.001$), but not at later time points (12 and 24 months [$p = 0.1$ for both]).

Table 37 PBK compared to CT: quality of life (SF-36) at 1 week to 24 months (RCTs)

Author; Year	Length of follow-up	PVP Mean ± SD	CT Mean ± SD	p value
Jin 2018[42]	12 months	78.1 ± 11.5	64.5 ± 20.3	0.02
Wardlaw 2009[168]	Baseline 1 month 3 months 6 months 12 months 24 months	26.0 ± 5.5 33.4 ± 5.6 35.6 ± 5.6 36.4 ± 5.6 35.9 ± 5.6 35.8 ± 5.6	25.5 ± 5.0 27.5 ± 5.6 31.1 ± 5.8 32.6 ± 5.7 33.8 ± 5.8 33.8 ± 5.8	NS < 0.0001 < 0.0001 0.001 0.1 0.1

Abbreviations
CT = conservative treatment, **PBK** = percutaneous balloon kyphoplasty, **NS** = not significant, **RCTs** = randomised controlled trials, **SD** = standard deviation, **SF-36** = short-form 36 questionnaire.

PBK vs CT, function (SOF-ADL), RCTs

No study reported this outcome.

Healthcare resource utilisation

PBK vs CT, length of hospitalisation, RCTs

One study reported the length of hospitalisation following PBK. In Wardlaw (2009) the median length of stay was 4.0 days (IQR 2.0–9.0 days) following PBK. Length of stay following CT was not reported.[168]

PBK vs CT, length of hospitalisation, non-RCTs

One study reported the length of hospitalisation following PBK. Edit-Koch (2011) noted the mean length of stay was 10.5 days in the PBK group and 17.9 days in the CT group (p = 0.006).[169]

PBK vs CT, number of doctor visits, non-RCTs

One trial reported the number of doctor visits following PBK and CT.[173] The data demonstrated statistically significant differences between PBK and CT groups at 12 months (p = 0.006) but not at 6 months (p = 0.17) (***Table 38***)

Table 38 PBK compared to CT: total doctor visits (non-RCTs)

Author; Year	Length of follow-up	PBK Mean ± SE	CT Mean ± SE	p value
Kasperk 2005[171 173]	6 months	7.8 ± 0.8	10.9 ± 2.7	0.17
	12 months	5.3 ± 0.7	11.6 ± 2.7	0.006

Abbreviations
CT = conservative treatment, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **SE** = standard error.

7.8 Results: Safety

7.8.1 PVP vs CT

Aside from exposure to local anaesthetic and vertebroplasty needles, patients in sham and CT arms received similar interventions (analgesics, bed rest and braces as needed). Therefore, these patients were pooled for the evaluation of safety-related outcomes. In the following section (***Section 7.8.1***), CT encapsulates both conservative treatment and sham arms.

Table 39 provides a summary of the absolute event rates for safety outcomes comparing PVP to CT. In the RCTs there were no statistically significant differences between PVP and CT for any safety outcomes. In the non-RCTs there was no difference in all-cause mortality or radiographic fracture between PVP and CT; the remaining outcomes were reported without statistical information. The rates of safety outcomes varied substantially between the included trials and likely reflected the small sample size in the non-RCTs.

Table 39 Summary of safety outcomes for PVP compared to CT

Outcomes	RCTs % (n/N)	Non-RCTs % (n/N)	Single-arm trials % (n/N)
All-cause mortality	PVP 4.8% (n = 31/648) CT 5.8% (n = 37/641)	PVP 17.0% (n = 15/88) CT 15.8% (n = 6/38)	NA
Severe adverse events	PVP 1.9% (n = 5/268) CT 1.9% (n = 5/263)	PVP 0.0% (n = 0/88) CT 0.0% (n = 0/38)	NA
Any adverse events	PVP 6.6% (n = 28/424) CT 5.5% (n = 24/438)	PVP 2.5% (n = 3/118) CT 0.0% (n = 0/68)	NA
Symptomatic fractures [a]	PVP 11.5% (n = 48/418) CT 7.4% (n = 31/422)	PVP 3.4% (n = 3/88) CT NR	NA
Radiographic fractures [a]	PVP 27.2% (n = 106/389) CT 23.8% (n = 88/369)	PVP 10.2% (n = 9/88) CT 10.5% (n = 4/38)	NA
Cement leakage [b]	PVP 55.1% (n = 343/623) [c]	NR	PVP 38.7% (n = 1,109/2,863)[d]

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NA** = not applicable, **NR** = not reported, **RCTs** = randomised controlled trials, **PVP** = percutaneous vertebroplasty, **RCTs** = randomised controlled trials.
Notes
a = fractures per patient
b = leaks per treated vertebrae reported, the leaks per patient can be found in ***Table 44*** and ***Table 45***.
c = 1 symptomatic leak, the remaining leaks were asymptomatic.
d = 4 symptomatic leaks, the remaining leaks were asymptomatic.

PVP vs CT, all-cause mortality, RCTs

Nine RCTs reported all-cause mortality and were included in the meta-analysis (***Figure 32***).[2 4 7 8 95 96 132 135 139] Overall, there was no statistically significant difference between PVP and CT (RR 0.84; 95% CI 0.52, 1.35; p = 0.47). The absolute risk for the PVP group was 4.8% (n = 31/648) and for the CT group 5.8% (n = 37/654). The Chi^2 test and I^2 statistic indicated low levels of heterogeneity and inconsistency (p = 0.89 and I^2 = 0%, respectively). All deaths were deemed unrelated to the intervention.

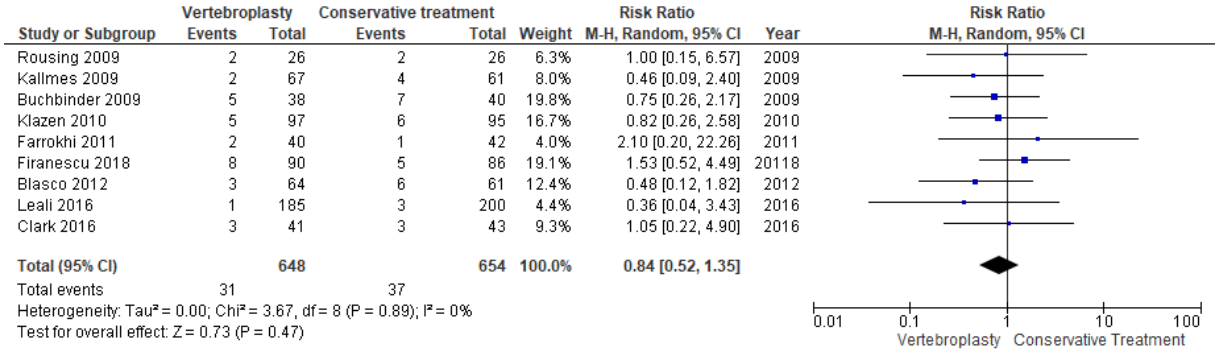


Figure 32 Forest plot indicating risk ratio of all-cause mortality for PVP compared to CT

PVP vs CT, all-cause mortality, non-RCTs

One non-RCT study provided evidence on all-cause mortality at 24 months (***Table 40***).[149] Overall, the incidence of mortality was similar between the PVP and CT arms (p = 0.89), however, incidence of fracture-related deaths differed between the 2 arms (p = 0.05). One fracture-related death was reported in the PVP arm and 4 fracture-related deaths in the CT arm. The remaining deaths were unrelated to the intervention.

Table 40 PVP compared to CT: all-cause and fracture-related mortality (non-RCTs)

Author; Year	Length of Follow-up	Mortality	PVP n/N (%)	CT n/N (%)	Hazard ratio (95% CI); p value
Diamond 2006[149]	24 months	All-cause	15/88 (17.0%)	6/38 (15.8%)	1.07 (0.42, 2.76) p = 0.89
Diamond 2006[149]	24 months	Fracture-related	1/15 (6.7%)	4/6 (66.7%)	0.11 (0.01, 0.96) p = 0.05

Abbreviations
CT = conservative treatment, **n** = number of patients with event; **N** = total number of patients; **PVP** = percutaneous vertebroplasty, **RCTs** = randomised controlled trials.

PVP vs CT, severe adverse events, RCTs

Four studies provided evidence on severe adverse events.[2 4 7 132] and all 4 were included in the meta-analysis (***Figure 33***). Overall, there were no statistically significant differences between PVP and CT groups (RR 0.99; 95% CI 0.29, 3.35; p = 0.99). The absolute risk for PVP was 1.9% (n = 5/268) and for the CT group it was 1.9% (n = 5/263). The Chi^2 and I^2 statistics indicated low levels of heterogeneity and inconsistency (P = 1.00 and I^2 = 0%, respectively).

Severe adverse events included respiratory arrest, humerus fracture and thecal sac injury in the PVP group, and spinal cord compression, idiopathic tachycardia and rigors in the CT group.

The reporting of severe adverse events differed between the included studies. Only 1 study provided definitions of what constitutes a severe adverse event.[2] Two studies reported the number of events per patient[4 132]; 1 did not specify whether the events were per patient or total events.[7]

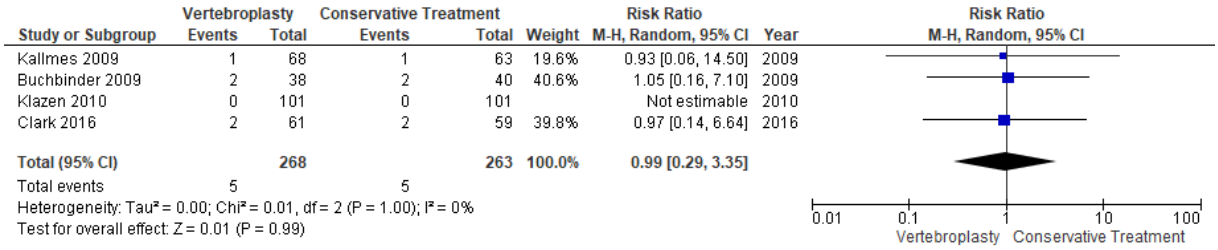


Figure 33 Forest plot indicating risk ratio of severe adverse events for PVP compared to CT

PVP vs CT, severe adverse event, non-RCTs

One study provided evidence on serve adverse events at 24 months (***Table 41***).[149] The reported data found no severe adverse events in either group.

Table 41 PVP compared to CT: severe adverse events (non-RCTs)

Author; Year	Follow-up	PVP n/N (%)	CT n/N (%)	p value
Diamond 2006[149]	24 months	0/88 (0.0%) patients	0/38 (0.0%) patients	NR

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **PVP** = percutaneous vertebroplasty, **RCTs** = randomised controlled trials.

PVP vs CT, any adverse events, RCTs

Seven studies provided evidence on any adverse events.[6 7 95 132 135 139 141] Six studies were included in the meta-analysis (***Figure 34***).[6 7 95 135 139 141] Overall, there was no statistically significant difference between PVP and CT groups (RR 1.68; 95% CI 0.57, 4.91; p = 0.35). The absolute risk for PVP was 6.6% (n = 28/424) and for the CT group it was 5.5% (n = 24/438). The Chi^2 and I^2 statistics indicated moderate levels of heterogeneity and inconsistency (P = 0.04, I^2 = 58.0%, respectively).

All 6 studies reported the number of patients with adverse events. One study reported adverse events in the PVP group but not in the CT group.[132] The study authors reported three perioperative adverse events: pain-induced vasovagal reaction (n = 2) and an asthma exacerbation (n = 1). The procedure was successfully completed in all patients who experienced adverse events.

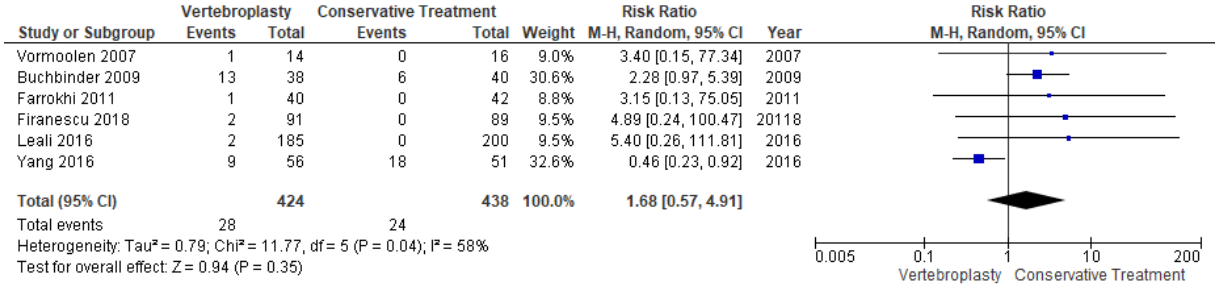


Figure 34 Forest plot indicating risk ratio of any adverse events for PVP compared to CT

PVP vs CT, any adverse event, non-RCTs

Two studies provided evidence on any adverse event (**Table 42**).[148 149] The absolute rate of adverse events was 2.5% (n = 3/118) in the PVP group and 0.0% (n = 0/68) in the CT group. The adverse events include a fracture of transverse processes (n = 2) and a psoas muscle haematoma (n = 1).

Table 42 PVP compared to CT: any adverse events (non-RCTs)

Author; Year	Length of follow-up	PVP	CT	p value
Andrei 2017[148]	12 months	0/30 (0.0%) patients	0/30 (0.0%) patients	NR
Diamond 2006[149]	24 months	3/88 (3.4%) patients	0/38 (0.0%) patients	NR
Absolute rate	12–24 months	3/118 (2.5%) patients	0/68 (0.0%) patients	

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NR** = not reported, **PVP** = percutaneous vertebroplasty, **RCTs** = randomised controlled trials.

PVP vs CT, new symptomatic vertebral fracture, RCTs

Six studies reported evidence on new symptomatic vertebral fractures (***Figure 35***) and all 6 were included in the meta-analysis.[6-8 134 135 139] Overall, there was no statistically significant difference between PVP and CT groups (RR 1.29; 95% CI 0.46, 3.62; p = 0.63). The absolute risk for PVP was 11.5% (n = 48/418) and for the CT group it was 7.3% (n = 31/422) (incidence per patient). The Chi^2 test and I^2 statistic indicated considerable levels of heterogeneity and inconsistency (P = 0.005 and I^2 = 70%).

Two studies noted that the new symptomatic fracture was adjacent to the initial fracture.[135 139] Four studies did not specify location of the new fracture in relation to the old fracture.[6-8 134]

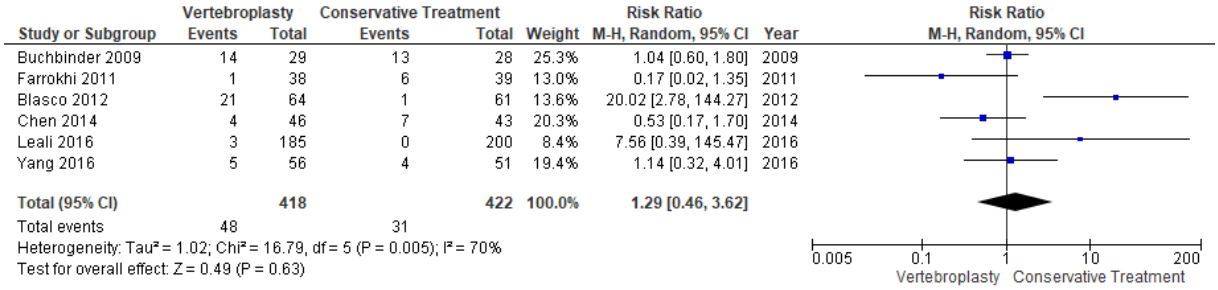


Figure 35 Forest plot indicating risk ratio of new clinical vertebral fractures for PVP compared to CT

PVP vs CT, new symptomatic vertebral fractures, non-RCTs

One study provided evidence on new symptomatic fractures.[150] Within 6 weeks of the procedure, three patients in the PVP group (3.4%) reported recurrent back pain attributable to new fractures. New fractures in the CT group were not reported. By 24 months, 18 new symptomatic fractures were reported, 11 of which were treated with vertebroplasty. The number of patients per treatment arm and location of the new fractures relative to the original fracture was not reported.

PVP vs CT, new radiographic vertebral fractures, RCTs

Seven studies reported evidence on new radiographic vertebral fractures and all 7 were included in the meta-analysis (***Figure 36***).[2 6-8 95 96 132] Overall, there was no statistically significant difference between PVP and CT groups (RR 1.18; 95% CI 0.70, 1.99; p = 0.54). The absolute risk for the PVP group was 27.2% (n = 106/389) and for the CT group it was 23.8% (n = 88/369). The Chi^2 test and I^2 statistic indicated moderate levels of heterogeneity and inconsistency (P = 0.002 and I^2 = 72%).

Four studies noted that the new symptomatic fracture was adjacent to the initial fracture.[7 8 96 132] Three studies did not specify new fracture location in relation to the old fracture.[2 6 95]

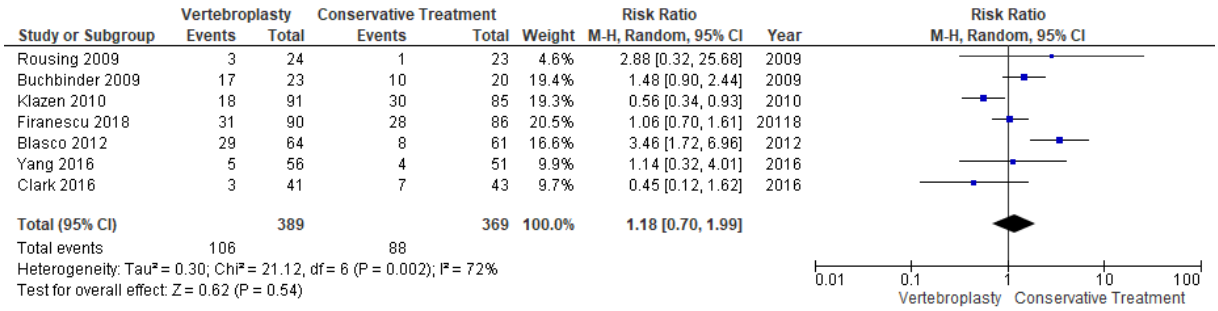


Figure 36 Forest plot indicating risk ratio of new radiographic vertebral fractures for PVP compared to CT

PVP vs CT, new radiographic vertebral fractures, non-RCTs

One study provided evidence on new radiographic fractures (***Table 43***).[149] The data demonstrated no statistically significant difference between PBK and CT groups (p = 0.52) with respect to new radiological vertebral fractures at 24 months.

Table 43 PVP compared to CT: new radiographic vertebral fractures (non-RCTs)

Author; Year	Follow-up	PVP n/N (%)	CT n/N (%)	p value
Diamond 2006[149]	24 months	9/88 (10.2%) patients	4/38 (10.5%) patients	0.52

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **PVP** = percutaneous vertebroplasty, **RCTs** = randomised controlled trials.

PVP vs CT, cement leakage, RCTs

Nine studies reported cement leakage following PVP (***Table 44***).[2 6-8 95 96 132 134 135] Studies reported the incidence of cement leak per vertebrae treated (k = 6)[6 8 95 132 134 135] or per patient (k = 2).[2 7] One study that reported no symptomatic leaks were identified.[96]

The absolute rate of cement leaks per treated vertebrae was 55.1% (n = 343/623). The range varied from 14.0% (n = 14/100)[135] to 91.3% (n = 105/115).[95] When assessed on a per patient basis, 39.4% of patients reported leaks (n = 39/99). The range varied from 34.4% (n = 21/61) to 47.4% (n = 18/38). One symptomatic leak resulted in extremity pain and weakness, necessitating an immediate bilateral laminectomy.[135] The remaining cases were asymptomatic.

Table 44 PVP compared to CT: cement leakage (RCTs)

Author; Year	Length of follow-up	Cement leakage per vertebral bodies treated or per patient n/N (%)	Symptomatic or Asymptomatic
Blasco 2012[8]	12 months	69/140 (49.0%) treated vertebrae	Asymptomatic
Buchbinder 2009[7]	24 months	18/38 (37.0%) patients	Asymptomatic
Chen 2014[134]	12 months	36/69 (52.0%) treated vertebrae	Asymptomatic
Clark 2016[2]	6 months	21/61 (34.4%) patients	Asymptomatic
Farrokhi 2011[135]	36 months	14/100 (14.0%) treated vertebrae	1 symptomatic case, remaining asymptomatic
Firanescu 2018[95]	12 months	105/115 (91.3%) treated vertebrae	Asymptomatic
Klazen 2010[132]	12 months	97/134 (72.0%) treated vertebrae	Asymptomatic
Rousing 2009[96]	12 months	Not reported	Asymptomatic
Yang 2016[6]	12 months	22/65 (33.8%) treated vertebrae	Asymptomatic
Absolute estimate	6–36 months	343/623 (55.0%) treated vertebrae 39/99 (39.4%) patients	

Abbreviations
n = number of patients with event, **N** = total number of patients, **RCTs** = randomised controlled trials.

PVP vs CT, cement leakage, non-RCTs

No study reported this outcome.

PVP vs CT, cement leakage, single-arm trials

Fifteen single-arm trials reported cement leakage following PVP (***Table 45***).[151-165] Studies reported the incidence of cement leak per vertebrae treated (k = 13), per patient (k = 1), or did not specify (k = 1). The absolute rate of cement leaks per treated vertebrae was 38.7% (n = 1,109/2,863), ranging from 11.7% (n = 15/128) to 81.6% (n = 164/201). On a per patient basis, 4.0% (n = 8/200) of patients reported leaks. There were 4 symptomatic leaks, which caused nerve root irritation and cement embolism. The remaining leaks were asymptomatic.

Table 45 PVP compared to CT: cement leakage (single-arm trials)

Author; Year	Length of follow-up	Cement leakage per vertebral bodies treated or per patient n/N (%)	Symptomatic or Asymptomatic
Al-Ali 2009[151]	12 months	219/660 (33.2%) treated vertebrae	Asymptomatic
Bae 2012 [a 152]	24 months	63.8% treated vertebrae	3 symptomatic patients (nerve root irritation), remaining asymptomatic
DePalma 2011[153]	24 months	29/163 (17.8%) treated vertebrae	Asymptomatic
Dohm 2014[164]	24 months	164/201 (81.6%) treated vertebrae	1 symptomatic (cement embolism), remaining asymptomatic
Fenoglio 2008[154]	20.4 months	7/52 (13.5%) treated vertebrae	NR
Kotwica 2011 [b 155]	24 months	8/200 (4.0%) patients	Asymptomatic
Masala 2012[157]	12 months	15/128 (11.7%) treated vertebrae	NR
Masala 2009[156]	36 months	4.8% [c]	Asymptomatic
Nieuwenhuijse 2012[159]	12 months	155/216 (71.8%) treated vertebrae	Asymptomatic
Nieuwenhuijse 2010[158]	12 months	99/125 (79.2%) of treated vertebrae [d]	Asymptomatic (one asymptomatic pulmonary cement embolism and cement spur)
Pitton 2008[160]	19.7 months	214/385 (55.6%) treated vertebrae	Asymptomatic
Santiago 2010[165]	12 months	14/69 (20.2%) treated vertebrae	NR
Saracen 2014[161]	24 months	83/594 (14.0%) treated vertebrae	NR
Voormolen 2006a[162]	12 months	79/168 (47.0%) treated vertebrae	Asymptomatic
Voormolen 2006b[163]	12 months	31/102 (30.4%) treated vertebrae	NR
Absolute rate	12–60 months	1,109/2,863 (38.7%) treated vertebrae 8/200 (4.0%) patients	

Abbreviations
n = number of patients with event, **N** = total number of patients, **NR** = not reported.
Notes
a = results of PMMA arm reported, absolute number of adjacent fractures could not be determined.
b = 200 were measured postoperatively and 80 patients were measured at 24 months.
c = not reported whether per patient or per vertebrae.
d = low and medium viscosity cement arms pooled.

7.8.2 PBK vs CT

Table 46 provides a summary of the absolute event rates for safety outcomes comparing PBK to CT. Overall, there were no statistically significant differences between PBK and CT with respect to radiographic fractures in either the RCTs or non-RCTs. The remaining outcomes were reported without statistical information (the rates in the arms were likely similar). The rates of safety outcomes varied substantially between the study designs and likely reflected the small sample size in the RCTs and non-RCTs.

Table 46 Summary of safety outcomes for PBK compared to CT

Outcomes	RCTs % (n/N)	Non-RCTs % (n/N)	Single-arm trials % (n/N)
All-cause mortality	PBK 6.0% (n = 9/149) CT 4.6% (n = 7/151)	PBK 2.5% (n = 1/40) CT 15.0% (n = 3/20)	NA
Severe adverse events	PBK 1.3% (n = 2/149) CT 0.0% (n = 0/151)	NR	NA
Any adverse events	PBK 131 events CT 131 events	PBK 0.0% (n = 0/40) CT 0.0% (n = 0/20)	NA
Symptomatic fractures [a]	PBK 14.1% (n = 21/149) CT NR	PBK 8.8% (n =3/34) CT NR	NA
Radiographic fractures [a]	PBK 33.0% (n = 38/115) CT 25.2% (n = 24/95)	PBK 6.5% (n = 3/46) CT 16.4% (n = 10/61)	NA
Cement leakage [b]	PBK 27.1% (n = 51/188)	PBK 11.3% (n = 11/97)	PBK 27.5% (n = 385/1,402)

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NA** = not applicable, **NR** = not reported, **RCTs** = randomised controlled trials, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.
Notes
a = fractures per patient.
b = leaks per treated vertebrae reported, the leaks per patient can be found in ***Table 55*** to ***Table 57***.

PBK vs CT, all-cause mortality, RCTs

One study provided evidence on all-cause mortality (***Table 47***).[168] By 12 months, there were 9 deaths in the PBK arm and 7 deaths in the CT arm. [168] All deaths were deemed unrelated to the intervention.

Table 47 PBK compared to CT: all-cause mortality (RCTs)

Author; Year	Length of follow-up	PBK n/N (%)	CT n/N (%)	p value
Wardlaw 2009[168]	12 months	9/149 (6.0%)	7/151 (4.6%)	NR

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, all-cause mortality, non-RCTs

One study provided evidence on all-cause mortality (***Table 48***).[172] There was 1 death in the PBK arm and 3 in the CT groups by 36 months. All deaths were deemed unrelated to the intervention.

Table 48 PBK compared to CT: all-cause mortality (non-RCTs)

Author; Year	Follow-up	PBK n/N (%)	CT n/N (%)	p value
Kasperk 2010[172]	36 months	1/40 (2.5%) patients	3/20 (15.0%) patients	NR

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, severe adverse events, RCTs

One study provided evidence on severe adverse events (***Table 49***).[168] By 12 months, there were 58 severe adverse events (as per the MEDRA classification) in the PBK group and 54 in the CT group. It was unclear whether patients experienced more than 1 adverse event. Of severe adverse events, the most common were cardiovascular and vascular disorders, back pain, and respiratory disorders. Infection, anaemia, neoplasms, and nervous system and psychiatric disorders were infrequent adverse events. Two events were attributed to PBK: a surgical site haematoma and a urinary tract infection. No severe adverse events were attributed to CT.

Table 49 PBK compared to CT: severe adverse event (RCTs)

Author; Year	Severe adverse event	PBK n/N (%)	CT n/N (%)	p value
Wardlaw 2009[168]	All events Events resulting in death Procedure-related	58/NR 9/149 (6.0%) 2/149 (1.3%)	54/NR 7/151 (4.6%) 0/151 (0.0%)	NR

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, severe adverse events, non-RCTs

No study reported this outcome.

PBK vs CT, any adverse events, RCTs

Three studies provided evidence on any adverse events (***Table 50***).[59 60 168] The results were not pooled as it was unclear whether the studies report the number of patients or the total number of adverse events.

There were no complications reported in Li (2019). Wardlaw (2009) reported 130 adverse events in the PBK group and 122 in the CT groups within 12 months. Liu (2019) reported 1 adverse event in the PBK group and 9 events in the CT groups, however, the length of follow-up was not reported, and it was not clear over what time span the events occurred.

Table 50 PBK compared to CT: any adverse event (RCTs)

Author; Year	Length of follow-up	PBK n	CT n	p value
Li 2017[60]	6 months	0	0	NR
Liu 2019[59]	NR	1	9	$p < 0.05$
Wardlaw 2009[168]	12 months	130	122	NS
Total		131	131	

Abbreviations
CT = conservative treatment, **n** = number of events, **NR** = not reported, **NS** = not significant, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PVP vs CT, any adverse event, non-RCTs

One study provided evidence on severe adverse events at 24 months (***Table 51***).[172] The reported data demonstrated that there were no adverse events in either arm.

Table 51 PBK compared to CT: severe adverse events (non-RCTs)

Author; Year	Follow-up	PVP n/N (%)	CT n/N (%)	p value
Kasperk 2010[172]	36 months	0/40 (0.0%) patients	0/20 (0.0%) patients	NR

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, new symptomatic vertebral fractures, RCTs

One study provided evidence on new symptomatic vertebral fractures at 12 months.[168] Twenty-one patients (14.1%) in the PBK group reported new symptomatic fractures, of which 12 received additional PBK procedures by 6 months. The incidence of new symptomatic fractures in the CT group was not reported. It was not clear whether the fractures were adjacent to the initial fracture.

PBK vs CT, new symptomatic vertebral fractures, non-RCTs

One study provided evidence on new symptomatic vertebral fractures (***Table 52***).[172] There were 7 new symptomatic fractures in three PBK patients in Kasperk (2010).[172] It was unclear whether fractures were adjacent to the initial fracture.

Table 52 PBK compared to CT: new symptomatic vertebral fractures (non-RCTs)

Author; Year	Follow-up	PBK n/N (%)	CT n/N (%)	p value
Kasperk 2010[172]	36 months	7 fractures in 3/34 (8.8%) patients	NR	NR

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **NR** = not reported, = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, new radiographic vertebral fractures, RCTs

One study provided evidence on new radiographic vertebral fractures at 12 months (***Table 53***).[168] The reported data demonstrated no statistically significant differences between PBK and CT groups (p = 0.20). It was unclear whether fractures were adjacent to the initial fracture.

Table 53 PBK compared to CT: new radiographic vertebral fractures (RCTs)

Author; Year	Follow-up	PBK n/N	CT n/N	p value
Wardlaw 2009[168]	12 months	38/115 (33.0%) patients	24/95 (25.2%) patients	0.20

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, new radiographic vertebral fractures, non-RCTs

Two studies provided evidence on new radiographic vertebral fractures (***Table 54***).[172 174] There were no significant differences between PBK and CT groups in the incidence of new radiographic fractures when assessed on a per treated vertebrae (p = 0.59)[172] or per patient basis (p = 0.12).[174]

Table 54 PBK compared to CT: new radiographic vertebral fractures (non-RCTs)

Author; Year	Follow-up	PBK n/N (%)	CT n/N (%)	p value
Kasperk 2010[172]	36 months	7/72 (9.7%) treated vertebrae	4/29 (13.8%) treated vertebrae	0.59
Movrin 2010[174]	12 months	3/46 (6.5%) patients	10/61 (16.4%) patients	0.12

Abbreviations
CT = conservative treatment, **n** = number of patients with event, **N** = total number of patients, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, cement leakage, RCTs

Two studies reported cement leakages (***Table 55***).[59 168] The absolute rate per treated vertebrae was 27.1% (n = 51/188). The rate of cement leakage per patient was 23.7% (n = 49/207), ranging from 1.7% (n = 1/58)[59] to 32.2% (n = 48/149).[168] In Wardlaw (2009),[168] all the cement leaks were asymptomatic. In Liu (2019),[59] it was not reported whether the leaks were symptomatic or asymptomatic.

Table 55 PBK compared to CT: cement leakage (RCTs)

Author; Year	Length of follow-up	PBK n/N (%)	Symptomatic or asymptomatic
Liu 2019[59]	NR	1/58 (1.7%) patients	NR
Wardlaw 2009[168]	12 months	51/188 (27.1%) treated vertebrae 48/149 (32.2%) patients	Asymptomatic
Absolute rate		49/207 (23.7%) patients 51/188 (27.1%) treated vertebrae	

Abbreviations
n = number of events, **N** = total number of patients/vertebrae, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, cement leakage, non-RCTs

Three studies reported cement leakages (***Table 56***)[170 172 174] by either the incidence per vertebrae treated (k = 2) or per patient (k = 1). The absolute rate per treated vertebrae was 11.3% (n = 11/97), ranging from 9.7% (n = 7/72)[172] to 16.0% (n = 4/25).[170] The rate per patient was 8.7% (n = 4/46).[174] The leaks were asymptomatic in 2 studies. In the remaining study it was not reported whether the leaks were symptomatic.

Table 56 PBK compared to CT: cement leakage (non-RCTs)

Author; Year	Length of follow-up	PBK n/N (%)	Symptomatic or asymptomatic
Kasperk 2010[172]	36 months	7/72 (9.7%) treated vertebrae	Asymptomatic
Giannotti 2012[170]	24 months	4/25 (16.0%) treated vertebrae	Asymptomatic
Movrin 2010[174]	12 months	4/46 (8.7%) patients	NR
Absolute rate	12–36 months	11/97 (11.3%) treated vertebrae 4/46 (8.7%) patients	

Abbreviations
n = number of events, **N** = total number of patients/vertebrae, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCTs** = randomised controlled trials.

PBK vs CT, cement leakage, single-arm trials

Six studies report cement leakages (***Table 57***)[63 164 165 175-177] by either the incidence of cement leak per vertebrae treated (k = 3), per patient (k = 2) or both (k = 1). The absolute rate per treated vertebrae was 27.5% (n = 385/1,402) ranging from 5.2% (n = 7/135) to 73.4% (n = 157/214). The rate per patient was 1.5% (n = 10/666) ranging from 0.5% (n = 3/564) to 6.9% (n = 7/102). Four symptomatic leaks resulted in cement embolism, hemiparesis, heart perforation and emergency surgery. The consequence of 1 cement embolism was not reported. The remaining cases were asymptomatic.

Table 57 PBK compared to CT: cement leakage (single-arm trials)

Author; Year	Length of follow-up	PBK n/N (%)	Symptomatic or asymptomatic
Dohm 2014[164]	24 months	157/214 (73.4%) treated vertebrae	1 symptomatic (cement embolism), remaining asymptomatic
Hillmeier 2004[175]	12 months	13/192 (6.8%) treated vertebrae	Asymptomatic
Hubschle 2014[63]	12 months	201/819 (24.5%) treated vertebrae	4 symptomatic, remaining asymptomatic
Prokop 2012[176]	6 months	3/564 (0.5%) patients 16%[a]	3 symptomatic (hemiparesis, cement embolism leading to heart perforation, and cement-filled stents requiring emergency surgery)
Robinson 2008[177]	6 months	7/102 (6.9%) patients 7/135 (5.2%) treated vertebrae	Asymptomatic
Santiago 2010[165]	12 months	7/42 (16.7%) treated vertebrae	Asymptomatic
Absolute rate	6–24 months	385/1,402 (27.5%) treated vertebrae 10/666 (1.5%) patients	

Abbreviations
n = number of events, **N** = total number of patients/vertebrae, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty.
Notes
a = not reported whether per patient or treated vertebrae.

PVP and PBK, radiation exposure, non-RCTs

Radiation exposure to the operator during vertebroplasty and kyphoplasty was reported in 1 study (***Table 58***).[185] Most of the radiation exposure during vertebroplasty occurred during needle/device placement rather than cement delivery. By contrast, radiation exposure during kyphoplasty was attributable to both needle/device placement and cement delivery. Overall, operators of vertebroplasty were exposed to less radiation than were operators of kyphoplasty ($p < 0.0001$). This was likely attributable to the different procedure times.

Table 58 PVP and PBK: mean radiation exposure rate

Author; Year	Outcome	PVP n = 20	PBK n = 87	p value
Ortiz 2005[185]	Needle/device placement mean ± SD	1.25 ± 1.3 µSv 3.9 ± 2.4 mins	4.1 ± 5.5 µSv 4.4 ± 1.4 mins	0.02 NS
Ortiz 2005[185]	Cement delivery mean ± SD	0.45 ± 0.94 µSv 1.5 ± 0.6 mins	4.5 ± 11.8 µSv 2.1 ± 0.9 mins	NS < 0.0001
Ortiz 2005[185]	Total exposure mean ± SD	1.7 ± 1.9 µSv 39.3 ± 8 mins	8.6 ± 13.9 µSv 55.7 ± 13mins	< 0.0001 < 0.0001

Abbreviation
n = number of patients, **NS** = not significant, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation, **µSv** = micro sieverts.
Notes
For reference, a dental x-ray results in exposure to 4 to 10 µSv.[186]

7.9 Results: Extended Assessment of Safety

The extended assessment of harms aimed to identify adverse events associated with PVP, PBK and CT which may have been missed from the RCTs owing to insufficient power or duration of follow-up. The assessment included existing databases evaluating PVP or PBK and meta-analyses or pooled analyses of comparators (NSAIDS, opioids and paracetamol). A summary of the extended assessment of harms is provided in ***Table 59***. The population enrolled in the RCTs and the database analyses differ with respect to age and presence of comorbidities. Patients enrolled in the database analyses tended to be older (mean age approximately 82 years) and were burdened by comorbidities (approximately 23% of patients reported a Charlson Comorbidity Index of 3 and above) more than patients in the RCTs. It was unclear to what extent the differences in demographic underscore the differences in safety-related outcomes between the RCTs and databases.

Table 59 Summary of safety-related outcomes evaluated in RCTs, database analyses and existing meta-analyses of CT

Intervention vs comparator	Mortality % or RR (95% CI)	Serious adverse event % or RR (95% CI)	Any adverse event % or RR (95% CI)
PVP vs CT *(database analyses)*	1yr +30% [a] 10yrs +8% [a]	1yr +20% [a] 10yrs +7%	1yr +2% [a] 10yrs +2%
PBK vs CT *(database analyses)*	1yr +55% [a] 10yrs +24%	1yr +19% [a] 10yrs +11%	1yr +1% [a] 10yrs +4%
PVP vs CT (RCTs)	PVP 4.8% CT 5.8% 0.84 (95% CI 0.52, 1.35)	PVP 1.9% CT 1.9% 0.99 (95% CI 0.29, 3.35)	PVP 6.6% CT 5.5% 1.68 (95% CI 0.57, 4.91)
PBK vs CT (RCTs)	PBK 6.0% CT 4.6%	PBK 1.3% CT 0.0%	PBK 131 events CT 131 events
NSAIDs vs placebo [b]	1.22 (95% CI 1.04, 1.44) *	1.37 (95% CI 1.14, 1.66) **	45.4% of patients
Opioids vs placebo [b]	NR	2.75 (95% CI 2.06, 3.67) **	1.42 (95% CI 1.22, 1.66) **
Paracetamol vs placebo [b c]	1.28 (95% CI 1.26, 1.30)	1.0 (95% CI 0.9, 1.1)	1.2 (95% CI 0.7, 2.1)

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **NR** = not reported, **NSAIDS** = nonsteroidal anti-inflammatory drug, **PBK** = percutaneous balloon kyphoplasty; **PVP** = percutaneous vertebroplasty, **RCT** = randomised controlled trial, **RR** = risk ratio.
Notes
* $p < 0.01$, ** $p < 0.0001$
a = the database analysis reported relative rates only. Therefore, a positive percentage indicated the event occurred more frequently in the CT arm than in the PVP or PBK arms. A negative rate implied the event occurred more frequently in the PVP or PBK arms.
b = the risk ratio for NSAIDs, opioids and paracetamol were relative to placebo. A positive risk ratio indicated the event occurred more frequently in the intervention group compared to placebo, a negative risk ratio indicated the event occurred less frequently in the intervention group.
c = The mortality associated with paracetamol was informed by observational studies. The incidence of serious and any adverse events were informed by RCTs.
The serious adverse event in the database analyses was represented by cardiac complications, and any adverse event was represented by UTI. The results from celecoxib was selected to be the representative NSAID. For NSAIDs and registry data, vascular/cardiac complications were selected as the serious adverse event.

7.9.1 PVP vs CT

PVP vs CT, mortality and adverse events, 30 days to 6 months

The rates of in-hospital mortality, readmission, bedsores and pneumonia at 30 days and 6 months statistically differed between PVP and CT groups ($p < 0.001$ for all outcomes) with lower rates in the PVP group (see ***Appendix C, Table 112***).[146] However, the incidence of deep vein thrombosis (DVT) and embolism was higher in the PVP arm ($p < 0.05$ at 30 days and $p < 0.001$ at 6 months). There were no statistical differences between the 2 groups for the rate of infection or neurological compromise ($p > 0.05$). Approximately 7.9% of PVP patients received additional vertebral augmentation procedures, suggesting that patients may have experienced a new fracture.

PVP vs CT, mortality and adverse events, 1 to 10 years

Over the 10 year follow-up period, the relative incidence of mortality and adverse events decreased (see ***Appendix C, Table 112***).[147] The highest incidence was observed during the first year of follow-up. By 10 years, the relative difference in event rate was less than 10% for all outcomes. However, the relative rates of mortality, cardiac complications, pneumonia, UTI and pulmonary embolism remained statistically different ($p < 0.001$ for all outcomes). The relative incidence of mortality, cardiac complications, pneumonia and urinary tract infection was higher in the CT group compared to the PVP group. In contrast, the incidence of pulmonary embolism was greater in the PVP group.

7.9.2 PBK vs CT

PBK vs CT, mortality and adverse events, 30 days to 6 months

The rates of in-hospital mortality, readmission, bedsores, pneumonia and pulmonary embolism at 30 days and 6 months were statistically different between PBK and CT groups ($p < 0.05$ or 0.001 for all outcomes), with a higher incidence observed in the CT group (see ***Appendix C, Table 113***).[146] The incidence of DVT was initially greater in the PBK arm however, by 6 months, patients treated with CT reported a higher incidence ($p < 0.001$). There were no differences between PBK and CT for the rate of infection or neurological compromise. Approximately 9.4% of PBK patients reported additional vertebral augmentation procedures during follow-up.

PBK vs CT, mortality and adverse events, 1 to 10 years

As for PBK, over the 10 year follow-up period the difference in the incidence of mortality and adverse event decreases, with the highest incidence observed during the first year of follow-up (see ***Appendix C, Table 113***).[147] Ten years post-treatment, the incidence of mortality, cardiac complication, DVT, pneumonia, pulmonary embolism, pulmonary complications and UTI was statistically different between

PBK and CT groups (p < 0.001). CT patients reported higher rates for mortality and all adverse events from 1 to 10 years.

7.9.3 CT: Opioids, NSAIDs and Paracetamol

NSAIDs

The safety of NSAIDs was informed by 1 meta-analysis and 1 pooled analysis of RCTs evaluating elderly patients (age > 65 years) (***Table 60***).[83] [178] Results of the meta-analysis indicated patients treated with coxibs (celecoxib, rofecoxib, etoricoxib, lumiracoxib and GW403681) reported a higher rate of mortality, major vascular events (including myocardial infarction, coronary death or stroke), heart failure and upper gastrointestinal adverse events compared to patients treated with placebo. Patients receiving diclofenac, ibuprofen and naproxen reported higher rates of heart failure and gastrointestinal adverse events compared to placebo, however, all-cause mortality was generally similar between the groups. Lastly, diclofenac but not ibuprofen or naproxen reported higher rates of major vascular events compared to placebo. Among elderly patients, 45.4–62.9% reported an adverse event and 20.1–33.6% reported a gastrointestinal adverse event following treatment with NSAIDs.

