

# NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE

## Overview

### **Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral fractures**

This overview is a summary of:

- the evidence and views submitted by the manufacturers, the consultees and their nominated clinical specialists and patient experts and
- the assessment report.

It highlights key issues for discussion at the first Appraisal Committee meeting and should be read with the full supporting documents for this appraisal. Please note that this document is a summary of the information available before comments on the assessment report have been received

## **Key issues for consideration**

### **Clinical effectiveness**

- Only 2 of the 9 randomised controlled trials identified were double blind and the Assessment Group noted that the quality of trials was variable. What is the Committee's view on the robustness of the evidence available from these trials?
- Only the 2 double-blind randomised controlled trials provided adequate information about the operating clinicians' training and experience, making it difficult to assess the extent to which study results may be replicable elsewhere. Does the Committee consider that the evidence available is generalisable to UK clinical practice?
- The Assessment Group noted that injection of local anaesthesia to the affected vertebral body was also considered a relevant comparator because this has been used as a 'sham' intervention in double-blind placebo-controlled trials of percutaneous vertebroplasty (PVP). Does the Committee agree that this is an appropriate comparator?

- None of the trials found any statistically significant differences in overall mortality between treatment groups. However, registry data indicate that survival is longer in patients with osteoporotic vertebral compression fractures who have vertebral augmentation than in patients who do not. What is the Committee's view on this?
- Does the Committee consider that subgroups based on baseline pain severity and time from fracture to intervention need to be considered separately?
- The Assessment Group noted that most of the complications were associated with leakage of bone cement outside the treated vertebra and that incidence of serious complications was rare, but also that the long-term implications of clinically silent cement leakages and pulmonary emboli were not clear. What is the Committee's view on the adverse reactions associated with PVP and percutaneous balloon kyphoplasty (BKP)?

### **Cost effectiveness**

- The Assessment Group stated that there was a difference in results based on whether the EQ-5D was taken directly from the trials or whether the mapping of visual analogue scale (VAS) scores from the network meta-analyses to EQ-5D was adopted. What is the Committee's view on the network meta-analyses and mapping carried out by the Assessment Group?
- The results for PVP have been estimated assuming the use of low-viscosity cement (including a component for using higher-viscosity cement). Does the Committee consider that it is appropriate to assume that it is low-viscosity cements that are mainly used in clinical practice? Does the Committee consider that any recommendation will explicitly exclude high-viscosity cement?
- The Assessment Group presented results for 6 scenarios based on differential mortality effects assumed for BKP and PVP and also based on EQ-5D data taken directly from the trials or mapping of stable VAS scores

from the network meta-analyses to EQ-5D. Which scenario does the Committee consider to be most appropriate?

- The Assessment Group highlighted that combination of all the sensitivity analyses may represent more plausible central estimates of the cost effectiveness of the interventions and should be given equal weight. What does the Committee consider the most plausible incremental cost-effectiveness ratio (ICER)?
- An exploratory analysis conducted by the Assessment Group indicated that, if all patients had a facet joint injection initially, the ICERs for BKP and PVP could be reduced by approximately a third. What is the Committee's view on this?

## **1 Background: clinical need and practice**

- 1.1 Vertebral fracture refers to a break in any of the bones (vertebrae) of the spinal column. Vertebral compression fractures usually occur when the front portion of the vertebral column is compressed. Vertebral compression fractures may be caused by trauma or a weakened vertebra, most commonly a result of osteoporosis. Osteoporotic vertebral compression fractures can be associated with curvature of the spine and loss of height and can result in pain, breathing difficulties, gastrointestinal problems and difficulties in performing activities of daily living. Osteoporotic vertebral compression fractures can also interfere with sleep, and people with these fractures can have side effects from high doses of analgesics. The symptoms of osteoporotic vertebral compression fractures can lead to deterioration in quality of life and loss of self-esteem.
- 1.2 The prevalence of vertebral fracture increases with age and is more likely in women. It is estimated that more than 2 million women in England and Wales have osteoporosis. Prevalence of