Table 60 NSAIDs compared to placebo: results from a meta-analysis and a pooled analysis of RCTs

Safety outcomes	Coxib vs placebo Adjusted rate ratio (99% CI)	Diclofenac vs placebo Adjusted rate ratio (99% CI)	Ibuprofen vs placebo Adjusted rate ratio (99% CI)	Naproxen vs placebo Adjusted rate ratio (99% CI)
Coxib and traditional NSAIDs Trialist collaboration 2013 [83]				
Any cause mortality	1.22 (1.04–1.44) p = 0.01	1.20 (0.94–1.54) p = 0.15	1.61 (0.90–2.88) p = 0.11	1.03 (0.71–1.49) p = 0.88
Major vascular events	1.37 (1.14–1.66) p = 0.0009	1.41 (1.12–1.78) p = 0.004	2.44 (0.89–2.33) p = 0.14	0.93 (0.69–1.27) p = 0.66
Heart failure	2.28 (1.62–3.20) p < 0.0001	1.85 (1.17–2.94) p = 0.009	2.49 (1.19–5.20) p = 0.02	1.87 (1.10–3.16) p = 0.02
Upper gastrointestinal adverse events	1.81 (1.17–2.81) p = 0.0070	1.89 (1.16–3.09) p = 0.01	3.97 (2.22–7.10) p < 0.0001	4.22 (2.71–6.56) p < 0.0001
Mallen 2011[178]				
Pooled analysis, 6–52 weeks	Celecoxib n = 5872	*Diclofenac* *n = 2334*	*Ibuprofen* *n = 151*	*Naproxen* *n = 1104*
Any adverse event	2665 (45.4%)	1139 (48.8%)	95 (62.9%)	647 (58.6%)
Gastrointestinal adverse event	1181 (20.1%)	564 (24.2%)	46 (30.5%)	371 (33.6%)

Abbreviation
CI = confidence interval, **n** = number of patients experiencing event.
Notes
A positive risk ratio indicates the event occurred more frequently in NSAIDs compared to placebo, a negative risk ratio indicates the event occurred less frequently.

Opioids

The safety of opioids was informed by 2 meta-analyses (***Table 61***).[82] [179] According to Els (2017), the number of deaths attributed to opioids was generally under-reported, with only 2 deaths reported among the included trials.[82] Compared to placebo or a pharmacological comparator, patients treated with opioids were more likely to withdraw due to adverse events, and experienced a higher rate of adverse events. Further, patients receiving opioids were at a higher risk of nausea, constipation, dizziness, drowsiness, pruritis and dry mouth compared to those receiving placebo or NSAIDs.

Table 61 Opioids compared to placebo or NSAIDs: results from meta-analyses of RCTs and observational trials

Safety outcomes	**Opioids vs placebo RR (95% CI)**	**Opioids vs active pharmacological comparator RR (95% CI)**
Els 2017[82]		
Deaths	Opioids n = 2 Placebo n = 0	Nil
Withdrawal due to adverse event	3.40 (3.02, 3.82) p < 0.0001 Absolute 25.0 vs 7.1%	3.23 (2.42, 4.30) p < 0.00001 Absolute = 15.4 vs 4.7%
Serious adverse event	2.75 (2.06, 3.67) p < 0.00001 Absolute = 7.5 vs 4.0%	5.0 (0.60, 41.39) p = 0.14 Absolute = 9.2 vs 1.8%
Any adverse event	1.42 (1.22, 1.66) p < 0.0001 Absolute = 78.2 vs 54.4%	1.21 (1.10, 1.33) p < 0.00001 Absolute = 57.8 vs 47.7%
Busse 2018[179]		
Specific adverse events	**Opioids vs placebo RR (95%CI)**	**Opioids vs NSAIDs RR (95%CI)**
Nausea (non-enriched trials)	3.17 (2.69, 3.73) p = NR Absolute rate = 25.9 vs 8.2%	2.51 (2.00, 3.15) p = NR Absolute rate = 7.6 vs 19.1%
Constipation	3.08 (2.65, 3.55) p = NR Absolute rate = 16.2 vs 5.3%	2.84 (1.82, 4.43) p = NR Absolute rate = 3.2 vs 9.0%
Dizziness (non-enrichment trials)	2.69 (2.33, 3.11) p = NR Absolute rate = 15.1 vs 5.6%	1.98 (1.47, 2.66) p = NR Absolute rate = 7.0 vs 14.0 %
Drowsiness (non-enriched trials)	3.59 (2.88, 4.47) p = NR Absolute rate =15.0 vs 4.2%	2.29 (1.52, 3.46) p = NR Absolute rate = 4.5 vs 10.3%
Headache	1.10 (0.99, 1.22) p = NR Absolute rate = 8.6 vs 7.8%	1.37 (1.10, 1.69) p = NR Absolute rate =9.8 vs 13.5%
Pruritis	2.59 (1.86, 3.62) p = NR Absolute rate = 9.7 vs 3.8%	4.01 (2.33, 6.89) p = NR Absolute rate = 1.5 vs 5.9%
Dry mouth	2.57 (1.98, 3.34) p = NR Absolute rate = 4.9 vs 1.9%	3.42 (1.73, 6.77) p = NR Absolute rate = 1.6 vs 5.6%

Abbreviation
CI = confidence interval, **n** = number of patients experiencing event, **NR** = not reported, **RR** = risk ratio.
Notes
A positive risk ratio indicates the event occurred more frequently in opioids compared to placebo/active comparator, a negative risk ratio indicates the event occurred less frequently.

Paracetamol

The safety of paracetamol was informed by 2 meta-analyses of RCTs and observational trials (***Table 62***).[180] [181] The meta-analyses of RCTs found no differences between placebo and paracetamol for adverse event, serious adverse event or withdrawal due to adverse event. In contrast, the results of the observational trials demonstrated an increased risk of mortality and gastrointestinal events in patients treated with paracetamol compared to placebo, however, the p value was not reported thus the statistical significance cannot be determined.

Table 62 Paracetamol compared to placebo: results from meta-analyses of RCTs and observational trials

Outcome	Overall risk ratio (95% CI), p value Paracetamol vs placebo
Machado 2015[181]	
RCT (follow-up <6 weeks to 10 years)	
Withdrawal due to adverse event	1.2 (0.9, 1.5) p = NS
Any adverse event	1.0 (0.9, 1.1) p = NS
Serious adverse event	1.2 (0.7, 2.1) p = NS
Abnormal liver function	3.8 (1.9, 7.4) p = significant [a]
Roberts 2015[180]	
Observational trials (follow-up 2–20 years)	
Mortality	1.28 (1.26, 1.30) p = NR
Cardiovascular adverse event Paracetamol consumed 1–4 days per month Paracetamol consumed >22 days per month	 0.95 (0.79, 1.14) p = NR 1.44 (1.27, 1.63) p = NR
Gastrointestinal adverse event	1.36 (1.31, 1.41) p = NR

Abbreviation
CI = confidence interval, **NR** = not reported, **NS** = not significant, **RR** = risk ratio.
Notes
a = exact p value not reported.
A positive risk ratio indicates the event occurred more frequently in paracetamol compared to placebo, a negative risk ratio indicates the event occurred less frequently.

7.10 GRADE Summary of Findings Table

Table 63 GRADE summary of findings: PVP compared to CT for OVCF (any fracture age)

Outcomes	Anticipated absolute effect* (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with CT	Risk with PVP				
Pain: VAS Follow up: 1 month	Mean pain in the CT group was 4.69	MD -1.52 mm lower (-2.86 lower to -0.17 lower)	-	384 (3 RCTs)	⊕⊕◯◯ LOW [a,b,c,d]	PVP statistically differed from CT at 1 month. The effect size was small and may translate to a clinically important difference.**
Function: ODI Follow up: 1 month	Mean ODI in the CT group was 52.46	MD -16.27 points lower (-23.53 lower to -9.01 lower)	-	196 (2 RCTs)	⊕⊕◯◯ LOW [a,b,c,d]	PVP statistically differed from CT at 1 month. The effect size was moderate and likely translates to a clinically important difference.
Function: RDQ Follow up: 1 month	Mean RDQ in CT group at 1 month was 13.52	MD -2.03 points lower (-3.06 lower to -1.01 lower	-	277 (2 RCTs)	⊕⊕◯◯ LOW [a,b,c,d]	PVP statistically differed from CT at 1 month. The effect size was small and may translate to a clinically important difference.**
QoL: EQ-5D Follow up: 1 month	Mean EQ-5D in CT group at 1 month was 0.50	MD 0.10 point lower (-0.10 lower to 0.31 higher)	-	188 (1 RCT)	⊕⊕◯◯ LOW [a,b,d,e,f]	PVP did not statistically differ from CT at 1 month.
QoL: QUALEFFO Follow up: 1 month	Mean QUALEFFO in CT group at 1 month was 53.70	MD -6.16 points lower (-15.84 lower to 3.52 higher)	-	295 (2 RCTs)	⊕⊕◯◯ LOW [a,b,d,f]	PVP did not statistically differ from CT at 1 month.
All-cause Mortality Follow up: up to 24 months	56 per 1,000	47 per 1,000 (29 to 76)	RR 0.84 (0.52 to 1.35)	1281 (9 RCTs)	⊕⊕⊕◯ MODERATE [d]	PVP did not statistically differ from CT (no effect).
Serious AE Follow up: up to 24 months	19 per 1,000	19 per 1,000 (6 to 64)	RR 0.81 (0.22 to 2.99)	531 (3 RCTs)	⊕⊕⊕◯ MODERATE [c,d,g]	PVP did not statistically differ from CT (no effect).

Abbreviations

AE = adverse event, **CI** = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **MD** = mean difference, **mm** = millimetres, **ODI** = Oswestry disability index, **PVP** = percutaneous vertebroplasty, **QoL** = quality of life, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RCTs** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **RR** = risk ratio, **VAS** = visual analogue scale.

Notes

***** = risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
****** = the effect size surpasses the lower bounds of identified MCIDs but not the upper bounds.
a = lack of blinding, incomplete accounting of patients and outcome events, **b** = considerable levels of heterogeneity as inferred by I^2 and Tau^2, **c** = 95% confidence interval around pooled estimates includes negligible effect and appreciable benefit/harm (depending on the MCID), **d** = low number of patients at evaluated timepoint (1 or 24 months) **e** = baseline difference between intervention arms, **f** = wide 95% confidence interval, **g** = trials had incomplete data.

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Table 64 GRADE summary of findings: PVP compared to sham for OCVF (any fracture age)

Outcomes	Anticipated absolute effect* (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with Sham	Risk with PVP				
Pain: NRS/VAS Follow up: 1 month	Mean pain in the sham group was 2.51	MD -0.76 mm lower (-1.21 lower to -0.31 lower)	-	449 (4 RCTs)	⊕⊕⊕◯ MODERATE [a]	PVP statistically differed from sham at 1 month. The effect size was small and was not clinically relevant.**
Function: ODI Follow up: NR	NR	NR	NR	NR	NR	NR
Function: RDQ Follow up: 1 month	Mean function in the sham group was 13	MD -0.28 point lower (-1.70 lower to 1.15 higher)	-	340 (3 RCTs)	⊕⊕◯◯ LOW [a,b,c]	PVP did not statistically differ from sham at 1 month.
QoL: EQ-5D Follow up: 1 month	Buchbinder 2009: no statistical difference at 1 month (p > 0.05) n = 73 Clark 2016: Significant difference at 1 month (p = 0.04) n = 98		-		⊕⊕◯◯ LOW [c,d]	The statistical effect was inconsistent; unclear whether PVP differed from sham.
QoL: QUALEFFO Follow up: 1 month	Mean function in the sham group was 50.3	MD -1.39 points lower (-3.24 lower to 0.47 higher)	-	352 (3 RCTs)	⊕⊕◯◯ LOW [a,b,c]	PVP did not statistically differ from sham at 1 month.
All-cause Mortality	See ***Table 63***					
Serious AE	See ***Table 63***					

Abbreviations

AE = adverse events**, CI** = confidence interval, **EQ-5D** = EuroQol 5 dimension questionnaire, **MD** = mean difference, **mm** = millimetres, **NR** = not reported, **NRS** = numerical rating scale, **ODI** = Oswestry disability index, **PVP** = percutaneous vertebroplasty, **QoL** = quality of life, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RCTs** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **VAS** = visual analogue scale.

Notes

***** = risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

****** = the effect size does not surpass the lower bounds of identified MCIDs.

a = considerable levels of heterogeneity as inferred by I^2 and Tau^2.

b = 95% Confidence interval around pooled estimates were wide.

c = low number of patients at 1 month.

d = direction of effect inconsistent between studies.

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

Table 65 GRADE summary of findings: PBK compared to CT for OVCF (mean fracture age ≥ 1 to 6 weeks)

Outcomes	Anticipated absolute effect* (95% CI)		Relative effect (95% CI)	Number of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with CT	Risk with PBK				
Pain: VAS Follow up: 1 month	Mean pain in the CT group was 4.15	MD -0.18 mm lower (-2.15 lower to 1.80 higher)		344 (2 RCTs)	⊕⊕◯◯ LOW [a,b,c,d]	PBK did not statistically differ from CT at 1 month, however there were statistical differences from 1 day to 1 week and at 3, 12 and 24 months. The effect size at early timepoints was large and likely translates to clinically important differences.**
Function: ODI Follow up: 6 months	PBK vs CT 15.1 ± 3.6 vs 18.7 ± 5.3 p < 0.05		-	80 (1RCT)	⊕⊕◯◯ LOW [a,d]	PBK statistically differed from CT at 1 month. The clinical impact is uncertain.
Function: RDQ Follow up: 1 month	PBK vs CT 10.9 ± 4.3 vs 15.1 ± 4.3 p < 0.0001		-	298 (1 RCT)	⊕⊕◯◯ LOW [a,d]	PBK statistically differed from CT at 1 month. The clinical impact is uncertain.
QoL: EQ-5D Follow up: 1 month	PBK vs CT 0.59 ± 1.07 vs 0.49 ± 1.04 p < 0.0001		-	298 (1 RCT)	⊕⊕◯◯ LOW [a,d]	PBK statistically differed from CT at 1 month. The clinical impact is uncertain.
QoL: QUALEFFO Follow up: NR	NR		-	-	-	-
All-cause Mortality Follow up: 24 months	46 per 1,000	0 per 1,000 (0 to 0)	not estimable	300 (1 RCT)	⊕⊕⊕◯ MODERATE [d]	PBK did not statistically differ from CT (no effect).
Serious AE Follow up: 24 months	PBK vs CT 58 vs 54 events 2 serious AE in the kyphoplasty group were treatment-related.			300 (1 RCT)	⊕⊕◯◯ LOW [d,e]	PBK did not statistically differ from CT (no effect).

Abbreviations

AE = adverse events, **CI** = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **MD** = mean difference, **mm** = millimetres, **NR** = not reported, **ODI** = Oswestry disability index, **PBK** = percutaneous balloon kyphoplasty, **QoL** = quality of life, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **RCTs** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **VAS** = visual analogue scale.

Notes

***** = risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).

** = the effect size surpasses the lower bounds of identified MCIDs.

a = lack of blinding, incomplete accounting of patients and outcome events, **b** = considerable levels of heterogeneity as inferred by I^2 and Tau^2, **c** = 95% Confidence interval around pooled estimates includes negligible effect and appreciable benefit/harm (depending on the MCID), **d** = low number of patients at evaluated timepoint (1 or 24 months), **e** = incomplete accounting of patients and outcome events.

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

8 Cost, Cost-Effectiveness and Budget Impact

8.1 Summary Statement Cost, Cost-Effectiveness and Budget Impact

A decision analytic model has been developed to quantify the cost-effectiveness of PVP and PBK compared to CT using incremental quality-adjusted life years (QALY), with univariate and probabilistic sensitivity analyses evaluating uncertainty and specific assumptions in the model. Results are presented as incremental cost-effectiveness ratios (ICER) and a hypothetical willingness-to-pay threshold was set at CHF100,000 per QALY gained.

The clinical evaluation found limited quality of life (EQ-5D) differences between PVP and CT/sham at 12 months, however, the results were subject to considerable heterogeneity – an effect potentially attributable to fracture age. For example, there were no significant differences in EQ-5D in trials evaluating fractures up to 1 year old (i.e. includes acute and sub-acute fractures, reflective of current Swiss reimbursement) (Buchbinder 2009). Thus, the intervention was not cost-effective for this patient population given the intervention had a higher cost when compared to CT.

Significant improvements in EQ-5D were observed at some time points following PVP in patients with acute fracture (i.e. less than 8 weeks old) (VERTOS II). Therefore, the cost-effectiveness model was developed in this sub-group. In fractures younger than 8 weeks, the estimated ICER for PVP vs CT was CHF19,669 per QALY at 1 year using adjusted baseline results of the VERTOS II trial, and was less than the hypothetical willingness-to-pay threshold at all evaluated timepoints.

The PBK model was informed by the FREE trial as it was the only RCT evaluating EQ-5D. The estimated ICER for PBK vs CT was CHF18,405 per QALY at 1 year.

Probabilistic sensitivity analyses (PSA) determined with 85% probability that PVP was superior (or cost-effective) compared to CT using the adjusted baseline results of VERTOS II over 12 months of follow-up. Results of the PBK PSA estimated an 87% probability that the intervention was superior compared to CT. Univariate sensitivity analyses indicated the ICERs for PVP and PBK were most impacted by the cost assumed for CT.

A budget impact analysis using three substitution scenarios (in which PVP and PBK were substituted with CT at different rates) was used to determine the financial implications of delisting the procedures. If PVP is delisted and all patients substitute to CT, then a net cost saving of CHF 6.5 million would occur in 2020. If PBK is delisted and all patients substitute to CT, then a net cost saving of CHF 3.8 million is estimated to occur in 2020. If both PVP and PBK were to be substituted by CT, CHF10.3 million would be saved in 2020, increasing to CHF13.5 million by 2024. Sensitivity analyses determined that the

financial impact was most sensitive to the substitution rate (i.e. 100%, 75% and 50%) and the number of physiotherapy visits included in CT. Inclusion of ambulatory PVP procedures did not significantly impact the analysis.

8.2 Methods

A decision analytic model was developed to quantify the cost-effectiveness of PVP and PBK com-pared to CT using incremental QALYs. The model was developed in TreeAge Pro (TreeAge Software, Inc, One Bank Street Williamstown, MA, 01267 USA).

Probabilistic sensitivity analysis was performed to account for uncertainty in the input parameters (See ***Table 68*** for assumptions). The analysis involved 10,000 iterations used to calculate 95% CI. The probability of the ICER being cost-effective was estimated using a hypothetical willingness-to-pay threshold of CHF100,000, and a cost-effectiveness acceptability curve was presented to demonstrate the probabilities of achieving a range of willingness-to-pay thresholds.

Annual costs for PVP, PBK and the comparator were taken from Swiss DRG costs.[187] ICERs for PVP and PBK versus CT were calculated using base case unit costs and health outcomes reported at 6, 12 and 24 months in a decision model. The base PVP versus CT comparison investigated the cost-effectiveness of the procedure in patients with acute fractures (less than 8 weeks) based on results of the VERTOS II trial.

The economic analysis of PBK utilised data from the 24 month FREE trial, which included patients with fractures of less than 3 months.[168] This is the only RCT of PBK that reported EQ-5D.

8.2.1 Economic Modelling Background

Review of economic literature

Economic studies were identified as part of the systematic literature searches outlined in ***Section 6.1***. In addition, supplementary literature searches were conducted to identify published economic analyses of PVP and PBK in EMBASE.com (EMBASE and MEDLINE) that may have been missed by the original searches. The supplementary search strategy involved search terms related to vertebroplasty, kyphoplasty and fractures (***Table 66***) and was undertaken on 22 January 2020. This resulted in 356 titles being identified. Results of the search are presented in ***Section 17.8*** (***Appendix H***) with ***Table 132*** listing economic publications relevant to PVP and PBK. Many identified titles referred to PVP and PBK but did not provide economic analysis and were consequently excluded.

Table 66 Search terms used for the identification of economic studies

Element of clinical question	Search terms
Population	spinal fractures OR spine OR vertebra OR spin* OR spine OR sacrum AND fractures OR fractur*
Intervention	polymethyl methacrylates OR bone cement OR polymethylmethacrylate OR pmma OR methylmethacrylate OR mma OR calcium phosphate OR glass polyalkenoate OR vertebroplasty OR kyphoplasty OR balloon
Comparator (if applicable)	Not applicable
Outcomes (if applicable)	Not applicable
Other	Health economics OR economic aspect OR economics OR biomedical technology assessment OR economic evaluation OR health care cost OR technology assessment OR cost effectiveness analysis OR cost minimisation analysis OR cost minimization analysis OR cost utility analysis OR quality adjusted life year OR QALY
Limits	Remove duplicates

The search identified 2 previous systematic reviews. The first, by Borgström (2015), was a review of PVP studies that included 4 economic analyses.[188] The authors indicated that many different models have been employed, and results were influenced by time horizon assumptions, quality of life improvements following treatment and impacts on length of hospital stay. Martelli (2015) undertook a systematic review using MEDLINE, PASCAL, COCHRANE and National Health Service Economic Evaluation database up to early 2014.[189] Twenty-one studies met the inclusion criteria, with the authors concluding that the level of evidence in economic evaluations of PVP and PBK was low.

The review presented in ***Table 132*** reflects the findings of Borgström (2015)[188] and Martelli (2015).[189] Four studies related to PVP, while most included PVP or PBK compared to CT. Most economic models referenced Strom's (2010) PBK Markov model developed for the UK.[190] Assumptions from this model were largely taken from the FREE study.[168] Adverse events were not included and quality of life differences between arms at 1 year were assumed to linearly decline over 2 years. Many of the company submissions (e.g. Medtronic) to Stevenson's (2014) review of vertebral fracture treatment (on behalf of NICE) utilised this model structure.[191]

Other economic models presented in ***Table 132*** undertook cost-effectiveness of trial data (eg. Japanese and USA studies). A number of cost analyses of PBK, PVP and CT were identified using insurer databases from USA, Germany and Austria. PBK was generally found to be more expensive, to have fewer adverse events and to involve patients with fewer comorbidities.

Time horizons assumed for the economic models listed in ***Table 132*** vary, with 2 years and lifetime analyses being the predominant period of analysis. Given most clinical trials were limited to 1 year of follow-up, many of the economic models have included extrapolation. For PVP, the VAPOUR trial had

a follow-up of 6 months[2] and the VERTOS II had a follow-up of 12 months.[132] The FREE trial (PBK vs CT) followed patients for 24 months.[168]

Most of the cost-effectiveness models were developed in Europe. The costs used in the Stevenson (2014) models were from the UK,[191] while the VERTOS II trial used costs Dutch cost data.[132] Length of hospital stay assumed in the Ström (2010), Svedbom (2013) and Stevenson (2014) models was one of the input variables that exerted the largest effect on cost-effectiveness.[54 190 191] The studies assumed surgery led to 6 fewer bed days than non-surgical management. When no difference in bed days was assumed, the cost-effectiveness ratio increased more than six-fold. For example, in Svedbom (2010), the cost per QALY gained rose from €3,337 to €21,649.[192]

The modelling presented in this HTA report used Swiss DRG costs. As this report has taken a payer perspective, the procedure was costed using the per-DRG rate, which is largely insensitive to bed days. The procedure is reimbursed per separation. Costs per DRG are subject to sensitivity analyses to gauge how robust the model results are to the constant cost per DRG assumption.

Modelling approaches differed as to whether interventions impacted mortality. Stevenson (2014) noted that there is no statistically significant difference with respect to mortality between PVP and optimal pain management at 12 months.[191] The study further reported, however, that PVP prolonged life compared with the comparator, but the order of this impact in the UK was uncertain. Given the uncertainty about mortality impacts, the base economic analysis in this assessment assumes no difference in mortality for the intervention and comparator.

Models reviewed by Stevenson (2014) also included differences in rates of refracture.[191] However, other assessments such as the 2011 Medical Services Advisory Committee (MSAC) assessment concluded subsequent vertebral fracture risk is poorly understood.[13] Further, Ström (2010) and Svedbom (2013) did not incorporate adverse events in their economic models, largely due to limited data.[190 192] The most common adverse event is cement leakage, which is mostly asymptomatic. Non-cement-related adverse events purportedly constitute less than 5% and 2% of all adverse events for PVP and PBK, respectively.[193 194] Given the uncertainty in both the frequency and consequences of adverse events, they are not included in the base case economic analyses. Rather, their impact is evaluated using sensitivity analyses.

In summary, the literature review identified a limited number of economic analyses of PVP and PBK. Most modelling studies reference the PBK Markov model of Ström (2010).[190] This was based on data from the FREE trial and utilised a 2 year time horizon.[168] Some models included differences in mortality, adverse events and refracture. The extrapolation assumed convergence, where the difference at the 1 year follow-up was assumed to converge by the end of year 2 and the mid-point utility difference included

in the economic analysis. Some models included differences in mortality, adverse events and refracture. The Stevenson (2014) model did not include adverse events in the base case but provided a sensitivity analysis.[191] The costs of PBK were generally higher than PVP, however, older patients with more comorbidities generally used PVP.

Overview of economic model

A decision analytic model (summarised in ***Table 67***) was developed to estimate the expected costs and QALYs associated with PVP and PBK compared with CT for an average patient with acute osteoporotic fracture (younger than 8 weeks for PVP; younger than 3 months for PBK). A model evaluating both acute and older fractures together was not undertaken as no quality of life improvement (as inferred by EQ-5D) was found in the Buchbinder (2009) study, which enrolled patients who had fractures for a maximum of 12 months before enrolment.[7] PVP would not be cost-effective using data from this trial given no incremental clinical benefit was identified and intervention costs were greater than those of the comparator.

Table 67 Summary of the economic evaluation

Perspective	This economic evaluation will be conducted from the perspective of the payer.
Patient population	The base analysis includes patients with acute vertebral fracture. PVP fracture age younger than 8 weeks, PBK younger than 3 months
Intervention	PVP and PBK
Comparator	Conventional treatment, or non-surgical treatments (including optimal medical therapy, physiotherapy or bracing)
Type of economic evaluation	Cost-utility analysis
Sources of evidence	Trials, Swiss DRG costs
Time horizon	6, 12 and 24 months.
Outcomes	Quality-adjusted life years/ life years gained
Methods used to generate results	Decision model
Software packages used	Excel, TreeAge Pro (PSA)

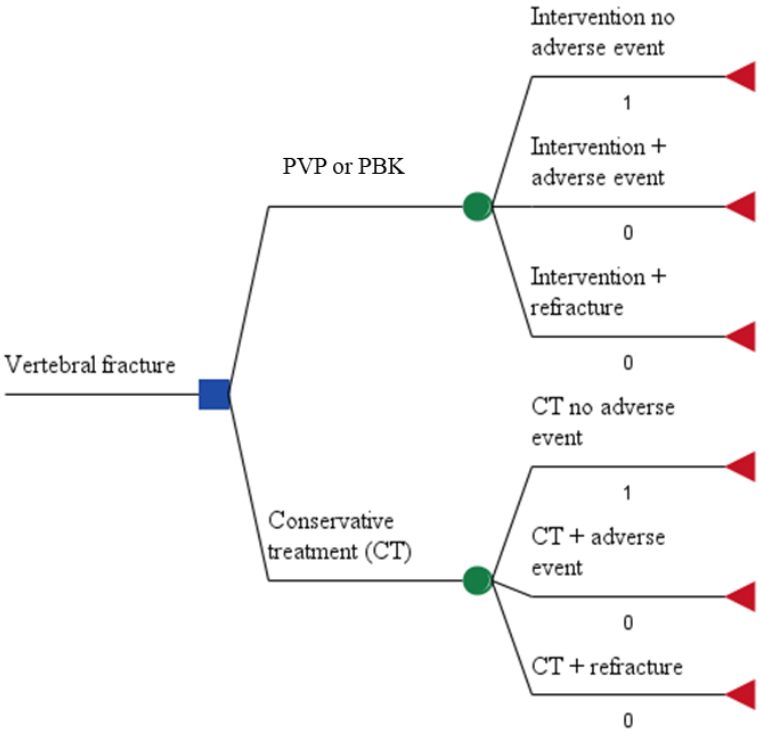
Abbreviations
DRG = diagnosis-related group, **PBK** = percutaneous balloon kyphoplasty, **PSA** = probabilistic sensitivity analysis, **PVP** = percutaneous vertebroplasty.

Type of economic evaluation

A decision analytic model was developed in Excel for the deterministic analysis, and in TreeAge for the PSA. The structure of the model is outlined in ***Figure 37***. PVP and PBK were compared to CT, without adverse events or refracture in the base model. Sensitivity analyses were provided where refracture, adverse events and different utility estimates were modelled. Costs and benefits were quantified for patients with painful OVCF younger than 8 weeks for PVP, and younger than 3 months for PBK.

Intervention and comparator

Effectiveness benefits of PVP and PBK versus CT were captured in the model as incremental QALYs. The comparator (CT) is a heterogeneous intervention involving bed-rest, pain medication and physiotherapy (***Section 3.2***). Swiss-DRG costs were derived from Schweizerische Operationsklassifikation (CHOP) codes for PBK, PVP and CT in Switzerland for the modelled interventions and comparator.[195]



Abbreviations
CT = conservative treatment, **PBK** = balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.

Figure 37 Decision tree for the cost utility model

Sources of evidence

The clinical evidence in ***Section 7*** compared PVP with CT (including optimal medical therapy, physiotherapy or bracing) or sham procedures, and PBK with CT. No studies were identified comparing PBK to a sham procedure. Outcomes for pain (VAS or NRS) and quality of life (QUALEFFO and EQ-5D) were reported.

For PVP, the trials of VERTOS II[132] and VAPOUR[2] provided outcomes used to measure incremental quality of life for fractures younger than 8 weeks. The same outcomes were reported for PBK based on the FREE trial.[168] The economic model estimated ICERs at 6 and 12 months for PVP, and at 6, 12 and 24 months for PBK. These timepoints correspond with follow-up in the key trials. The base case assumed no refracture or adverse event impacts, given non-significant differences were found in the clinical evaluation (see ***Section 7.8***). Sensitivity analyses are presented to gauge the robustness of model results to changes in this assumption.

Methods used to generate results

The economic analysis took the perspective of the payer. This perspective was recommended in the Swiss FOPH HTA assessment template. Health service costs were valued at 100% of Swiss-DRG items.

Univariate and probabilistic sensitivity analyses were undertaken. High and low utility values for the univariate analysis were taken from the VERTOS II trial for PVP[132] and the FREE trial for PBK.[168] PVP, PBK and CT costs were varied using standard deviations derived from Swiss DRG costs, and an alternate formulation of CT costs based on 9 physiotherapy visits per patient. The univariate analysis includes scenarios where refracture and adverse events are included.

The probabilistic sensitivity analysis was based on parameters and distribution assumptions included in ***Table 68***. Cost assumptions were included as normal distributions (average and standard deviations based on Swiss DRGs), while utility estimates were included as beta distributions for PBK and triangular distributions for PVP.

Outcomes

Health-related quality of life outcomes were reported using EQ-5D. EQ-5D values were taken from VAPOUR[2] and VERTOS II trials[132] to determine the cost-effectiveness of PVP, and from the FREE trial for PBK.[168]

8.3 Evidence Table

Model assumptions were derived for costs and QALY health outcomes, and are summarised in ***Table 68*** along with sources and the derivation of each assumption.

Table 68 Summary of evidence for the economic evaluation

Assumption	Value			Source of Evidence and Comments
Cost				
PVP	**Base**	**Standard deviation**		
PVP DRG weight in the base case.	11,163	5,284		The base case uses the "Other interventions on the spine, age> 15 years" Code I10C weight from Swiss DRG Datenspiegel 8.0. Accessed 30 October 2020.[187]
PBK				
PBK DRG weight in the base case.	11,163	5,284		The base case uses the "Other interventions on the spine, age> 15 years" Code I10C weight from Swiss DRG Datenspiegel 8.0. Accessed 30 October 2020.[187]
CT				
CT DRG weight in the base case.	9,039	6,343		The base case uses the "Bone diseases and arthropathies, age> 15 years and more than 1 day of occupancy" Code I69B weight from Swiss DRG Datenspiegel 8.0. Accessed 30 October 2020.[187]
Incremental utility outcome for baseline analysis				
PVP vs CT	**Base**	**High**	**Low**	
6 months	0.03	-	-	Incremental QALYs calculated using the VERTOS II study for PVP.[132] The 1 year incremental utility estimate reported by the authors included an adjustment for unequal baseline values. It was used as the 1 year incremental QALY estimate in base calculations and included as a triangular distribution in the PSA. The authors did not present EQ-5D estimates, standard errors or confidence intervals for adjusted intervention and comparator arms of the study, so beta distributions could not be estimated. High and low values were included in a triangular distribution, corresponding with the 95% CI.
12 months	0.11	0.18	0.04	
PBK vs CT				
6 months	0.06	-	-	Incremental QALYs for PBK are derived from FREE.[168] The study presented average treatment effect over 24 months as 0.10 (0.05, 0.15) in Table 3, p. 974.143 EQ-5D outcomes at each period of follow-up are included as beta distributions in the PSA and incremental QALYs over 12 months estimated. High and low estimates of incremental gains correspond with 95% CI (0.08, 0.15).
12 months	0.12	0.15	0.08	
24 months	0.21	-	-	

Abbreviations
CT = conservative therapy**, CHF** = Swiss franc, **DRG** = diagnosis-related group, **EQ-5D** = EuroQol 5 dimensions questionnaire, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty**, PSA** = probabilistic sensitivity analysis, **QALY** = quality-adjusted life year, **FOPH** = Federal Office of Public Health.

8.3.1 Applicability of Trials

Characteristics of patients comprising the clinical evidence are compared with circumstances of use in Switzerland (summarised in ***Table 69***).

Table 69 Features of PVP and PBK patient populations across the included studies

Parameter	Value	Sources/Comments
Demographics	73-81 years predominantly women	Key PVP trials include VAPOUR and VERTOS II. The VAPOUR trial included patients with an average age 80 in the PVP arm and 81 in the sham arm, and approximately 68–79% were women.[2] The VERTOS II trial included patients with an average age of 75.2 in the PVP arm and 75.4 in the CT arm, and 69% were women.[132] Key PBK trials include the FREE trial. These patients were average age 72.2 in the PBK arm and 74.1 in the CT arm, and 77% were women.[168]
Clinical characteristics	Back pain score ≥ 4 points for FREE, NRS pain score ≥ 7 for VAPOUR[2], VAS pain > 5 for VERTOS II[132] Fracture age < 3 months for FREE, < 6 weeks for VAPOUR[2] and VERTOS II[132]	VERTOS II included patients with vertebral fracture (at least 15% height loss) and pain for up to 6 weeks (VAS ≥ 5).[132] Similar criteria were used for VAPOUR, with fracture age younger than 6 weeks and NRS pain >7.[2] Buchbinder (2009) included patients with back pain duration up to 12 months.[7] FREE was undertaken at 21 sites in 8 countries, February 2003–December 2005.[168] Patients with 1–3 vertebral fractures were eligible for enrolment (confirmed by MRI and 1 with 15% loss of height). Participants needed a back-pain score of ≥ 4 points on 0–10 scale and fracture age younger than 3 months.
PVP	Hospital inpatients 57% of all patients in VAPOUR[2] [222]	PVP involved the use of 11-gauge or 13-gauge PVP needle using unipedicular or bipedicular technique with fluoroscopic guidance in the VAPOUR trial.[2] Patients were provided intravenous midazolam and fentanyl. VERTOS II involved a similar procedure, with 2 11- or 13-gauge bone-biopsy needles placed transpedicularly in the fractured vertebrae.[132]
PBK	Inpatient	Delivery of PBK in the FREE trial undertaken with introducer instruments, inflatable bone tamps, and PMMA bone cement using a percutaneous, bilateral, transpedicular or extrapedicular approach.[168] Most procedures done under general anaesthesia.
CT	Standard practice in participating centres.	The VAPOUR placebo procedure was designed to simulate PVP, with participants receiving usual medical care directed by their attending physicians.[2] The VERTOS II control arm involved typical CT. Analgesia was optimised and all patients were prescribed bisphosphonates, calcium supplementation and vitamin D.[132] All patients in the FREE trial were provided with analgesics, bed rest, back braces, physiotherapy, rehabilitation programmes and walking aids according to standard practice of participating hospitals.[168]

Abbreviations
CT = conservative therapy**, MRI** = magnetic resonance imaging, **NRS** = numerical rating scale, **PBK** = percutaneous balloon kyphoplasty, **PMMA** = polymethyl methacrylate, **PVP** = percutaneous vertebroplasty**, VAS** = visual analogue scale.

Demographics (age and gender)

Vertebral fracture incidence data by age and gender, as outlined by Svedbom (2014),[54] and the age structure of the Swiss population in 2020 are combined in ***Figure 38***.[196] The average age for trial participants in key studies was 73–81 years, and the procedure was predominately performed in women.

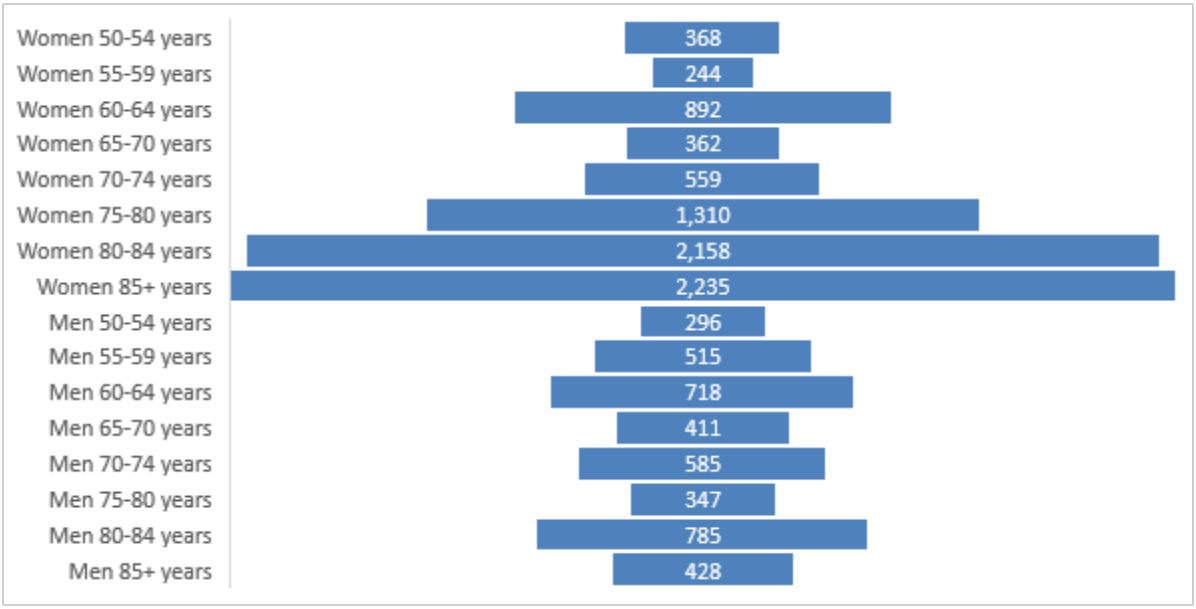


Figure 38 Estimated incidence of vertebral fracture cases in Switzerland by gender/age, 2020

Source
Compiled from Swiss population data and fracture incidence data.[54] [196]

Age and gender profiles for the key trials used for estimating health outcomes in the economic analysis were similar to what could be expected in Switzerland. It is evident that women 75 years and older accounted for most vertebral fracture cases. The Swiss DRG "I10C – other interventions on the spine, age > 15 years" had utilisation of 19.4%, 23.6% and 13.2% for the age groups 60– 69, 70–79 and 80+ years, respectively.[195]

Clinical characteristics

PVP is currently reimbursed without restriction in Switzerland.[15] The Buchbinder (2009) trial included patients with fractures ages of up to 1 year, which may reflect the current Swiss patient population for PVP given there are no fracture age restrictions.[7] The PVP base economic model included data from the VAPOUR and VERTOS II trials where patients had fractures younger than 8 weeks.[2] [132]

Patients were eligible for inclusion in the FREE trial if they had a back-pain score of 4 points, but the trial excluded those with fractures older than 3 months.[168] This was the only PBK trial, so no sub-group analysis is provided for fracture age.

Setting

Most PVP procedures, and all PBK procedures, are delivered in a hospital in-patient setting in Switzerland. Sensitivity analyses were undertaken using a range of cost weights to account for lower cost ambulatory delivery of PVP.

8.3.2 Utility Measures in PVP and PBK Trials

Health outcomes reported in the clinical evaluation included quality of life measures such as EQ-5D, back-specific functional status questionnaires, pain reporting, analgesic use, all-cause mortality, adverse events (mainly cement leakage) and the incidence of vertebral fractures. Health outcomes reported in PVP and PBK studies outlined in the clinical evidence review are summarised in ***Table 70*** and the length of follow-up in ***Table 71***. The rationales for selecting the FREE trial for PBK[168]; along with VERTOS II[132] for the PVP economic model are described.

QALY was the main outcome included in the cost-utility analysis, typically estimated using the results of the EQ-5D questionnaire. The reporting of this outcome across PVP and PBK studies is presented in ***Table 70***.

The FREE trial reported EQ-5D, which can be used to estimate utility and QALY differences for PBK and CT.[168]

PVP studies that included patients with factures of less than 8 weeks, and that reported EQ-5D, included VAPOUR[2], Rousing (2009)[96] and VERTOS II;[132] however, Stevenson (2014)[191] noted that Rousing (2009) did not report this outcome for all participants, with scores presented for only 58% of the intervention group and 71% of the control group.[96] Further, the EQ-5D values at baseline were different, with scores of 0.356 for the PVP arm and 0.083 for the CT arm ($p = 0.04$).[96] Given these discrepancies, Rousing (2009) was not included in the utility calculations for PVP vs CT.

Utility outcomes included in the base economic model for the PVP intervention were taken from the VERTOS II trial, which included patients with fractures younger than 8 weeks.[132] The 6 month utility gain for this study was similar to VAPOUR at 6 months,[2] noting VERTOS II also reported utility at 12 months.[132] At 6 months, the unadjusted utility difference was 0.03 for PVP compared to CT in VERTOS II, and a difference of 0.02 was estimated for VAPOUR over the same period. A sensitivity analysis is presented in ***Section 8.4*** using the VAPOUR utility results.

Buchbinder (2009)[7] and Kallmes (2009)[4] reported EQ-5D data for PVP on fractures up to 1 year of age. Stevenson (2014)[191] noted that Buchbinder (2009) added this outcome to their protocol in June 2005 to allow comparison with the Kallmes (2009) study,[4] therefore EQ-5D scores were available for only 79% of PVP participants and 73% of controls.[7] Buchbinder (2009) found no difference in average utility between PVP and comparator arms. Based on this result, PVP was more costly and of similar

effectiveness. The dominance of the comparator in patient populations with fractures of greater than 8 weeks is discussed in the conclusions section.

Table 70 Key outcomes assessed in RCTs evaluating PVP or PBK

	PBK		**PVP (fracture age <8 weeks)**			**PVP (fracture age >8 weeks)**			
Source	FREE [168]	Liu (2010)[197]	VERTOS II[132]	VAPOUR[2]	Rousing (2009)[96]	Buchbinder (2009)[7]	Kallmes (2009)[4]	Blasco (2012)[8]	Farrokhi (2011)[135]
Pain	VAS pain	VAS pain	VAS pain	NRS, VAS pain	VAS pain	0–10 scale	0–10 scale	VAS pain	0–10 scale
Analgesic use	X	-	X	X	-	X	X	X	-
Incidence of new fractures	X	-	X	-	X	X	-	X	X
Back-specific functional status	RDQ	-	-	RDQ	-	RDQ	RDQ	-	OW-LBP scale
Other QoL	SF-36 PCS, EQ-5D	-	EQ-5D, QUALE FFO	EQ-5D, QUALE FFO	SF-36 PCS MCS, DPQ, EQ-5D, Barthel Index, MMSE	QUALE FFO, AQoL, EQ-5D	SF-36 PCS, EQ-5D	QUALE FFO-41	-
All-cause mortality	X	-	X	X	x	-	-	X	X
AE	Leaks	-	Leaks	X	Leaks	Leaks	X	Leaks	Leaks

Abbreviations

AE = adverse event**, AQoL**, = Australian quality of life, **DPQ** -= Dallas Pain Questionnaire**, EQ-5D** = EuroQol 5 dimension questionnaire**, PBK** = balloon kyphoplasty, **MCS** = mental component score, **MMSE** = mini-mental state examination, **NRS** = Numeric Rated Scale (NRS) back pain, **OW-LBP** = Oswestry lower back pain, **PBK** = percutaneous balloon kyphoplasty, **PCS** = physical component score, **PVP** = percutaneous vertebroplasty, **QoL** = quality of life, **QUALEFFO-41** = quality of life questionnaire of the European Foundation for Osteoporosis–41 questions, **RDQ** = Roland–Morris disability questionnaire, **SF-36** = short form 36 questionnaire, **VAS** = visual analogue scale.

Length of follow-up

Duration of follow-up in the PVP and PBK studies varied. For example, Kallmes (2009) reported outcomes for 1 month[4] and Farrokhi (2011) reported outcomes for 3 years.[135] The FREE trial for PBK followed patients for 2 years,[168] while Liu (2010) was limited to 6 months.[197] Attrition bias was an issue in some trials. Stevenson (2014)[191] noted the studies of Buchbinder (2009),[7] Farrokhi (2011),[135] Kallmes (2009),[4] Rousing (2009)[96] and VERTOS II[132] had at least 80% of participants at the last follow-up, while Blasco (2012) and the FREE trial had 76%[8] and 78% of participants[168] at 12 months, respectively. Further, there were differences in the number of PBK and CT patients at the final follow-up in the FREE trial – 83% and 74%, respectively.[168]

Economic model results were calculated at 6, 12 and 24 months based on the follow-up of key trials,[2] [132] [168] and outcomes were subject to sensitivity analysis to gauge how robust model results are to different lengths of follow-up.

Table 71 Length of follow-up across RCTs evaluating PVP or PBK

	PBK		PVP (fracture age < 8 weeks)			PVP (fracture age > 8 weeks)			
Trial	FREE [168]	Liu (2010)[197]	VERTOS II[132]	VAPOUR[2]	Rousing (2009)[96]	Buchbinder (2009)[7]	Kallmes (2009)[4]	Blasco (2012)[8]	Farrokhi (2011)[135]
<2 weeks		X	X	X	X	X	X	X	X
1 month	X		X	X	X	X	X		
2 months								X	X
3 months	X		X	X	X	X			
6 months	X	X	X	X	X	X		X	X
12 months	X		X					X	X
24 months	X								X
36 months									X

Abbreviations
PBK = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.

PVP utility

VERTOS II[132] and VAPOUR[2] trials evaluated PVP in fractures younger than 8 weeks of age. The VERTOS II trial recruited patients with back pain for 6 weeks or less from hospitals in the Netherlands and Belgium.[132] The EQ-5D scores from VERTOS II are presented in ***Table 72***. The incremental QALYs of 0.03 and 0.08 were estimated for 6 and 12 months for PVP, respectively. Different baseline utilities were also reported in VERTOS II. An average utility of 0.27 was reported for PVP (0.03) compared to 0.38 (0.03) for CT,[132] noting the difference was less than in Rousing (2009).[96] The authors used regression analysis to adjust for baseline differences, and estimated PVP accumulated an additional 0.01 (95% CI 0.014, 0.006) QALY at 1 month and 0.108 (95% CI 0.177, 0.040) by 1 year. The 0.108 incremental QALY gained is used in base economic calculations at 12 months for the PVP versus CT analysis, noting the value was calculated using adjusted utilities. A univariate sensitivity analysis was conducted using the 0.08 unadjusted mean gain at 12 months for PVP (***Table 72***).

Table 72 EQ-5D score and corresponding quality-adjusted life years from the VERTOS II trial

	Follow-up duration	EQ-5D		QALYs			Calculation
		PVP Mean ± SD	CT Mean ± SD	Time Weight	PVP	CT	
A	Baseline	0.27 ± 0.03	0.38 ± 0.03	-	-	-	
B	2 weeks	0.6 ± 0.3	0.5 ± 0.3	0.04	0.02	0.02	EQ-5D [A+B] / 2* time weight
C	1 month	0.6 ± 0.2	0.5 ± 0.3	0.05	0.03	0.02	EQ-5D [B+C] / 2* time weight
D	3 months	0.6 ± 0.3	0.6 ± 0.3	0.16	0.10	0.09	EQ-5D [C+D] / 2* time weight
E	6 months	0.7 ± 0.3	0.6 ± 0.3	0.25	0.16	0.15	EQ-5D [D+E] / 2* time weight
F	12 months	0.7 ± 0.3	0.6 ± 0.3	0.50	0.35	0.30	EQ-5D [E+F] / 2* time weight
G	Cumulative QALY 6 months (PVP-CT = 0.03)				0.30	0.28	QALY [B+C+D+E]
H	Cumulative QALY 12 months (PVP-CT = 0.08)				0.65	0.58	QALY [B+C+D+E+F]

Abbreviations
EQ-5D = EuroQol 5 dimension questionnaire, **CT** = conservative treatment**, PVP** = percutaneous vertebroplasty, **QALY** = quality-adjusted life year.
Source
Calculated using VERTOS II data.[132]

Results from the VAPOUR trial are presented in ***Table 73***. Baseline EQ-5D scores for PVP and CT were similar at 0.60 and 0.59.[2] However, there were more patients lost to follow-up in the PVP arm than the CT arm, which may bias results toward PVP. The VAPOUR incremental QALY gain of 0.02 (***Table 73***) is similar to the unadjusted baseline gain of VERTOS II of 0.03 at 6 months (***Table 75***). VERTOS II

results are used in the base economic model (See ***Table 67***) given the longer follow-up of the trial. An ICER was estimated using the VAPOUR trial at 6 months and compared to VERTOS II results.

Table 73 EQ-5D score and corresponding quality-adjusted life years from the VAPOUR trial

		EQ-5D			QALYS			Calculation
	Length of follow-up	**PVP**	**CT**	**Mean Difference (95%CI)**	**Time weight**	**PVP**	**CT**	
A	Baseline	0.60	0.59	-	-	-	-	-
B	3 days	0.69	0.65	-0·03 (-0·07, 0·05)	0.01	0.01	0.01	EQ-5D [A+B] / 2* time weight
C	2 weeks	0.69	0.68	-0·01 (-0·06, 0·03)	0.03	0.02	0.02	EQ-5D [C+B] / 2* time weight
D	1 month	0.75	0.70	-0·05 (-0·09, 0.0)	0.05	0.03	0.03	EQ-5D [D+C] / 2* time weight
E	3 months	0.75	0.71	-0·03 (-0·08, 0·01)	0.17	0.13	0.12	EQ-5D [D+E] / 2* time weight
F	6 months	0.80	0.74	-0·06 (-0·10, -0·01)	0.25	0.19	0.18	EQ-5D [E+F] / 2* time weight
G	Total Cumulative QALY 6 months (PVP-CT = 0.02)					0.38	0.35	QALY [B+C+D+E+F]

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **QALY** = quality-adjusted life year, **PVP** = percutaneous vertebroplasty.
Source
VAPOUR trial, Clark (2016).[2]

The EQ-5D results in Buchbinder (2009) were used to infer the QALY of acute and older (greater than 8 weeks) fractures following PVP.[7] A total of 78 participants were enrolled, with outcomes being assessed at 1 week to 6 months. No EQ-5D differences were reported at any time of follow-up (see ***Table 74***), consequently the incremental QALY was zero.

Table 74 EQ-5D score from the Buchbinder (2009) trial

Length of follow-up	PVP Mean ± SD	CT Mean ± SD	Mean difference (95% CI)	p value
Baseline	0.3 ± 0.3	0.3 ± 0.3		
Change 1 week	0.1 ± 0.3	0.1 ± 0.3	0.0 (-0.1 to 0.2)	NR
Change 1 month	0.1 ± 0.3	0.1 ± 0.3	0.0 (-0.1 to 0.1)	NR
Change 3 months	0.2 ± 0.3	0.2 ± 0.4	0.0 (-0.1 to 0.2)	NR
Change 6 months	0.2 ± 0.4	0.2 ± 0.4	0.0 (-0.1 to 0.2)	NR

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **QoL** = quality of life, **PVP** = percutaneous vertebroplasty, **SD** = standard deviation.
Source
Buchbinder (2009)[7]

Mortality

Six of the studies presented in ***Table 70*** report mortality.[2 8 96 132 135 168] None found any statistically significant differences. Stevenson (2014)[191] noted that data from the Blasco (2012)[8], Rousing (2009)[96] and VERTOS II[132] studies were combined by meta-analysis with studies reporting mortality at different timepoints. However, statistical significance was still not achieved when the data were pooled. The clinical evidence presented in ***Section 7.8*** reports similar findings, with no significant differences observed between PVP and CT, that is, 29 PVP deaths in 640 participants (4.5%), 36 CT deaths in 641 participants (5.6%) (RR 0.80; 95% CI 0.50, 1.29). Likewise, there were no differences in mortality across the PBK studies (9/149 patients and 7/151 patients in PBK and CT arms, respectively), therefore differences in mortality were not included in the economic model.

Adverse events

Stevenson (2014) noted that cement leakage associated with PVP ranged from nil in VERTOS to 72% in VERTOS II.[191] The clinical evidence presented in ***Section 7.8*** noted the incidence of cement leakage following PVP ranged from 1% to 37% of patients (weighted mean proportion of 16.7%). There were no significant differences in other adverse events when comparing PVP to CT or sham. The FREE trial reported cement extravasation in 51 of 188 (27%) of vertebrae treated with PBK in 48 patients.[168] The Stevenson (2014) economic model applied a QALY decrement for serious adverse events of 0.02.[191] This value was estimated assuming that the rates of mortality and morbidity were 1 in 1,000 and 1 in 100, respectively. A univariate sensitivity analysis was included for both the PVP and PBK models, where 16.7% of patients suffer a utility decrease of 0.02.[191] The impact on the estimated ICER was minor.

Refracture

There was no significant difference between PVP and CT with respect to the incidence of new clinical vertebral fractures (see ***Section 7.8***). The PVP group had 48 fractures in 418 participants (11.5%) and the CT group had 31 fractures in 422 participants (7.3%), resulting in a RR of 1.29 (95% CI 0.46, 3.62). The same observation was made for the incidence of radiographic fractures, with no evidence of important differences between groups in 7 studies. The key PBK study, FREE, did not report vertebral fracture for CT.[168]

Some of the economic models discussed earlier included an allowance for refracture. The Stevenson (2014) economic model included refracture risk based on the patients' bone mineral density T-score.[191] If bisphosphonate therapy was being used, a relative risk of 0.58 was applied. Svedbom (2013) included a sensitivity analysis – where PBK was assumed to increase the risk of the first additional fracture by 50%, although patients using bisphosphonates also had a risk reduction.[192]

A sensitivity analysis was included in the PVP and PBK models, where the interventions were associated with a 5% increase in fracture risk. An additional CT unit cost is applied for refracture cases. The inclusion of this assumption has minimal impact on the estimated ICER.

PBK utility

The FREE trial and Liu (2010) studies investigated health outcomes for PBK.[168] [197] Liu (2010) included patients with fractures younger than 43 days (mean duration 16–17 days).[197] There was no statistical difference in VAS pain scores between PVP and PBK groups at any stage. The study only compared PVP to PBK, so CT is not included in the economic analysis.

The FREE trial included 300 patients with similar baseline SF-36, PCS, EQ-5D, RDQ and back pain scores between PBK and CT arms.[168] EQ-5D responses throughout the trial are outlined in ***Table 75***, along with estimated incremental QALYs (change since baseline) for each arm. Utility differences were 0.13 at 6 months, 0.1 at 1 year and 0.08 at 2 years.[168] At 6, 12 and 24 months, incremental QALY gains for PBK are estimated to be 0.06, 0.12 and 0.21, respectively, which are included in the economic model.

Table 75 PBK vs CT: EQ-5D differences reported in the FREE trial

		EQ-5D		QALYs			Calculation
	Follow-up time	PBK	CT	Time weight	PBK	CT	
A	Baseline	0.16 (0.11–0.22)	0.17 (0.12–0.22)	-	-	-	
B	1 month	0.54 (0.49–0.60)	0.37 (0.31–0.42)	0.08	0.03	0.02	EQ-5D [A+B] / 2* time weight
C	3 months	0.59 (0.53–0.65)	0.49 (0.44–0.55)	0.17	0.09	0.07	EQ-5D [B+C] / 2* time weight
D	6 months	0.63 (0.57–0.68)	0.50 (0.45–0.56)	0.25	0.15	0.12	EQ-5D [C+D] / 2* time weight
E	12 months	0.61 (0.56–0.67)	0.51 (0.45–0.57)	0.50	0.31	0.25	EQ-5D [D+E] / 2* time weight
F	24 months	0.61 (0.56–0.67)	0.53 (0.47–0.59)	1.00	0.61	0.52	EQ-5D [F+E] / 2* time weight
G	Cumulative QALY 6 months (PBK-CT = 0.06)				0.28	0.22	Sum QALY [B+C+D]
H	Cumulative QALY 12 months (PBK-CT = 0.12)				0.59	0.47	Sum QALY [B+C+D+E]
I	Cumulative QALY 24 months (PBK-CT = 0.21)				1.20	0.99	Sum QALY [B+C+D+E+F]

Abbreviations
EQ-5D = EuroQol 5 dimension questionnaire, **PBK** = balloon kyphoplasty, **QALY** = quality-adjusted life year.
Source
Calculated using FREE trial data, Table 3, p. 974.[167]

8.3.3 Costs input of PVP, PBK and CT

Activities associated with PVP, PBK and CT are outlined in ***Table 76***. Costs associated with follow-up include general doctor consultations, prescriptions and specialist consultations. These were included in the MSAC 2011 assessment, along with an allowance for follow-up community care.[13] Pain management associated with vertebral fracture included simple analgesics such as paracetamol, NSAIDs, or the use of opiates for uncontrolled pain. All follow-up activities except pain management were assumed to have similar costs in the MSAC 2011 assessment. Differences in pain management costs were limited to AUD1,471 to 1,368 (CHF855–795) for PVP and CT.[13]

Table 76 Activities associated with PVP, PBK and CT

Work-up and staging	PVP	PBK	CT
Initial GP consultation	X	X	X
GP for prescriptions	X	X	X
Radiologist consultation	X	X	X
Bone densitometry	X	X	X
MRI	X	X	X
Pathology	X	X	X
Specialist	X	X	X
Procedure			
PVP kit	X		
PBK kit		X	
Intraoperative imaging	X	X	
Operator	X	X	
Anaesthetist	X	X	
Postoperative imaging	X	X	
Hospitalisation – AR-DRG I69B or AR-DRG I24Z	X	X	
Hospitalisation – AR-DRG I24Z	X	X	X
Analgesia, care and follow-up costs			
Follow-up GP consultation	X	X	X
Follow-up rheumatologist consultation	X	X	X
Community supportive care	X	X	X
Analgesia	X	X	X

Abbreviations
CT = conservative treatment, **GP** = general practitioner, **MRI** = magnetic resonance imaging, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.
Source
MSAC 2011, Table 43, p. 181[13]

The costs included in this report were calculated using Swiss-DRG costs (summarised in ***Table 77***). The PVP and PBK DRG weight in the base case was 0.988 (using the Swiss DRG "Other interventions on the spine, age > 15 years" I10C, from Swiss DRG costs presented on Datenspiegel). The CT DRG weight was 0.743 (using the "Bone diseases and arthropathies, age > 15 years and more than 1 day of occupancy" I69B, from Swiss DRG costs presented on Datenspiegel).[187]

The average hospital length of stay for each of the DRGs was 5.4 days for I10C and 6.9 days for I69B. The difference of 1.5 days was less than the 5.5-day difference in average length of stay in the VAPOUR trial,[2] and 6 days less than the stay assumed for PBK in the Strom (2010) economic model.[190] Costs were subject to sensitivity analysis, presented in the next section.

Table 77 Costs of PVP, PBK and CT

Intervention	Weight	Base (CHF)	Standard deviation (CHF)	Source
PVP cost				
PVP DRG using the I10C weight from Swiss DRG-Version 9.0. (2018/2020)	0.988	11,163	5,284	The base case uses the "Other interventions on the spine, age> 15 years" Code I10C weight from Datenspiegel.[187] The PSA uses a normal distribution and standard deviation.
PBK cost				
PBK DRG using the I10C weight from Swiss DRG-Version 9.0. (2018/2020).	0.988	11,163	5,284	The base case uses the "Other interventions on the spine, age> 15 years" Code I10C weight from Datenspiegel.[187]. The PSA uses a normal distribution and standard deviation.
CT cost				
CT DRG weight using I69B weight from Swiss DRG-Version 9.0. (2018/2020)	0.743	9,039	6,343	The base case uses the "Bone diseases and arthropathies, age> 15 years and more than 1 day of occupancy" Code I69B weight from Datenspiegel.[187]. The PSA uses a normal distribution and standard deviation.

Abbreviations
CHF = Swiss franc, **CT** = conservative treatment, **DRG** = diagnosis-related group, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.
Sources
187

8.4 Results: Cost-Effectiveness

8.4.1 PVP vs CT

ICER

The incremental cost and effectiveness of the PVP versus CT comparison at 6 and 12 months (using results from the VERTOS II trial[132]) are presented in ***Table 78***. The ICER for PVP was less than a hypothetical willingness-to-pay of CHF100,000 using results from VERTOS II. ICERs were CHF19,669 at 12 months and CHF84,847 per incremental QALY at 6 months.

Table 78 Incremental cost-effectiveness of PVP compared to CT

	Cost in CHF	Incremental cost	Incremental QALYs	ICER
6 months				
PVP	11,163	-	-	-
CT	9,039	2,124	0.03	84,847
12 months				
PVP	11,163	-	-	-
CT	9,039	2,124	0.11	19,669

Abbreviations
CT = conservative treatment**, ICER** = incremental cost-effectiveness ratio, **PVP** = percutaneous vertebroplasty, **QALY** = quality-adjusted life year.

Less incremental QALYs were accumulated at 6 months when compared to 12 months in the VERTOS II trial, therefore the ICER is lower at 12 months given the cost difference at 6 and 12 months is estimated to be the same. The VERTOS II incremental QALY gain for PVP was 0.03, which is similar to the VAPOUR trial which reported an incremental gain of 0.02. The ICERs at 6 months were CHF84,847 per QALY and CHF95,361 for VERTOS II and VAPOUR, respectively.

Univariate sensitivity analysis

Sensitivity of the results to different model assumptions was explored in univariate sensitivity analysis for the PVP and PBK models.

A tornado graph for PVP compared to CT at 12 months (using VERTOS II data[132]) is presented in ***Figure 39***. ICER estimates were most affected by the assumption that CT includes 9 physiotherapy treatments, upper and lower DRG costs (based on the standard deviations) and high and low utility gains from VERTOS II. The inclusion of refracture, adverse events and unadjusted EQ-5D scores at baseline had minimal impact on the results.

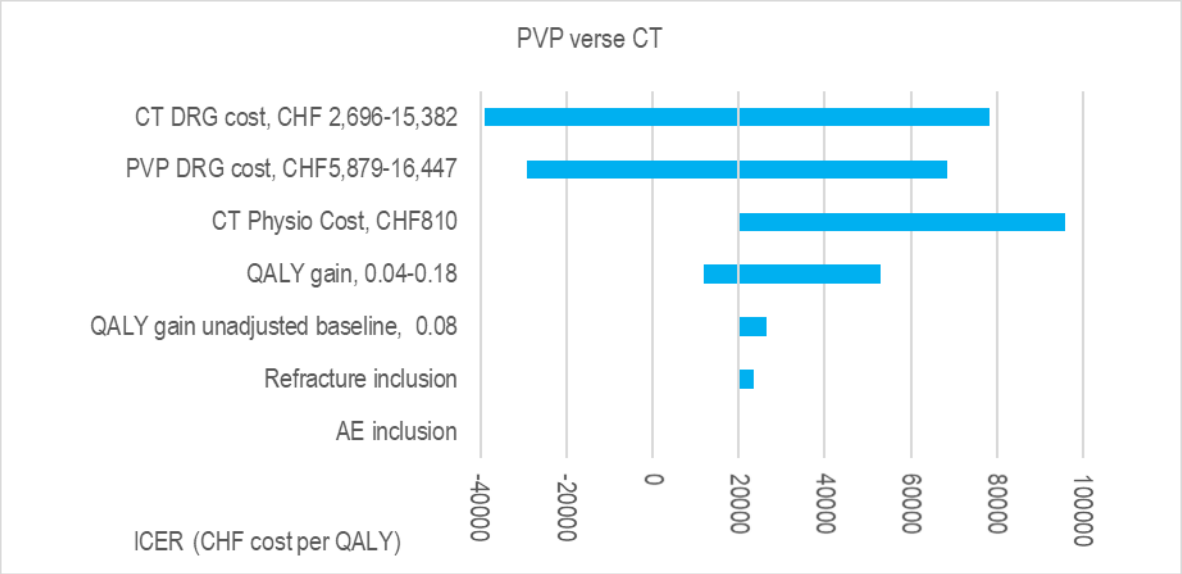


Figure 39 PVP compared to CT: incremental 12 month cost-effectiveness tornado graph

Abbreviations
CHF = Swiss franc, **CT** = conservative treatment, **ICER** = incremental cost-effectiveness ratio, **PVP** = percutaneous vertebroplasty, **QALY** = quality-adjusted life year.

Probabilistic sensitivity analysis

Inputs were specified as distributions (described in ***Table 68***). A mean expected ICER of CHF19,965 per QALY (95% CI, from PSA, CHF-147,805, CHF197,521, ***Figure 40***) was estimated for PVP compared to CT. Using a hypothetical willingness-to-pay threshold of CHF100,000 per QALY, PVP reported an 85% probability of being cost effective when compared with CT (incremental effectiveness >0, incremental cost >0).

The cost-effectiveness acceptability curve is presented for the PVP versus CT comparison at 12 months in ***Figure 41***. The graph presents the probability that the PVP will be cost-effective against the willingness-to-pay thresholds on the horizontal axis. It is evident that PVP has more than an 85% chance of being cost-effective using adjusted results from the VERTOS II trial at willingness-to-pay thresholds greater than CHF100,000.

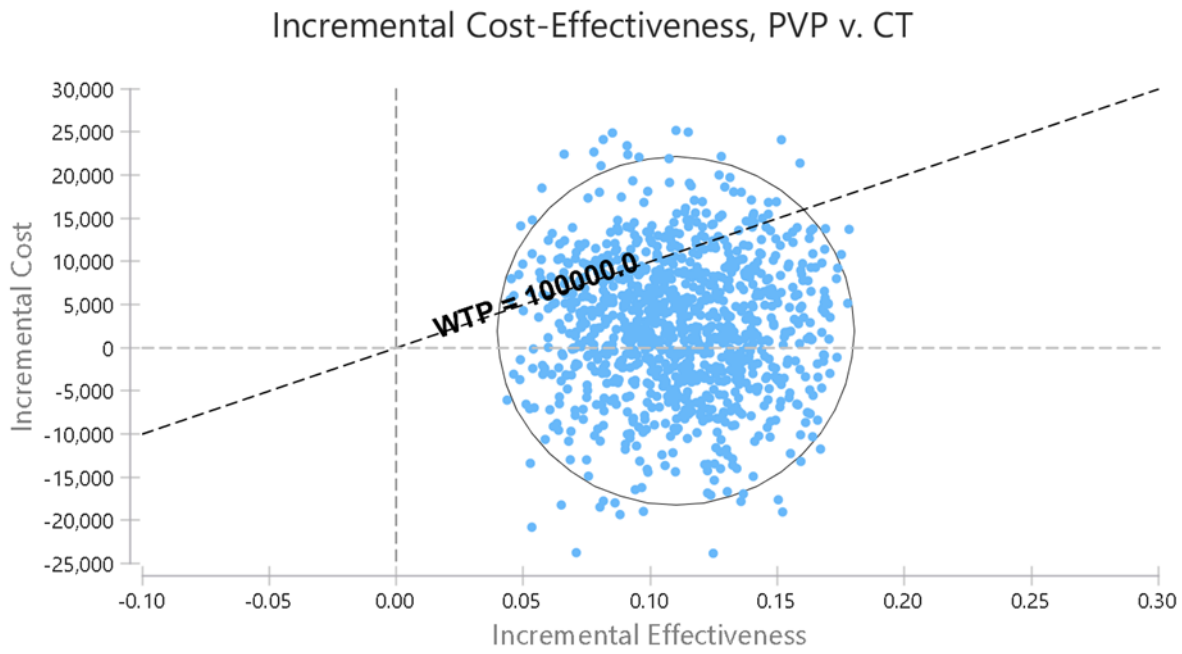


Figure 40 PVP compared to CT at 12 months: cost-effectiveness plane using adjusted VERTOS II results

Abbreviations
CT = conservative treatment, **PVP** = percutaneous vertebroplasty, **WTP** = willingness-to-pay (CHF).

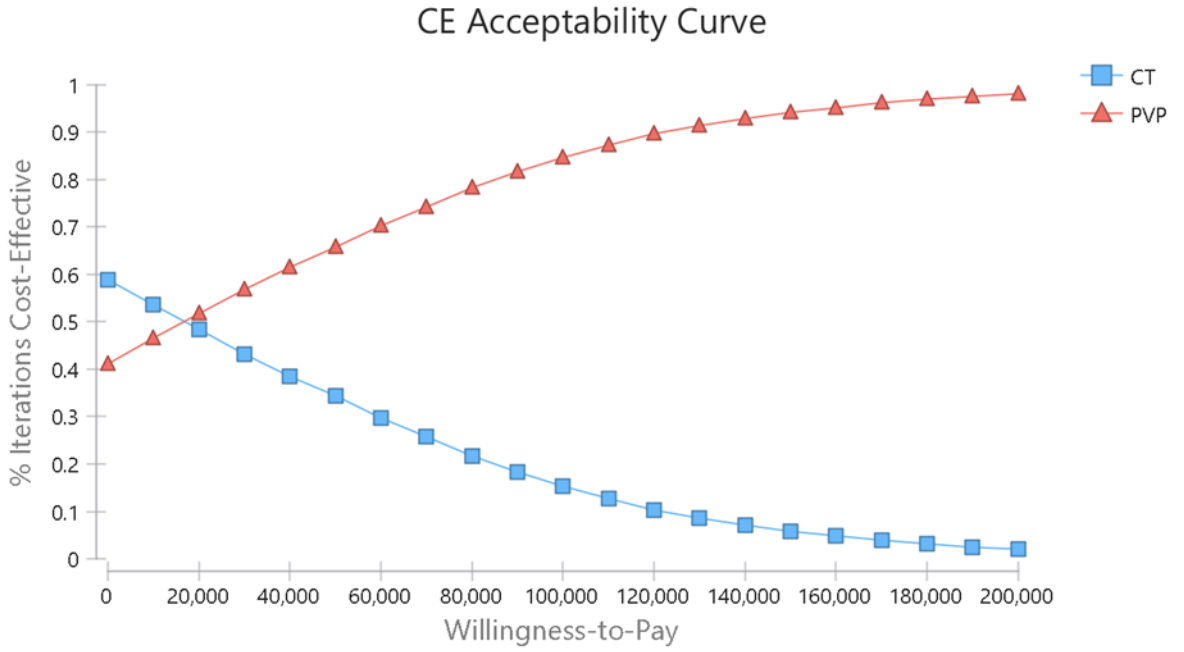


Figure 41 PVP compared to CT at 12 months: cost-effectiveness acceptability using adjusted VERTOS II results

Abbreviations
CE = cost effectiveness, **CT** = conservative treatment, **PVP** = percutaneous vertebroplasty.

8.4.2 PBK vs CT

ICER

The incremental cost and the incremental effectiveness of substituting PBK with CT at 6, 12 and 24 months (using results from the FREE trial[168]) are presented in ***Table 79***. The ICERs for the intervention of CHF10,341 at 24 months, CHF18,405 at 12 months and CHF36,678 at 6 months per incremental QALY were less than a hypothetical willingness-to-pay of CHF100,000.

Table 79 Incremental cost-effectiveness of PBK compared to CT

	Cost in CHF	Incremental cost	Incremental QALYs	ICER
6 months				
PBK	11,163	-	-	-
CT	9,039	2,124	0.06	36,678
12 months				
PBK	11,163	-	-	-
CT	9,039	2,124	0.12	18,405
24 months				
PBK	11,163	-	-	-
CT	9,039	2,124	0.21	10,341

Abbreviations
CT = conservative treatment**, ICER** = incremental cost-effectiveness ratio, **PBK** = percutaneous balloon kyphoplasty, **QALY** = quality-adjusted life year.

Univariate sensitivity analysis

Results for PBK vs CT at 12 months are presented in ***Figure 42***. ICER estimates were most affected by the assumption that CT includes 9 physiotherapy treatments, upper and lower DRG costs (based on standard deviations) and assumed utility gain has moderate to large effects. The inclusion of adverse events had minimal impact on the results.

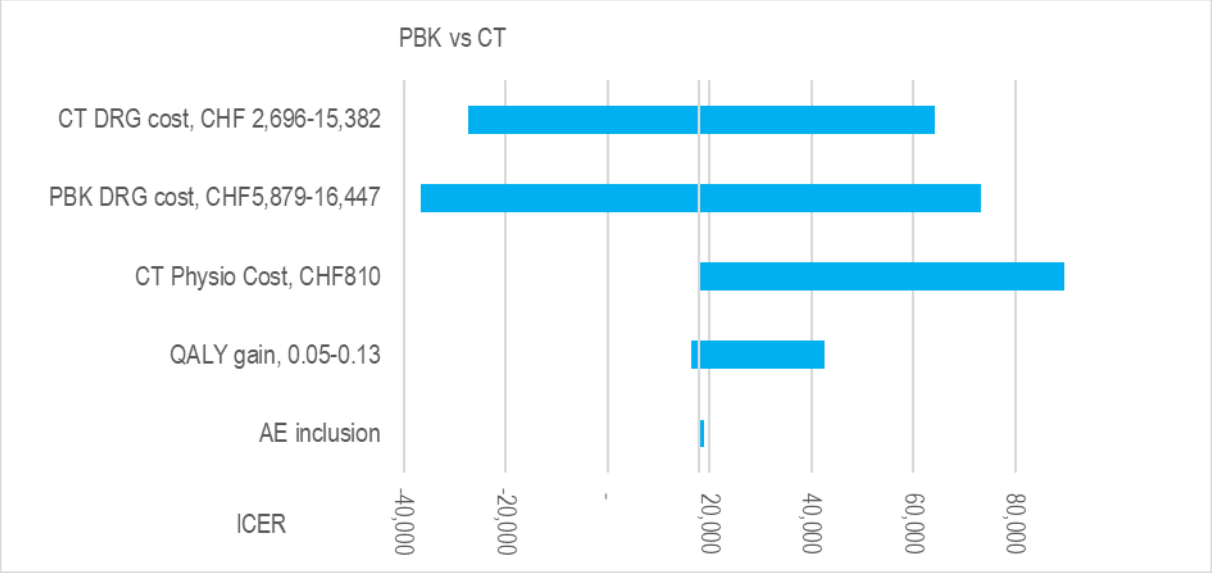


Figure 42 PBK compared to CT: incremental 12 month cost-effectiveness tornado graph

Abbreviations
CHF = Swiss franc, **CT** = conservative treatment, **ICER** = incremental cost-effectiveness ratio, **PBK** = percutaneous balloon kyphoplasty, **QALY** = quality-adjusted life year.

Probabilistic sensitivity analysis

Inputs were specified as distributions (described in ***Table 68***). A mean expected ICER of CHF18,183 per QALY (95% CI from PSA, CHF-130,211, CHF168,532, ***Figure 43***) was estimated for PBK compared to CT at 12 months. Using a hypothetical willingness-to-pay threshold of CHF100,000/QALY, there was an 87% probability that PBK is cost effective when compared with CT (Incremental effectiveness > 0, Incremental cost > 0).

The cost-effectiveness acceptability curve is presented for the PBK versus CT comparison at 12 months in ***Figure 44***. It is evident that PBK has an 87% chance of being cost-effective at willingness-to-pay thresholds of greater than 100,000 CHF.

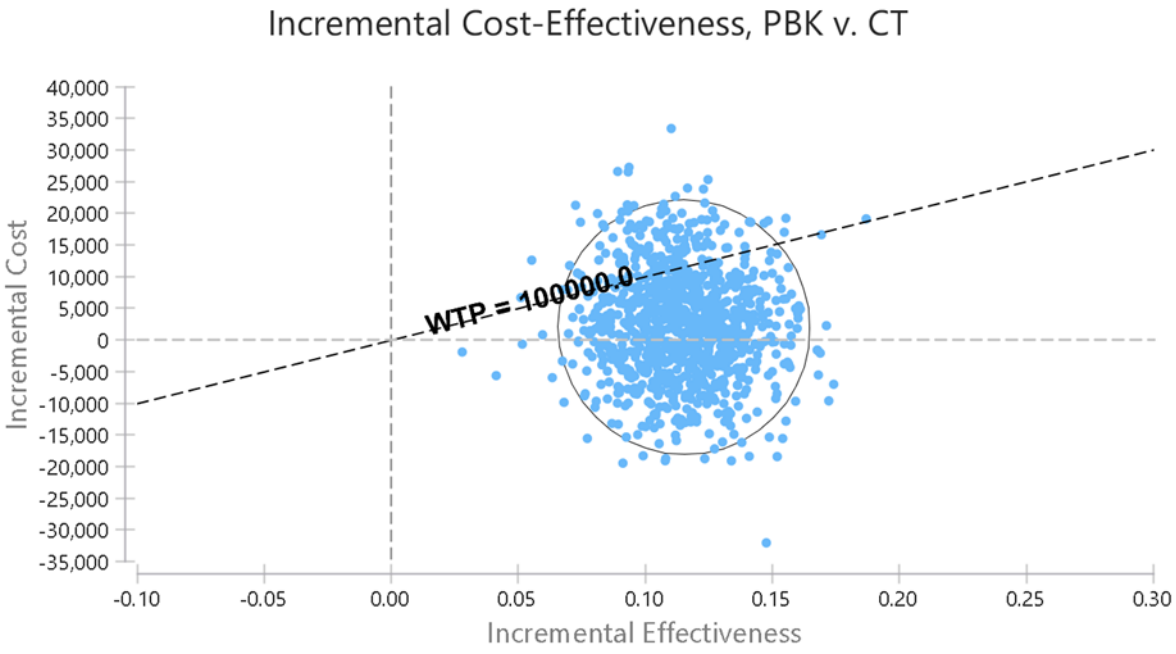


Figure 43 PBK compared to CT: 12 month cost-effectiveness plane

Abbreviations
CT = conservative treatment, **PBK** = percutaneous balloon kyphoplasty, **WTP** = willingness-to-pay (CHF).

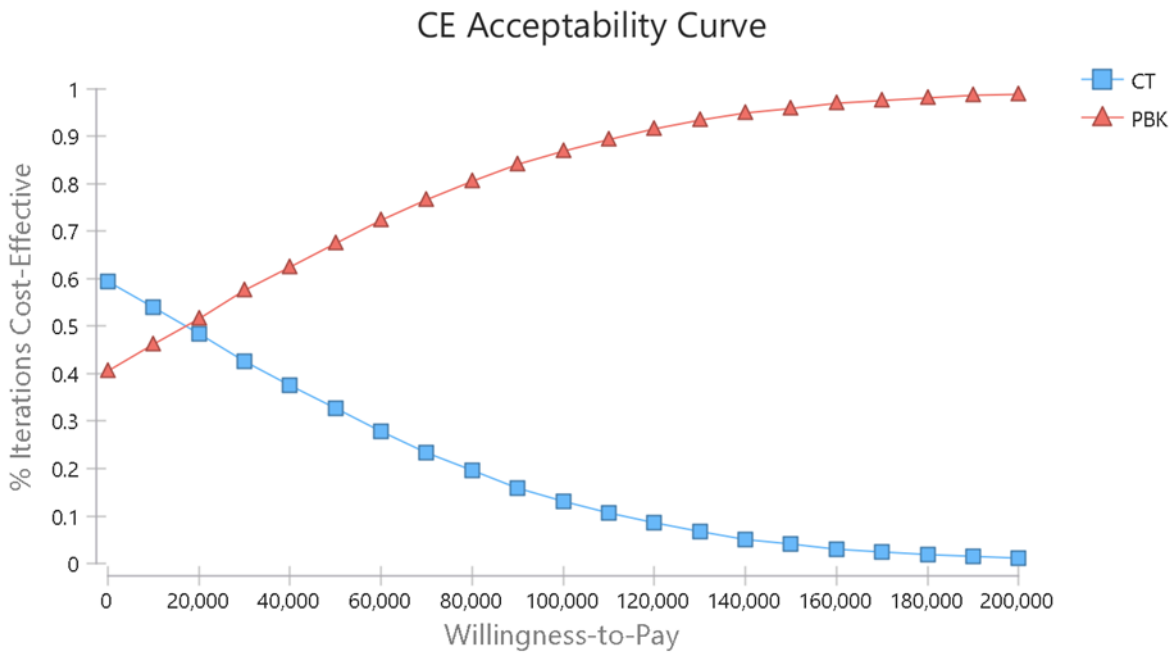


Figure 44 PBK compared to CT: incremental 12 month cost-effectiveness acceptability curve

Abbreviations
CE = cost effectiveness, **CT** = conservative treatment, **PBK** = percutaneous balloon kyphoplasty.

8.4.3 Key Drivers of the Economic Model

Univariate sensitivity analyses demonstrated that the assumption about costs was a key driver of model value for the PVP analysis using VERTOS II results,[132] and PBK using FREE data.[168] Key drivers are summarised in ***Table 80***. It should be noted that these results apply to patients with acute fractures, which are likely to be a sub-population of those availing these interventions in Switzerland. If the patient population included in Buchbinder (2009) - which includes those with fractures of up to 12 months – is deemed more representative of the Swiss context, then PVP is not cost-effective, as no improvement in quality of life (EQ-5D score) was reported.[7]

Table 80 Key drivers of the economic model

Description	Method/Value	Impact
Inclusion of adverse events or refracture	The Stevenson (2014) model included refracture and a sensitivity analysis for adverse events[191], while the Strom (2010) model also included a sensitivity analysis for adverse events.[190] They were included as PBK and PVP univariate sensitivity analyses for the 12 month economic models	**Low** The clinical evidence found no important differences in fracture and adverse events. Sensitivity analysis with 5% of the intervention arm receiving an additional CT unit cost and 16.7% having a 0.02 disutility for adverse events had limited impact.
Utility differences between intervention and comparator	The high and low QALY gains for PBK were included using 95% CI for FREE.[168] For PVP, results from the VERTOS II adjusted baseline analysis were used along with low and high estimates based on the incremental gain 95% CIs.[132]	**Moderate** Use of the high and low 95% CI utility results produced relatively moderate changes in the estimated ICER, however, the ICER remained below the hypothetical willingness-to-pay threshold
Cost differences	Weights for the intervention and comparator were taken from Swiss DRGs. An alternate costing of CT was included that assumed 9 physiotherapy visits.	**High-Moderate** Inclusion of differing hospital costs (high and low costs based on standard deviations) had a high impact on the estimated ICER. The inclusion of an alternate costing of CT to include physiotherapy treatment had a large impact, given the cost was much lower than the DRG cost used in base assumptions.

Abbreviations
CHF = Swiss franc, **CI** = confidence interval, **CT** = conservative treatment, **DRG** = diagnosis-related group, **PVP** = vertebroplasty, **ICER** = incremental cost-effectiveness ratio, **QALY** = quality-adjusted life year.

8.5 Results: Budget Impact

The financial implications of delisting PVP and PBK were examined using budget impact analysis from a payer perspective. Three scenarios of CT substituting for PVP or PBK are presented. In the first scenario, it was assumed 100% of current PVP and PBK procedures will be substituted by CT. The second scenario assumed 50% of current PVP and PBK will be substituted by CT, and the third scenario assumed 75% of current PVP and PBK will be substituted by CT. Those patients not substituting to CT were assumed to avail PBK and PVP using out-of-pocket and other non-insurance sources of finance. Data sources used to project the number of PVP and PBK procedures in Switzerland for the next 5 years are provided in ***Table 81***.

8.5.1 Assumptions for Budgetary Impact Analysis

Svedbom (2014) utilised prevalence and cost data to describe the epidemiology and economic burden of osteoporotic fractures in Switzerland (for 2010).[54] As previously mentioned (***Section 2.2***), it was estimated there were 1,381,000 Swiss men and 1,660,000 Swiss women at risk of osteoporosis in 2010 (at risk defined as those aged 50 years and older). By combining data from Lippuner (2009) it was calculated that approximately 74,000 fractures occurred in Switzerland in 2010, of which 11,000 were vertebral fractures.[198]

The incidence of vertebral fracture, as reported in Svedbom (2014),[54] have been combined with Swiss national population projections[196] to generate estimates of vertebral fracture for the years 2010, 2015, 2020 and 2025. Estimates for 2020–2024 are presented in ***Table 81***. It is estimated there will be 13,505 vertebral fractures in 2024. These estimates are similar to those by Svedbom (2014), where vertebral fractures were estimated to increase from 10,963 to 14,151 between 2020 and 2025.[54]

Table 81 Swiss vertebral fracture, treatment cost and uptake assumptions

Description	2020	2021	2022	2023	2024	Source
Vertebral fracture treatment projections						
Total Swiss vertebral fractures	12,215	12,537	12,860	13,183	13,505	See ***Section 17.9***
PVP procedures per year	3,040	3,246	3,458	3,677	3,902	Calculated
PBK procedures per year	1,789	1,950	2,115	2,287	2,465	Calculated
PVP rate (% of fractures)	25%	26%	27%	28%	29%	Calculated and projected at growth rate
PBK rate (% of fractures)	15%	16%	16%	17%	18%	Calculated and projected at growth rate

Description	2020	2021	2022	2023	2024	Source
Baseline vertebral fracture treatment (PBK, PVP and CT) costs						
PVP (CHF)	33,941,410	36,237,573	38,605,775	41,046,015	43,558,294	FOPH data assume 2018 volume indexed at population growth
PBK (CHF)	19,975,968	21,763,267	23,615,401	25,532,369	27,514,173	FOPH data assume 2018 volume indexed at population growth
Total (CHF)	**53,917,378**	**58,000,840**	**62,221,176**	**66,578,384**	**71,072,467**	FOPH data assume 2018 volume indexed at population growth
PVP, PBK and CT uptake assumptions						
Change in payer-supported PVP patients per year	-3,040	-3,246	-3,458	-3,677	-3,902	Delist procedures
Change in payer-supported PBK patients per year	-1,789	-1,950	-2,115	-2,287	-2,465	Delist procedures
Change in payer-supported CT patients per year	4,830	5,196	5,574	5,964	6,367	Calculated (PVP + PBK delisted procedures)

Abbreviations
CHF = Swiss franc, **CT** = conservative treatment, **FOPH** = Federal Office of Public Health, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.

The number of PVP and PBK inpatient procedures up to 2018 (provided by the FOPH) is presented in ***Figure 45***, along with a projection for 2019–2025 based on estimated PVP and PBK procedure rates as a percentage of vertebral fractures. FOPH data reported 1,461 inpatient PVP procedures in 2010, increasing to 2,674 by 2018. Similarly, there were 566 inpatient PBK procedures in 2010, increasing to 1,501 by 2018. As a proportion of estimated vertebral fractures, PVP was estimated to have a procedural rate of 15% in 2010, which increased to 23% in 2018. The procedural rate of PBK was 6% in 2010, which increased to 13% by 2018.

Schwenkglenks (2005) estimated the hospitalisation rate for vertebral fractures in Switzerland using patient-level inpatient data (from ICD-10 S72.0–S72.2 cases).[199] They calculated a 30% probability that a vertebral fracture would result in clinical presentation,[200-202] and a 33% probability of hospitalisation.[200 203] This estimate is similar to the combined PVP and PBK hospitalisation rate estimated in 2018.

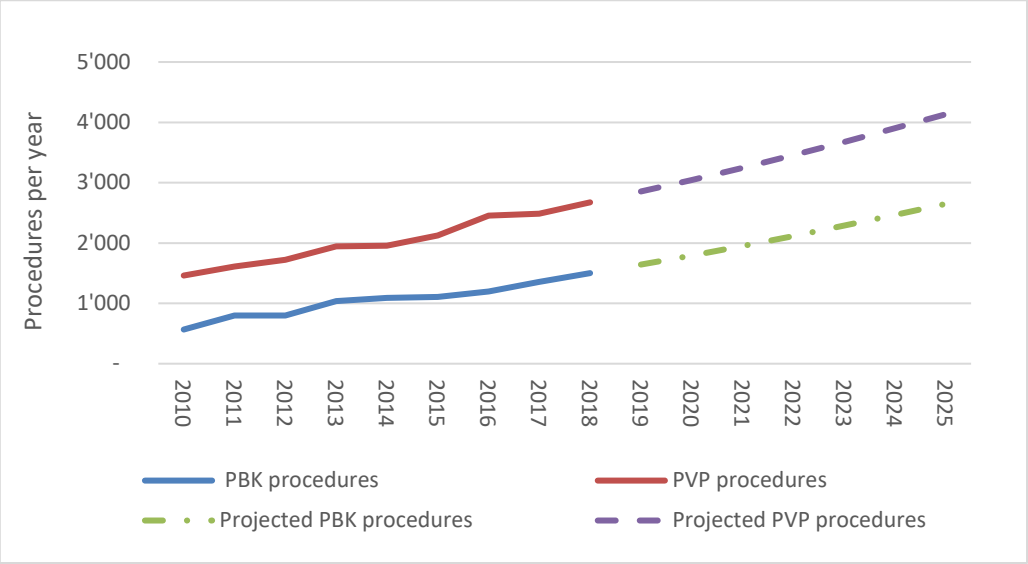


Figure 45 Number of PVP and PBK procedures per year in Switzerland

Abbreviations
PBK = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.

Lippuner (2011) estimated spine-related hospitalisations increased by 4.5% per year for women and 3.8% for men between 2000 and 2007, with a hospital cost of CHF34.9 million and CHF8.4 million in 2007, respectively.[204] The rate of annual increase seems to have accelerated in recent years compared to 2000 to 2007, with the rate of increase for PVP and PBK averaging 10% per year between 2010 and 2018 (FOPH inpatient data). This increase has been driven by factors such as the ageing population and increasing numbers of fractures, along with higher procedural rates. The proportion of estimated vertebral fractures treated with PVP and PBK has increased by 1% per year between 2010 and 2018 using FOPH data.

This rate of growth was used to estimate PVP and PBK procedures for 2020–2024 in the budget impact analysis, in the event these procedures are not delisted. It was estimated that 3,902 PVP and 2,465 PBK procedures will be performed in 2024 based on the assumed growth rate (***Figure 45***). Using current Swiss-DRG costs (see ***Table 77***), the net payer cost was estimated to increase from CHF33.9 million to 43.6 million between 2020–2024 for PVP and from CHF20.0 million to 27.5 million for PBK, in the event they remain listed. This corresponds to a total increase of CHF53.9 million to 71.1 million from 2020 to 2024.

8.5.2 Financial Implications

The 5-year budget impact of delisting PVP and PBK (with complete substitution by CT) is presented in ***Table 82***.

PVP substituted by CT

If PVP were to be delisted, there would be a cost saving for the payer of CHF33.9 million in 2020. If all patients substituted to CT, then a cost saving of CHF6.5 million would occur in 2020, increasing to CHF8.3 million by 2024.

PBK substituted by CT

If PBK were to be delisted, there would be a cost saving for the payer of CHF20.0 million in 2020. If all patients substituted to CT, then a cost saving of CHF3.8 million would occur in 2020, increasing to CHF5.2 million by 2024.

PVP and PBK substituted by CT

If all PBK and PVP patients substituted to CT, then a cost saving of CHF10.3 million would be realised in 2020, increasing to CHF13.5 million by 2024.

Table 82 Changed treatment costs in Switzerland, 2020–2024

		2020	2021	2022	2023	2024	Source
Change in patient medicine usage							
Change in OVCF patients using payer-supported PVP	Patients	-3,040	-3,246	-3,458	-3,677	-3,902	Assumption
Change in OVCF patients using payer-supported PBK	Patients	-1,789	-1,950	-2,115	-2,287	-2,465	Assumption
Change in OVCF patients using payer-supported CT not PVP	Patients	3,040	3,246	3,458	3,677	3,902	Assumption
Change in OVCF patients using payer-supported CT not PBK	Patients	1,789	1,950	2,115	2,287	2,465	Assumption
Changed treatment costs							
Change in payer PVP costs	CHF	-33,941,410	-36,237,573	-38,605,775	-41,046,015	-43,558,294	Calculated
Change in payer costs for CT instead of PVP	CHF	27,482,594	29,341,813	31,259,362	33,235,241	35,269,451	Calculated
Net PVP treatment costs	**CHF**	**-6,458,816**	**-6,895,760**	**-7,346,413**	**-7,810,774**	**-8,288,843**	**Calculated**
Change in payer PBK costs	CHF	-19,975,968	-21,763,267	-23,615,401	-25,532,369	-27,514,173	Calculated
Change in payer costs for CT instead of PBK	CHF	16,174,679	17,621,867	19,121,553	20,673,735	22,278,415	Calculated
Net PBK treatment costs	**CHF**	**-3,801,289**	**-4,141,400**	**-4,493,848**	**-4,858,634**	**-5,235,757**	**Calculated**
Net change in overall treatment costs	**CHF**	**-10,260,105**	**-11,037,160**	**-11,840,261**	**-12,669,408**	**-13,524,600**	**Calculated**

Abbreviations
CHF = Swiss franc, **CT** = conservative treatment, **OVCF** = osteoporotic vertebral compression fracture, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.

Sensitivity analysis

Base assumptions were subject to a sensitivity analysis using the Swiss-DRG cost for CT (***Table 83***). An alternate costing based on 9 physiotherapy visits at CHF90 per visit was included. Given the large difference between the base CT unit cost of CHF9,039 and CHF810 for physiotherapy, this assumption had the largest impact on financial analysis results. If the cost per PVP or PBK procedure were to decrease over the next 5 years, less savings would result through delisting. A scenario was modelled which included 5% per annum unit cost decreases.

The budget impact was also sensitive to the number of patients who would switch to this procedure. The lower the proportion of patients switching from PBK and PVP to CT results in larger cost savings for the payer, as those not substituting are assumed to avail PBK and PVP using out-of-pocket and other sources of non-insurance finance. There was considerable uncertainty around this rate of substitution, which had a large impact on estimated financial estimates.

A small proportion of Swiss PVP procedures are delivered in an ambulatory setting. TARMED data provided by FOPH indicated 238 PVP were delivered during 2019 in ambulatory settings. Only inpatient-delivered services were included in the base budget impact analysis, which understates total PVP costs. A sensitivity analysis was included where total PVP procedures are increased by 8% to reflect ambulatory delivery. Increases in the number of PVP procedures to include ambulatory services had a minor impact on budget calculations, compared to other scenarios. Changes in the assumed annual procedure growth rate did not significantly impact costs.

Table 83 Net payer cost sensitivity analysis (CHF)

	2020	2021	2022	2023	2024
Base	**-10,260,105**	**-11,037,160**	**-11,840,261**	**-12,669,408**	**-13,524,600**
Proportion OVCF treated by PVP has zero growth rate	-9,793,048	-10,318,068	-10,856,796	-11,409,233	-11,975,378
Proportion OVCF treated by PBK has zero growth rate	-9,741,153	-10,238,169	-10,747,523	-11,269,214	-11,803,242
50% of current OVCF patients using PVP substitute with CT	-24,001,402	-25,708,067	-27,469,942	-29,287,028	-31,159,326
75% of current OVCF patients using PVP substitute with CT	-17,130,754	-18,372,613	-19,655,101	-20,978,218	-22,341,963
50% of current OVCF patients using PBK substitute with CT	-18,347,445	-19,848,094	-21,401,037	-23,006,275	-24,663,808
75% of current OVCF patients using PBK substitute with CT	-14,303,775	-15,442,627	-16,620,649	-17,837,841	-19,094,204
CT assumed to be 9 physiotherapy visits	-50,005,142	-53,792,308	-57,706,417	-61,747,468	-65,915,460
Base PVP patients increase by 8% for ambulatory services	-10,776,811	-11,588,821	-12,427,974	-13,294,269	-14,187,708
PBK cost per procedure 5% annual decrease	-10,260,105	-9,948,997	-11,402,111	-11,976,445	-12,553,400
PVP cost per procedure 5% annual decrease	-10,260,105	-9,225,281	-8,076,198	-6,815,220	-5,444,809

Abbreviations
CHF = Swiss franc, **CT** = conservative treatment, **OVCF** = osteoporotic vertebral compression fracture, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty.

9 Legal Issues

9.1 Summary Statement Legal Issues

Disinvestment of PVP or PBK is unlikely to result in any potential legal issues, as outlined in the EUnetHTA Core Model.

9.2 Methods

The scoping reported noted there were no legal issues related to the disinvestment of PVP and PBK based on the sub-questions in the EUnetHTA Core Model 3.0. In addition, a non-systematic search was performed to identify any further issues. As this was not a systematic search, a PRISMA chart is not provided.

9.3 Results

No legal issues were identified from the systematic and non-systematic searches.

10 Social Issues

10.1 Summary Statement Social Issues

There was limited literature addressing the social issues associated with PVP and PBK. In the absence of direct evidence, literature addressing existing osteoporosis treatments was used to supplement the results, noting the applicability of these studies was uncertain.

Patients had limited knowledge of interventional radiology procedures (including PVP and PBK) and osteoporosis treatments more broadly. However, of patients who received PVP and PBK, most held a positive perception of the treatment.

PVP and PBK potentially reduced the burden of care compared to CT because more patients were discharged home without assistance and reported greater levels of independence.

Older males with osteoporosis represented an at-risk group, as they are often under-treated owing to traditional beliefs regarding the disease.

10.2 Methods

Sub-questions used to frame patient and social aspects of PVP and PBK are outlined in ***Section 17.2*** (***Appendix B***). To address these aspects, literature identified from a systematic and non-systematic search were used (detailed in ***Section 6***). The non-systematic search involved targeted searches of Psych Info and PubMed using the following terms: “vertebroplasty”, “kyphoplasty”, “expectations”, “outlook”, “perception”, “osteoporosis” and “burden”. The non-systematic searches were conducted by a single reviewer who identified an additional 19 studies. A PRISMA chart was not provided owing to the use of systematic and non-systematic searches. The results of the literature searches were summarised using narrative synthesis.

10.3 Evidence Tables

Nineteen studies were included in the assessment of social issues (***Table 84***).[205-223] The studies consisted of primary (k = 15) and secondary research (k = 4). Primary research studies were mostly performed in North America (k = 6), Europe (k = 5) and Australia (k = 3). Two studies were conducted in Singapore and China, and their applicability was uncertain owing to different population demographics and healthcare systems. Medical practitioners (general practitioners and nurses) and patients with osteoporosis were the most studied populations. Participants were generally recruited using convenience sampling methods from hospitals, primary care or community centres with sample sizes ranging from 13 to 1,407 participants.

Secondary research studies included analysis of medical databases (k = 2), websites (k = 1) or published literature (k = 2). The literature reviews analysed available evidence regarding patient perception of osteoporosis, with emphasis placed on male patients. The analysis of Medicare databases aimed to determine the incidence of mortality and health resource utilisation following PVP, PBK and CT in the USA. The analysis of a pharmaceutical database in Australia aimed to determine how news media affects patients' utilisation of osteoporosis health resources. The analysis of vertebroplasty-based websites sought to assess the type of information presented to patients. Studies providing epidemiological information were not included because they did not specifically address social issues.

There was limited primary and secondary research addressing social issues associated with PVP and PBK. Rather, most studies considered the broader context of osteoporosis treatments. Therefore, in the absence of literature addressing PVP or PBK, studies pertaining to osteoporosis treatments were presented, noting the generalisability of the results is uncertain.

Table 84 Characteristics of included studies for social issues

Author; Year; Country	Indication; Sample size	Design; Setting	Interview/survey topics
Baerlocher 2007[205] Canada	Patients undergoing interventional radiology procedure n = 100	Survey Interventional radiology clinic	Patient awareness of interventional radiology, satisfaction with procedure
Lindsay 2016[206] USA	Patients self-diagnosed with osteoporosis n = 1,407	Survey NA	Patient knowledge of osteoporosis and treatment barriers
Merele 2019[207] France	Patients with osteoporosis n = 98	Focus group NA	Patient knowledge and attitude towards osteoporosis
Naik-Panvelkar 2020[208] Australia	General practitioners n = 13	Interview Telephone interview	General practitioner knowledge and attitude towards osteoporosis
Otmar 2012[209] Australia	General practitioner and nurses n = 16	Survey Community practices	General practitioner knowledge and attitude towards osteoporosis

Author; Year; Country	Indication; Sample size	Design; Setting	Interview/survey topics
Ribeiro 2000[210] Canada	Patients with osteoporosis n = 185	Interview Church group	Patient knowledge of osteoporosis
Sale 2014[211] Canada	Patients with osteoporosis n = 25	Interview Hospitals	Patient knowledge of osteoporosis
Werner 2005[212] Israel	Patients with osteoporosis n = NA	Review NA	Patient knowledge and attitude towards osteoporosis
Williams 2002[213] UK	Patients with osteoporosis n = 163	Interview Local primary schools	Patient knowledge of osteoporosis
Carr 2018[214] USA	Patients undergoing kyphoplasty n = 151	Survey Telephone interview	Patient satisfaction with procedure
Rapan 2017[215] Croatia	Patients undergoing vertebroplasty n = 50	Survey Inpatient stay at hospital	Patient satisfaction with procedure
Sambrook 2010[216] Australia	NA	Analysis of pharmaceutical benefits scheme database NA	Analysis of database for prescription utilisation
Sullivan 2014[217] USA	NA	Analysis of websites NA	Websites offering information regarding vertebroplasty
Kaffashian 2011[218] Canada	Patients with osteoporosis fracture NA	Questionnaire CaMos cohort	Utilisation of post-care services
Klezl 2012[222] UK	Patients undergoing kyphoplasty n = 53	Questionnaire Teaching hospital	Utilisation of post-care services
Siddiqui 2010[223] Singapore	Caregivers of patients with vertebral fracture NA	Interview Tertiary hospital	Caregiver stress due to assisting relative with fracture
Xie 2015[219] China	Patients with osteoporosis fracture n = 123	Questionnaire Hospital	Utilisation of post-care services or caregivers
Adler 2014[220] USA	NA	Review of literature	Overview of osteoporosis in men
Nielsen 2011[221] Denmark	Patients with osteoporosis n = 16	Focus group Hospital outpatient department	Patient knowledge and attitude towards osteoporosis

Abbreviations
CaMos = Canadian Multicentre Osteoporosis Study, **n** = number of participants, **NA** = not applicable, **UK** = United Kingdom, **USA** = United States of America.

10.4 Results

10.4.1 Patient Expectations of PVP and PBK

Patients appeared to have limited knowledge of PVP and interventional radiology procedures more broadly. For example, a survey of patients undergoing interventional radiology procedures noted 72% of patients were unaware of what constitutes interventional radiology, and 87% had not received information regarding their procedure. Further, 98% believed other patients are also unaware of interventional radiology procedures.[205] Given patients were unaware of interventional radiology (which includes PVP and PBK), it was unclear what expectation or wishes patients had regarding the technology.

There was greater evidence regarding treatment expectations when considering osteoporosis treatments more broadly, noting the applicability of these issues to this HTA was uncertain.

Patients' expectations of osteoporosis treatments were limited owing to lack of knowledge about the disease.[212] Only a small number of patients recalled receiving advice on treatment strategies from their GP.[210 213] Furthermore, several patients reported they had received incorrect or unclear information from their physician.[211] Of patients with knowledge of osteoporosis, reasons for not utilising treatments included the belief that osteoporosis was not a serious condition warranting treatment and fears of medication-related side-effects.[206] This suggested that patients may have limited or negative expectations regarding current osteoporosis treatments.

Patients often derive treatment expectations and medical knowledge from physicians,[212] therefore it is likely that the 2 share similar views. However, among physicians, there was conflicting expectations regarding osteoporosis treatments with surveys reporting that some GPs believed osteoporosis medications were efficacious while others believed they had limited efficacy.[208 209] Lack of perceived efficacy was attributed to uncertainty regarding the extent of fracture prevention and the substantial lag between treatment initiation and reduction in fracture risk.[208] A lack of knowledge of suitable treatments,[207] concerns regarding long-term safety and cost of medication,[207 209] and a perceived lack of urgency in treating osteoporosis[208 209] also contributed to the poor treatment expectations among GPs.

10.4.2 Patient Perceptions of PVP and PBK

There was limited information regarding patient perception of PVP and PBK. Patients' perception of the procedure may be skewed owing to the way it is portrayed on the internet. Websites do not provide sufficient levels of detail and are potentially misleading because they are more likely to report the benefits of PVP rather than the risks.[217] This finding was not unique to PVP or internet media, with a study demonstrating that negative, poorly-informed television news media led to a reduction in

osteoporosis medication prescriptions, which in turn increased the number of patients experiencing fractures.[216]

Once treated, however, patients held a positive perception of PVP and PBK. For example, in a survey of patients undergoing PBK for acute compression fractures—of which 72% were osteoporotic or spontaneous in nature—85% noted the procedure was tolerable and 55% indicated they received adequate pain relief. Of these patients, 66% noted they would have the same procedure again.[214] A similar finding was observed for patients undergoing PVP.[215]

More broadly, patients undergoing interventional radiology procedures (including PVP) reported a mean satisfaction rate of 8.8/10 (10 being the highest) following the procedure. Further, most patients were conformable during the procedure, believed they would be able to return to previous levels of activity quickly, and would choose interventional radiology procedures over surgery for future treatments.[205] Overall, the results reinforced the concept that patients hold a positive perception of PVP and interventional radiology procedures.

10.4.3 Caregiver Burden

Symptomatic OVCFs often impose a substantial burden on the individual and those around them. Patients with fractures often require assistance to complete basic daily activities (see Melton [2003][224] for review) and additional support provided by caregivers. Of elderly patients with vertebral fractures, 34–50% were discharged to a nursing facility, 11–15% were discharged home with formal support, and 24–38% went home without support.[146]

For those sent home without formal support, assistance is typically provided by caregivers. For example, 38.6% of patients with vertebral fracture report receiving informal care for an average of 37 days (SD = 26) following hospitalisation. Approximately a third (36%) of informal caregivers had a paying job, of which 25% had taken at least 1 day away from their job to care for their relative.[218] In China, the caregiver sacrificed an average of 16.4 days (± 26.0 days) corresponding to ¥3,233 ± 4,184 (CHF448.56 ± 580.51) in lost income.[219] The applicability of this finding to Switzerland was uncertain, however, it suggested that providing care for patients with OVCF placed significant stress on the caregiver, mainly attributed to increased financial strain.

There was limited information comparing the burden of care following PVP or PBK. Chen (2013) noted patients undergoing PBK or PVP were more likely to be discharged to home than to a nursing facility when compared to CT patients.[146] This may increase the immediate burden on caregivers, however, CT patients also require support, thus the burden may only be delayed. Klezl (2012) observed that patients who received CT were less independent (as inferred by activities of daily living) and required more carer assistance 1 year later than patients treated with PBK.[222] Collectively, the results suggested PVP and

PBK patients may be discharged earlier, potentially increasing immediate caregiver burden, however, CT patients required more overall caregiver support compared to PVP and PBK.

10.4.4 Patient Groups with Poor Access to PVP and PBK

There was limited information regarding patient groups without access to PVP or PBK.

When considering osteoporosis more broadly, older males represented a potentially under-treated group. Despite accounting for approximately 20–30% of patients with osteoporosis-related fractures[225] [226], males are traditionally overlooked by the medical community, often leading to inadequate treatment. For example, older males are less likely to be evaluated and treated for osteoporosis[220] and are less frequently hospitalised for vertebral fractures compared to females (noting younger males and females exhibit similar hospitalisation rates).[55] [147] Once hospitalised, males with vertebral fractures exhibit greater hospital-related mortality,[147] an effect that persists up to 1 year following discharge.[225] Some evidence suggests the mortality rate post-PVP is higher among males than females.[227]

It was unclear why sex influences osteoporotic- and hospital-related mortality. Existing attitudes towards osteoporosis were likely a barrier to accessing appropriate healthcare resources. For example, a survey of GPs indicated they consider osteoporosis to primarily affect females,[209] as osteoporosis and related fractures generally occur later in males compared to females.[220] This attitude was echoed by male patients, who consider osteoporosis to mainly affect females.[221] Further, older male patients were less likely to engage with healthcare resources owing to traditional attitudes regarding masculinity and perception of health.[221] Owing to increasing life expectancy, the number of males with osteoporosis and fractures is projected to increase in Switzerland,[199] [204] creating a growing unmet medical need.

11 Ethical Issues

11.1 Summary Statement Ethical Issues

The ethical impact of disinvesting PVP and PBK was uncertain. Ethical concerns included whether disinvestment may prevent patients from accessing a potentially safe and effective treatment for vertebral fractures. This concern would primarily affect older adults, who are disproportionally affected by vertebral fractures; however, it was unclear whether the effects of PVP and PBK were attributable to a "treatment" or placebo effect. Thus, it was unclear whether the risks outweigh the benefits derived from these procedures.

11.2 Methods

Sub-questions used to frame the ethical aspects of PVP and PBK are outlined in ***Section 17.2*** (***Appendix B***)***, Table 94***. To address the questions, literature from a systematic (detailed in ***Section 6***) and non-systematic search were used. The non-systematic search involved targeted searches of PubMed and Ethmed using a combination of "vertebroplasty", "kyphoplasty", "burden", "placebo" and "autonomy". Non-systematic searches were conducted by a single reviewer who identified an additional three studies. A PRIMSA chart was not provided owing to the use of systematic and non-systematic searches. For questions in which there was an absence of PVP or PBK literature, general ethical concerns (in line with a principlist approach) relating to osteoporosis treatments and the elderly were used. The results of the literature searches were summarised using narrative synthesis.

11.3 Evidence Tables

Three studies evaluating ethical issues associated with PVP and PBK were included (***Table 85***).[222 228 229] Two studies assessed the discharge location of patients undergoing PVP or PBK. Klezl (2012) surveyed patients who underwent PBK at a teaching hospital in the UK;[222] Crouser (2018) analysed the American College of Surgeons–National Surgical Quality Improvement Program database for procedural and discharge codes relating to PVP.[228] The sample size (53–2,361) and duration of follow-up (30 days–1 year) varied between the 2 studies. The third study reviewed the ongoing debate surrounding the effectiveness of PVP and PBK.[229]

Additional studies providing epidemiological information were not included in the table because they did not specifically address ethical issues.

Table 85 Characteristics of included studies for ethical issues

Author; Year; Country	Indication; Sample size	Study design	Interview/survey topics
Crouser 2018[228] USA	Patients undergoing vertebroplasty n = 2,361	Analysis of ACS-NSQIP database	Utilisation of post-care services
Klezl 2012[222] UK	Patients undergoing kyphoplasty n = 53	Questionnaire Teaching hospital	Utilisation of post-care services
Miller 2013[229] Australia	NA	Commentary	Risks and benefits of vertebroplasty and kyphoplasty

Abbreviations
ACS-NSQIP = American College of Surgeons–National Surgical Quality Improvement Program, **NA** = not applicable, **UK** = United Kingdom, **USA** = United States of America.

11.4 Results

11.4.1 Symptoms and Burden of Disease Attributable to OVCF

Osteoporotic vertebral fractures account for approximately 59% of all hospitalised fractures in Switzerland and are most common among females age 70 and above.[54 55] However, most vertebral fractures are asymptomatic or slightly symptomatic[32 80] and are consequently under-reported and under-diagnosed.[230] This is concerning given the increased mortality associated with vertebral fractures.[35]

Symptomatic vertebral fractures impair an individual's quality of life as they are often confined to bed and have limited functional capacity.[231 232] Vertebral fractures also lead to a reduction in vertebral height and kyphosis[233] which can worsen in pulmonary function and reduce appetite.[32 233 234] The fracture can cause nerve irritation, resulting in further pain,[80] and increase the risk of subsequent fracture.[33 235]

Treatments for vertebral fracture (e.g. bed rest and immobilisation) may cause further harm as bone mineral density and muscular strength are reduced, increasing the risk of subsequent falls.[71] Further, immobilisation increases the risk of developing bedsores and DVT. All these factors likely contribute to the increased risk of mortality following vertebral fracture.[35]

11.4.2 Perceived Benefits and Harms of PVP and PBK

Non-maleficence: a norm of avoiding causation of harm

A key ethical concern when considering an intervention is the avoidance or minimisation of harm. In this context, harm included adverse physical and psychological consequences of PVP, PBK and CT.

Results from ***Section 7*** indicated PVP and PBK had an equivalent or superior safety profile relative to CT (depending on the type of study). Common adverse events included cement leakage and

subsequent vertebral fracture. Most cases of cement leakage were asymptomatic. However, severe adverse events attributable to cement leakage, including pulmonary embolism, were reported (see ***Section 7.8***). Similarly, vertebral fractures were either asymptomatic or caused pain. Severe treatment-related adverse events during or following PVP and PBK were infrequent and were more common in the CT group when larger database analyses were considered.

CT for the management of OVCF includes analgesic medication (NSAIDs and opioids), bed rest and bracing.[236] Prolonged periods of immobility can lead to complications such as deconditioning (muscle and bone loss), which was not adequately captured by the included safety studies. Deconditioning among older adults is concerning as the subsequent loss of balance and increased risk of falls causes further harm and limits quality of life.[236] Adverse events associated with NSAIDs and opioids include heart failure, thromboembolism and respiratory depression, which further increase the risk of mortality.[82] [83] These adverse events may be minimised if patients underwent PVP and PBK instead of CT.[147]

Any medical treatment may cause distress if individuals' expectations are not met. Alternatively, for patients who believe PVP or PBK improves their condition, removing the reimbursement may impede access to the technology, possibly resulting in psychological distress if the desired medication cannot be obtained. This harm is minimised because PVP and PBK will remain available as a fee-paying service if disinvested. However, individuals with limited financial means may be unable to afford the procedure if the service becomes fee-paying. Thereby limiting their access to the procedure.

Beneficence: a group of norms for providing benefits and balancing benefit against risks and costs

Vertebral fractures can result in pain that lasts approximately 2 to 3 months following the injury. [236] Pain significantly impairs an individual's quality of life by limiting their ability to perform daily activities and act independently. The loss of independence and pain can further lead to the emergence of mental health disorders.[237] Results from ***Section 7*** indicated PVP and PBK significantly improved function and reduced pain, length of hospital stay and discharge to nursing homes compared to CT. (However, they were unlikely to reduce pain and function compared to sham.) This suggested patients receiving PVP and PBK may exhibit greater functionality, independence and overall quality of life. Further, an individual may avoid harms associated with CT such as opioid- or NSAID-related adverse events.[82] [83]

However, owing to the lack of clinically relevant differences between sham and intervention arms (see ***Section 7.7***), was is currently unclear whether PVP and PBK exhibit true effectiveness or whether effects were attributable to a strong placebo response.[229] The treatment ritual (or confounding variables such as the local anaesthetic or nerve block) rather than the injection of cement may cause the observed benefits associated with PVP.[229] This raises concerns regarding the overall risk-benefit of the procedure.

If the effects are attributable to a placebo response, it is presently unclear whether there is sufficient benefit (or avoidance of harm) to compensate for potential adverse events associated with PVP and PBK (however rare they may be). In contrast, if the procedure is truly effective, then disinvesting may result in harm to patients because an effective treatment is no longer available.

11.4.3 Benefits and Harms of PVP and PBK to Relatives, Caregivers, Commercial Entities or Society

There was limited literature addressing the benefits and harms of PVP and PBK for relatives, other patients, commercial entities and societies.

Beneficence: a group of norms for providing benefits and balancing benefit against risk and cost

Klezl (2012) observed that, 1 year after treatment, individuals who received PBK were more independent (as inferred by activities of daily living) and required less carer support than patients who received CT.[222] This was likely to benefit relatives who frequently fulfil the role of carers, and staff at nursing homes or rehabilitation centres.

Non-maleficence: a norm of avoiding causation of harm

The decreased length of hospital stay observed in the databases analyses, suggested the burden of care is transferred from hospital staff to caregivers, rehabilitation centre and nursing home staff.[146] Vertebroplasty patients discharged to an inpatient facility (skilled-care facilities and rehabilitation units) reported higher complication rates and 30-day mortality compared to vertebroplasty patients discharged home.[228] However, rates of readmission to hospital and reoperation were similar between the 2 groups. This suggests inpatient facilities were likely to experience more burden because patients discharged to their facility had increased needs. It was unclear whether complication rates and 30-day mortality rates in inpatient facilities differed between PVP and CT patients.

11.4.4 Use of PVP and PBK in Vulnerable Patient Groups

There was limited primary and secondary research addressing the use of PVP and PBK in vulnerable patient groups. Consequently, the results were broadened to patients with osteoporotic fractures and older adults.

Older adults (age 70 and above) have the highest incidence of osteoporotic vertebral fractures in Switzerland.[55] However, the symptoms associated with fractures, namely pain, is often under-diagnosed and under-treated in older adults[238], suggesting these individuals are a vulnerable patient group. Under-treating older adults is likely due to the complexity associated with their medical management. For example, ageing increases the risk of developing comorbidities, with approximately 50% of Swiss adults age 65 and above reporting at least 1 chronic illness. Further, 25% of adults age 65–79 and 41% of

adults older than 80 years have several comorbidities.[49] The presence of multiple comorbidities increases an individual's risk of mortality and disability and is associated with poorer quality of life.[239] The presence of multiple comorbidities also increases utilisation of polypharmacy.[240] Polypharmacy complicates medical care as physicians need to determine whether a drug could interact with another medication or health state. Further, polypharmacy is taxing on the individual. It is associated with increased risk of developing an adverse event or "geriatric syndrome", and increases medication non-adherence.[241] Polypharmacy is also associated with increased risk of falls among older adults[241], and increases financial hardship.[242] This is particularly concerning because older adults note medication cost as a key reason for non-adherence.[243]

From a social perspective, older adults are at greater risk of losing independence and becoming socially isolated.[32] This problem is likely amplified in older patients with OVCF who are often bedridden for prolonged durations. Older adults often have poorer health literacy compared to younger adults[244], which is associated with a higher use of emergency services, healthcare costs and overall mortality.[245] [246] Using PVP and PBK may alleviate problems caused by OVCF and the potentially reduce the burden associated with polypharmacy.

11.4.5 PVP and PBK Influence on Patient Decision-Making and Autonomy

Medication and technology can directly or indirectly influence patients' decision-making capabilities thereby affecting their autonomy.[247] This is particularly concerning among vulnerable populations with attributes that undermine their ability to provide informed consent.[248] The 2 relevant vulnerable population in the context of this HTA were older adults and those receiving opioids for pain.[248]

Unrelieved pain significantly reduces quality of life as it takes a physical, emotional, cognitive and socioeconomic toll, thereby impacting individual autonomy. For example, pain reduces attention, concentration and short-term memory.[249] Deficits in memory and concentration can affect retention of information for patients confronted with complex medical information. Consequently, patients may be unable to provide informed consent or reach an appropriate decision because they could not evaluate the risks or consider alternatives appropriately.[248]

Providing pain relief can potentially protect a person's autonomy by alleviating many of the above concerns.[250] However, medications that alleviate pain may also hinder autonomy. There was conflicting evidence regarding the effect of opioids on cognitive function.[251] Studies have demonstrated an increase, a decrease or no change in cognitive problems such as forgetfulness or inattention. Of the studies noting an improved cognitive capability, the authors postulated it may relate to a reduction in pain. Further, opioids can induce states of addiction, physical dependence and hyperalgesia[252], which may influence a person's ability to act autonomously. Individuals may be unable to appraise the risks

and benefits of opioids appropriately as the try to control affective states such as pain and addiction.[252] If PVP or PBK alleviate pain they may improve a patient's ability to act autonomously because the effects of pain and opioid-induced cognitive impairment are minimised.

12 Organisational Issues

12.1 Summary Statement Organisational Issues

There were limited organisational issues associated with PVP and PBK. Compared to CT, these procedures were associated with reduced lengths of stay, complications, nursing home admissions and engagement in community health services (e.g. GP). If PVP and PBK were disinvested and more patients were managed conservatively it may increase the burden on these healthcare resources.

12.2 Methods

Sub-questions framing the organisational aspects of PVP and PBK are outlined in ***Section 17.2*** (***Appendix B***)***, Table 95***. To address the questions, literature identified from a systematic (detailed in ***Section 6***) and non-systematic search were used. The non-systematic search involved targeted searches of PubMed using the following terms: “vertebroplasty”, “kyphoplasty”, “length of stay”, “adverse events”, “benefits” and “burden”. The non-systematic searches were conducted by a single reviewer who identified an additional 4 studies. A PRIMSA chart was not provided owing to the use of systematic and non-systematic searches. The results of the literature searches were summarised using narrative synthesis.

12.3 Evidence Tables

Seven studies evaluating organisational issues associated with PVP and PBK were included (***Table 86***).[146 147 171 222 229 253 254] Three studies provided commentaries discussing the risks and benefits of PVP and PVK. Two studies analysed the United States Centres for Medicare & Medicaid Services Medicare Provider Analysis and Review File database to determine healthcare resource utilisation, readmission and complications following PVP, PVK and CT. One prospective study measured the number of GP visits and 1 study assessed patient independence following PBK.

Table 86 Characteristics of included studies for organisational issues

Author; Year; Country	Indication; Sample size	Design; Follow-up; Setting	Interview/survey topics
Chen 2013[146] USA	Patients undergoing PVP, PBK and CT n = 68,752	Medicare database	Utilisation of post-care services
Ong 2018[147] USA	Patients undergoing PVP, PBK and CT n = 2,077,944	Medicare database	Readmission rate
Klezl 2012[222] UK	Patients undergoing kyphoplasty n = 53	Questionnaire Teaching hospital	Utilisation of post-care services
Grafe 2005[171] Germany	Osteoporotic vertebral fracture n = 60	Prospective cohort, 12 months follow-up	Utilisation of post-care services
Buchbinder 2019[253] Australia	NA	Commentary	Risks and benefits of vertebroplasty and kyphoplasty
Clark 2016[254] Australia	NA	Commentary	Benefits of vertebroplasty
Miller 2013[229] Australia	NA	Commentary	Risks and benefits of vertebroplasty and kyphoplasty

Abbreviations
ACS-NSQIP = American College of Surgeons–National Surgical Quality Improvement Program, **CT** = conservative treatment, **NA** = not applicable, **PVP** = percutaneous vertebroplasty, **PBK** = percutaneous kyphoplasty.

12.4 Results

12.4.1 PVP and PBK Impact on Work Processes

Vertebroplasty and kyphoplasty may reduce hospital and community health resource use. Results from the RCTs and analyses of Medicare databases determined patients who received PVP and PBK had reduced lengths of stay and hospital readmission rates compared to CT-treated patients.[2 146 147 226] Further, a greater proportion of PVP and PBK patients were discharged home compared to CT patients.[147] CT patients were less independent (as inferred by activities of function outcomes, see ***Section 7***), required more carers and saw GPs more often than did patients treated with kyphoplasty.[171 222] Collectively, the results suggested PVP and PBK led to a reduction in healthcare resource use because patients had a shorter length of stay and required fewer healthcare services during follow-up compared to CT patients.

12.4.2 Impact of Delisting PVP and PBK on Patient Engagement with Healthcare Resources

No literature was identified to answer this research question.

12.4.3 PVP and PBK Impact on Other Healthcare Resources

As previously mentioned, PVP and PBK reduced length of stay and the number of patients discharged to a nursing facility.[146] This likely liberates hospital and nursing facility resources such as staff, beds and CT needs (braces and medication). In addition, the (potentially) reduced complications following these procedures may further reduce hospital resource demands. However, PVP and PBK will increase demand for radiographic equipment (x-ray, fluoroscopy, MRI and computed tomography) and operating staff (interventional radiologist, surgeons, nursing and support staff) associated with these procedures. It is unclear whether demand for rehabilitation staff such as physiotherapists would change.

12.4.4 Management Problems Caused by Removing PVP and PBK

Conservative treatment involves bed rest, bracing and analgesic medication. Removing PVP and PBK would likely increase demand for hospital and nursing facility resources owing to the increased length of stay associated with CT.[146] Further, analgesic medication such as opioids and NSAIDs could cause adverse events that would necessitate further staff and resources to be managed accordingly. Given there is a limited number of hospital and nursing facility beds, the wait time for treatment may increase. Once discharged, CT patients utilise community health services more, which may add further strain to the system.[171]

12.4.5 Are PVP and PBK Accepted?

There is ongoing debate amongst individual clinicians (for example Buchbinder [2019][253], Clark [2016][254]) and organisations (see ***Section 13***) regarding the effectiveness of PVP and PBK. Some believe the procedures safe and efficacious, while others do not. Those who believe in the efficacy purport that it is mostly effective in specific sub-groups of populations, namely those with fractures younger than 8 weeks because they are not fully healed and highlight flaws in RCTs that evaluate both populations[254] Clinicians and organisations who do not support these procedures suggest there is an inconsistent direction of effect (with respect to pain and function) and procedure is mostly explained by the placebo effect.[229] Collectively, it was unclear how accepted PVP and PBK are among the medical community. No patient organisations were identified to address this question.

13 Additional Issues

There was discordance in the literature regarding the clinical utility of PVP and PBK. Three clinical practice guidelines[76 93] and 5 consensus statements were identified from the literature (***Table 87***).[9 11 12 39] Five position statements and 2 clinical practice guideline were in favour of PVP and PBK. One guideline was against the procedures. The organisations were from Europe or North America.

Table 87 Summary of clinical guidelines and recommendations regarding PVP and PBK

Organisation	Recommendation	Strength of recommendation
Guidelines		
American Academy of Orthopaedic Surgeons (2011)[76]	PVP not recommended for patients with OVCF who are neurologically intact PBK an option for patients with OVCF who are neurologically intact	Strong Weak
American College of Radiology (2018)[93]	PVA indicated for new OVCF without malignancy or "red flags" PVA indicated for asymptomatic spinal fracture PVA indicated for pathologic spinal fracture with severe pain, deformity or pulmonary dysfunction	Usually appropriate May be appropriate Usually appropriate
Centre Hospitalier Universitaire Vaudois (2020)[92]	PVP and PBK is recommended for painful OVCF refractory to medical therapy (within 2–6 weeks of fracture)	NR
Position statements		
American Academy of Family Physicians (2016)[39]	PVP and PBK recommended for painful VCF refractory to medical therapy	C *
Cardiovascular and Interventional Radiological Society of Europe (2017)[11]	PVP and PBK recommended for painful VCF refractory to medical therapy	NR
NICE (2016)[12]	PVP and PBK recommended for painful OVCF refractory to medical therapy	NR
Society of Interventional Radiology, American Association of Neurological Surgeons, Congress of Neurological Surgeons, American College of Radiology, American Society of Neuroradiology, American Society of Spine Radiology, Canadian Interventional Radiology Association, Society of Neurointerventional Surgery (2018)[9]	PVP and PBK recommended for painful neoplastic VCF or OVCF refractory to medical therapy	NR
Schweizerische Vereinigung gegen die Osteoporose (2015)[255]	PVP and PBK recommended for painful OVCF refractory to medical therapy	NR

Abbreviations
NICE = National Institute for Health and Care Excellence guidelines, **NR** = not reported, **OVCF** = osteoporotic vertebral fracture, **PBK** = percutaneous balloon kyphoplasty, **PVA** = percutaneous vertebral augmentation (includes vertebroplasty and kyphoplasty), **PVP** = percutaneous vertebroplasty, **VCF** = vertebral compression fracture.
Notes
* = strength of recommendations ranges from A (recommendations based on the highest levels of evidence) to C (recommendations based on the lowest level of evidence [consensus, expert opinion, usual practice or case studies]).

14 Discussion

The objective of this HTA is to evaluate the clinical and economic effectiveness of PVP and PBK and to consider social, legal, ethical and organisational issues relating to their disinvestment. To address the clinical effectiveness of PVP and PBK a systematic search of published literature was undertaken. The search identified 12 RCTs, 2 non-RCTs, 2 database analyses and 15 single-arm trials for PVP, and 4 RCTs, 4 non-RCTs, 2 database analyses and 6 single-arm trials for PBK. The studies varied with respect to eligibility requirements, length of follow-up, comparator and risk of bias. PVP compared to CT was the most commonly studied comparison, with fewer studies comparing PVP to sham and PBK to CT. No study compared PBK to sham. The quality of the evidence ranges from low to moderate.

14.1 Findings of the Clinical Evaluation

The comparative safety of PVP and PBK, as informed by the RCTs, suggests the incidence of mortality, serious adverse events, any adverse events and new fractures is similar to CT. However, the results are uncertain for several reasons: they are likely underpowered to detect group differences; the RCT participants may reflect a comparatively healthier population of adults compared to the Swiss population undergoing the procedures; and the limited follow-up may miss known adverse events associated with NSAIDs and opioids – an effect more problematic in older adults.[256] Consequently, the evidence base was expanded to include analyses of the US Medicare Database. The results from the database analyses contrast with those from the RCTs. PVP and PBK exhibited lower rates of mortality and adverse events compared to CT. For most safety outcomes, the relative difference between PVP or PBK and CT was approximately 2–10%, noting that the absolute rates are not reported and therefore the total number of patients is uncertain. Importantly, adverse events commonly associated with prolonged periods of bed rest and medication use such as bedsores, and cardiac and respiratory adverse events are lower following PVP and PBK, which may account for the reduction in mortality. Interestingly, the reduction in mortality and adverse events is significantly greater following PBK than PVP. Further, patients undergoing PVP have a greater likelihood of reporting pulmonary embolism than those undergoing PBK and CT. While the cause of this is not reported, cement leakage and its subsequent migration to the lungs is a causative factor in pulmonary embolism.[257 258] Thus the increase in prevalence may reflect the increased incidence of cement leakage following PVP compared to PBK (see ***Section 7.8***), noting the interventions are not statistically compared. The lower incidence of cement leakage following PBK is thought to reflect the lower injection pressure required to perform the procedure. This enables more precise placement of cement compared to PVP.[259]

It is important to note, the database analyses are at serious risk of bias, and it is unclear to what extent differences in patient demographics (within the databases and compared to RCTs) underscore the differences in adverse events between treatment arms.

For clinical effectiveness outcomes, there are statistical and clinically significant differences in pain- (VAS) and function-related (ODI and RDQ) outcomes between PVP and CT groups in the short-term. The reduction in ODI and RDQ scores is generally consistent across most timepoints, noting that by 12 months the results remained statistically but not clinically significant. In contrast, the reduction in pain decays over time, with the greatest reductions observed 1 day to 1 month post-intervention and scores decreasing thereafter. Again, by 12 months the score is not clinically significant. However, the clinical interpretation of pain is uncertain. Pain scores surpassed the lowest identified MCID, but if other identified MCIDs are used the results are no longer clinically relevant. Furthermore, pain scores (and most other outcomes) reported at 36 months are uncertain because they are informed by only 1 study with limited patient numbers. Interestingly, the utilisation of analgesic medication generally does not differ between the 2 groups, suggesting while subjective measures of pain decreased, more objective measures did not. This issue may relate to the lack of blinding among the RCTs, which predisposes them to outcome bias because participants have knowledge of the assigned intervention. Outcome bias may lead to over- or under-estimation of the true effect.[260] The effect of PVP on quality of life outcomes is inconsistent. For example, there are limited between-group differences in EQ-5D, and QUALEFFO differed only at 1 week. It is unclear whether the QUALEFFO scores surpass MCID thresholds owing to the lack of published literature addressing this outcome.

There are statistically significant differences between PVP and sham groups for pain (VAS/NRS) from 1 to 12 months post-intervention. Again, analgesic use does not differ between the 2 groups at most timepoints, noting the lack of difference may reflect the pooling of different analgesic classes (NSAIDs and opioids). It is unclear whether NSAIDs or opioids are differentially reduced following PVP because several studies did not provide this level of information. Further, inconsistent statistical differences are observed for QUALEFFO, with PVP and sham groups reporting differences at 3 and 12 months but not at any other timepoint. RDQ statistically differs at 3 months. The effect sizes for most outcomes were small, subject to considerable heterogeneity, and unlikely to translate to clinically meaningful differences. Similar inconsistencies were observed for EQ-5D, with 1 study reporting a statistically significant difference between PVP and sham groups, and another study reporting no difference between the study arms.

Sub-group analyses were used to investigate potential sources of heterogeneity and inconsistency. Acute fractures appear slightly more responsive to PVP at earlier timepoints as the reductions in pain (when compared to CT and sham) and EQ-5D (when compared to sham) are greater compared to older

fractures. At later timepoints, EQ-5D differences persist, however, the reduction in pain is similar between the sub-groups and the effects do not surpass MCIDs. Additional function or quality of life outcomes were unable to be compared via sub-group analysis as the methods of analysing the sub-groups differ.

A small number of trials informed the evidence base comparing PBK to CT. Pain—the only pooled outcome—reports statistically and clinically meaningful differences between PBK and CT in the short-term (one day to 3 months). Like PVP however, the improvement in pain decreases over time and the difference between groups is not clinically meaningful by 12 months. ODI, RDQ and EQ-5D statistically differ between PBK and CT groups, noting most of the outcomes are informed by only 1 study, adding further uncertainty to the results. The differences between groups persisted from 1 to 12 months, although whether they translate to clinical improvements is uncertain. Like PVP, the PBK analysis is subject to considerable outcome bias. Further, sub-group analysis based on fracture age could not be performed owing to the small number of studies identified.

Collectively, PVP and PBK may result in an immediate, clinically relevant short-term improvement in pain, function and some quality of life measures, compared to CT. The clinical relevancy tends to dissipate at later timepoints, albeit the results remain statistically significant. When compared to sham treatments, PVP statistically differs with respect to pain and some quality of life measures. However, the results are not clinically relevant, are inconsistent, and there are generally no functional improvements. For all comparisons the quality of reported outcomes ranged from low to moderate, further limiting the certainty of the results.

14.2 Comparison to Previous Literature

The results from the current meta-analyses are broadly congruent with recently published studies.[94 261] The effect sizes and measures of heterogeneity are mostly similar for pain, QUALEFFO, RDQ and EQ-5D scores for comparisons between PVP and CT/sham. There are slight deviations at 1 and 3 months for pain and QUALEFFO outcomes. This is likely attributable to the different methods of analysis and the pooling of timepoints in previous publications. For example, the meta-analysis in Buchbinder (2018) treats each timepoint as an independent event and consequently 2 and 3 months are pooled together.[94] By contrast, the current HTA utilises longitudinal meta-analyses which calculates the dependence between longitudinal effect sizes. This likely results in a more precise estimate. However, a problem with the longitudinal meta-analysis is that equal weighting is given to each timepoint, consequently timepoints with only 1 study (typically later timepoints) readily influence the model.

The effect sizes and GRADE score in Buchbinder (2018) differ for the PVP and sham comparison.[94] The differences in pain, RDQ and QUALEFFO effect sizes reflect the inclusion of an additional

unpublished trial; the pooling and transformation of ODI scores to RDQ scores; and the imputation of SDs from other trials in Buchbinder (2018).[94] The difference in GRADE scores is reflective of the increased heterogeneity and inconsistency measures—owing to the different statistical approach and smaller sample sizes—due to the absence of pooling ODI and RDQ. Lastly, the current HTA's analysis of fracture age produces differing results to Buchbinder (2018). The authors in Buchbinder (2018) delineated patients from Buchbinder (2009)[7] into acute and sub-acute fractures, thus breaking randomisation.

The results of the PBK meta-analysis are broadly congruent with existing literature owing to the limited RCT evidence base.[262] Existing meta-analyses differ from the current HTA by including non-RCTs or utilising Bayesian statistics.[262 263] Nevertheless, the results remain similar, that is PBK is superior to CT when considering pain, function and quality of life measures.

14.3 Quality and Applicability of the Clinical Evidence

The quality of the reported outcomes is low to moderate as inferred by GRADE. Common sources of downgrading in PVP trials relate to risk of bias and inconsistency. The main bias concerns include the lack of blinding in CT trials, and the large losses to follow-up that may result in an enriched patient population. The considerable levels of heterogeneity and inconsistency add to the uncertainty, which further lowers the quality of outcomes. The inconsistency of effects relates to the small sample sizes (particularly at later timepoints) and the opposing direction of effect in 2 key sham trials.[2 7] The 2 studies have similar risk of bias scores so the difference between the 2 is unlikely methodological in nature. The limited sub-groups analysis suggests the difference may be attributable to fracture age (in the short-term) however additional studies are required to verify this.

Indirectness is the main concern among PBK trials, as the eligibility criteria is not entirely reflective of Swiss reimbursement guidelines. In Switzerland, PBK is currently reimbursed for patients with thoracolumbar fractures younger than 8 weeks old who are unresponsive to analgesics, have pain (VAS ≥ 5), and vertebral deformation (i.e. thoracic kyphosis >15°, lumbar kyphosis >10°, and/or vertebral body height reduction of more than one third compared to adjacent bodies). The eligibility requirements of the included trials generally reflect some but not all of these specifications. Consequently, the evaluated patient population may reflect comparatively healthier (or unhealthier) populations than those currently receiving PBK in Switzerland. Lastly, the pivotal PBK trial (FREE trial) notes the sponsor had an input into the design, monitoring or reporting of results.

The PBK evidence is unlikely to be addressed in the near future. A search of clinical trials databases did not find any ongoing clinical trials evaluating PBK compared to CT, whereas there are 7 ongoing clinical trials evaluating PVP, of which 3 are anticipated to be completed by 2021. The additional trials,

along with the Swiss Implant Registry - SIRIS Spine, will aid in addressing areas of uncertainty associated with PBK and PVP in Switzerland

The clinical interpretation of the evidence is limited by the absence of vertebral fracture-specific MCIDs. Only 2 vertebral fracture-specific MCIDs were identified, both relating to RDQ. The remaining MCIDs generally pertain to chronic back pain requiring surgery. Back pain has a different clinical profile to OVCF with respect to patient demographics, symptomology and treatment expectations. Therefore, the applicability of identified MCIDs is uncertain, as those specific to chronic back pain may over- or under-estimate clinically meaningful thresholds for OVCF.

14.4 Limitations of the Economic Analysis

There are several limitations associated with the health economics analysis performed in this report. For example, there is considerable heterogeneity in EQ-5D scores across key trials comparing PVP to sham or CT.[2 7 132] Factors such as fracture age and nature of the comparator potentially underpin this variation. Further, there are no restrictions on PVP reimbursement in Switzerland. Therefore, patients with differing severities of pain, durations of fracture or degrees of kyphosis can use the subsidised procedure. This potential variation is partly reflected in the Buchbinder (2009) trial, in which 40% (22/55) of patients had acute fractures (younger than 8 weeks) and 60% (33/55) had sub-acute fractures (noting there is limited data on which to base this conclusion).[7] If we consider the Buchbinder (2009) to be representative of the Swiss context, then PVP is not cost-effective because no improvement in quality of life (EQ-5D score) was reported.[7]

If adjusted baseline results from the VERTOS II[132] trial are used at 12 months of follow-up, the procedure is cost-effective, as the ICER of CHF19,669 /QALY is less than the hypothetical willingness-to-pay threshold of CHF100,000/QALY. This result is similar to the cost-effectiveness results presented by Klazen (2010),[132] where an adjusted trial-based ICER of €22,685/QALY gained was calculated. This suggests that the procedure is cost effective in patients with acute fractures (i.e. those younger than 8 weeks). The results are not so clear when the 6 month results are used. The ICER at 6 months using results of the VAPOUR[2] trial was CHF95,361, which is approaching a hypothetical willingness-to-pay threshold of CHF100,000; the ICER estimated using VERTOS II incremental QALYs at 6 months was far higher at CHF84,847 when compared to the ratio at 12 months.

Only the FREE study reported quality of life scores for PBK, thus only 1 study is used to estimate cost-effectiveness of PBK.[168] An ICER of CHF18,405/QALY is estimated, which is similar to Strom's (2010) (£8,800 per QALY, approximately CHF10,361) cost-effectiveness result for PBK in the UK.[190]

Swiss-DRG costs are used to cost PBK, PVP and CT procedures in the economic model as the analysis takes a payer perspective. PVP and PBK are assigned the same Swiss DRG weight for costing.

However, the total costs of the procedures are not reflected in DRG costs. The scoping report noted PBK involved the insertion of a balloon tamp, longer operating times, potential overnight stay and a more expensive delivery system (additional US$3,000 for PBK).[61] The use of DRGs presumes the services are all provided in hospital settings. This is appropriate for PBK as all procedures are performed in an inpatient setting in Switzerland. However, a small proportion of PVP services are provided in ambulatory settings. The use of DRG costs to assess cost-effectiveness of the procedure in this setting leads to uncertainty because reimbursed TARMED costs differ when compared to hospital-delivered PVP.

14.5 Legal, Social, Ethical and Organisational Considerations

There are several social, ethical and organisational considerations associated with the disinvestment of PVP and PBK. From a social and organisational perspective, patients undergoing PVP and PBK have shorter hospital stays and a greater proportion of patients are discharged home than to an assisted living facility. However, the earlier discharge may increase the immediate burden on caregivers. It is unclear if caregiver burden is simply delayed in CT patients who will likely be discharged home at a later timepoint. There is some evidence to suggest that PBK reduces caregiver burden compared to CT at 1 year post-procedure. Collectively, PVP and PBK may reduce the utilisation of healthcare resources and caregiver burden however, additional studies are required to confirm this finding in a Swiss context.

Key ethical concerns relate to the primary demographic of PVP and PBK, that is, older adults. Older adults often receive inadequate treatment owing to complex medical needs. For example, they often have multiple comorbidities, are at a greater risk of a medication-related adverse events, and are particularly susceptible to the dangers associated with prolonged immobilisation (including muscle and bone wastage). Therefore, a procedure that theoretically provides adequate pain relief without additional medication or prolonged periods of bed rest is ideal. However, it is unclear whether PVP has a true “treatment effect”. The lack of clinically meaningful differences between PVP and sham arms suggests the results of PVP are potentially attributable to a confounding treatment (such as local anaesthesia) or a placebo effect. Whether the benefit associated with the procedure (compared to CT) sufficiently outweighs the risks remains to be fully determined.

15 Conclusions

The clinical effectiveness and safety of PVP for OVCF are informed by a mid-sized evidence base of moderate to high quality trials. PVP reports greater reductions in pain (VAS) and function (ODI and RDQ) at 1 month compared to CT. However, the effects do not persist at later timepoints, they are subject to considerable heterogeneity, and are informed by a limited number of patients. There are no differences in quality of life measures (QUALEFFO or EQ-5D) or analgesic use. To investigate causes of heterogeneity, fracture age underwent sub-group analysis. In the short-term, acute fractures are potentially more responsive to PVP with greater reductions in pain compared to older fractures. Again, the effects do not persist at later timepoints. Overall, the quality of reported outcomes, as inferred by GRADE, ranges from moderate to low.

At 1 month, there are statistical differences between PVP and sham with respect to pain (VAS/NRS) but not quality of life (QUALEFFO), function (RDQ) or analgesic use. By 12 months, there are statistical differences in pain and QUALEFFO. The results are subject to substantial heterogeneity, inconsistent differences, small sample sizes and do not surpass identified MCIDs. Acute fractures report slightly greater reductions in pain and some quality of life measures (EQ-5D) compared to older fractures at 1 month. The differences persist at 6 months for EQ-5D but not pain. Overall, the quality of reported outcomes ranges from moderate to low.

PVP has a comparable safety profile to CT when considering the RCT and non-RCT evidence base. However, when considering larger databases with longer follow-up, PVP is potentially safer as the relative incidence of mortality and adverse events are reduced (noting the effect sizes are small to moderate, the absolute rates cannot be determined, and the studies are of serious risk of bias).

The clinical effectiveness and safety of PBK for OVCF are informed by a small-sized evidence base of moderate to low quality trials. PBK significantly improves pain, function and quality of life compared to CT. The improvement in pain reduces over time and is below identified MCIDs at 6 months. However, the difference in quality of life and function persist at later timepoints in 1 study. The quality of reported outcomes, as inferred by GRADE, ranges from low to moderate with key concerns relating to the unclear applicability of included trials.

PBK and CT have comparable safety profiles across the RCT and non-RCT evidence base. When considering larger databases with longer follow-up, PBK is a potentially safer intervention as inferred by lower mortality and adverse events rates (noting the effect size is small to moderate, the absolute rates cannot be determined, and the studies are of serious risk of bias).

The clinical evaluation found limited EQ-5D differences between PVP and CT/sham. However, there is considerable heterogeneity in results. For example, there are no EQ-5D differences in studies that enrolled patients with acute and subacute fractures. Consequently, PVP is not cost-effective in this combined population. However, EQ-5D differences are observed for acute fractures. Therefore, an economic model was developed to ascertain the cost-effectiveness of PVP compared to CT for acute fractures. PVP is estimated to be cost-effective for fractures younger than 8 weeks using 12 months of data from the VERTOS II trial. Results are not as clear for 6 months of follow-up. Similarly, PBK was found to be cost-effective, noting the estimates are derived from 1 trial only. Delisting would result in a net cost saving for the payers, particularly for PVP because this procedure is more widespread. The total cost saving is mainly driven by the assumed substitution rates for PVP and PBK.

The main social and organisational issues relate to the shift in healthcare resources following PVP and PBK as more patients are likely to be discharged home. From an ethical perspective, it is unclear whether PVP and PBK have true “effectiveness” or whether their impact is attributable to a placebo or confounding effect owing to the lack of consistent differences between sham and intervention arms.

16 References

1.Scheuter C, Wertli MM, Haynes AG, et al. Unwarranted regional variation in vertebroplasty and kyphoplasty in Switzerland: A population-based small area variation analysis. *PloS one* 2018;13(12):e0208578. doi: 10.1371/journal.pone.0208578.

2.Clark W, Bird P, Gonski P, et al. Safety and efficacy of vertebroplasty for acute painful osteoporotic fractures (VAPOUR): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet* 2016;388(10052):1408-16.

3.Firanescu CE, de Vries J, Lodder P, et al. Percutaneous Vertebroplasty is no Risk Factor for New Vertebral Fractures and Protects Against Further Height Loss (VERTOS IV). *Cardiovasc Intervent Radiol* 2019; 42:991–1000.

4.Kallmes DF, Comstock BA, Heagerty PJ, et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. *New England Journal of Medicine* 2009;361(6):569-79.

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

6.Yang EZ, Xu JG, Huang GZ, et al. Percutaneous Vertebroplasty Versus Conservative Treatment in Aged Patients With Acute Osteoporotic Vertebral Compression Fractures: A Prospective Randomized Controlled Clinical Study. *Spine (Phila Pa 1976)* 2016;41(8):653-60.

7.Buchbinder R, Osborne RH, Ebeling PR, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med* 2009;361(6):557-68.

8.Blasco J, Martinez-Ferrer A, Macho J, et al. Effect of vertebroplasty on pain relief, quality of life, and the incidence of new vertebral fractures: a 12-month randomized follow-up, controlled trial. *J Bone Miner Res* 2012;27(5):1159-66.

9.American College of Radiology. ACR–ASNR–ASSR–SIR–SNIS practice parameter for the performance of vertebral augmentation 2018, Available from: https://www.acr.org/-/media/ACR/Files/Practice-Parameters/VerebralAug.pdf?la=en, accessed 10 August 2019.

10.McGuire R. AAOS Clinical Practice Guideline: the Treatment of Symptomatic Osteoporotic Spinal Compression Fractures. *The Journal of the American Academy of Orthopaedic Surgeons* 2011;19(3):183-4.

11.Tsoumakidou G, Too CW, Koch G, et al. CIRSE Guidelines on Percutaneous Vertebral Augmentation. *Cardiovasc Intervent Radiol* 2017;40(3):331-42. doi: 10.1007/s00270-017-1574-8.

12.National Institute of Health and Care Excellence (NICE). Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures London2016, Available from: https://www.nice.org.uk/guidance/ta279/chapter/1-Guidance8 May 2019.

13.Medical Services Advisory Committee. 27.1 Review of Kyphoplasty for the Treatment of Vertebral Compression Fracture and Review of Interim Funded Service: Vertebroplasty for the Treatment of Vertebral Compression Fracture 2011, Available from: http://webarchive.nla.gov.au/gov/20160615053438/http://www.msac.gov.au/internet/msac/publishing.nsf/Content/27.1-public.

14.Medical Advisory Secretariat. Public Summary Document: Application No. 1466 – Vertebroplasty for severely painful osteoporotic vertebral fractures: MSAC, 2020.

15.Das Eidgenössische Departement des Innern(EDI). Verordnung des EDI über Leistungen in der obligatorischen Krankenpflegeversicherung: KLV 2019, Available from: https://www.admin.ch/opc/de/classified-compilation/19950275/index.html, accessed 20 April 2019.

16.Johnell O, Kanis J. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Metabolic Bone Diseases* 2006;17(12):1726-33. doi: 10.1007/s00198-006-0172-4.

17.World Health Organization. WHO Scientific Group on the Assessment of Osteoporosis at Primary Care Level, 2004.

18.Chavassieux P, Seeman E, Delmas PD. Insights into material and structural basis of bone fragility from diseases associated with fractures: how determinants of the biomechanical properties of bone are compromised by disease. *Endocrine reviews* 2007;28(2):151. doi: 10.1210/er.2006-0029.

19.Parfitt M. Skeletal Heterogeneity and the Purposes of Bone Remodeling: Implications for the Understanding of Osteoporosis. Osteoporosis. 4 ed2013:855-72.

20.Manolagas SC, Jilka RL. Bone Marrow, Cytokines, and Bone Remodeling — Emerging Insights into the Pathophysiology of Osteoporosis. *The New England Journal of Medicine* 1995;332(5):305-11. doi: 10.1056/NEJM199502023320506.

21.Kenkre JS, Bassett J. The bone remodelling cycle. *Ann Clin Biochem* 2018;55(3):308-27. doi: 10.1177/0004563218759371.

22.Sipos W, Pietschmann P, Rauner M, et al. Pathophysiology of osteoporosis. *Wiener Medizinische Wochenschrift* 2009;159(9):230-34. doi: 10.1007/s10354-009-0647-y.

23.Consensus development conference: Diagnosis, prophylaxis, and treatment of osteoporosis. *The American Journal of Medicine* 1993;94(6):646-50. doi: 10.1016/0002-9343(93)90218-E.

24.Hernlund E, Svedbom A, Ivergard M, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Archives of osteoporosis* 2013;8:136. doi: 10.1007/s11657-013-0136-1.

25.Warriner AH, Patkar NM, Yun H, et al. Minor, major, low-trauma, and high-trauma fractures: what are the subsequent fracture risks and how do they vary? *Curr Osteoporos Rep* 2011;9(3):122-8. doi: 10.1007/s11914-011-0064-1.

26.National Institutes of Health Osteoporosis and Related Bone Diseases. Osteoporosis Overview. In: Health NIo, ed. Bethesda, MD: National Institute of Health, 2018:5.

27.Nesterova AP, Klimov EA, Zharkova M, et al. Chapter 12 - Diseases of the musculoskeletal system. In: Nesterova AP, Klimov EA, Zharkova M, et al., eds. Disease Pathways: Elsevier 2020:533-68.

28.Sozen T, Ozisik L, Basaran NC. An overview and management of osteoporosis. *Eur J Rheumatol* 2017;4(1):46-56. doi: 10.5152/eurjrheum.2016.048.

29.Heshmati HM, Khosla S. Idiopathic osteoporosis: a heterogeneous entity. *Ann Med Interne (Paris)* 1998;149(2):77-81.

30.Mirza F, Canalis E. Management of endocrine disease: Secondary osteoporosis: pathophysiology and management. *Eur J Endocrinol* 2015;173(3):R131-51. doi: 10.1530/EJE-15-0118.

31.Poole K, Compston J. Osteoporosis and its management. *British Medical Journal* 2006;333(7581):1251-56.

32.Old JL, Calvert M. Vertebral compression fractures in the elderly. *Am Fam Physician* 2004;69(1):111-6.

33.Alexandru D, So W. Evaluation and management of vertebral compression fractures. *The Permanente journal* 2012;16(4):46. doi: 10.7812/TPP/12-037.

34.Rajasekaran S, Kanna RM, Schnake KJ, et al. Osteoporotic Thoracolumbar Fractures—How Are They Different?—Classification and Treatment Algorithm. *Journal of Orthopaedic Trauma* 2017;31.

35.Kado DM, Browner WS, Palermo L, et al. Vertebral fractures and mortality in older women: a prospective study. Study of Osteoporotic Fractures Research Group. *Arch Intern Med* 1999;159(11):1215-20. doi: 10.001/archinte.159.11.15.

36.Nazrun AS, Tzar MN, Mokhtar SA, et al. A systematic review of the outcomes of osteoporotic fracture patients after hospital discharge: morbidity, subsequent fractures, and mortality. *Ther Clin Risk Manag* 2014;10:937-48.(doi):10.2147/TCRM.S72456.

37.Garshick E, Kelley A, Cohen SA, et al. A prospective assessment of mortality in chronic spinal cord injury. *Spinal Cord* 2005;43(7):408-16. doi: 10.1038/sj.sc.3101729.

38.Angeli A, Guglielmi G, Dovio A, et al. High prevalence of asymptomatic vertebral fractures in post-menopausal women receiving chronic glucocorticoid therapy: A cross-sectional outpatient study. *Bone* 2006;39(2):253-59. doi: 10.1016/j.bone.2006.02.005

39.McCarthy J, Davis A. Diagnosis and Management of Vertebral Compression Fractures. *American family physician* 2016;94(1):44-50.

40.Lenchik L, Rogers LF, Delmas PD, et al. Diagnosis of osteoporotic vertebral fractures: importance of recognition and description by radiologists. *AJR American journal of roentgenology* 2004;183(4):949. doi: 10.2214/ajr.183.4.1830949

41.Hanrahan CJ, Shah LM. MRI of spinal bone marrow: part 2, T1-weighted imaging-based differential diagnosis. *AJR American journal of roentgenology* 2011;197(6):1309. doi: 10.2214/AJR.11.7420

42.Jin C, Xu G, Weng D, et al. Impact of Magnetic Resonance Imaging on Treatment-Related Decision Making for Osteoporotic Vertebral Compression Fracture: A Prospective Randomized Trial. *Med Sci Monit* 2018;24:50-57.

43.Mauch JT, Carr CM, Cloft H, et al. Review of the Imaging Features of Benign Osteoporotic and Malignant Vertebral Compression Fractures. *AJNR American journal of neuroradiology* 2018;39(9):1584. doi: 10.3174/ajnr.A5528

44.Harvard Health Publishing. Treating osteoporotic fractures of the spine: Harvard University; 2008, Available from: https://www.health.harvard.edu/newsletter_article/Treating_osteoporotic_fractures_of_the_spine.

45.Riggs BL, Melton LJ, 3rd. The worldwide problem of osteoporosis: insights afforded by epidemiology. *Bone* 1995;17(5 Suppl):505s-11s.

46.Bouza C, Lopez-Cuadrado T, Almendro N, et al. Safety of balloon kyphoplasty in the treatment of osteoporotic vertebral compression fractures in Europe: a meta-analysis of randomized controlled trials. *Eur Spine J* 2015;24(4):715-23. doi: 10.1007/s00586-014-3581-7.

47.Panda A, Das CJ, Baruah U. Imaging of vertebral fractures. *Indian J Endocrinol Metab* 2014;18(3):295-303. doi: 10.4103/2230-8210.131140.

48.Suzuki N, Ogikubo O, Hansson T. The prognosis for pain, disability, activities of daily living and quality of life after an acute osteoporotic vertebral body fracture: its relation to fracture level, type of fracture and grade of fracture deformation. *European Spine Journal* 2009;18(1):77-88. doi: 10.1007/s00586-008-0847-y.

49.Chastonay P, Weber D, Mattig T. The Health of Older People in Switzerland. *Journal of Public Health International* 2018;1(2):5-8. doi: 10.14302/issn.2641-4538.jphi-18-2426.

50.Keller TS, Harrison DE, Colloca CJ, et al. Prediction of osteoporotic spinal deformity. *Spine (Phila Pa 1976)* 2003;28(5):455-62. doi: 10.1097/01.Brs.0000048651.92777.30.

51.Mayo Clinic Staff. Overview: Mayo Foundation for Medical Education and Research (MFMER); 2020, Available from: https://www.mayoclinic.org/diseases-conditions/kyphosis/symptoms-causes/syc-20374205, accessed 20 February 2020.

52.Hall SE, Criddle RA, Comito TL, et al. A Case–Control Study of Quality of Life and Functional Impairment in Women with Long–Standing Vertebral Osteoporotic Fracture. *Osteoporosis International* 1999;9(6):508-15. doi: 10.1007/s001980050178.

53.Quraishi NA, Sahota O, Fairbank J. Aetiology: BMJ Publishing Group; 2019, Available from: https://bestpractice-bmj-com.proxy.library.adelaide.edu.au/topics/en-gb/819, accessed 20 February 2020.

54.Svedbom A, Ivergard M, Hernlund E, et al. Epidemiology and economic burden of osteoporosis in Switzerland. *Archives of osteoporosis* 2014;9:187. doi: 10.1007/s11657-014-0187-y.

55.Lippuner K, Golder M, Greiner R. Epidemiology and direct medical costs of osteoporotic fractures in men and women in Switzerland. *Osteoporos Int* 2005;16(Suppl 2):S8-S17. doi: 10.1007/s00198-004-1696-0.

56.Denaro V, Longo UG, Maffulli N, et al. Vertebroplasty and kyphoplasty. *Clinical cases in mineral and bone metabolism : the official journal of the Italian Society of Osteoporosis, Mineral Metabolism, and Skeletal Diseases* 2009;6(2):125. doi: 10.1016/j.bone.2009.01.119.

57.Wang H, Sribastav SS, Ye F, et al. Comparison of Percutaneous Vertebroplasty and Balloon Kyphoplasty for the Treatment of Single Level Vertebral Compression Fractures: A Meta-analysis of the Literature. *Pain physician* 2015;18(3):209-22.

58.Gu CN, Brinjikji W, Evans AJ, et al. Outcomes of vertebroplasty compared with kyphoplasty: a systematic review and meta-analysis. *Journal of neurointerventional surgery* 2016;8(6):636-42.

59.Liu Q, Cao J, Kong JJ. Clinical effect of balloon kyphoplasty in elderly patients with multiple osteoporotic vertebral fracture. *Nigerian journal of clinical practice* 2019;22(3):289-92.

60.Li Y, Zhu J, Xie C. A comparative study of percutaneous kyphoplasty and conservative therapy on vertebral osteoporotic compression fractures in elderly patients. *International Journal of Clinical and Experimental Medicine* 2017;10(5):8139-45.

61.Mathis JM, Ortiz AO, Zoarski GH. Vertebroplasty versus kyphoplasty: a comparison and contrast. *AJNR American journal of neuroradiology* 2004;25(5):840-5.

62.Nussbaum DA, Gailloud P, Murphy K. A Review of Complications Associated with Vertebroplasty and Kyphoplasty as Reported to the Food and Drug Administration Medical Device Related Web Site. *Journal of Vascular and Interventional Radiology* 2004;15(11):1185-92. doi: 10.1097/01.RVI.0000144757.14780.E0.

63.Hubschle L, Borgstrom F, Olafsson G, et al. Real-life results of balloon kyphoplasty for vertebral compression fractures from the SWISSspine registry. *Spine J* 2014;14(9):2063-77.

64.Schupfner R, Stoevelaar HJ, Blattert T, et al. Treatment of Osteoporotic Vertebral Compression Fractures: Applicability of Appropriateness Criteria in Clinical Practice. *Pain physician* 2016;19(1):E113-20.

65.Benneker LM, Hoppe S. Percutaneous cement augmentation techniques for osteoporotic spinal fractures. *European Journal of Trauma and Emergency Surgery* 2013;39(5):445-53. doi: 10.1007/s00068-013-0265-7.

66.Röllinghoff M, Zarghooni K, Schlüter-Brust K, et al. Indications and contraindications for vertebroplasty and kyphoplasty. *Archives of Orthopaedic and Trauma Surgery* 2010;130(6):765-74. doi: 10.1007/s00402-010-1083-6.

67.Swiss Society of Spinal Surgery. Schweizerische Gesellschaft für Spinale Chirurgie 2004, Available from: https://www.bag.admin.ch/dam/bag/de/dokumente/kuv-leistungen/referenzdokumente-klv-anhang-1/01.3-richtlinien-vom-23.09.2004-betreffend-ballon-kyphoplastie-zur-behandlung-von-wirbelkoerperfrakturen.pdf.download.pdf/01.3%20Richtlinien%20der%20Schweizerische%20Gesellschaft%20f%C3%BCr%20Spinale%20Chirurgie%20vom%2023.09.2004%20betreffend%20Ballon-Kyphoplastie%20zur%20Behandlung%20von%20Wirbelk%C3%B6rperfrakturen%20(in%20englisch).pdf, accessed 20 April 2010.

68.Canadian Agency for Drugs and Technologies in Helath (CADTH). Balloon Kyphoplasty for the Treatment of Vertebral Compression Fractures: A Review of the Guidelines and Clinical and Cost-Effectiveness 2008, Available from: https://www.cadth.ca/balloon-kyphoplasty-

treatment-vertebral-compression-fractures-adults-clinical-and-cost-effectiveness, accessed 10 May 2019.

69.Eckel TS, Olan W. Vertebroplasty and Vertebral Augmentation Techniques. *Techniques in Vascular and Interventional Radiology* 2009;12(1):44-50.

70.Nerve Blockade and Interventional Therapy. 1st ed. 2019. ed. Tokyo: Springer Japan 2019.

71.Lamy O, Uebelhart B, Aubry-Rozier B. Risks and benefits of percutaneous vertebroplasty or kyphoplasty in the management of osteoporotic vertebral fractures. *Osteoporos Int* 2014;25(3):807-19.

72.Vogl TJ, Pflugmacher R, Hierholzer J, et al. Cement directed kyphoplasty reduces cement leakage as compared with vertebroplasty: results of a controlled, randomized trial. *Spine (Phila Pa 1976)* 2013;38(20):1730-6. doi: 10.1097/BRS.0b013e3182a14d15.

73.Zhang H, Xu C, Zhang T, et al. Does percutaneous vertebroplasty or balloon kyphoplasty for osteoporotic vertebral compression fractures increase the incidence of new vertebral fractures? A meta-analysis. . *Pain physician* 2017;20(1):E13-28.

74.ISPM and Obsan. Schweizer Atlas der Gesundheitsversorgung - Vertebroplastik Bern: Schweizer Akutspitälern; 2017, Available from: http://versorgungsatlas.ch/index.php/de/VERT/, accessed 6 May 2019.

75.ISPM and Osban. Schweizer Atlas der Gesundheitsversorgung - Kyphoplastik Bern: Schweizer Akutspitälern; 2017, Available from: http://versorgungsatlas.ch/index.php/de/KYPH-2/, accessed 6 May 2019.

76.Esses SI, McGuire R, Jenkins J, et al. American Academy of Orthopaedic Surgeons clinical practice guideline on: the treatment of osteoporotic spinal compression fractures. *The Journal of bone and joint surgery American volume* 2011;93(20):1934-6. doi: 10.2106/JBJS.9320ebo.

77.Cha KH, Cho TG, Kim CH, et al. Comparative study on the period of absolute bed rest of vertebral compression fracture. *Korean Journal of Spine* 2013;10(3):144. doi: 10.14245/kjs.2013.10.3.144.

78.Barr JD, Jensen ME, Hirsch JA, et al. Position statement on percutaneous vertebral augmentation: a consensus statement developed by the Society of Interventional Radiology (SIR), American Association of Neurological Surgeons (AANS) and the Congress of Neurological Surgeons (CNS), American College of Radiology (ACR), American Society of Neuroradiology (ASNR), American Society of Spine Radiology (ASSR), Canadian Interventional Radiology Association (CIRA), and the Society of NeuroInterventional Surgery (SNIS). *J Vasc Interv Radiol* 2014;25(2):171-81. doi: 10.1016/j.jvir.2013.10.001.

79.Sinaki M, Itoi E, Wahner HW, et al. Stronger back muscles reduce the incidence of vertebral fractures: a prospective 10 year follow-up of postmenopausal women. *Bone* 2002;30(6):836-41. doi: 10.1016/S8756-3282(02)00739-1.

80.Prather H, Hunt D, Watson JO, et al. Conservative care for patients with osteoporotic vertebral compression fractures. *Physical medicine and rehabilitation clinics of North America* 2007;18(3):577-91, xi. doi: 10.1016/j.pmr.2007.05.008.

81.Megale RZ, Pollack A, Britt H, et al. Management of vertebral compression fracture in general practice: BEACH program. *PloS one* 2017;12(5):e0176351-e51. doi: 10.1371/journal.pone.0176351.

82.Els C, Jackson TD, Kunyk D, et al. Adverse events associated with medium- and long-term use of opioids for chronic non-cancer pain: an overview of Cochrane Reviews. *Cochrane Database Syst Rev* 2017;10:CD012509.(doi):10.1002/14651858.CD012509.pub2.

83.Coxib and traditional NSAID Trialists' (CNT) Collaboration, Bhala N, Emberson J, et al. Vascular and upper gastrointestinal effects of non-steroidal anti-inflammatory drugs: meta-analyses of individual participant data from randomised trials. *Lancet* 2013;382(9894):769-79. doi: 10.1016/S0140-6736(13)60900-9.

84.Zuo X-H, Zhu X-P, Bao H-G, et al. Network meta-analysis of percutaneous vertebroplasty, percutaneous kyphoplasty, nerve block, and conservative treatment for nonsurgery options

of acute/subacute and chronic osteoporotic vertebral compression fractures (OVCFs) in short-term and long-term effects. *Medicine (Baltimore)* 2018;97(29):e11544-e44. doi: 10.1097/MD.0000000000011544.

85. Chandler G, Dalley G, Hemmer Jr J, et al. Gray ramus communicans nerve block: novel treatment approach for painful osteoporotic vertebral compression fracture. *Southern medical journal* 2001;94(4):387-93.

86. Pehora C, Pearson AM, Kaushal A, et al. Dexamethasone as an adjuvant to peripheral nerve block. *Cochrane Database Syst Rev* 2017;11(11):CD011770-CD70. doi: 10.1002/14651858.CD011770.pub2.

87. Chakravarthy K, Richter H, Christo PJ, et al. Spinal Cord Stimulation for Treating Chronic Pain: Reviewing Preclinical and Clinical Data on Paresthesia-Free High-Frequency Therapy. *Neuromodulation : journal of the International Neuromodulation Society* 2018;21(1):10-18. doi: 10.1111/ner.12721.

88. Pang D. Current experience of spinal neuromodulation in chronic pain: Is there a role in children and young people? *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society* 2017;21(1):56-66. doi: 10.1016/j.ejpn.2016.07.001.

89. Bell J, Blacker N, Edwards S, et al. Osteoporosis Pharmacological prevention and management in older people. *Australian Family Physician* 2012;41:110-18.

90. Black DM, Bauer DC, Schwartz AV, et al. Continuing bisphosphonate treatment for osteoporosis--for whom and for how long? *N Engl J Med* 2012;366(22):2051-3. doi: 10.1056/NEJMp1202623.

91. Gambacciani M, Levancini M. Management of postmenopausal osteoporosis and the prevention of fractures. *Panminerva medica* 2014;56(2):115-31.

92. Aubry-Rozier B, Ecker T, Saliou G, et al. Vertebroplasty: the new CHUV consensus. *Rev Med Suisse* 2020;16(685):492-97.

93. Shah LM, Jennings JW, Kirsch CFE, et al. ACR Appropriateness Criteria((R)) Management of Vertebral Compression Fractures. *J Am Coll Radiol* 2018;15(11S):S347-S64. doi: 10.1016/j.jacr.2018.09.019.

94. Buchbinder R, Johnston RV, Rischin KJ, et al. Percutaneous vertebroplasty for osteoporotic vertebral compression fracture. *Cochrane Database Syst Rev* 2018;11:Cd006349. doi: 10.1002/14651858.CD006349.pub4.

95. Firanescu CE, de Vries J, Lodder P, et al. Vertebroplasty versus sham procedure for painful acute osteoporotic vertebral compression fractures (VERTOS IV): randomised sham controlled clinical trial. *Bmj* 2018;361:k1551.

96. Rousing R, Andersen MO, Jespersen SM, et al. Percutaneous vertebroplasty compared to conservative treatment in patients with painful acute or subacute osteoporotic vertebral fractures: three-months follow-up in a clinical randomized study. *Spine (Phila Pa 1976)* 2009;34(13):1349-54.

97. Müller KR. Switzerland - Medical Devices and Diagnostics: ISPOR; 2012, Available from: https://tools.ispor.org/HTARoadMaps/SwitzerlandMDD.asp#3, accessed 6 March 2020.

98. Joimax GmbH. Kyphoplasty: Joimax GmbH; 2019, Available from: https://www.joimax.com/en/products/vba-management-program/kyphoplasty/, accessed 25 February 2020.

99. Medtronic. Kyphon V Premium Vertebroplasty: Medtronic; 2020, Available from: https://www.medtronic.com/us-en/healthcare-professionals/products/spinal-orthopaedic/vertebroplasty/kyphon-v-vertebroplasty.html, accessed 25 February 2020.

100. Medtronic. Kyphon: Medtronic; 2019, Available from: https://www.medtronic.com/us-en/healthcare-professionals/products/spinal-orthopaedic/balloon-kyphoplasty/kyphon-balloon-kyphoplasty.html accessed 25 February 2020.

101. Merit Medical Systems. StabiliT® VP Vertebroplasty System: Merit Medical Systems; 2020, Available from: https://www.merit.com/merit-interventional/vertebral-compression-

fracture/vcf/stabiliit-vp-vertebral-augmentation-system/#toggle-id-2 accessed 22 February 2020.

102.Merit Medical Systems. Osseoflex® SN Steerable Needle: Merit Medical Systems; 2020, accessed February 25 2020.

103.Joimax GmbH. Vertebroplasty: Joimax GmbH; 2019, Available from: https://www.joimax.com/en/products/vba-management-program/kyphoplasty/ accessed 25 February 2020.

104.Health Information and Quality Authority. Vertebroplasty and kyphoplasty for osteoporotic vertebral compression fractures. Health Technology Assessment of Scheduled Procedures: Health Information and Quality Authority, 2013.

105.National Institute of Health and Care Excellence (NICE). Vertebroplasty and kyphoplasty for spinal compression fractures caused by osteoporosis: National Institute of Health and Care Excellence (NICE), 2013.

106.Chandra RV, Meyers PM, Hirsch JA, et al. Vertebral augmentation: report of the Standards and Guidelines Committee of the Society of NeuroInterventional Surgery. *Journal of NeuroInterventional Surgery* 2014;6(1):7. doi: 10.1136/neurintsurg-2013-011012.

107.Kovacs FM, Abraira V, Royuela A, et al. Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. *BMC Musculoskelet Disord* 2008;9:43.(doi):10.1186/471-2474-9-43.

108.Maughan EF, Lewis JS. Outcome measures in chronic low back pain. *Eur Spine J* 2010;19(9):1484-94. doi: 10.007/s00586-010-1353-6.

109.Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa* 2008;33(1):90-4. doi: 10.1097/BRS.0b013e31815e3a10.

110.Parker SL, Adogwa O, Mendenhall SK, et al. Determination of minimum clinically important difference (MCID) in pain, disability, and quality of life after revision fusion for symptomatic pseudoarthrosis. *Spine J* 2012;12(12):1122-8. doi: 10.016/j.spinee.2012.10.006.

111.Parker SL, Godil SS, Shau DN, et al. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. *J Neurosurg Spine* 2013;18(2):154-60. doi: 10.3171/2012.10.SPINE12312.

112.Lee MK, Yost KJ, McDonald JS, et al. Item response theory analysis to evaluate reliability and minimal clinically important change of the Roland-Morris Disability Questionnaire in patients with severe disability due to back pain from vertebral compression fractures. *Spine J* 2017;17(6):821-29. doi: 10.1016/j.spinee.2017.01.002.

113.Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa* 2000;25(24):3115-24. doi: 10.1097/00007632-200012150-00006.

114.Copay AG, Glassman SD, Subach BR, et al. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J* 2008;8(6):968-74. doi: 10.1016/j.spinee.2007.11.006.

115.Johnsen LG, Hellum C, Nygaard OP, et al. Comparison of the SF6D, the EQ5D, and the oswestry disability index in patients with chronic low back pain and degenerative disc disease. *BMC Musculoskelet Disord* 2013;14:148.(doi):10.1186/471-2474-14-148.

116.Lauridsen HH, Hartvigsen J, Manniche C, et al. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskelet Disord* 2006;7:82.(doi):10.1186/471-2474-7-82.

117.ICH Expert Working Group. Structure and content of clinical study reports E3: International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use, 1995.

118.Trout AT, Kallmes DF, Kaufmann TJ. New fractures after vertebroplasty: adjacent fractures occur significantly sooner. *AJNR American journal of neuroradiology* 2006;27(1):217-23.

119.EUnetHTA Joint Action 2 Work Package 8. HTA Core Model ® version 3.0 (PDF) 2016, Available from: www.htacoremodel.info/BrowseModel.aspx.

120.Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews* 2016;5(1) doi: 10.1186/s13643-016-0384-4.

121.Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.(doi):10.1136/bmj.l4898.

122.Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;355:i4919.(doi):10.1136/bmj.i4919.

123.Institute of Health Economics (IHE). Quality Appraisal of Case Series Studies Checklist Edmonton (AB): Institute of Health Economics; 2014, Available from: http://www.ihe.ca/research-programs/rmd/cssqac/cssqac-about.

124.Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.(doi):10.1136/bmj.j4008.

125.GRADEproGDT. GRADEpro Guideline Development Tool [Software]: McMaster University (developed by Evidence Prime, Inc.); 2015, Available from: gradepro.org.

126.The Cochrane Collaboration. Review Manager (RevMan) [Computer program] Version 5.3. In: Centre TNC, ed.: Copenhagen, DK, 2014.

127.Ishak KJ, Platt RW, Joseph L, et al. Meta-analysis of longitudinal studies. *Clin Trials* 2007;4(5):525-39 doi: 10.1177/1740774507083567.

128.Aue A, Hormann S, Horvath L, et al. Break Detection in the Covariance Structure of Multivariate Time Series Models. *The Annals of Statistics* 2009;37(6B):4046-87.

129.Bahreini M, Jalili M, Moradi-Lakeh M. A comparison of three self-report pain scales in adults with acute pain. *J Emerg Med* 2015;48(1):10-8. doi: 10.1016/j.jemermed.2014.07.039.

130.Gajasinghe S, Wijayaratna M, Abayadeera A. Correlation between numerical rating scale (NRS) and visual analogue scale (VAS) in assessment of pain in post operative patients. *Sri Lankan Journal of Anaesthesiology* 2010;18(2):81-83.

131.Higgins JPT., Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011.

132.Klazen CA, Lohle PN, de Vries J, et al. Vertebroplasty versus conservative treatment in acute osteoporotic vertebral compression fractures (Vertos II): an open-label randomised trial. *Lancet* 2010;376(9746):1085-92. doi: 10.1016/s0140-6736(10)60954-3.

133.International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. Clinical Safety Data Management: Definitions and Standards for Expediated Reporting E2A. 1994.

134.Chen D, An ZQ, Song S, et al. Percutaneous vertebroplasty compared with conservative treatment in patients with chronic painful osteoporotic spinal fractures. *J Clin Neurosci* 2014;21(3):473-7. doi: 10.1016/j.jocn.2013.05.017.

135.Farrokhi MR, Alibai E, Maghami Z. Randomized controlled trial of percutaneous vertebroplasty versus optimal medical management for the relief of pain and disability in acute osteoporotic vertebral compression fractures. *J Neurosurg Spine* 2011;14(5):561-9. doi: 10.3171/2010.12.SPINE10286.

136.Klazen CA, Venmans A, de Vries J, et al. Percutaneous vertebroplasty is not a risk factor for new osteoporotic compression fractures: results from VERTOS II. *AJNR Am J Neuroradiol* 2010;31(8):1447-50. doi: 10.3174/ajnr.A2148.

137.Venmans A, Klazen CA, Lohle PN, et al. Percutaneous vertebroplasty and pulmonary cement embolism: results from VERTOS II. *AJNR Am J Neuroradiol* 2010;31(8):1451-3. doi: 10.3174/ajnr.A2127.

138.Venmans A, Klazen CA, van Rooij WJ, et al. Postprocedural CT for perivertebral cement leakage in percutaneous vertebroplasty is not necessary--results from VERTOS II. *Neuroradiology* 2011;53(1):19-22. doi: 10.1007/s00234-010-0705-6.

139.Leali PT, Solla F, Maestretti G, et al. Safety and efficacy of vertebroplasty in the treatment of osteoporotic vertebral compression fractures: a prospective multicenter international randomized controlled study. *Clin Cases Miner Bone Metab* 2016;13(3):234-36. doi: 10.11138/ccmbm/2016.13.3.234.

140.Rousing R, Hansen KL, Andersen MO, et al. Twelve-months follow-up in forty-nine patients with acute/semiacute osteoporotic vertebral fractures treated conservatively or with percutaneous vertebroplasty: a clinical randomized study. *Spine (Phila Pa 1976)* 2010;35(5):478-82. doi: 10.1097/BRS.0b013e3181b71bd1.

141.Voormolen MH, Mali WP, Lohle PN, et al. Percutaneous vertebroplasty compared with optimal pain medication treatment: short-term clinical outcome of patients with subacute or chronic painful osteoporotic vertebral compression fractures. The VERTOS study. *AJNR American journal of neuroradiology* 2007;28(3):555-60.

142.Kroon F, Staples M, Ebeling PR, et al. Two-year results of a randomized placebo-controlled trial of vertebroplasty for acute osteoporotic vertebral fractures. *J Bone Miner Res* 2014;29(6):1346-55. doi: 10.002/jbmr.2157.

143.Staples MP, Howe BM, Ringler MD, et al. New vertebral fractures after vertebroplasty: 2-year results from a randomised controlled trial. *Archives of osteoporosis* 2015;10:229.(doi):10.1007/s11657-015-0229-0.

144.Firanescu C, Lohle PN, de Vries J, et al. A randomised sham controlled trial of vertebroplasty for painful acute osteoporotic vertebral fractures (VERTOS IV). *Trials* 2011;12:93.(doi):10.1186/745-6215-12-93.

145.Comstock BA, Sitlani CM, Jarvik JG, et al. Investigational vertebroplasty safety and efficacy trial (INVEST): patient-reported outcomes through 1 year. *Radiology* 2013;269(1):224-31. doi: 10.1148/radiol.13120821.

146.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. doi: 10.2106/JBJS.K.01649.

147.Ong KL, Beall DP, Frohbergh M, et al. Were VCF patients at higher risk of mortality following the 2009 publication of the vertebroplasty "sham" trials? *Osteoporos Int* 2018;29(2):375-83. doi: 10.1007/s00198-017-4281-z.

148.Andrei D, Popa I, Brad S, et al. The variability of vertebral body volume and pain associated with osteoporotic vertebral fractures: conservative treatment versus percutaneous transpedicular vertebroplasty. *Int Orthop* 2017;41(5):963-68.

149.Diamond TH, Bryant C, Browne L, et al. Clinical outcomes after acute osteoporotic vertebral fractures: a 2-year non-randomised trial comparing percutaneous vertebroplasty with conservative therapy. *Med J Aust* 2006;184(3):113-7.

150.Diamond TH, Champion B, Clark WA. Management of acute osteoporotic vertebral fractures: a nonrandomized trial comparing percutaneous vertebroplasty with conservative therapy. *Am J Med* 2003;114(4):257-65.

151.Al-Ali F, Barrow T, Luke K. Vertebroplasty: what is important and what is not. *AJNR American journal of neuroradiology* 2009;30(10):1835-9.

152.Bae H, Hatten HP, Jr., Linovitz R, et al. A prospective randomized FDA-IDE trial comparing Cortoss with PMMA for vertebroplasty: a comparative effectiveness research study with 24-month follow-up. *Spine (Phila Pa 1976)* 2012;37(7):544-50.

153.DePalma MJ, Ketchum JM, Frankel BM, et al. Percutaneous vertebroplasty for osteoporotic vertebral compression fractures in the nonagenarians: a prospective study evaluating pain reduction and new symptomatic fracture rate. *Spine (Phila Pa 1976)* 2011;36(4):277-82.

154.Fenoglio L, Cena P, Migliore E, et al. Vertebroplasty in the treatment of osteoporosis vertebral fractures: report on 52 cases. *J Endocrinol Invest* 2008;31(9):795-8.

155.Kotwica Z, Saracen A. Early and long-term outcomes of vertebroplasty for single osteoporotic fractures. *Neurol Neurochir Pol* 2011;45(5):431-5.

156.Masala S, Mammucari M, Angelopoulos G, et al. Percutaneous vertebroplasty in the management of vertebral osteoporotic fractures. Short-term, mid-term and long-term follow-up of 285 patients. *Skeletal Radiol* 2009;38(9):863-9.

157.Masala S, Nano G, Marcia S, et al. Osteoporotic vertebral compression fractures augmentation by injectable partly resorbable ceramic bone substitute (Cerament™|SPINE SUPPORT): a prospective nonrandomized study. *Neuroradiology* 2012;54(6):589-96.

158.Nieuwenhuijse MJ, Muijs SP, van Erkel AR, et al. A clinical comparative study on low versus medium viscosity polymethylmetacrylate bone cement in percutaneous vertebroplasty: viscosity associated with cement leakage. *Spine (Phila Pa 1976)* 2010;35(20):E1037-44.

159.Nieuwenhuijse MJ, van Erkel AR, Dijkstra PD. Percutaneous vertebroplasty for subacute and chronic painful osteoporotic vertebral compression fractures can safely be undertaken in the first year after the onset of symptoms. *J Bone Joint Surg Br* 2012;94(6):815-20.

160.Pitton MB, Herber S, Bletz C, et al. CT-guided vertebroplasty in osteoprotic vertebral fractures: incidence of secondary fractures and impact of intradiscal cement leakages during follow-up. *Eur Radiol* 2008;18(1):43-50.

161.Saracen A, Kotwica Z. Treatment of multiple osteoporotic vertebral compression fractures by percutaneous cement augmentation. *Int Orthop* 2014;38(11):2309-12.

162.Voormolen MH, Lohle PN, Juttmann JR, et al. The risk of new osteoporotic vertebral compression fractures in the year after percutaneous vertebroplasty. *J Vasc Interv Radiol* 2006;17(1):71-6.

163.Voormolen MH, Lohle PN, Lampmann LE, et al. Prospective clinical follow-up after percutaneous vertebroplasty in patients with painful osteoporotic vertebral compression fractures. *J Vasc Interv Radiol* 2006;17(8):1313-20.

164.Dohm M, Black CM, Dacre A, et al. A randomized trial comparing balloon kyphoplasty and vertebroplasty for vertebral compression fractures due to osteoporosis. *AJNR American journal of neuroradiology* 2014;35(12):2227-36.

165.Santiago FR, Abela AP, Alvarez LG, et al. Pain and functional outcome after vertebroplasty and kyphoplasty. A comparative study. *Eur J Radiol* 2010;75(2):e108-13.

166.Jensen MP, Chen C, Brugger AM. Interpretation of visual analog scale ratings and change scores: a reanalysis of two clinical trials of postoperative pain. *J Pain* 2003;4(7):407-14. doi: 10.1016/s1526-5900(03)00716-8.

167.Van Meirhaeghe J, Bastian L, Boonen S, et al. A randomized trial of balloon kyphoplasty and nonsurgical management for treating acute vertebral compression fractures: vertebral body kyphosis correction and surgical parameters. *Spine (Phila Pa 1976)* 2013;38(12):971-83. doi: 10.1097/BRS.0b013e31828e8e22.

168.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(9668):1016-24. doi: 10.1016/s0140-6736(09)60010-6.

169.Eidt-Koch D, Greiner W. Quality of life results of balloon kyphoplasty versus non surgical management for osteoporotic vertebral fractures in Germany. *Health Economics Review* 2011;1(1):1-7.

170.Giannotti S, Carmassi F, Bottai V, et al. Comparison of 50 vertebral compression fractures treated with surgical (kyphoplasty) or non surgical approach. *Clin Cases Miner Bone Metab* 2012;9(3):184-6.

171.Grafe IA, Da Fonseca K, Hillmeier J, et al. Reduction of pain and fracture incidence after kyphoplasty: 1-year outcomes of a prospective controlled trial of patients with primary osteoporosis. *Osteoporos Int* 2005;16(12):2005-12.

172.Kasperk C, Grafe IA, Schmitt S, et al. Three-year outcomes after kyphoplasty in patients with osteoporosis with painful vertebral fractures. *J Vasc Interv Radiol* 2010;21(5):701-9.

173.Kasperk C, Hillmeier J, Noldge G, et al. Treatment of painful vertebral fractures by kyphoplasty in patients with primary osteoporosis: a prospective nonrandomized controlled study. *J Bone Miner Res* 2005;20(4):604-12.

174.Movrin I, Vengust R, Komadina R. Adjacent vertebral fractures after percutaneous vertebral augmentation of osteoporotic vertebral compression fracture: a comparison of balloon kyphoplasty and vertebroplasty. *Arch Orthop Trauma Surg* 2010;130(9):1157-66.

175.Hillmeier J, Grafe I, Da Fonseca K, et al. [The evaluation of balloonkyphoplasty for osteoporotic vertebral fractures. An interdisciplinary concept]. *Orthopade* 2004;33(8):893-904.

176.Prokop A, Koukal C, Dolezych R, et al. [Kyphoplasty in the treatment of osteoporotic spine fractures. Experience in over 500 patients]. *Z Gerontol Geriatr* 2012;45(8):756-60.

177.Robinson Y, Tschoke SK, Stahel PF, et al. Complications and safety aspects of kyphoplasty for osteoporotic vertebral fractures: a prospective follow-up study in 102 consecutive patients. *Patient Saf Surg* 2008;2:2.(doi):10.1186/754-9493-2-2.

178. Mallen SR, Essex MN, Zhang R. Gastrointestinal tolerability of NSAIDs in elderly patients: a pooled analysis of 21 randomized clinical trials with celecoxib and nonselective NSAIDs. *Curr Med Res Opin* 2011;27(7):1359-66. doi: 10.185/03007995.2011.581274.

179.Busse JW, Wang L, Kamaleldin M, et al. Opioids for Chronic Noncancer Pain: A Systematic Review and Meta-analysis. *JAMA* 2018;320(23):2448-60. doi: 10.1001/jama.2018.18472.

180.Roberts E, Delgado Nunes V, Buckner S, et al. Paracetamol: not as safe as we thought? A systematic literature review of observational studies. *Ann Rheum Dis* 2016;75(3):552-9. doi: 10.1136/annrheumdis-2014-206914.

181.Machado GC, Maher CG, Ferreira PH, et al. Efficacy and safety of paracetamol for spinal pain and osteoarthritis: systematic review and meta-analysis of randomised placebo controlled trials. *BMJ* 2015;350:h1225.(doi):10.1136/bmj.h225.

182.von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of clinical epidemiology* 2008;61(4):344-9. doi: 10.1016/j.jclinepi.2007.11.008.

183.Federal Statistical Office (FSO). Population by migration status 2018, Available from: https://www.bfs.admin.ch/bfs/en/home/statistics/population/migration-integration/by-migration-status.html, accessed 27 March 2020.

184.Federal Statistical Office (FSO). International migration 2019, Available from: https://www.bfs.admin.ch/bfs/en/home/statistics/population/migration-integration/international-migration.html, accessed 27 March 2020.

185.Ortiz AO, Buonocore BJ, Zoarski GH. Height restoration following kyphoplasty for treatment of painful osteoporotic vertebral compression fractures. *Journal of Women's Imaging* 2005;7(3):102-10.

186.NSW Agency for Clinical Innovation. Radiology Clinician Fact Sheet ACI, 2012:2.

187.SwissDRG AG. Datenspiegel, SwissDRG 8.0 2019, Available from: https://datenspiegel80.swissdrg.org/drgs?locale=de, accessed 30 October 2020.

188.Borgstrom F, Beall DP, Berven S, et al. Health economic aspects of vertebral augmentation procedures. *Osteoporos Int* 2015;26(4):1239-49. doi: 10.1007/s00198-014-2953-5.

189.Martelli N, Devaux C, van den Brink H, et al. A Systematic Review of the Level of Evidence in Economic Evaluations of Medical Devices: The Example of Vertebroplasty and Kyphoplasty. *PloS one* 2015;10(12):e0144892. doi: 10.1371/journal.pone.0144892.

190.Strom O, Leonard C, Marsh D, et al. Cost-effectiveness of balloon kyphoplasty in patients with symptomatic vertebral compression fractures in a UK setting. *Osteoporos Int* 2010;21(9):1599-608. doi: 10.1007/s00198-009-1096-6.

191.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(17):1-290. doi: 10.3310/hta18170.

192.Svedbom A, Alvares L, Cooper C, et al. Balloon kyphoplasty compared to vertebroplasty and nonsurgical management in patients hospitalised with acute osteoporotic vertebral compression fracture: a UK cost-effectiveness analysis. *Osteoporos Int* 2013;24(1):355-67. doi: 10.1007/s00198-012-2102-y.

193.Hulme PA, Krebs J, Ferguson SJ, et al. Vertebroplasty and kyphoplasty: a systematic review of 69 clinical studies. *Spine (Phila Pa 1976)* 2006;31(17):1983-2001. doi: 10.1097/01.brs.0000229254.89952.6b.

194.Lee MJ, Dumonski M, Cahill P, et al. Percutaneous treatment of vertebral compression fractures: a meta-analysis of complications. *Spine (Phila Pa 1976)* 2009;34(11):1228-32. doi: 10.1097/BRS.0b013e3181a3c742.

195.Swiss DRG AG. DRG I10B - Other spine procedures with certain procedures or halo traction or age <16 years or complex diagnosis 2019, Available from: https://datenspiegel80.swissdrg.org/drgs/5ba4b5e23bba1e5e8056aff7?locale=de, accessed 27 March 2020.

196.Federal Statistical Office (FSO). National projections 2015, Available from: https://www.bfs.admin.ch/bfs/en/home/statistics/population/population-projections/national-projections.html, accessed 27 March 2020.

197.Liu JT, Liao WJ, Tan WC, et al. Balloon kyphoplasty versus vertebroplasty for treatment of osteoporotic vertebral compression fracture: a prospective, comparative, and randomized clinical study. *Osteoporos Int* 2010;21(2):359-64. doi: 10.1007/s00198-009-0952-8.

198.Lippuner K, Johansson H, Kanis JA, et al. Remaining lifetime and absolute 10-year probabilities of osteoporotic fracture in Swiss men and women. *Osteoporos Int* 2009;20(7):1131-40. doi: 10.1007/s00198-008-0779-8.

199.Schwenkglenks M, Lippuner K, Hauselmann HJ, et al. A model of osteoporosis impact in Switzerland 2000-2020. *Osteoporos Int* 2005;16(6):659-71. doi: 10.1007/s00198-004-1743-x.

200.Bonjour JP, Burckhardt P, Dambacher M, et al. [Epidemiology of osteoporosis]. *Schweiz Med Wochenschr* 1997;127(16):659-67.

201.Cooper C, Atkinson EJ, Kotowicz M, et al. Secular trends in the incidence of postmenopausal vertebral fractures. *Calcif Tissue Int* 1992;51(2):100-4. doi: 10.1007/bf00298496.

202.Melton LJ, 3rd, Lane AW, Cooper C, et al. Prevalence and incidence of vertebral deformities. *Osteoporos Int* 1993;3(3):113-9. doi: 10.1007/bf01623271.

203.Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 2002;359(9319):1761-7. doi: 10.1016/s0140-6736(02)08657-9.

204.Lippuner K, Popp AW, Schwab P, et al. Fracture hospitalizations between years 2000 and 2007 in Switzerland: a trend analysis. *Osteoporos Int* 2011;22(9):2487-97. doi: 10.1007/s00198-010-1487-8.

205.Baerlocher MO, Asch MR, Puri G, et al. Awareness of interventional radiology among patients referred to the interventional radiology department: a survey of patients in a large Canadian community hospital. *J Vasc Interv Radiol* 2007;18(5):633-7. doi: 10.1016/j.jvir.2007.02.034.

206.Lindsay BR, Olufade T, Bauer J, et al. Patient-reported barriers to osteoporosis therapy. *Archives of osteoporosis* 2016;11:19.(doi):10.1007/s11657-016-0272-5.

207.Merle B, Dupraz C, Haesebaert J, et al. Osteoporosis prevention: where are the barriers to improvement in a French general population? A qualitative study. *Osteoporos Int* 2019;30(1):177-85. doi: 10.1007/s00198-018-4720-5.

208.Naik-Panvelkar P, Norman S, Elgebaly Z, et al. Osteoporosis management in Australian general practice: an analysis of current osteoporosis treatment patterns and gaps in practice. *BMC Fam Pract* 2020;21(1):32. doi: 10.1186/s12875-020-01103-2.

209.Otmar R, Reventlow SD, Nicholson GC, et al. General medical practitioners' knowledge and beliefs about osteoporosis and its investigation and management. *Archives of osteoporosis* 2012;7:107-14.(doi):10.1007/s11657-012-0088-x.

210.Ribeiro V, Blakeley J, Laryea M. Women's knowledge and practices regarding the prevention and treatment of osteoporosis. *Health Care Women Int* 2000;21(4):347-53. doi: 10.1080/073993300245195.

211.Sale JE, Beaton D, Bogoch E. Secondary prevention after an osteoporosis-related fracture: an overview. *Clin Geriatr Med* 2014;30(2):317-32. doi: 10.1016/j.cger.2014.01.009.

212.Werner P. Knowledge about osteoporosis: assessment, correlates and outcomes. *Osteoporos Int* 2005;16(2):115-27. doi: 10.1007/s00198-004-1750-y.

213.Williams B, Cullen L, Barlow JH. "I never realised how little I knew!": a pilot study of osteoporosis knowledge, beliefs, and behaviours. *Health Care Women Int* 2002;23(4):344-50. doi: 10.1080/0739933029008919.

214.Carr AJ, Thompson PW, Cooper C. Factors associated with adherence and persistence to bisphosphonate therapy in osteoporosis: a cross-sectional survey. *Osteoporos Int* 2006;17(11):1638-44. doi: 10.007/s00198-006-0166-2.

215.Rapan S, Batrnek J, Rapan V, et al. Quality of Life in Patients Following Vertebroplasty. *Open Access Maced J Med Sci* 2017;5(1):42-47. doi: 10.3889/oamjms.2017.018.

216.Sambrook PN, Chen JS, Simpson JM, et al. Impact of adverse news media on prescriptions for osteoporosis: effect on fractures and mortality. *Med J Aust* 2010;193(3):154-6.

217.Sullivan TB, Anderson JT, Ahn UM, et al. Can Internet information on vertebroplasty be a reliable means of patient self-education? *Clin Orthop Relat Res* 2014;472(5):1597-604. doi: 10.007/s11999-013-3425-5.

218.Kaffashian S, Raina P, Oremus M, et al. The burden of osteoporotic fractures beyond acute care: the Canadian Multicentre Osteoporosis Study (CaMos). *Age Ageing* 2011;40(5):602-7. doi: 10.1093/ageing/afr085.

219.Xie Z, Burge R, Yang Y, et al. Posthospital Discharge Medical Care Costs and Family Burden Associated with Osteoporotic Fracture Patients in China from 2011 to 2013. *J Osteoporos* 2015;2015:258089.(doi):10.1155/2015/258089.

220.Adler RA. Osteoporosis in men: a review. *Bone Res* 2014;2:14001.(doi):10.1038/boneres.2014.1.

221.Nielsen DS, Brixen K, Huniche L. Men's experiences of living with osteoporosis: focus group interviews. *Am J Mens Health* 2011;5(2):166-76. doi: 10.1177/1557988310372800.

222.Klezl Z, Bhangoo N, Phillips J, et al. Social implications of balloon kyphoplasty: prospective study from a single UK centre. *Eur Spine J* 2012;21(9):1880-6. doi: 10.007/s00586-012-2262-7.

223.Siddiqui MQ, Sim L, Koh J, et al. Stress levels amongst caregivers of patients with osteoporotic hip fractures - a prospective cohort study. *Ann Acad Med Singapore* 2010;39(1):38-42.

224.Melton LJ, 3rd. Adverse outcomes of osteoporotic fractures in the general population. *J Bone Miner Res* 2003;18(6):1139-41. doi: 10.359/jbmr.2003.18.6.1139.

225.Kanis JA, Oden A, Johnell O, et al. Excess mortality after hospitalisation for vertebral fracture. *Osteoporos Int* 2004;15(2):108-12. doi: 10.1007/s00198-003-1516-y.

226.Maravic M, Taupin P, Roux C. Hospital burden of vertebral fractures in France: influence of vertebroplasty. *Osteoporos Int* 2013;24(7):2001-6. doi: 10.1007/s00198-012-2264-7.

227.McDonald RJ, Achenbach SJ, Atkinson EJ, et al. Mortality in the vertebroplasty population. *AJNR Am J Neuroradiol* 2011;32(10):1818-23. doi: 10.3174/ajnr.A2616.

228.Crouser N, Malik AT, Jain N, et al. Discharge to Inpatient Care Facility After Vertebroplasty/Kyphoplasty: Incidence, Risk Factors, and Postdischarge Outcomes. *World Neurosurg* 2018;118:e483-e488.(doi):10.1016/j.wneu.2018.06.221.

229.Miller FG, Kallmes DF, Buchbinder R. Vertebroplasty and the placebo response. *Radiology* 2011;259(3):621-5. doi: 10.1148/radiol.11102412.

230.Delmas PD, van de Langerijt L, Watts NB, et al. Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study. *J Bone Miner Res* 2005;20(4):557-63. doi: 10.1359/JBMR.041214.

231.Jung HJ, Park YS, Seo HY, et al. Quality of Life in Patients with Osteoporotic Vertebral Compression Fractures. *J Bone Metab* 2017;24(3):187-96. doi: 10.11005/jbm.2017.24.3.187..

232.Fink HA, Ensrud KE, Nelson DB, et al. Disability after clinical fracture in postmenopausal women with low bone density: the fracture intervention trial (FIT). *Osteoporos Int* 2003;14(1):69-76. doi: 10.1007/s00198-002-1314-y.

233.Kado DM, Miller-Martinez D, Lui LY, et al. Hyperkyphosis, kyphosis progression, and risk of non-spine fractures in older community dwelling women: the study of osteoporotic fractures (SOF). *J Bone Miner Res* 2014;29(10):2210-6. doi: 10.1002/jbmr.2251.

234.van der Jagt-Willems HC, de Groot MH, van Campen JP, et al. Associations between vertebral fractures, increased thoracic kyphosis, a flexed posture and falls in older adults: a prospective cohort study. *BMC Geriatr* 2015;15:34. doi:10.1186/s12877-015-0018-z.

235.Melton LJ, 3rd, Atkinson EJ, Cooper C, et al. Vertebral fractures predict subsequent fractures. *Osteoporos Int* 1999;10(3):214-21. doi: 10.1007/s001980050218.

236.Kondo KL. Osteoporotic vertebral compression fractures and vertebral augmentation. *Semin Intervent Radiol* 2008;25(4):413-24. doi: 10.1055/s-0028-1103000.

237.Gold DT. The nonskeletal consequences of osteoporotic fractures. Psychologic and social outcomes. *Rheum Dis Clin North Am* 2001;27(1):255-62. doi: 10.1016/s0889-857x(05)70197-6.

238.Connelly P. The Management of Chronic Pain in Older Persons. *Journal of American Geriatrics Society* 1998;46(5):635-51. doi: 10.1111/j.1532-5415.1998.tb01084.x.

239.Gijsen R, Hoeymans N, Schellevis FG, et al. Causes and consequences of comorbidity: a review. *J Clin Epidemiol* 2001;54(7):661-74. doi: 10.1016/s0895-4356(00)00363-2.

240.Morin L, Johnell K, Laroche ML, et al. The epidemiology of polypharmacy in older adults: register-based prospective cohort study. *Clin Epidemiol* 2018;10:289-298.(doi):10.2147/CLEP.S153458.

241.Maher RL, Hanlon J, Hajjar ER. Clinical consequences of polypharmacy in elderly. *Expert Opin Drug Saf* 2014;13(1):57-65. doi: 10.1517/14740338.2013.827660.

242.Hovstadius B, Petersson G. The impact of increasing polypharmacy on prescribed drug expenditure-a register-based study in Sweden 2005-2009. *Health Policy* 2013;109(2):166-74. doi: 10.1016/j.healthpol.2012.09.005.

243.Soumerai SB, Pierre-Jacques M, Zhang F, et al. Cost-related medication nonadherence among elderly and disabled medicare beneficiaries: a national survey 1 year before the medicare drug benefit. *Arch Intern Med* 2006;166(17):1829-35. doi: 10.001/archinte.166.17.29.

244.Cutilli CC. Health literacy in geriatric patients: An integrative review of the literature. *Orthop Nurs* 2007;26(1):43-8. doi: 10.1097/00006416-200701000-00014.

245.Baker DW, Wolf MS, Feinglass J, et al. Health literacy and mortality among elderly persons. *Arch Intern Med* 2007;167(14):1503-9. doi: 10.001/archinte.167.14.03.

246.Weiss BD, Palmer R. Relationship between health care costs and very low literacy skills in a medically needy and indigent Medicaid population. *J Am Board Fam Pract* 2004;17(1):44-7. doi: 10.3122/jabfm.17.1.44.

247.Saarni SI, Hofmann B, Lampe K, et al. Ethical analysis to improve decision-making on health technologies. *Bull World Health Organ* 2008;86(8):617-23. doi: 10.2471/blt.08.051078.

248.Tait RC. Vulnerability in clinical research with patients in pain: a risk analysis. *J Law Med Ethics* 2009;37(1):59-72. doi: 10.1111/j.748-720X.2009.00351.x.

249.Dick B, Eccleston C, Crombez G. Attentional functioning in fibromyalgia, rheumatoid arthritis, and musculoskeletal pain patients. *Arthritis Rheum* 2002;47(6):639-44. doi: 10.1002/art.10800.

250.Carvalho AS, Martins Pereira S, Jacomo A, et al. Ethical decision making in pain management: a conceptual framework. *J Pain Res* 2018;11:967-976.(doi):10.2147/JPR.S162926.

251.Cheatle MD, Savage SR. Informed consent in opioid therapy: a potential obligation and opportunity. *J Pain Symptom Manage* 2012;44(1):105-16. doi: 10.1016/j.jpainsymman.2011.06.015.

252.Savage S. The patient-centered opioid treatment agreement. *Am J Bioeth* 2010;10(11):18-9. doi: 10.1080/15265161.2010.519426.

253.Buchbinder R, Busija L. Why we should stop performing vertebroplasties for osteoporotic spinal fractures. *Intern Med J* 2019;49(11):1367-71. doi: 10.111/imj.14628.

254.Clark WA. Vertebroplasty is not a do-not-do treatment. *Med J Aust* 2016;204(2):71-71.e1. doi: 10.5694/mja15.01201.

255.Schweizerische Vereinigung gegen die Osteoporose. Osteoporose Diagnostik - Prävention - Behandlun, 2015.

256.Wongrakpanich S, Wongrakpanich A, Melhado K, et al. A Comprehensive Review of Non-Steroidal Anti-Inflammatory Drug Use in The Elderly. *Aging Dis* 2018;9(1):143-50. doi: 10.14336/AD.2017.0306.

257.Sinha N, Padegal V, Satyanarayana S, et al. Pulmonary cement embolization after vertebroplasty, an uncommon presentation of pulmonary embolism: A case report and literature review. *Lung India* 2015;32(6):602-5. doi: 10.4103/0970-2113.168119.

258.Kim YJ, Lee JW, Park KW, et al. Pulmonary cement embolism after percutaneous vertebroplasty in osteoporotic vertebral compression fractures: incidence, characteristics, and risk factors. *Radiology* 2009;251(1):250-9. doi: 10.1148/radiol.2511080854.

259.Zhan Y, Jiang J, Liao H, et al. Risk Factors for Cement Leakage After Vertebroplasty or Kyphoplasty: A Meta-Analysis of Published Evidence. *World Neurosurg* 2017;101:633-642.(doi):10.1016/j.wneu.2017.01.124.

260.Hrobjartsson A, Thomsen AS, Emanuelsson F, et al. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *CMAJ* 2013;185(4):E201-11. doi: 10.1503/cmaj.120744.

261.Mattie R, Laimi K, Yu S, et al. Comparing Percutaneous Vertebroplasty and Conservative Therapy for Treating Osteoporotic Compression Fractures in the Thoracic and Lumbar Spine: A Systematic Review and Meta-Analysis. *J Bone Joint Surg Am* 2016;98(12):1041-51. doi: 10.2106/JBJS.15.00425.

262.Rodriguez AJ, Fink HA, Mirigian L, et al. Pain, Quality of Life, and Safety Outcomes of Kyphoplasty for Vertebral Compression Fractures: Report of a Task Force of the American Society for Bone and Mineral Research. *J Bone Miner Res* 2017;32(9):1935-44. doi: 10.002/jbmr.3170.

263.Zhu RS, Kan SL, Ning GZ, et al. Which is the best treatment of osteoporotic vertebral compression fractures: balloon kyphoplasty, percutaneous vertebroplasty, or non-surgical treatment? A Bayesian network meta-analysis. *Osteoporos Int* 2019;30(2):287-98. doi: 10.1007/s00198-018-4804-2.

264.Gautschi OP, Stienen MN, Corniola MV, et al. Assessment of the Minimum Clinically Important Difference in the Timed Up and Go Test After Surgery for Lumbar Degenerative Disc Disease. *Neurosurgery* 2017;80(3):380-85. doi: 10.1227/NEU.0000000000001320.

265.Masala S, Ciarrapico AM, Konda D, et al. Cost-effectiveness of percutaneous vertebroplasty in osteoporotic vertebral fractures. *Eur Spine J* 2008;17(9):1242-50. doi: 10.1007/s00586-008-0708-8.

266.Takura T, Yoshimatsu M, Sugimori H, et al. Cost-Effectiveness Analysis of Percutaneous Vertebroplasty for Osteoporotic Compression Fractures. *Clin Spine Surg* 2017;30(3):E205-e10. doi: 10.1097/BSD.0b013e3182aa4c29.

267.Mehio AK, Lerner JH, Engelhart LM, et al. Comparative hospital economics and patient presentation: vertebroplasty and kyphoplasty for the treatment of vertebral compression fracture. *AJNR American journal of neuroradiology* 2011;32(7):1290-4. doi: 10.3174/ajnr.A2502.

268.Medical Advisory Secretariat. Percutaneous vertebroplasty for treatment of painful osteoporotic vertebral compression fractures: an evidence-based analysis. *Ont Health Technol Assess Ser* 2010;10(19):1-45.

269.Fritzell P, Ohlin A, Borgstrom F. Cost-effectiveness of balloon kyphoplasty versus standard medical treatment in patients with osteoporotic vertebral compression fracture: a Swedish multicenter randomized controlled trial with 2-year follow-up. *Spine (Phila Pa 1976)* 2011;36(26):2243-51. doi: 10.1097/BRS.0b013e3182322d0f.

270.Chen C, Li DW, Wang Q, et al. [The cost effectiveness analysis of minimally invasive surgery and conservative treatment in elderly osteoporotic spinal fracture]. *Zhongguo Gu Shang* 2016;29(7):614-18. doi: 10.3969/j.issn.1003-0034.2016.07.006.

271.Edidin AA, Ong KL, Lau E, et al. Cost-effectiveness analysis of treatments for vertebral compression fractures. *Appl Health Econ Health Policy* 2012;10(4):273-84. doi: 10.2165/11633220-000000000-00000.

272.Eidt D, Greiner W. PMS30 Cost Analysis Of Balloon Kyphoplasty Versus Non Surgical Management For Osteoporotic Vertebral Fractures In Germany. *Value in Health* 2009;12(7):A438-A39. doi: 10.1016/S1098-3015(10)75170-6.

273.Joestl J, Lang N, Bukaty A, et al. Osteoporosis associated vertebral fractures-Health economic implications. *PloS one* 2017;12(5):e0178209. doi: 10.1371/journal.pone.0178209.

274.Takahashi S, Hoshino M, Yasuda H, et al. Cost-effectiveness of Balloon Kyphoplasty for Patients With Acute/Subacute Osteoporotic Vertebral Fractures in the Super-Aging Japanese Society. *Spine (Phila Pa 1976)* 2019;44(5):E298-e305. doi: 10.1097/brs.0000000000002829.

275.Becker S, Pfeiffer KP, Ogon M. Comparison of inpatient treatment costs after balloon kyphoplasty and non-surgical treatment of vertebral body compression fractures. *Eur Spine J* 2011;20(8):1259-64. doi: 10.1007/s00586-011-1692-y.

276.Goz V, Errico TJ, Weinreb JH, et al. Vertebroplasty and kyphoplasty: national outcomes and trends in utilization from 2005 through 2010. *Spine J* 2015;15(5):959-65. doi: 10.1016/j.spinee.2013.06.032.

277.Lange A, Kasperk C, Alvares L, et al. Survival and cost comparison of kyphoplasty and percutaneous vertebroplasty using German claims data. *Spine (Phila Pa 1976)* 2014;39(4):318-26. doi: 10.1097/brs.0000000000000135.

17 Appendices

17.1 Appendix A: Source of Literature (databases and websites)

Table 88 Databases searched and number of search results

Source	Location	Initial search (up to 4 April 2019)	Updated search (up to 13 December 2019)*
PubMed	https://www.ncbi.nlm.nih.gov	2,773	2,880
Embase	https://www.embase.com/	4,696	2,762
The Cochrane Library	https://www.cochranelibrary.com/	453	10**
Cinahl	https://www.ebscohost.com/nursing/products/cinahl-databases/cinahl-complete	472	793
York CRD (inc. HTA, NHS, EED, DARE)	https://www.crd.york.ac.uk/CRDWeb/	106	0**
CEA Registry	http://healtheconomics.tuftsmedicalcenter.org/cear4/home.aspx	5	0**
Econlit	https://www.aeaweb.org/econlit/	8	0**
ETHMED	http://www.ethicsweb.eu/search_ets	10	12
Total		8,523	886 after de-duplication with initial search

Notes
* = Additional key words were added to the updated search strategy. Results from both searches were combined and de-duplicated, resulting in 886 new citations.
** = New citations published after 4 April 2019 using the original search terms, plus citations identified using the additional keywords (no search limits).

Table 89 Grey literature sources

HTA websites	
International	
National Information Centre of Health Services Research and Health Care Technology (NICHSR)	https://www.nlm.nih.gov/nichsr/db.html
National Library of Medicine Health Services/Technology Assessment Texts (HSTAT)	https://www.ncbi.nlm.nih.gov/books/NPBK16710/
International Information Network on New and Emerging Health Technologies (EuroScan International Network)	https://www.euroscan-network.global/index.php/en/47-public-features/761-database-home
Australia	
Adelaide Health Technology Assessment (AHTA)	https://www.adelaide.edu.au/ahta/pubs/
Centre for Clinical Effectiveness, Monash University	http://monashhealth.org/health-professionals/cce/
Centre for Health Economics, Monash University	https://www.monash.edu/business/che
National Health and Medical Research Council	https://www.nhmrc.gov.au/

HTA websites	
Austria	
Institute of Technology Assessment / HTA unit	https://www.oeaw.ac.at/ita/publikationen/
Ludwig Boltzmann Institute for Health Technology Assessment (LBI-HTA)	https://hta.lbg.ac.at/page/publikationen/en
Canada	
Institute of Health Economics (IHE)	http://www.ihe.ca
Institut National d'Excellence en Santé et en Services (INESSS)	https://www.inesss.qc.ca/en/home.html
Alberta Heritage Foundation for Medical Research (AHFMR)	http://www.ahfmr.ab.ca/
Alberta Institute of Health Economics	http://www.ihe.ca/
The Canadian Agency for Drugs and Technologies in Health (CADTH)	http://www.cadth.ca/
The Canadian Association for Health Services and Policy Research (CAHSPR)	https://www.cahspr.ca/
Centre for Health Economics and Policy Analysis (CHEPA), McMaster University	http://www.chepa.org/
Centre for Health Services and Policy Research (CHSPR), University of British Columbia	http://www.chspr.ubc.ca/
Institute for Clinical and Evaluative Studies (ICES)	http://www.ices.on.ca/
Saskatchewan Health Quality Council (Canada)	http://www.hqc.sk.ca/
Denmark	
Danish National Institute of Public Health	https://www.sdu.dk/en/sif/forskning
Finland	
Finnish National Institute for Health and Welfare	https://thl.fi/en/web/thlfi-en/publications
France	
French National Authority for Health (Haute Autorité de Santé; HAS)	http://www.has-sante.fr/
Germany	
German Institute for Medical Documentation and Information (DIMDI)	https://www.dimdi.de/
Institute for Quality and Efficiency in Health Care (IQWiG)	http://www.iqwig.de
The Netherlands	
Health Council of the Netherlands (Gezondheidsraad)	https://www.gezondheidsraad.nl/
Zorginstituut Nederland (ZIN)	https://www.zorginstituutnederland.nl/
Norway	
The Norwegian Institute of Public Health (NIPHNO)	http://www.fhi.no/
Singapore	
Agency for Care Effectiveness (ACE)	http://www.ace-hta.gov.sg/
Spain	
Agencia de Evaluación de Tecnologías Sanitarias, Instituto de Salud "Carlos III" / Health Technology Assessment Agency (AETS)	http://publicaciones.isciii.es/
Andalusian Agency for Health Technology Assessment	http://aetsa.org/produccion-cientifica/
Catalan Agency for Health Technology Assessment (CAHTA)	http://www.gencat.cat

HTA websites	
Sweden	
Center for Medical Health Technology Assessment	http://www.cmt.liu.se/?l=en&sc=true
Swedish Council on Technology Assessment in Health Care (SBU)	http://www.sbu.se/en/
Switzerland	
Swiss Network on Health Technology Assessment (SNHTA)	http://www.snhta.ch/
United Kingdom (UK)	
National Health Service Health Technology Assessment (UK) / National Coordinating Centre for Health Technology Assessment (NCCHTA)	https://www.nihr.ac.uk/funding-and-support/funding-for-research-studies/funding-programmes/health-technology-assessment/
NHS Quality Improvement Scotland	http://www.nhshealthquality.org/
National Institute for Clinical Excellence (NICE)	http://www.nice.org.uk/
United States of America (USA)	
Agency for Healthcare Research and Quality (AHRQ)	https://www.ahrq.gov/research/findings/index.html
Harvard School of Public Health	http://www.hsph.harvard.edu/
Institute for Clinical and Economic Review (ICER)	http://www.icer-review.org/
Institute for Clinical Systems Improvement (ICSI)	http://www.icsi.org
Minnesota Department of Health (US)	http://www.health.state.mn.us/
Office of Health Technology Assessment Archive (US)	http://ota.fas.org/
U.S. Blue Cross / Blue Shield Association Technology Evaluation Center (Tec)	https://www.bcbs.com/news/press-releases/blue-cross-blue-shield-association-launches-evidence-street-website-streamline
Veteran's Affairs Research and Development Technology Assessment Program (US)	http://www.research.va.gov/default.cfm
Clinical trial registries	
ClinicalTrials.gov	https://clinicaltrials.gov/
Cochrane Central Register of Controlled Trials	https://www.cochranelibrary.com/central
EU Clinical Trials Registry	https://www.clinicaltrialsregister.eu/ctr-search/search
WHO International Clinical Trials Registry Platform (ICTRP)	http://www.who.int/ictrp/en/
Current Controlled Trials MetaRegister	http://www.isrctn.com
Australian New Zealand Clinical Trials Registry	http://www.anzctr.org.au/
Specialty websites	
Geneva Medical Association	https://www.amge.ch/
Eular	https://www.eular.org/index.cfm
European Geriatric Medicine Society	https://www.eugms.org/home.html
Australia and New Zealand Society for Geriatric Medicine	http://www.anzsgm.org/
Swiss Orthopaedic Association	http://www.swissorthopaedics.ch/de/
American Orthopaedic Association	http://www.aoassn.org/aoaimis/aoanew
Australian Orthopaedic Association	https://www.aoa.org.au/
Australian Society of Orthopaedic Surgeons	http://www.asos.org.au/
British Orthopaedic Association	https://www.boa.ac.uk/
Canadian Orthopaedic Association	http://coa-aco.org/
Swiss Society for Neuroscience	https://www.swissneuroscience.ch/

HTA websites	
Neurosurgical Society of Australasia	http://www.nsa.org.au/
Swiss Society of Spinal Surgery	https://www.spinesociety.ch/
North American Spine Society	https://www.spine.org/
International Osteoporosis Foundation	https://www.iofbonehealth.org/
Osteoporosis Australia	https://www.osteoporosis.org.au/
Society of Interventional Radiology	https://www.sirweb.org/
Clinical practice guidelines	
Guidelines International Network (GIN)	https://www.g-i-n.net/library/international-guidelines-library
Association of Scientific Medical Societies (AWMF)	https://www.awmf.org/awmf-online-das-portal-der-wissenschaftlichen-medizin/awmf-aktuell.html
National Guideline Clearinghouse	https://www.ahrq.gov/gam/index.html
Scottish Intercollegiate Guidelines Network	http://www.sign.ac.uk/guidelines/published/
Other sources	
New York Academy of Medicine Grey Literature Report	http://www.greylit.org
EMA	http:///www.ema.europa.eu/
NHS National Institute for Health Research (NIHR)	http://www.nets.nihr.ac.uk

17.1.1 Scoping Report Searches

Table 90 Search strategy – Medline (Inception to 4 April 2019)

Number	Query	Results
1.	Spinal fractures[Text Word]	NR
2.	Spinal fractures[MeSH Terms]	NR
3.	Osteoporotic fractures[Text Word]	NR
4.	Osteoporotic fractures[MeSH Terms]	NR
5.	Compression fracture[Text Word]	NR
6.	Compression fracture[MeSH Terms]	NR
7.	Spinal fracture[Text Word]	NR
8.	Spinal fracture[MeSH Terms]	NR
9.	Spinal tumor[Text Word]	NR
10.	Spinal tumor[MeSH Terms]	NR
11.	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10	NR
12.	Vertebroplasty[Text Word]	NR
13.	Vertebroplasty[MeSH Terms]	NR
14.	Kyphoplasty[Text Word]	NR
15.	Kyphoplasty[MeSH Terms]	NR
16.	Sarcoplasty[Text Word]	NR
17.	Cementoplasty[Text Word]	NR
18.	Cementoplasty[MeSH Terms]	NR
19.	#12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18	2,773

Abbreviations
NR = not reported

Table 91 Search Strategy – Embase (Inception to 4 April 2019)

Number	Query	Results
1.	Kyphoplasty/exp or Kyphoplasty.mp.	3,370
2.	Sarcoplasty.mp	3
3.	Vertebroplasty.mp.	5,760
4.	Pediculoplasty.mp.	11
5.	Cementoplasty.mp. or Cementoplasty/exp	6,832
6.	Percutaneous vertebroplasty.mp. or Percutaneous vertebroplasty/exp	6,678
7.	#1 OR #2 OR #3 OR #4 OR #5 OR #6	7,529
8.	Spinal fractures.mp. or Spine fracture/exp	23,439
9.	Osteoporotic fractures.mp. or Fragility fracture/exp	18,967
10.	Fractures, compression.mp. or Compression fracture/exp	5,366
11.	Compression fracture.mp.	5,991
12.	Spinal fracture.mp. or Spine fracture/exp	22,970
13.	Spinal tumor/exp	7,958
14.	#8 OR #9 OR #10 OR #11 OR #12 OR #13	46,531
15.	#7 AND #14	4,696

Table 92 Search Strategy – Cochrane (Inception to 4 April 2019)

Number	Query	Results
1.	MeSH descriptor: [Vertebroplasty] explode all terms	121
2.	(vertebroplasty);ti,ab,kw	334
3.	#1 OR #2	363
4.	MeSH descriptor: [Kyphoplasty] explode all trees	49
5.	(kyphoplasty):ti,ab,kw	218
6.	#4 OR #5	218
7.	#3 AND #6	453

Table 93 Search Strategy – CINAHL (Inception to 5 April 2019)

Number	Query	Results
1.	TX Vertebroplasty	1,441
2.	TX Kyphoplasty	1,386
3.	TX Cementoplasty	0
4.	TX Sarcoplasty	0
5.	TX Percutaneous vertebroplasty	687
6.	#1 OR #2 OR #3 OR #4 OR #5	2,073
7.	TX Spinal fracture	12,057
8.	TX Osteoporotic fractures	5,877
9.	TX Compression fractures and osteoporosis	1,786
10.	TX Compression fracture of the spine	2,834
11.	TX Compression fracture pain	4,183
12.	#7 OR #8 OR #9 OR #10 OR #11	16,240
13.	#6 and #12	472

Table 94 Search Strategy – YORK CRD (including DARE, NHS EED, HTA) (Inception to 8 April 2019)

Number	Query	Results
1.	Vertebroplasty[Any field]	91
2.	Kyphoplasty[Any field]	73
3.	#1 OR #2	106

Table 95 Search Strategy – CEA Registry (Inception to 8 April 2019)

Number	Query	Results
1.	TX Vertebroplasty	4
2.	X Kyphoplasty	4
3.	#1 OR #2	5

Notes
(All but one was also captured in PubMed search)

Table 96 Search Strategy – Econlit (Inception to 8 April 2019)

Number	Query	Results
1.	TX Vertebroplasty	8
2.	X Kyphoplasty	2
3.	#1 OR #2	8

Table 97 Search Strategy – Ethicsweb (Inception to 8 April 2019)

Number	Query	Results
1.	TX Vertebroplasty	10
2.	X Kyphoplasty	2
3.	#1 OR #2	10

17.1.2 Updated Searches

Table 98 Search strategy – Medline (April 2019 – 13 December 2019)

Number	Query	Results
1.	Spinal [Text word]	384,651
2.	Spine [Text word]	133,576
3.	Vertebra* [Text word]	250,958
4.	((Spinal [Text word] OR Spine [Text word]) OR Vertebra* [Text word])	600,690
5.	Osteoporosis [Text word]	84,265
6.	Osteoporotic [Text word]	20,481
7.	Compression [Text word]	119,480
8.	((Osteoporosis [Text word] OR Osteoporotic [Text word]) OR Compression [Text word])	205,801
9.	Fractur*	300,785
10.	((Spinal [Text word] OR Spine [Text word]) OR Vertebra* [Text word]) AND ((Osteoporosis [Text word] OR Osteoporotic [Text word]) OR Compression [Text word]) AND Fractur*	18,921
11.	Vertebroplasty [Text word]	3,674
12.	Kyphoplasty [Text word]	1,937
13.	Sarcoplasty [Text word]	140
14.	Cementoplasty [Text word]	285
15.	(((Vertebroplasty [Text word] OR Kyphoplasty [Text word]) OR Sarcoplasty [Text word]) OR Cementoplasty [Text word])	4,778
16.	(((Vertebroplasty [Text word] OR Kyphoplasty [Text word]) OR Sarcoplasty [Text word]) OR Cementoplasty [Text word]) AND ((Spinal [Text word] OR Spine [Text word]) OR Vertebra* [Text word]) AND ((Osteoporosis [Text word] OR Osteoporotic [Text word]) OR Compression [Text word]) AND Fractur*	2,880

Abbreviations
NR = not reported

Table 99 Search Strategy – Embase (April 2019 – 13 December 2019)

Number	Query	Results
1.	Spinal	449,915
2.	Spine	314,287
3.	Vertebr*	264,299
4.	#1 OR #2 OR #3	809,175
5.	Osteoporosis'	158,127
6.	'Osteoporosis'/exp	126,995
7.	Osteoporotic	27,725
8.	Compression	175,331
9.	#5 OR #6 OR #7 OR #8	333,326
10.	Fractur*	400,163
11.	#4 AND #9 AND #10	35,615
12.	Kyphoplasty	3,489
13.	Sarcoplasty	203
14.	Vertebroplasty	5,921
15.	'Vertebroplasty'/exp	6,724
16.	Cementoplasty	480
17.	'Cementoplasty'/exp	7,002
18.	#12 OR #13 OR #14 OR #15 OR #16 OR # OR #17	7,898
19.	#11 AND #18	4,652
20.	#19 AND ('article'/it OR 'article in press'/it)	2,762

Table 100 Search Strategy – Cochrane (April 2019 – 13 December 2019)

Number	Query	Results
1.	MeSH descriptor: [Vertebroplasty] explode all terms	130
2.	(vertebroplasty);ti,ab,kw	348
3.	#1 OR #2	382
4.	MeSH descriptor: [Kyphoplasty] explode all trees	55
5.	(kyphoplasty):ti,ab,kw	239
6.	#4 OR #5	239
7.	#3 AND #6 (Publication date April 2019 – 13 December 2019)	10

Table 101 Search Strategy – CINAHL (April 2019 – 13 December 2019)

Number	Query	Results
1.	TX Spinal	74,957
2.	TX Spine	41,623
3.	S1 OR S2	103,471
4.	TX Osteoporosis	37,368
5.	TX Osteoporosis	6,927
6.	TX Compression	37,782
7.	S4 OR S5 OR S6	74,690
8.	TX Fractur*	96,927
9.	S3 AND S7 AND S8	6,666
10.	TX Vertebroplasty	1,465
11.	TX Kyphoplasty	1,420
12.	TX Cementoplasty	146
13.	TX Sarcoplasty	74
14.	S10 OR S11 OR S12 OR S13	2,274
15.	S9 AND S14	793

Table 102 Search Strategy – Ethicsweb (April 2019 – 13 December 2019)

Number	Query	Results
1.	(spin* fracture) AND (Spine OR Spinal) AND (fracture OR fractures OR compression) AND (osteoporosis	12

17.2 Appendix B: HTA Key Questions

Sub-Questions: efficacy, effectiveness and safety

Relevant sub-questions on safety and effectiveness from the EUnetHTA core model (Version 3.0) are outlined in ***Table 103*** and ***Table 104***.

Table 103 Sub-Questions: safety

Topic	Research Question	Element ID
Patient safety	How safe is the technology in comparison to the comparator(s)?	C0008
Patient safety	Are there susceptible patient groups that are more likely to be harmed through the use of the technology?	C0005
Patient safety	Are the technology and comparator(s) associated with user-dependent harms?	C0007
Occupational safety	What kind of occupational harms can occur when using the technology?	C0020

Table 104 Sub-Questions: effectiveness

Topic	Research Question	Element ID
Mortality	Is there an expected beneficial effect of the technology on mortality?	D0001
Morbidity	How does the technology affect symptoms and findings (severity, frequency) of the disease or health condition?	D0005
Morbidity	How does the technology affect progression (or recurrence) of the disease or health condition?	D0006
Function	What is the effect of the technology on body functions of patients?	D0011
Function	What is the effect of the technology on work ability?	D0014
Function	What is the effect of the technology on return to previous living conditions?	D0015
Function	How does the use of technology affect activities of daily living?	D0016
Health-related quality of life	What is the effect of the technology on generic health-related quality of life?	D0012
Health-related quality of life	What is the effect of the technology on disease-specific quality of life?	D0013
Change in management	How does the technology modify the need for hospitalisation?	D0010

Sub-Questions: costs, cost-effectiveness and budget impact

Key questions related to costs, budget impact and cost-effectiveness relevant to PVP and PBK are outlined in **Table 105**.

Table 105 Sub-Questions: costs, budget impact and cost-effectiveness

Topic	Research Question	Element ID
Resource utilisation	What types of resources are used when delivering the assessed technology and its comparators (resource-use identification)?	E0001
Resource utilisation	What amounts of resources are used when delivering the assessed technology and its comparators (resource-use measurement)?	E0002
Resource utilisation	What were the measured and/or estimated costs of the assessed technology and its comparator(s) (resource-use valuation)?	E0009
Resource utilisation	How does the technology modify the need for other technologies and use of resources?	D0023
Resource utilisation	What are the likely budget impacts of implementing the technologies being compared?	G0007
Measurement and estimation of outcomes	What is (are) the measured and/or estimated health-related outcome(s) of the assessed technology and its comparator(s)?	E0005
Examination of costs and outcomes	What are the estimated differences in costs and outcomes between the technology and its comparator(s)?	E0006
Characterising uncertainty	What are the uncertainties surrounding the costs and economic evaluation(s) of the technology and its comparator(s)?	E0010
Characterising heterogeneity	To what extent can differences in costs, outcomes, or cost-effectiveness be explained by variations between any sub-groups using the technology and its comparator(s)?	E0011
Validity of the model(s)	What methodological assumptions were made in relation to the technology and its comparator(s)?	E0013
Validity of the model(s)	To what extent can the estimates of costs, outcomes or economic evaluation(s) be considered as providing valid descriptions of the technology and its comparator(s)?	E0012

Sub-Questions: legal, social and ethical issues

Sub-questions related to patient, social and ethical aspects relevant to PVP and PBK are outlined in ***Table 106*** and ***Table 107***.

Table 106 Sub-Questions: patient and social aspects

Topic	Research Question	Element ID
Patient perspective	What expectations and wishes do patients have with regard to the technology and what do they expect to gain from the technology?	H0100
Patient perspective	What is the burden on caregivers?	H0002
Social group aspects	Are there groups of patients who currently don't have good access to available therapies?	H0201

Table 107 Sub-Questions: ethical aspects

Topic	Research Question	Element ID
Autonomy	Is the technology used for individuals that are especially vulnerable?	F0005

Sub-Questions: organisational issues

Key questions related to organisational aspects relevant to PVP and PBK are outlined in ***Table 108***.

Table 108 Sub-Questions: organisational aspects

Topic	Research Question	Element ID
Health delivery process	How does the technology affect the current work processes?	G0001
Health delivery process	What kind of patient/participant flow is associated with removing the technology from basic health insurance?	G0100
Process-related costs	How does the technology modify the need for other technologies and use of resources?	D0023
Management	What management problems and opportunities will removing the technology cause?	G0008

17.3 Appendix C: Risk of Bias and Study Information from Extended Assessment of Harms

Table 109 Extended assessment of harms: risk of bias summary for the meta-analyses

AMSTAR 2 Questions	Busse 2018[179]	CNT Collaboration 2013[83]	Els 2017[82]	Machado 2015[181]	Mallen 2011[178]	Roberts 2015[180]
Did the research questions and inclusion criteria for the review include the components of PICO?	Yes	Yes	Yes	Yes	Yes	Yes
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	Partial yes	Partial yes	Yes	No	No
Did the review authors explain their selection of the study designs for inclusion in the review?	Yes	Yes	Yes	Yes	Yes	Yes
Did the review authors use a comprehensive literature search strategy?	Partial yes	Partial yes	Yes	Partial yes	No	Partial yes
Did the review authors perform study selection in duplicate?	Yes	Yes	Yes	Yes	No	Yes
Did the review authors perform data extraction in duplicate?	Yes	Yes	Yes	Yes	No	No
Did the review authors provide a list of excluded studies and justify the exclusions?	No	No	Yes	No	No	Yes
Did the review authors describe the included studies in adequate detail?	Partial yes	Yes	Yes	Yes	Yes	Partial yes
Did the review authors use a satisfactory technique for assessing the risk of bias in individual studies that were included in the review?	Yes	Yes	Yes	Yes	No	Partial yes
Did the review authors report on the sources of funding for the studies included in the review?	Yes	Yes	Yes	Yes	Yes	Yes
If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results?	Yes	Yes	Yes	Yes	No	Yes
If meta-analysis was performed, did the review authors assess the potential impact of risk of bias in individual studies on the results of the meta-analysis or other evidence synthesis?	Yes	Yes	No	Yes	No	No
Did the review authors account for risk of bias in individual studies when interpreting/ discussing the results of the review?	Yes	No	No	Yes	No	Yes

AMSTAR 2 Questions	Busse 2018[179]	CNT Collaboration 2013[183]	Els 2017[182]	Machado 2015[181]	Mallen 2011[178]	Roberts 2015[180]
Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?	No	Yes	No	Yes	No	No
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	No	No	Yes	No	No
Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?	Yes	Yes	Yes	Yes	Yes	Yes

Table 110 PVP: risk of bias summary for safety outcomes in the single-arm trials

	Al-Ali 2009[151]	Bae 2012[152]	DePalam 2011[153]	Dohm 2014[164]	Fenoglio 2008[154]	Kotwica 2011[155]	Masala 2012[157]	Masala 2009[156]	Niuewenhuijse 2012[159]	Niuewenhuijse 2010[158]	Pitton 2008[160]	Saracen 2014[161]	Santiago 2010[165]	Voormolen 2006a[162]	Voormolen 2006b[163]
Study objective															
1. Objective clearly stated	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Study design															
2. Prospective	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NK	Y	Y	Y
3. Multicentre	N	Y	N	Y	N	NK	N	N	N	N	N	NK	NK	N	N
4. Consecutive recruitment	Y	NK	Y	Y	Y	Y	Y	Y	NK	Y	NK	Y	Y	Y	Y
Study population															
5. Were patient characteristics included?	P	Y	Y	P	P	Y	P	Y	P	Y	P	P	P	Y	Y
6. Eligibility criteria clearly stated	P	Y	Y	Y	Y	N	Y	Y	N	Y	P	N	Y	P	Y
7. Did patient enter the study at a similar point in the disease	NK	N	N	Y	NK	N	NK	Y	Y	Y	NK	NK	NK	NK	Y
Intervention and co-intervention															
8. Was the intervention of interest clearly described?	Y	Y	Y	P	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

	Al-Ali 2009[151]	Bae 2012[152]	DePalam 2011[153]	Dohm 2014[164]	Fenoglio 2008[154]	Kotwica 2011[155]	Masala 2012[157]	Masala 2009[156]	Niuewenhuijse 2012[159]	Niuewenhuijse 2010[158]	Pitton 2008[160]	Saracen 2014[161]	Santiago 2010[165]	Voormolen 2006a[162]	Voormolen 2006b[163]
9. Were additional interventions clearly described?	N	N	N	N	Y	N	Y	Y	N	N	P	N	N	Y	N
Outcome measure															
10. Were relevant outcome measures established a priori?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
11. Were outcome assessors blinded to the intervention?	NK	Y	NK	Y	NK	NK	NK	NK	NK	NK	NK	NK	Y	NK	NK
12. Were the outcomes measured using appropriate objective methods?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
13. Were the relevant outcome measures made before and after the intervention?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Statistical analysis															
14. Were the statistical tests used to assess the relevant outcomes appropriate?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Results and conclusions															
15. Was follow-up long enough for	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

	Al-Ali 2009[151]	Bae 2012[152]	DePalam 2011[153]	Dohm 2014[164]	Fenoglio 2008[154]	Kotwica 2011[155]	Masala 2012[157]	Masala 2009[156]	Niuewenhuijse 2012[159]	Niuewenhuijse 2010[158]	Pitton 2008[160]	Saracen 2014[161]	Santiago 2010[165]	Voormolen 2006a[162]	Voormolen 2006b[163]
important events and outcomes to occur?															
16. Were losses to follow-up reported?	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y
17. Did study provide estimates of random variability in the data analysis of relevant outcomes?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
18. Were the adverse events reported?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
19. Were the conclusions supported by results?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Competing interest and sources of support															
20. Were both competing interests and sources of support for the study reported?	P	Y	Y	Y	N	P	P	N	P	P	N	P	N	P	P

Abbreviations
N = no, **NA** =not applicable, **NK** = not known, **P** = partial, **Y** = yes.

Table 111 PBK: risk of bias summary for safety outcomes in the single-arm trials

Domain	Dohm 2014[164]	Hubschle 2014[63]	Prokop 2012[176]	Santiago 2010[165]	Robinson 2008[177]	Hillmeier 2004[175]
Study objective						
1. Objective clearly stated	Y	Y	N	Y	Y	Y
Study design						
2. Prospective	Y	N	N	Y	Y	Y
3. Multicentre	Y	Y	Y	NK	NK	Y
4. Consecutive recruitment	Y	NK	NK	Y	Y	NK
Study population						
5. Were patient characteristics included?	P	P	N	P	P	P
6. Eligibility criteria clearly stated?	Y	N	N	Y	N	Y
7. Did patient enter the study at a similar point in the disease?	NK	N	Y	NK	Y	N
Intervention and co-intervention						
8. Was the intervention of interest clearly described?	P	N	P	Y	N	Y
9. Were additional interventions clearly described?	N	N	Y	N	Y	N
Outcome measure						
10. Were relevant outcome measures established a priori?	Y	Y	N	Y	Y	Y
11. Were outcome assessors blinded to the intervention?	Y	NK	NK	Y	NK	NK
12. Were the outcomes measured using appropriate objective methods?	Y	Y	Y	Y	Y	Y
13. Were the relevant outcome measures made before and after the intervention?	Y	Y	Y	Y	Y	Y
Statistical analysis						
14. Were the statistical tests used to assess the relevant outcomes appropriate?	NA	NA	NA	NA	NA	NA
Results and conclusions						

Domain	Dohm 2014[164]	Hubschle 2014[63]	Prokop 2012[176]	Santiago 2010[165]	Robinson 2008[177]	Hillmeier 2004[175]
15. Was follow-up long enough for important events and outcomes to occur?	Y	Y	Y	Y	Y	Y
16. Were losses to follow-up reported?	Y	Y	Y	N	N	Y
17. Did study provide estimates of random variability in the data analysis of relevant outcomes?	NA	NA	NA	NA	NA	NA
18. Were the adverse events reported?	Y	Y	Y	Y	Y	Y
19. Were the conclusions supported by results?	Y	Y	P	Y	Y	Y
Competing interest and sources of support						
20. Were both competing interests and sources of support for the study reported?	Y	Y	P	N	Y	P

Abbreviations
N = no, **NA** =not applicable, **NK** = not known, **P** = partial, **Y** = yes.

Table 112 PVP compared to CT: mortality and adverse events (database analyses)

Follow-up	Mortality [a]	Readmission	Bedsores	Cardiac complications	DVT	Infection	Neurological compromise	Pneumonia	Pulmonary embolism	Pulmonary/respiratory complications	UTI	Additional vertebral augmentation	Readmission within 30 days
30 days	PVP, 0.53% CT, 1.72% **	PVP, 52.4% CT, 61.9% **	PVP, 0.71% CT, 1.25% **	NR	PVP, 3.29% CT, 2.49% **	PVP, 0.15% CT, 0.11%	PVP, 0.03% CT, 0.02%	PVP, 3.70% CT, 3.87% **	PVP, 0.42% CT, 0.30% *	NR	NR	NR	PVP 52.4%** CT 61.9%
6 months	NR	NR	PVP, 3.02% CT, 4.41% **	NR	PVP, 6.61% CT, 5.52% **	PVP, 0.73% CT, 0.79%	PVP, 0.12% CT, 0.07%	PVP, 11.1% CT, 12.95% **	PVP, 1.86% CT, 1.42% **	NR	NR	PVP, 7.89% CT, NA	NR
1 year	30% **	-1.0% **	NR	20% **	5% **	6% **	NR	10% **	-3% *	1% **	2% **	NR	NR
2 years	29% **	-1.0% *	NR	13% **	3% **	1%	NR	7% **	-7% **	1% *	3% **	NR	NR
5 years	12% **	NR	NR	9% **	0%	1%	NR	6% **	-6% **	0%	3% **	NR	NR
8 years	9% **	NR	NR	7% **	0%	1%	NR	5% **	-6% **	0%	2% **	NR	NR
10 years	8% **	NR	NR	7% **	0%	1%	NR	5% **	-6% **	0%	2% **	NR	NR

Abbreviations

CT = conservative treatment, **DVT** = deep vein thrombosis, **NA** = not applicable, **NR** = not reported, **PVP** = percutaneous vertebroplasty, **UTI** = urinary tract infection.

Notes

For long-term outcomes (1–10 years), a positive percentage implies the event occurred more frequently in nonsurgical management group compared to vertebroplasty or balloon kyphoplasty. A negative percentage implies the event occurred more frequently in the vertebroplasty or kyphoplasty groups compared to nonsurgical management.

a = in-hospital mortality reported, * = $p < 0.05$, ** = $p < 0.001$.

Source [134 147]

Table 113 PBK compared to CT: mortality and adverse events (database analyses)

Follow-up	Mortality [a]	Readmission	Bedsores	Cardiac complications	DVT	Infection	Neurological compromise	Pneumonia	Pulmonary embolism	Pulmonary/ respiratory complications	UTI	Additional vertebral augmentation	Readmission within 30 days
30 days	PBK, 0.35% CT, 1.72% **	PBK, 35.2% CT, 61.9% **	PBK, 0.52% CT, 1.25% **	NR	PBK, 2.74% CT, 2.49% **	PBK, 0.06% CT, 0.11%	PBK, 0.04% CT, 0.02%	PBK, 1.73% CT, 3.87% **	PBK, 0.23% CT, 0.30% *	NR	NR		PBK 35.2%** CT 61.9%
6 months	NR	NR	PBK, 2.39% CT, 4.41% **	NR	PBK, 5.32% CT, 5.52% **	PBK, 0.66% CT, 0.79%	PBK, 0.11% CT, 0.07%	PBK, 8.05% CT, 12.95% **	PBK, 1.27% CT, 1.42% **	NR	NR	PBK, 9.41% CT, NA	NR
1 year	55% **	9.0% *	NR	19% **	2% **	2% *	NR	23% **	7% **	2% *	1% *	NR	NR
2 years	37% **	8.0% *	NR	15% **	2% *	2% *	NR	19% **	5% **	3% **	3% **	NR	NR
5 years	26% **	NR	NR	11% **	3% **	1%	NR	15% **	3% **	3% **	3% **	NR	NR
8 years	24% **	NR	NR	11% **	3% **	0%	NR	14% **	3% **	4% **	4% **	NR	NR
10 years	24% **	NR	NR	11% **	3% **	0%	NR	14% **	3% **	4% **	4% **	NR	NR

Abbreviations
CT = conservative treatment, **DVT** = deep vein thrombosis, **NA** = not applicable, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **UTI** = urinary tract infection.
Notes
For long-term outcomes (1–10 years), a positive percentage implies the event occurred more frequently in nonsurgical management group compared to vertebroplasty or balloon kyphoplasty. A negative percentage implies the event occurred more frequently in the vertebroplasty or kyphoplasty groups compared to nonsurgical management.
a = in-hospital mortality reported, * = $p < 0.05$, ** = $p < 0.001$.
Source [134 147]

17.4 Appendix D: Sub-group Analysis Results

17.4.1 Clinical Effectiveness: PVP vs CT

Pain-related outcomes

Table 114 PVP vs CT: sub-group analyses of the visual analogue scale

Length of follow-up	Number of studies	Heterogeneity	PVP number of participants (n/N)	CT number of participants (n/N)	Mean difference (95% CI)
Fracture < 8 weeks (k = 3)					
1 day	2[6 132]	t^2 = 0.00 I^2 = 18.19%	154/165	145/167	-3.03 (-3.24, -2.82) p < 0.0001
1 week	2[6 132]	t^2 = 0.44 I^2 = 97.35%	153/165	144/167	-2.38 (-3.25, -1.50) p < 0.0001
1 month	2[6 132]	t^2 = 0.04 I^2 = 21.92%	152/165	143/167	-2.41 (-2.61, -2.22) p < 0.0001
3 months	3[6 96 132]	t^2 = 0.30 I^2 = 93.15%	172/190	160/191	-1.36 (-2.01, -0.72) p < 0.0001
6 months	2[6 132]	t^2 = 0.22 I^2 = 96.15%	145/165	132/165	-1.19 (-1.78, -0.61) p = 0.0001
12 months	3[6 96 132]	t^2 = 0.09 I^2 = 88.30%	164/190	150/191	-1.26 (-1.65, -0.88) p < 0.0001
Fracture > 8 weeks (k = 4)					
1 week	2[134 135]	t^2 = 13.36 I^2 = 99.79%	86/86	85/85	-0.57 (-4.20, 3.07) p = 0.76
2 weeks	2[8 141]	t^2 = 1.68 I^2 = 98.37%	69/82	75/77	-0.84 (-2.14, 0.46) p = 0.21
1 month	1[134]	t^2 = 0.08 I^2 = 73.76%	46/46	43/43	-1.12 (-1.53, -0.71) p < 0.0001
2 months	2[8 135]	t^2 = 1.14 I^2 = 97.62%	94/104	98/103	-2.16 (-3.24, -1.09) p < 0.0001
3 months	1[134]	t^2 = 0.08 I^2 = 73.76%	46/46	43/43	-1.32 (-1.73, -0.90) p = 0.0002
6 months	3[8 134 135]	t^2 = 1.18 I^2 = 97.70%	136/150	136/146	-1.17 (-2.25, -0.08) p = 0.04

Length of follow-up	Number of studies	Heterogeneity	PVP number of participants (n/N)	CT number of participants (n/N)	Mean difference (95% CI)
12 months	3[8 134 135]	t^2 = 0.90 I^2 = 97.01%	131/150	130/146	-1.28 (-2.24, -0.33) p = 0.01
24 months	1[135]	t^2 = 0.08 I^2 = 73.76%	38/40	39/42	-0.71 (-1.14, -0.28) p = 0.001
36 months	1[135]	t^2 = 0.08 I^2 = 73.76%	37/40	39/42	-1.71 (-2.15, -1.26) p < 0.0001

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **PVP** = percutaneous vertebroplasty.

Function-related outcomes

Table 115 PVP vs CT: sub-group analyses of the Oswestry disability index

Length of follow-up	Number of studies	Heterogeneity	PVP number of participants (n/N)	CT number of participants (n/N)	Mean difference (95% CI)
Fracture < 8 weeks (k = 1) [a]					
1 week	1[6]	NA	56/64	51/66	-17.74 (-20.74, -14.74) p < 0.05
1 month	1[6]	NA	56/64	51/66	-24.15 (-27.41, -20.89) p < 0.05
3 months	1[6]	NA	56/64	51/66	-25.28 (-28.50, -22.07) p < 0.05
6 months	1[6]	NA	56/64	51/66	-18.11 (-21.04, -15.18) p < 0.05
12 months	1[6]	NA	56/64	51/66	-7.92 (-10.64, -5.20) p < 0.05
Fracture > 8 weeks (k = 2)					
1 day	1[134]	t^2 = 0.40 I^2 = 83.90%	46/46	43/43	-13.75 (-14.77, -12.73) p < 0.0001
1 week	2[134 135]	t^2 = 0.60 I^2 = 88.66%	86/86	85/85	-14.45 (-15.63, -13.27) p < 0.0001
1 month	1[134]	t^2 = 0.40 I^2 = 83.90%	46/46	43/43	-12.95 (-13.96, -11.93) p < 0.0001
2 months	1[135]	t^2 = 0.40 I^2 = 83.90%	40/40	42/42	-15.45 (-16.51, -14.39) p < 0.0001
3 months	1[134]	t^2 = 0.40 I^2 = 83.90%	46/46	43/43	-15.34 (-16.41, -14.27) p < 0.0001
6 months	2[134 135]	t^2 = 18.51 I^2 = 99.59%	86/86	85/85	-14.05 (-20.03, -8.06) p < 0.0001
12 months	1[135]	t^2 = 0.40 I^2 = 83.90%	38/40	39/42	-12.46 (-13.52, -11.40) p < 0.0001
24 months	1[135]	t^2 = 0.40 I^2 = 83.90%	38/40	39/42	-12.46 (-13.53, -11.39) p < 0.0001
36 months	1[135]	t^2 = 0.40	37/40	39/42	-14.46

Length of follow-up	Number of studies	Heterogeneity	PVP number of participants (n/N)	CT number of participants (n/N)	Mean difference (95% CI)
		I^2 = 83.90%			(-15.56, -13.35) p < 0.0001

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **PVP** = percutaneous vertebroplasty.

Notes

a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed. Statistical significance was based on analysis performed in the study, indicating statistically significant differences at 1 week to 12 months.

Table 116 PVP vs CT: sub-group analyses of Roland-Morris disability questionnaire

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	CT number of patients (n/N)	Mean difference (95%CI)
Fracture < 8 weeks (k = 1) [a]					
1 week	1[132]	NA	97/101	93/101	-2.00 (-3.44, 0.56)
1 month	1[132]	NA	96/101	92/101	-1.50 (-3.22, 0.22)
3 months	1[132]	NA	92/101	86/101	-2.40 (-4.29, -0.52)
6 months	1[132]	NA	89/101	81/101	-1.70 (-3.72, 0.32)
12 months	1[132]	NA	86/101	77/101	-1.90 (-4.01, 0.21)
Fracture > 8 weeks (k = 1) [b]					
1 day	1[134]	NA	46/46	43/43	-2.50 (-3.14, -1.86) p < 0.05
1 week	1[134]	NA	46/46	43/43	-2.10 (-2.63, -1.57) p < 0.05
1 month	1[134]	NA	46/46	43/43	-2.60 (-3.26, -1.94) p < 0.05
3 months	1[134]	NA	46/46	43/43	-1.80 (-2.17, -1.43) p < 0.05
6 months	1[134]	NA	46/46	43/43	-2.60 (-2.98, -2.22) p < 0.05

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **PVP** = percutaneous vertebroplasty.
Notes
a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed. Statistical significance was not reported in the study; 95% CI is used to delineate statistical significance.
b = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed. Statistical significance was based on analysis performed in the study, indicating statistically significant differences at 1 day to 6 months.

Quality of life-related outcomes

Table 117 PVP vs CT: sub-group analyses of questionnaire of the European Foundation for Osteoporosis

Length of follow-up	Number of studies	Heterogeneity	PVP number of participants (n/N)	CT number of participants (n/N)	Mean difference (95% CI)
Fracture < 8 weeks (k = 2)					
1 week	2[6 132]	t^2 = 15.50 I^2 = 99.92%	153/165	144/167	-6.68 (-12.15, -1.22) p = 0.02
1 month	2[6 132]	t^2 = 64.14 I^2 = 99.98%	152/165	143/167	-9.86 (-20.97, 1.24) p = 0.08
3 months	2[6 132]	t^2 = 44.29 I^2 = 99.97%	148/165	137/167	-9.31 (-18.53, -0.08) p = 0.05
6 months	2[6 132]	t^2 = 52.36 I^2 = 99.98%	145/165	132/167	-8.52 (-18.55, 1.52) p = 0.10
12 months	2[6 132]	t^2 = 17.01 I^2 = 99.93%	142/165	128/167	-5.42 (-11.14, 0.30) p = 0.06
Fracture > 8 weeks [a]					
2 weeks	1[8]	NA	51/64	59/61	3.33 (-3.61, 10.28) p > 0.05
2 months	1[8]	NA	54/64	56/61	2.38 (-4.56, 9.32) p > 0.05
6 months	1[8]	NA	50/64	54/61	-2.06 (-4.78, 8.89) p > 0.05
12 months	1[8]	NA	47/64	48/61	2.54 (-5.06, 10.14) p > 0.05

Abbreviations
CI = confidence interval, **CT** = conservative treatment, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **PVP** = percutaneous vertebroplasty.
Notes
a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed. Statistical significance was not reported in the study; 95% CI is used to delineate statistical significance.

17.4.2 Clinical Effectiveness: PVP vs Sham

Pain-related outcomes

Table 118 PVP vs sham: sub-group analyses of pain (NRS and VAS)

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	Sham number of patients (n/N)	Mean difference (95%CI)
Fracture < 8 weeks (k = 2)					
1 day	1[95]	t^2 = 0.02 I^2 = 66.88%	90/90	86/86	0.33 (0.05, 0.60) p = 0.02
3 days	1[2]	t^2 = 0.02 I^2 = 66.88%	58/61	55/59	-1.61 (-1.93, -1.28) p < 0.0001
1 week	1[95]	t^2 = 0.02 I^2 = 66.88%	90/90	86/86	0.02 (-0.25, 0.29) p = 0.90
2 weeks	1[2]	t^2 = 0.02 I^2 = 66.88%	55/61	57/59	-1.10 (-1.41, -0.79) p < 0.0001
1 month	2[2 95]	t^2 = 0.48 I^2 = 98.32%	145/151	143/145	-0.90 (-1.88, 0.07) p = 0.07
3 months	2[2 95]	t^2 = 0.58 I^2 = 98.61%	143/151	138/145	-0.75 (-1.82, 0.32) p = 0.17
6 months	2[2 95]	t^2 = 0.40 I^2 = 98.02%	141/151	137/145	-0.84 (-1.74, 0.06) p = 0.07
12 months	1[95]	t^2 = 0.02 I^2 = 66.88%	90/90	86/86	-0.54 (-0.81, -0.28) p < 0.0001
Fracture > 8 weeks (k = 2)					
3 days	1[4]	t^2 = 0.00 I^2 = 4.44%	58/59	31/31	0.32 (0.02, 0.63) p = 0.04
1 week	1[7]	t^2 = 0.00 I^2 = 4.44%	37/38	37/40	0.58 (0.25, 0.90) p = 0.0005
2 weeks	1[4]	t^2 = 0.00 I^2 = 4.44%	56/59	30/31	-0.18 (-0.47, 0.12) p = 0.25
1 month	2[4 7]	t^2 = 0.01 I^2 =22.03%	93/97	68/71	-0.66 (-0.91, -0.41) p < 0.0001
3 months	2[4 7]	t^2 = 0.00 I^2 = 0.20%	91/97	66/71	-0.70 (-0.93, -0.48) p < 0.0001

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	Sham number of patients (n/N)	Mean difference (95%CI)
6 months	2[4 7]	t^2 = 0.09 I^2 = 80.59%	88/97	64/71	-0.53 (-1.00, -0.06) p = 0.03
12 months	2[4 7]	t^2 = 0.10 I^2 = 81.88%	86/97	57/71	-0.74 (-1.24, -0.25) p = 0.003
24 months	1[7]	t^2 = 0.00 I^2 = 4.44%	29/38	28/40	-1.14 (-1.51, -0.77) p < 0.0001
VAS (k = 1) [a]					
1 day	1[95]	NA	90/90	86/86	0.42 (-0.31, 1.15)
1 week	1[95]	NA	90/90	86/86	0.11 (-0.63, 0.85)
1 month	1[95]	NA	90/90	86/86	-0.41 (-1.15, 0.33)
3 months	1[95]	NA	90/90	86/86	-0.21 (-0.97, 0.54)
6 months	1[95]	NA	90/90	86/86	-0.39 (-1.16, 0.38)
12 months	1[95]	NA	90/90	86/86	-0.45 (-1.23, 0.34)
NRS (k = 3)					
3 days	2[2 4]	t^2 = 2.38 I^2 = 99.37%	116/120	86/90	-0.04 (-1.82, 1.74) p = 0.97
1 week	1[7]	t^2 = 0.02 I^2 = 58.60%	37/38	37/40	0.49 (0.14, 0.85) p = 0.01
2 weeks	2[2 4]	t^2 = 0.52 I^2 = 97.18%	111/120	87/90	-0.41 (-1.25, 0.44) p = 0.35
1 month	3[2 4 7]	t^2 = 0.18 I^2 = 92.94%	143/158	125/130	-0.91 (-1.42, -0.40) p = 0.0005
3 months	3[2 4 7]	t^2 = 0.11 I^2 = 87.75%	144/158	118/130	-0.90 (-1.32, -0.49) p < 0.0001
6 months	3[2 4 7]	t^2 = 0.24 I^2 = 93.99%	139/158	115/130	-0.79 (-1.36, -0.21) p = 0.01
12 months	2[4 7]	t^2 = 0.29 I^2 = 95.14%	86/97	57/71	-1.03 (-1.68, -0.37) p = 0.002
24 months	1[7]	t^2 = 0.02 I^2 = 58.60%	29/38	28/40	-1.26 (-1.65, -0.87)

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	Sham number of patients (n/N)	Mean difference (95%CI)
					p < 0.0001

Abbreviations
CI = confidence interval, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **NRS** = numerical rating scale, **PVP** = percutaneous vertebroplasty, **VAS** = visual analogue scale.
Notes
a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed.

Table 119 PVP vs sham: sub-group analyses of analgesic use

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	Sham number of patients (n/N)	Risk Ratio (RR, 95% CI)
Fracture < 8 weeks (k = 2)					
1 day	2[2,95]	Chi^2 = 11.72 P < 0.0006 I^2 = 91%	127/148	113/143	1.07 (0.77, 1.50) p = 0.68
1 week	2[2,95]	Chi^2 = 2.51 P =0.09 I^2 = 60%	123/144	116/142	1.03 (0.88, 1.20) p = 0.44
1 month	2[2,95]	Chi^2 = 1.11 P = 0.29 I^2 = 10%	59/154	78/148	1.09 (0.88, 1.35) p = 0.72
3 months	2[2,95]	Chi^2 = 2.60 P = 0.11 I^2 = 61%	85/138	91/133	0.88 (0.67, 1.17) p = 0.39
6 months	2[2,95]	Chi^2 = 1.30 P = 0.25 I^2 = 23%	72/133	84/134	0.85 (0.68, 1.07) p = 0.16
12 months	1[95]	NA	44/79	37/79	1.19 (0.88, 1.62) p = 0.27
Fracture > 8 weeks (k = 1) [a]					
1 month	1[4]	NA	37/68	27/63	1.27 (0.88, 1.62) p = 0.19

Abbreviations
CI = confidence interval, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **PVP** = percutaneous vertebroplasty.
Notes
a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed.

Function-related outcomes

Table 120 PVP vs sham: sub-group analyses of Roland-Morris disability questionnaire

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	Sham number of patients (n/N)	Mean difference (95%CI)
Fracture < 8 weeks (k = 1) [a]					
1 week	1[95]	NA	90/90	86/86	0.82 (-1.04, 2.68) p < 0.05
1 month	1[95]	NA	90/90	86/86	-1.12 (-2.96, 0.72) p < 0.05
3 months	1[95]	NA	90/90	86/86	-0.61 (-2.46, 1.24) p < 0.05
6 months	1[95]	NA	90/90	86/86	-0.88 (-2.76, 1.00) p < 0.05
12 months	1[95]	NA	90/90	86/86	-0.01 (-1.94, 1.92) p < 0.05
Fracture > 8 weeks (k = 2)					
1 day	1[4]	t^2 = 0.05 I^2 = 66.22%	58/59	31/31	0.46 (-0.08, 1.01) p = 0.09
1 week	1[7]	t^2 = 0.05 I^2 = 66.22%	35/38	38/40	-2.26 (-2.84, -1.69) p < 0.0001
2 weeks	1[4]	t^2 = 0.05 I^2 = 66.22%	56/59	30/31	-0.01 (-0.53, 0.51) p = 0.97
1 month	2[4 7]	t^2 = 1.82 I^2 = 98.52%	96/97	68/71	0.15 (-1.73, 2.03) p = 0.88
3 months	2[4 7]	t^2 = 0.09 I^2 = 76.81%	91/97	66/71	-1.32 (-1.80, -0.84) p < 0.0001
6 months	2[4 7]	t^2 = 3.84 I^2 = 99.29%	88/97	64/71	-0.77 (-3.50, 1.96) p = 0.58
12 months	2[4 7]	t^2 = 1.45 I^2 = 98.15%	86/97	57/71	-1.15 (-2.84, 0.54) p = 0.18
24 months	1[7]	t^2 = 0.05 I^2 = 66.22%	29/38	28/40	-0.05 (-0.61, 0.51) p = 0.86

Abbreviations

CI = confidence interval, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **PVP** = percutaneous vertebroplasty.

Notes

a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed. Statistical significance was based on analysis performed in the study, indicating no statistically significant differences at any timepoint.

Quality of life-related outcomes

Table 121 PVP vs sham: sub-group analysis of questionnaire of the European Foundation for Osteoporosis

Length of follow-up	Number of studies	Heterogeneity	PVP number of patients (n/N)	Sham number of patients (n/N)	Mean difference (95%CI)
Fracture < 8 weeks (k = 2)					
1 week	1[95]	t^2 = 0.80 I^2 = 99.02%	90/90	86/86	0.60 (-0.66, 1.85) p = 0.35
2 weeks	1[2]	t^2 = 0.80 I^2 = 99.02%	48/61	54/59	-5.37 (-6.64, -4.10) p < 0.0001
1 month	2[2 95]	t^2 = 1.05 I^2 = 99.26%	138/151	138/145	-2.27 (-3.71, -0.84) p = 0.0002
3 months	1[95]	t^2 = 0.80 I^2 = 99.02%	90/90	86/86	-1.36 (-2.61, -0.11) p = 0.03
6 months	2[2 95]	t^2 = 29.32 I^2 =99.97%	136/151	134/145	-3.17 (-10.68, 4.34) p = 0.41
12 months	1[95]	t^2 = 0.80 I^2 = 99.02%	90/90	86/86	-1.31 (-2.57, -0.06) p = 0.04
Fracture > 8 weeks (k = 1)[a]					
1 week	1[7]	NA	35/38	38/40	-4.10 (-7.95, -0.24) p < 0.05
1 month	1[7]	NA	38/38	38/40	0.40 (-4.50, 5.30) p > 0.05
3 months	1[7]	NA	36/38	37/40	-0.10 (-5.54, 5.34) p > 0.05
6 months	1[7]	NA	35/38	36/40	0.30 (-5.93, 6.53) p > 0.05
12 months	1[7]	NA	33/38	34/40	-2.10 (-8.21, 4.01) p > 0.05
24 months	1[7]	NA	29/38	28/40	1.30 (-5.44, 8.04) p > 0.05

Abbreviations

CI = confidence interval, **k** = number of studies, **n** = number of patients at timepoint, **N** = total number of patients, **NA** = not applicable, **PVP** = percutaneous vertebroplasty.

Notes

a = There was only 1 study in the sub-group, so a longitudinal meta-analysis was not performed. Statistical significance was based on analysis performed in the study, indicating a statistically significant difference at 1 week but no other timepoints.

17.5 Appendix E: Minimum Clinically Important Differences and Improvements for Outcomes of Interest

A non-systematic search was conducted to identify MCIDs and minimum clinically important improvements (MCIIs) for the outcomes of interest. The identified MCIDs were intended to act as a guide and did not result from a comprehensive assessment of the literature.

There were a limited number of published MCIDs specifically examining patients with vertebral fracture. Consequently, the search was expanded to include any spine-related pathology. The MCIDs generally relate to patients with chronic lower back pain or lumbar degenerative disc disease, with measures of pain and function the most frequently reported outcomes. The different MCIDs values are attributable to the method of determining the MCID (anchor or distribution based) and patient demographics.

Given much of the MCIDs and MCIIs are in different populations from that assessed in the HTA, their applicability to the current report is uncertain.

No MCIDs were identified for QUALEFFO, SOF-ADL or analgesic use.

Table 122 Minimum clinically important differences/improvements for outcomes of interest

MIC/MCID/MCII	Study type	Population demographics	Reference
Vertebral fractures *Roland-Morris disability questionnaire*			
Distribution-based [a] 2–8 (MCID)	Cohort study	PVP or PBK Age: 75 Sex: 66.4% female	Lee 2017[112]
2–3 (scoring range 0–23) (MCID)	SR	NR	Roland 2000[113]
Other patient populations			
EuroQol 5 dimension questionnaire			
0.24 (MCID)	Cohort study	Patients with cervical radiculopathy	Parker 2013[111]
0.17 (MIC)	Cohort study	Patients with chronic back pain undergoing surgery or rehabilitation Age: 41 Sex: 2.6% Female Duration of symptoms: 2 years	Johnsen 2013[115]
Numerical rating scale			
Anchor-based 4.0 Distribution-based 0.86 (MCID)	Cohort study	Patients with chronic lower back pain undergoing physical therapy Age: 52 Sex: 67% female Duration of symptoms: 6m	Maughan & Lewis 2010[108]
2.0 or 30% from baseline	Systematic review and	Patients with chronic lower back	Ostelo 2008[109]

MIC/MCID/MCII	Study type	Population demographics	Reference
[b] 1–4.5 [c] (MIC)	panel input	pain	
Average 4 (95% CI, 3.4, 5.0) (MDC) 1.5 (MCII)	Cohort study	Patients seeking treatment for neck pain Age: 54.1 Sex: 77.5% Female Duration of pain: 541.7 days	Kovacs 2008[107]
Oswestry disability index			
Distribution-based 12.81 (scoring range 0–50) (MCID)	Cohort study	Patients undergoing spinal surgery	Copay 2008[114]
Anchor-based 7.5 Distribution-based 6.06 (MCID)	Cohort study	Patients with chronic lower back pain undergoing physical therapy Age: 52 Sex: 67% female Duration of symptoms: 6 months	Maughan & Lewis 2010[108]
10 or 30% from baseline [b] 4–15.0 [c] (MIC)	Systematic review and panel input	Patients with chronic lower back pain	Ostelo 2008[109]
Roland-Morris disability questionnaire			
Anchor-based 3.5 Distribution-based 1.78 (MCID)	Cohort study	Patients with chronic lower back pain undergoing physical therapy Age: 52 Sex: 67% female Duration of symptoms: 6 months	Maughan & Lewis 2010[108]
5% or 30% from baseline [b] 2.0–8.6 [c] (MIC)	Systematic review and panel input	Patients with chronic lower back pain	Ostelo 2008[109]
Short form 36 questionnaire			
3 (MCID)	Cohort study	Patients with chronic back pain Age: 44–47 Sex: Female 53–54% Duration of symptoms: <30 days 12–73%	Lauridsen 2006[116]
1.16 (scoring scale 1–10)	Cohort study	Patients undergoing spinal surgery	Copay 2008[114]
Timed up-and-go			
3.4 seconds (MCID)	Cohort study	Patients with lumbar degenerative disc disease undergoing microdiscectomy, fusion or decompression Age: 56.2 Sex: 43% female	Gautschi 2017[264]
Visual analogue scale			

MIC/MCID/MCII	Study type	Population demographics	Reference
15 points or 30% from baseline [b] 2.0–29 [c] (MIC)	Systematic review and panel input	Patients with chronic lower back pain	Ostelo 2008[109]
2.6 (MCID)	Cohort study	Patients with cervical radiculopathy	Parker 2013[111]
Back pain 4–6 Leg pain 3.9–6 (MCID)	Cohort study	Patients with lumbar degenerative disc disease undergoing laminectomy/foraminotomy Age: 56.3 Sex: Female 66%	Parker 2012[110]

Abbreviations

NR = not reported, **MDC** = minimum detectable change, **MIC** = minimum important change, **MCID** = minimum clinically important difference, **MCII** = minimum clinically important improvements, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty, **SR** = systematic review.

Notes

a = Distribution-based refers to standard error of measurement, **b** = estimates based on literature search, **c** = estimates derived from expert group.

17.6 Appendix F: GRADE Evidence Profile Tables

17.6.1 Clinical Effectiveness

Table 123 GRADE evidence profile table for PVP vs CT for osteoporotic vertebral fractures at 1 month

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PVP	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS												
3	randomised trials	serious [a]	serious [b]	not serious	serious [c,d]	none	198	186	-	MD 1.52 mm lower (2.86 lower to 0.17 lower)	⊕⊕◯◯ LOW	CRITICAL
Pain: analgesic use (number of patients taking analgesics)												
2	randomised trials	serious [a]	serious [b]	serious [e]	serious [f]	none	56/110 (50.9%)	71/104 (68.3%)	RR 0.53 (0.10, 2.69)	321 fewer per 1,000 (from 614 fewer to 1,000 more)	⊕⊕◯◯ LOW	IMPORTANT
Function: ODI												
2	randomised trials	serious [a]	serious [b]	not serious	serious [c,d]	none	102	94	-	MD 16.27 points lower (23.53 lower to 9.01 lower)	⊕⊕◯◯ LOW	CRITICAL
Function: RDQ												

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PVP	Sham	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials	serious [a]	serious [b]	not serious	serious [c,d]	none	142	135	-	MD 2.03 points lower (3.06 lower to 1.01 lower)	⊕⊕◯◯ LOW	CRITICAL
Function: timed up-and-go												
1	randomised trial	serious [a]	not serious	not serious	serious [d]	none [g]	PVP vs CT, mean ± SD 3m 16.0 ± 5.5 vs 17.0 ± 9.7, p = 0.75				⊕⊕◯◯ LOW	IMPORTANT
QoL: EQ-5D												
1	randomised trial	serious [a]	serious [b]	not serious	serious [d,f,h]	serious [a]	96	92	-	MD 0.10 points higher (0.11 lower to 0.31 higher)	⊕⊕◯◯ LOW	CRITICAL
QoL: QUALEFFO												
2	randomised trials	serious [a]	serious [b,c]	not serious	serious [f]	none	152	143	-	MD 6.16 points lower (15.84 lower to 3.52 lower)	⊕⊕◯◯ LOW	CRITICAL
QoL: SF-36												
1	randomised trials	serious [a]	not serious	not serious	serious [d,f]	none [g]	PVP vs CT, mean ± SD 3m 34.0 ± 9.5 vs 29.3 ± 11.0, p = 0.12				⊕⊕◯◯ LOW	CRITICAL

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**PVP**	**Sham**	**Relative (95% CI)**	**Absolute (95% CI)**		
Length of hospital stay												
1	randomised trial	not serious	not serious	not serious	serious [d]	none	PVP vs CT mean (95%CI) 7.6 (5.8, 9.3) days vs 11.7 (9.1, 14.3) days, p = 0.01				⊕⊕⊕◯ MODERATE	IMPORTANT

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **m** = months, **MD** = mean difference, **ODI** = Oswestry disability index, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **QoL** = quality of life, **RDQ** = Roland-Morris disability questionnaire, **RR** = risk ratio, **SD** = standard deviation, **SF-36** = short form 36 questionnaire, **VAS** = visual analogue scale.

Notes

a = lack of blinding, incomplete accounting of patients and outcome events; **b** = considerable levels of heterogeneity as inferred by I^2 and Tau^2 (or Chi^2); **c** = 95% CI around pooled estimates includes negligible effect and appreciable benefit/harm (depending on the MCID); **d** = low number of patients at evaluated timepoint; **e** = indirect marker of pain; **f** = wide 95% CI; **g** = study protocol unavailable, cannot determine impact; **h** = baseline differences between intervention arms.

Table 124 GRADE evidence profile table for PVP vs sham for osteoporotic vertebral fractures at 1 month

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsist-ency	Indirectness	Imprecision	Other considerations	PVP	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain: NRS/VAS												
4	randomised trials	not serious	serious [a]	not serious	not serious	none	238	211	-	MD 0.76 points lower (1.21 lower to 0.31 lower)	⊕⊕⊕◯ MODERATE	CRITICAL
Pain: analgesic use (number of patients using analgesics)												
3	randomised trials	not serious	serious [b,c]	serious [d]	serious [e]	not serious	130/209 (62.2%)	128/205 (62.4%)	RR 0.99 (0.79, 1.24)	6 fewer per 1,000 (from 131 fewer to 150 more)	⊕⊕◯◯ LOW	IMPORTANT
Function: timed up-and-go												
1	randomised trial	not serious	serious [b]	not serious	serious [e]	none	PVP vs Sham; mean ± SD 12m, -2.6 ± 12.2 vs 4.3 ± 13.4, p = NR				⊕⊕◯◯ LOW	IMPORTANT
Function: RDQ												
3	randomised trials	not serious	serious [a]	not serious	serious [b,e]	none	186	154	-	MD 0.28 points lower (1.70 lower to	⊕⊕◯◯ LOW	CRITICAL

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsist-ency	Indirectness	Imprecision	Other considerations	PVP	Sham	Relative (95% CI)	Absolute (95% CI)		
										1.15 higher)		
QoL: EQ-5D												
3	randomised trials	not serious	serious [f]	not serious	serious [e]	none	PVP vs Sham; mean ± SD Buchbinder (2009) 0.10 ± 0.30 vs 0.10 ± 0.30, p = NS Clark (2016) 0.75 ± 0.11 vs 0.70 ± 0.11, p = 0.04 Kallmes (2009) 0.70 ± 0.18 vs 0.64 ± 0.20, p = 0.13				⊕⊕◯◯ LOW	CRITICAL
QoL: QUALEFFO												
3	randomised trials	not serious	serious [a]	not serious	serious [b,e]	none	176	176	-	MD 1.39 points lower (3.24 lower to 0.47 higher)	⊕⊕◯◯ LOW	CRITICAL
QoL: SF-36												
1	randomised trial	not serious	serious [b]	not serious	serious [e]	none	PVP vs Sham; mean ± SD 29.7 ± 9.6 vs 28.7 ± 8.0, p = 0.45				⊕⊕◯◯ LOW	CRITICAL
QoL: SOF-ADL												
1	randomised trial	not serious	serious [b]	not serious	serious [e]	none	PVP vs Sham; mean ± SD				⊕⊕◯◯ LOW	CRITICAL

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsist-ency	Indirectness	Imprecision	Other considerations	PVP	Sham	Relative (95% CI)	Absolute (95% CI)		
							7.7 ± 3.7 vs 8.2 ± 3.6, p = 0.51					
Length of hospital stay												
1	randomised trial	not serious	not serious	not serious	serious [e]	none	PVP vs Sham; median (IQR) 8.5 (4–13) days vs 14 (7–22) days				⨁⨁◯◯ LOW	CRITICAL

Abbreviations

CI = confidence interval, **EQ-5D** = EuroQol 5 dimension questionnaire, **IQR** = interquartile range, **m** = months, **MD** = mean difference, **ODI** = Oswestry disability index, **PVP** = percutaneous vertebroplasty, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **QoL** = quality of life, **RDQ** = Roland-Morris disability questionnaire, **RR** = risk ratio, **SD** = standard deviation, **SF-36** = short form 36 questionnaire, **SOF-ADL** = study of osteoporotic fractures–activities of daily living, **NR** = not reported, **NRS** = numerical rating scale, **NS** = not significant, **VAS** = visual analogue scale.

Notes

a = considerable levels of heterogeneity as inferred by I^2 and Tau^2 (or Chi^2), **b** = wide 95% CI or SD, **c** = moderate levels of heterogeneity as inferred by I^2 and Chi^2, **d** = indirect measure of pain, **e** = low number of patients at evaluated timepoint, **f** = direction of effect inconsistent between studies.

Table 125 GRADE evidence profile table for PBK vs CT for osteoporotic vertebral fractures (RCTs)

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsist-ency	Indirectness	Imprecision	Other considerations	PBK	CT	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS (follow-up: 1 month)												
2	randomised trials	serious [a]	serious [b]	not serious [c]	serious [d,e]	none	176	168	-	MD 18 mm lower (2.15 lower to 1.80 higher)	⊕⊕◯◯ LOW	CRITICAL
Pain: analgesic use (number of patients taking analgesics) (follow-up: 1 month)												
1	randomised trial	serious [a]	not serious	serious [c,f]	serious [d]	none	81/144 (56.3%)	105/115 (91.3%)	not estimable		⊕⊕◯◯ LOW	IMPORTANT
Function: ODI (follow-up: 1 month)												
1	randomised trial	serious [a]	not serious	not serious [c]	serious [d]	none	PBK vs CT; mean ± SD 15.1 ± 3.6 vs 18.7 ± 5.3, p < 0.05				⊕⊕◯◯ LOW	CRITICAL
Function: RDQ (follow-up: 1 month)												
1	randomised trial	serious [a]	not serious	not serious [c]	serious [d]	none	PBK vs CT; mean ± SD 10.9 ± 4.3 vs 15.1 ± 4.3, p < 0.0001				⊕⊕◯◯ LOW	CRITICAL
Function: timed up-and-go (follow-up: 1 month)												
1	randomised trial	serious [a]	not serious	not serious [c]	serious [d]	none	PBK vs CT; mean ± SD 12.7 ± 6.7 vs 18.7 ± 6.9, p = 0.09				⊕⊕◯◯ LOW	IMPORTANT
QoL: EQ-5D (follow-up: 1 month)												
1	randomised trial	serious [a]	not serious	not serious [c]	serious [d]	none	PBK vs CT; mean ± SD 0.59 ± 1.07 vs 0.49 ± 1.04, p < 0.0001				⊕⊕◯◯ LOW	CRITICAL
QoL: SF-36 (follow-up: 1 month)												
1	randomised trial	serious [a]	not serious	not serious [c]	serious [d]	none	PBK vs CT; Mean ± SD 33.4 ± 5.6 vs 27.5 ± 5.6, p < 0.0001				⊕⊕◯◯ LOW	CRITICAL

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	**Study design**	**Risk of bias**	**Inconsist-ency**	**Indirectness**	**Imprecision**	**Other considerations**	**PBK**	**CT**	**Relative (95% CI)**	**Absolute (95% CI)**		
Length of hospital stay												
1	randomised trial	not serious	not serious	not serious	serious [d]	none	median length of stay 4 days (IQR 2–9 days) following PBK				⨁⨁⨁◯ MODERATE	IMPORTANT

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **EQ-5D** = EuroQol 5 dimension questionnaire, **IQR** = interquartile range, **MD** = mean difference, **ODI** = Oswestry disability index, **PBK** = percutaneous balloon kyphoplasty, **QUALEFFO** = quality of life questionnaire of the European Foundation for Osteoporosis, **QoL** = quality of life, **RCT** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **RR** = risk ratio, **SD** = standard deviation, **SF-36** = short form 36 questionnaire, **VAS** = visual analogue scale.

Notes

a = lack of blinding and concealment, complete accounting of patients or outcome events; **b** = considerable levels of heterogeneity as inferred by I^2 and Tau^2, **c** = demographics broadly congruent with Swiss population, however indication for kyphoplasty differs slightly; **d** = low number of patients at evaluated timepoint; **e** = 95% CI around pooled estimates includes negligible effect and appreciable benefit/harm (depending on the MCID); **f** = indirect marker of pain.

Table 126 GRADE evidence profile table for PBK vs CT for osteoporotic vertebral fractures (non-RCTs)

Certainty assessment							Number of patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PBK	CT	Relative (95% CI)	Absolute (95% CI)		
Pain: VAS (follow-up: 3 months)												
1	observational study	serious [a]	serious [b]	not serious	serious [c]	none	PBK vs CT; mean ± SD 42.4 ± 17.9 vs 33.9 ± 18.4, p = 0.012				⊕⊕◯◯ LOW	CRITICAL
Pain: analgesic use (number of patients taking analgesics) (follow-up: NR)												
1	observational study	not serious	not serious	serious [d]	serious [c]	none	22/40 (55.0%)	13/20 (65.0%)	not estimable	-	⊕⊕◯◯ LOW	CRITICAL
Function: RDQ (follow-up: 1 month)												
1	observational study	serious [a]	serious [e]	not serious	serious [c]	none	PBK vs CT; mean ± SD 10.3 ± NR vs 14.4 ± NR, p = 0.004				⊕⊕◯◯ LOW	CRITICAL

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **NR** = not reported, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **RDQ** = Roland-Morris disability questionnaire, **SD** = standard deviation, **VAS** = visual analogue scale.

Notes

a = concerns regarding blinding and unclear modification to measurement tool; baseline VAS scores significantly differed; **b** = large estimates of variance; **c** = low number of patients at evaluated timepoint; **d** = indirect measure of pain; **e** = measures of variance not reported.

1 month follow-up was presented. For outcomes without 1 month data the closest follow-up duration was reported.

17.6.2 Safety

Table 127 GRADE evidence profile table for PVP vs CT and sham for osteoporotic vertebral fractures (RCTs)

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PVP	Sham and CT	Relative (95% CI)	Absolute (95% CI)		
All-cause mortality (follow-up: 6–36 months)												
9	randomised trials	not serious	not serious	not serious	serious [a]	none	31/640 (4.8%)	37/641 (5.8%)	RR 0.84 (0.52, 1.35)	11 fewer per 1,000 (from 27 fewer to 20 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Serious adverse events (follow-up: 24 months)												
3	randomised trials	not serious [b]	not serious	not serious	serious [a,c]	none	5/268 (1.9%)	5/263 (1.9%)	RR 0.99 (0.29, 3.35)	4 fewer per 1,000 (from 13 fewer to 45 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Any adverse events (follow-up: 6–36 months)												
7	randomised trials	not serious [b]	serious [d]	not serious	serious [a,c]	none	28/424 (6.6%)	24/438 (5.5%)	RR 1.68 (0.57, 4.91)	4 fewer per 1,000 (from 38 fewer to 344 more)	⊕⊕◯◯ LOW	IMPORTANT
Cement leakage (follow-up: 6–36 months)												
9	randomised trials	not serious [b]	serious [d]	not serious	serious [a,c]	none	Absolute rate of cement leaks per treated vertebrae 55.1% (n = 343/623), ranging from 14.0% (n =				⊕⊕◯◯ LOW	IMPORTANT

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PVP	Sham and CT	Relative (95% CI)	Absolute (95% CI)		
							14/100) to 91.3% (n = 105/115).					
New fractures (symptomatic) (follow-up: 24 months)												
6	randomised trials	not serious	serious [d]	not serious	serious [a,f]	not serious	48/418 (11.5%)	31/422 (7.3%)	RR 1.29 (0.46, 3.62)	21 more per 1,000 (from 40 fewer to 192 more)	⨁⨁◯◯ LOW	CRITICAL
New fractures (radiological evidence) (follow-up: 6–36 months)												
7	randomised trials	not serious	serious [d]	not serious	serious [a,f]	none	106/389 (27.2%)	88/369 (23.8%)	RR 1.18 (0.70, 1.99)	45 more per 1,000 (from 6 fewer to 109 more)	⨁⨁◯◯ LOW	IMPORTANT

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **n** = number of patients, **PVP** = percutaneous vertebroplasty, **RCT** = randomised controlled trials, **RR** = risk ratio.

Notes

a = low number of patients at evaluated timepoint; **b** = all trials had incomplete data, could influence event rate; **c** = 95% CI around pooled estimates includes negligible effect and appreciable benefit/harm; **d** = moderate to considerable levels of heterogeneity as measured by I2 and Chi squared test; **e** = point estimates vary substantially; **f** = likely includes negligible effect and appreciable harm estimates.

Table 128 GRADE evidence profile table for PVP vs CT for osteoporotic vertebral fractures (database analyses, non-RCTs and single-arm trials)

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PVP	CT	Relative (95% CI)	Absolute (95% CI)		
All-cause mortality (Non-RCTs) (follow-up: 24 months)												
1	observational study	not serious	not serious	not serious	serious [a]	none	15/88 (17.0%)	6/38 (15.8%)	not pooled	not pooled	⊕⊕⊕◯ MODERATE	CRITICAL
Serious adverse events (Non-RCTs) (follow-up: 24 months)												
1	observational study	serious [b]	not serious	not serious	serious [a]	none	0/88 (0.0%)	0/38 (0.0%)	not pooled	not pooled	⊕⊕◯◯ LOW	CRITICAL
Any adverse events (Non-RCTs) (follow-up: 24 months)												
2	observational studies	serious [c]	not serious	not serious	serious [a]	none	3/118 (2.5%)	0/118 (0.0%)	not pooled	not pooled	⊕⊕◯◯ LOW	IMPORTANT
New fracture (radiographic) (Non-RCTs) (follow-up: 24 months)												
1	observational study	serious [b]	not serious	not serious	serious [a]		21/88 (23.9%)	9/38 (23.7%)	not pooled	not pooled	⊕⊕◯◯ LOW	CRITICAL
All-cause mortality (databases) (follow-up: 30 days to 10 years)												
2	observational studies	not serious	not serious	serious [d]	not serious	none	PVP vs CT 30 days, 0.53% vs 1.72%, p < 0.001 1 year, 30% decrease compared to CT 10 years, 8% decrease compared to CT				⊕⊕⊕◯ MODERATE	CRITICAL
Any adverse events (databases) (follow-up: 30 days to 10 years)												
2	observational studies	not serious	not serious	serious [d]	not serious	none	Incidence of adverse event generally lower in PVP group compared to CT (1–20% difference reflecting the adverse event)				⊕⊕⊕◯ MODERATE	IMPORTANT
Cement leakage (single-arm trials) (follow-up: 12–60 months)												
15	observational studies	not serious	serious [e]	not serious	serious [a]	none	Absolute rate of cement leaks per treated vertebrae 38.5% (n = 1101/2863), ranging from 11.7% (n =				⊕⊕◯◯ LOW	IMPORTANT

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PVP	CT	Relative (95% CI)	Absolute (95% CI)		
							15/128) to 81.6% (n = 164/201) treated vertebrae.					

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **n** = number of patients, **PVP** = percutaneous vertebroplasty, **RCT** = randomised controlled trials.

Notes

a = low number of patients at evaluated time-point, **b** = intervention characteristics not well defined, trials had incomplete data, **d** = the codes used to identify relevant patients may also include non-osteoporotic fractures, **e** = range of effects varies, **f** = likely includes negligible effect and appreciable harm estimates.

Table 129 GRADE evidence profile table for PBK vs CT for osteoporotic vertebral fractures (RCTs)

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PBK	CT	Relative (95% CI)	Absolute (95% CI)		
All-cause mortality (follow-up: 24 months)												
1	randomised trial	not serious	not serious	not serious [a]	serious [b]	none	9/149 (6.0%)	7/151 (4.6%)	not pooled	not pooled	⊕⊕⊕◯ MODERATE	CRITICAL
Serious adverse events (follow-up: 24 months)												
1	randomised trial	serious [c]	not serious	not serious [a]	serious [b]	none	PBK vs CT, 58 vs 54 events				⊕⊕◯◯ LOW	CRITICAL
Any adverse events (follow-up: 24 months)												
1	randomised trial	serious [c]	not serious	not serious [a]	serious [b]	none	PBK vs CT, 130 vs 122 events				⊕⊕◯◯ LOW	IMPORTANT
Cement leakage (follow-up: 24 months)												
1	randomised trial	serious [c]	not serious	not serious [a]	serious [b]	none	51/188 cement leakage per vertebral bodies treated				⊕⊕◯◯ LOW	IMPORTANT
New fractures (clinical per person) (follow-up: 12 months)												
1	randomised trials	serious [c]	not serious	not serious [a]	serious [b]	none	21/149 (14.1%)	9/151 (6.0%)	not pooled	not pooled	⊕⊕◯◯ LOW	CRITICAL
New fractures (radiographic per person) (follow-up: 12 months)												
1	randomised trials	serious [c]	not serious	not serious [a]	serious [b]	none	38/115 (33.0%)	24/95 (25.3%)	not pooled	not pooled	⊕⊕◯◯ LOW	IMPORTANT

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials.

Notes

a = demographics broadly congruent with Swiss population, however indication for kyphoplasty differs slightly; **b** = low number of patients/events at evaluated timepoint; **c** = incomplete accounting of patients and outcome events.

Table 130 GRADE evidence profile table for PBK vs CT for osteoporotic vertebral fractures (database analyses, non-RCTs and single-arm trials)

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PBK	CT	Relative (95% CI)	Absolute (95% CI)		
Mortality (non-RCTs) (follow-up: 36 months)												
1	observational study	not serious	not serious	not serious	serious [a]	none	1/40 (2.5%)	3/20 (15.0%)	not pooled	not pooled	⊕⊕⊕◯ MODERATE	CRITICAL
Any adverse event (non-RCTs) (follow-up: 36 months)												
1	observational study	not serious	not serious	not serious	serious [a]	none	0/40 (0.0%)	0/20 (0.0%)	not pooled	not pooled	⊕⊕⊕◯ MODERATE	IMPORTANT
New fracture (non-RCTs) (follow-up: 12–24 months)												
3	observational studies	not serious	serious	not serious	serious [a]	none	19/105 (18.1%)	25/100 (25.0%)	not pooled	not pooled	⊕⊕◯◯ LOW	CRITICAL
Cement leakage (non-RCTs) (follow-up: 12–36 months)												
3	observational studies	not serious	serious [b]	not serious	serious [a]	none	Absolute rate per treated vertebrae 11.3% (n = 11/97), ranging from 9.7% (n = 7/72) to 16.0% (n = 4/25). Rate per patient 8.7% (n = 4/46).				⊕⊕◯◯ LOW	IMPORTANT
All-cause mortality (databases) (follow-up: 30 days to 10 years)												
2	observational studies	not serious	not serious	serious [c]	not serious	none	PBK vs CT 30 days, 0.35% vs 1.72%, p < 0.001 1 year, 55% decrease compared to CT 10 years, 24% decrease compared to CT				⊕⊕⊕◯ MODERATE	CRITICAL
Any adverse events (databases) (follow-up: 30 days to 10 years)												
2	observational studies	not serious	not serious	serious [c]	not serious	none	Incidence of adverse event generally lower in PBK group compared to CT (1–23% difference depending on adverse event)				⊕⊕⊕◯ MODERATE	IMPORTANT
Cement leakage (single-arm trials) (follow-up: 6–36 months)												
6	observational	not	serious [d]	not serious	serious [e]	none	Absolute rate per treated vertebrae 27.5% (n = 385/1,402) ranging from 5.2% (n = 7/135) to 73.4% (n =				⊕⊕◯◯	IMPORTANT

Certainty Assessment							Number of Patients		Effect		Certainty	Importance
Number of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PBK	CT	Relative (95% CI)	Absolute (95% CI)		
	studies	serious					157/214). Rate per patient 2.3% (n = 18/768) ranging from 0.5% (n = 3/564) to 7.1% (n = 8/102)				LOW	

Abbreviations

CI = confidence interval, **CT** = conservative treatment, **PBK** = percutaneous balloon kyphoplasty, **RCT** = randomised controlled trials, **yr** = years.

Notes

a = low number of patients/events at evaluated timepoint, **b** = inconsistent direction of effect, **c** = codes used to identify relevant patients may also include non-osteoporotic fractures, **d** = range varies between studies, **e** = likely includes negligible effect and appreciable harm estimates.

17.7 Appendix G: Ongoing Clinical Trials

A list of identified ongoing clinical trials is presented in ***Table 131***. There are 4 RCTs comparing PVP to CT or sham procedure and three non-randomised trials. Follow-up times range from 3 to 24 months. No ongoing PBK trials were identified.

Table 131 Ongoing clinical trials

Trial registry ID	Indication; Target sample size	Design	Intervention; Comparator(s)	Primary outcomes	Expected completion date; Status
NCT01677806	Acute (clinical onset <6 weeks) OVCF in patients aged >50 140 participants	RCT	PVP CT	Pain at 1 month. Function, quality of life, and incident fractures at 1, 3, 6 and 12 months	December 2014 Last update September 2014 Unknown
NCT01963039 (VERTOS V)	Acute OVCF 180 participants	RCT	PVP Sham procedure	Pain with VAS up to 12 months	July 2018 Unknown
NCT03360383	Acute OVCF 400 participants	RCT	PVP CT	Change in WHO-classified pain status up to 12 months	June 2020 Not yet recruiting
NCT03617094	Acute (<10 days) vertebral fracture in patients aged >50 58 participants	RCT	PVP CT	Difference in kyphotic angle at 3 months. Improvement in VAS pain up to 3 months	December 2020 Recruiting
NCT03692143	Women with OVCF 90 participants	non-RCT	PVP without teriparatide PVP with daily injection teriparatide.	QoL up to 2 years with SF-36 up to 24 months	December 2030 Active, not recruiting
ChiCTR1800016493	OVCF of thoracolumbar spine 900 participants	non-RCT	PVP Kyphoplasty Physical therapy and TCM	Back pain incidence up to 2 years. VAS & ODI up to 6 months	November 2021 Recruiting
NCT03330340	Osteoporosis 106 participants	non-RCT	PVP Conservative management	Incidence of vertebral refracture up to 12 months	December 2019 Not yet recruiting

Abbreviations

CT = conservative treatment, **NCT** = ClinicalTrials.gov identifier, **ODI** = Oswestry disability index, **OVCF** = osteoporotic vertebral compression fracture, **PVP** = percutaneous vertebroplasty, **QoL** = quality of life, **RCT** = randomised controlled trial; **SF-36** = short form 36 questionnaire, **TCM** = traditional Chinese medicine, **VAS** = visual analogue scale, **WHO** = World Health Organization.

17.8 Appendix H: Review of PVP and PBK Economic Studies

Table 132 lists the publications included in the review of economic evaluations of PVP and PBK. Many identified titles referred to PVP and PBK but did not provide economic analysis and were subsequently excluded.

Table 132 Overview of existing, relevant economic evaluations of PVP

Study	Method	Relevance
PVP		
MSAC (2011)[13]	A budget impact analysis was conducted to review interim publicly funded access to the procedure, for patients with painful OVCF. The committee concluded, "*future use of vertebroplasty is uncertain and contingent upon several factors: the increase in demand for vertebroplasty due to the increase in vertebral fractures in an ageing community; the increase of vertebroplasty provision in Australian States currently offering low levels of service; and, the extent to which kyphoplasty, if funded, will replace vertebroplasty as a treatment for painful vertebral fractures*" (p.17).	Cost-effectiveness analysis was not undertaken as "*the balance of benefit and harm in the evidence base does not favour vertebroplasty relative to conservative management in the treatment of osteoporotic patients with painful vertebral fracture. The effectiveness of kyphoplasty relative to conservative management in osteoporotic patients with painful vertebral fracture could not be established without a placebo control"* (p. 17).
Klazen 2010[132]	The authors undertook a cost-utility analysis of PVP using results of VERTOS II. Patients undertook EQ-5D across the 1 year follow-up and medical costs were calculated.	An intention-to-treat approach was used, with PVP estimated to result in an additional 0.108 QALYs with a corresponding ICER of €22,685 per additional QALY.
Masala 2008 [265]	The economic analysis was based on a retrospective study of conservative treatment vs PVP. Cost-effectiveness calculated as the average cost per patient per reduction of 1 point in VAS, ambulation, or activity of daily living scale.	At 1 week, PVP was significantly more cost effective for all measured outcomes compared to medical management ($p<0.05$), however, at 12 months no significant differences in cost-effectiveness where found.
Takura 2017[266]	The cost-effectiveness analysis used data from an open-label non-randomised single-arm study in Japan. 163 patients (76 years) were followed up over 52 weeks. Health-related QoL was measured using EQ-5D, RDQ, 8-item short form health survey, and VAS. Medical costs were taken from hospital and Japanese health insurance records.	QALY gain of 0.162 was found over 1 year. The estimated lifetime gain of QALYs was 1.421. ICER of US$2,154 was calculated.
PBK and/or PVP		
NICE Assessment Group Model (Stevenson 2014)[191]	The NICE Assessment Group developed an economic model for PVP, PBK, optimal pain management and placebo. The analysis was conducted over 50 years and allowed for refracture. Six scenarios were modelled to capture	The ICER was less than £16,000 per QALY gained for PVP and PBK. Results varied with assumptions about mortality, hospitalisation costs, adverse event rates, and the assumption of convergence for EQ-

Study	Method	Relevance
	differences in mortality and derivation of utility data. Adverse events were included in a sensitivity analysis by assuming that adverse effects led to a 0.02 QALY reduction.	5D.
Johnson and Johnson's model (see Stevenson 2014)[191]	Johnson and Johnson presented a model to the Stevenson review outlined in the paper. The model included PVP, BKP, "invasive control procedure" and "non-invasive management". It took a UK NHS perspective and 1 year timeframe. Key assumptions were derived from meta-analysis of VAS scores and EQ-5D data.	An ICER of £4,392 per QALY gained was calculated for PVP compared to optimal pain management. PVP dominated PBK.
Svedbom 2013 [192]	PBK and PVP were assessed against CT in a Markov tunnel model (derived from Ström [2010]) using cost data from UK. Mortality impacts were quantified using US Medicare claims data mortality hazard ratios from Edidin (2011). EQ-5D were used for utility from the FREE[168] and VERTOS II trials.[132]	The study concluded "PBK may be a cost-effective strategy for the treatment of patients hospitalised with acute OVCF in the UK compared to non-surgical management and PVP." (p. 355). Surgery was assumed to reduce hospital stay by 6 days and results were sensitive to assumptions about mortality.
Medtronic (see Stevenson 2014)[191]	The Medtronic model followed Strom and had a UK NHS perspective. Interventions included PBK, PVP and optimal pain management. Patients entered the model at 70 years, which included 6 month cycles over a lifetime time horizon. Utility values were taken from the VERTOS II trial for PVP and FREE trial for PBK. The model included reduced mortality impacts using hazard ratios from US Medicare registry data. Recurrent vertebral fracture was included.	ICER of £2,167 per QALY gained for PVP was estimated compared with optimal pain management. PBK vs PVP was £2,510.
Strom 2010[190]	The authors developed a model for PBK compared to non-surgical management in the UK setting. Assumptions were largely taken from the FREE study.[168] It had a 6 month cycle and lifetime projection, with sub-states for recovery, refracture and fracture-related mortality. Adverse events were not included. The quality of life difference at 1 year was assumed to linearly decline over 2 years. PBK was associated with 6 fewer bed days compared to non-surgical management.	PBK was associated with QALY gains of 0.17 and cost/QALY gains at £8,800, which is below UK willingness-to-pay. Sensitivity analyses showed that the results were most sensitive to assumptions about avoided length of hospital stay and persistence of PBK-related benefits. The model was used for many of the company submissions to the NICE review.
Mehio 2011[267]	The authors reviewed hospital discharge and billing records from the Premier Perspective database for a retrospective cohort (2007–2008) across 600 hospitals in the USA. Differences in total hospital cost for PVP and PBK were assessed using analysis of covariance. Total of 3,617 patients received PVP (64% inpatient, 36% outpatient), and 8,118	Average inpatient costs were US$9,837 for PVP compared to US$13,187 for PBK (p <0.0001). Outpatient PVP costs were US$3,319 compared to US$8,100 for PBK (p <0.0001). Lower PVP costs were largely due to differences in hospital supplies and surgery.

Study	Method	Relevance
	received PBK (54% inpatient, 46% outpatient). Approximately 75% were women.	
Ontario Health Technology Assessment Series 2010[268]	The study reviewed the volumes and costs of PVP and PBK medicines in Ontario (May 2010) using data from the Ministry of Health and Long-Term Care Health Analytics Branch. Pharmacotherapy typically consisted of analgesic opiate agonists NSAIDs.	Cost differences were reported for PVP and PBK medicines
Fritzell 2011[269]	The authors undertook a cost-effectiveness analysis using data from 67 Swedish patients included in the FREE trial. Mean age 72 years in the PBK group (71% female), and mean age 75 years in the control group (78% female).	The difference in QALYs gained over 24 months was 0.085 in favour of PBK. The cost per QALY gained was in the base case calculated at €101,626.
Chen 2016[270]	The authors undertook a cost-effectiveness analysis using the clinical data of 152 patients from the 309th Hospital of PLA in China from October 2013 to July 2014. Patients who received CT (51 cases), PVP (50 cases) and PBK (51 cases) were included in the analysis.	The average hospitalisation of the PVP and PKP group was 3.4 days, while the conservative group had an average of 14 days. PVP was superior to PBK and CT.
Edidin 2012[271]	PVP and PBK cost-effectiveness was assessed using 858,978 US Medicare patients' data. Life expectancy was estimated using a Weibull survival model. Median payer costs were identified for each treatment group for up to 3 years following VCF diagnosis.	PBK was found to be cost effective. The cost per life-year gained for PBK and PVP patients was US$1,863–6,687 and US$2,452–13,543 respectively, compared to non-surgical patients. The cost per life-year gained for PBK compared with PVP ranged from US$-4,878 (cost saving) to US$2,763.
Eidt & Greiner 2009[272]	Costs for PBK were estimated from a prospective, non-randomised sample in 8 study centres across Germany 2005–2008. Data was recorded by questionnaires at baseline and by phone at follow-up. Resource usage was valued.	Total costs for the PBK patients were higher than for the NSM patients. Compared to NSM, PBK-treated patients had significantly shorter hospital stays for both the initial and follow-up hospitalisation (9.6 versus 14.7, $p < 0.001$).
Joestl 2017[273]	The authors assessed the costs of vertebral fractures over 10 years in Austria. Patients had an average age of 75.6 years, with most being women. A total of 694 were treated conservatively and 25 patients (4%) underwent surgery.	A total of 384 (57%) of the 669 conservatively managed patients were treated as outpatients, and 285 (43%) as inpatients; 47% received infusions with NSAIDs, and 5% were given opioid analgesic patches.
Takahashi 2019[274]	The authors presented a cost-effectiveness analysis of PBK and NSM. QALYs were evaluated using SF-6D. The analysis was performed across 71 matched cases using the Strom Markov model with a life-time horizon, average age 78 years and 6 month cycle length. It was assumed that all additional OVFs were treated with NSM.	PBK procedure was ¥402,988 more than NSM, and QALY at 6 month follow-up was 0.153. ICERs for 3 and 20 years were ¥4,404,158 and ¥2,416,406, respectively. PBK was found to be a cost-effective treatment in Japan.
Becker 2011[275]	The authors estimated the costs	There were no statistical differences

Study	Method	Relevance
	following CT or PBK 2002–2005 in Austria. Number of readmissions, length of hospital stays, and DRG-related costs were calculated for the surgical and CT groups.	in mortality rates, but readmissions were 1.62 times higher (p = 0.039) and length of stay 1.09 times higher (p = 0.046) in the CT group. No difference in DRG scores were found (p = 0.11).
Goz 2015[276]	The study investigated trends in utilisation of PVP and PBK 2005–2010 using the US National Inpatient Sample database (63,459 inpatient admissions included). PVP had higher mortality (0.93% vs 0.60%, p < 0.001), longer length of stay (6.78 vs 5.05 days, p < 0.0001), and lower total cost (US$42,154 vs US$46,101, p < 0.0001). PVP had a higher rate of postoperative anaemia secondary to acute bleeding and higher rate of venous thromboembolic events.	PBK was associated with lower complication rates, shorter length of stay, and a higher total direct cost compared with PVP. Utilisation rates showed a significant decrease since 2009 in both PVP and PBK. Patients undergoing PVP were on average older (76.7 vs. 77.8 years, p < 0.0001) and had more comorbidities.
Lange 2014[277]	The authors examined survival and treatment costs from a third party-payer perspective for PVP and CT patients in Germany. Claims data from a major health insurance fund were used. Mortality risk differences between operated (PBK, PVP) and CT cohorts were assessed by Cox regression.	The surgical cohort was 43% less likely to suffer mortality (hazard ratio = 0.57; p < 0.001). Painkiller consumption varied for PVP: €3,321 vs PBK: €2,224.

Abbreviations

CT = conservative therapy, **DRG** = diagnosis-related group, **EQ-5D** = EuroQol 5 dimension questionnaire, **ICER** = incremental cost-effectiveness ratio, **NHS** = National Health Service, **NICE** = National Institute for Health and Care Excellence, **NSAID** = Non-Steroidal anti-inflammatory drug, **NSM** = non-surgical management, **ODB** = Ontario drug benefit, **OVCF** = osteoporotic vertebral compression fracture, **PBK** = percutaneous balloon kyphoplasty, **PVP** = percutaneous vertebroplasty, **QALY** = quality-adjusted life year, **QoL** = quality of life, **RDQ** = Roland-Morris disability questionnaire, **SF-6D** = short form 6 dimension questionnaire, **USD** = United States dollar, **VAS** = visual analogue scale, **VCF** = vertebral compression fracture.

17.9 Appendix I: Swiss Vertebral Fracture Estimates, 2010, 2015, 2020 and 2025

Description	2010	2015	2020	2025
Swiss population (thousands)[196]				
Women 50–54 years	279.6	324.4	326.0	301.7
Women 55–59 years	243.4	277.8	320.6	321.8
Women 60–64 years	233.0	237.9	272.1	313.0
Women 65–70 years	205.3	224.6	230.3	263.0
Women 70–74 years	164.3	196.0	215.0	221.0
Women 75–80 years	145.6	152.1	181.7	200.3
Women 80–84 years	119.8	125.3	132.0	159.6
Women 85+ years	123.0	140.2	160.8	180.6
Men 50–54 years	286.7	334.5	332.7	310.3
Men 55–59 years	243.5	284.0	330.3	329.1
Men 60–64 years	226.8	233.3	274.1	317.9
Men 65–70 years	191.9	210.9	220.0	258.2
Men 70–74 years	139.1	176.1	195.6	205.2
Men 75–80 years	110.4	121.6	155.8	174.6
Men 80–84 years	74.6	86.9	98.2	128.3
Men 85+ years	55.5	68.0	88.9	110.2
Vertebral fracture incidence assumptions (per 100,000)[54]				
Women 50–54 years	113	113	113	113
Women 55–59 years	76	76	76	76
Women 60–64 years	328	328	328	328
Women 65–70 years	157	157	157	157
Women 70–74 years	260	260	260	260
Women 75–80 years	721	721	721	721
Women 80–84 years	1,635	1,635	1,635	1,635
Women 85+ years	1,390	1,390	1,390	1,390
Men 50–54 years	89	89	89	89
Men 55–59 years	156	156	156	156
Men 60–64 years	262	262	262	262
Men 65–70 years	187	187	187	187
Men 70–74 years	299	299	299	299
Men 75–80 years	223	223	223	223
Men 80–84 years	799	799	799	799
Men 85+ years	482	482	482	482
Vertebral fractures				
Women 50–54 years	316	367	368	341
Women 55–59 years	185	211	244	245
Women 60–64 years	764	780	892	1,027
Women 65–70 years	322	353	362	413
Women 70–74 years	427	510	559	575
Women 75–80 years	1,050	1,097	1,310	1,444

Description	2010	2015	2020	2025
Women 80–84 years	1,959	2,049	2,158	2,609
Women 85+ years	1,710	1,949	2,235	2,510
Men 50–54 years	255	298	296	276
Men 55–59 years	380	443	515	513
Men 60–64 years	594	611	718	833
Men 65–70 years	359	394	411	483
Men 70–74 years	416	527	585	614
Men 75–80 years	246	271	347	389
Men 80–84 years	596	694	785	1,025
Men 85+ years	268	328	428	531
Total	**9,847**	**10,880**	**12,215**	**13,828**

17.10 Appendix J: List of Excluded Trials

The following list provides a summary of key excluded trials. It is not a comprehensive list owing to the number of publications screened by full text.

Wang B, Guo H, Yuan L, et al. A prospective randomized controlled study comparing the pain relief in patients with osteoporotic vertebral compression fractures with the use of vertebroplasty or facet blocking. Eur Spine J 2016;25(11):3486-94. * *only 1 study reported facet blocking, therefore, to focus the report on sham and conservative treatments, the study was excluded.*

Edidin A, Ong K, Lau E, Kurtz S. Mortality risk for operated and non-operated vertebral fracture patients in the Medicare population. J Bone Miner Res 2011;26(7): 1617-1626. *Mortality and morbidity in the Medicare population sufficiently captured by Chen 2013 and Ong 2018.*[134 147]

McCullough B, Comstock B, Deyo R, et al. Major medical outcomes with spinal augmentation vs conservative therapy. JAMA Intern Med 2013;173(16): 1514-1521. *Mortality and morbidity in the Medicare population sufficiently captured by Chen 2013 and Ong 2018.*[134 147]

Tsai Y, Hsiao F, Kao C, 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): 41-47. *Mortality and morbidity sufficiently captured by Chen 2013 and Ong 2018.* [134 147]

Edidin AA, Ong KL, Lau E, Kurtz SM. Morbidity and Mortality After Vertebral Fractures: Comparison of Vertebral Augmentation and Non-operative Management in the Medicare Population. Spine (Phila Pa 1976). 2015;40(15):1228-41. *Mortality and morbidity in the Medicare population sufficiently captured by Chen 2013 and Ong 2018.* [134 147]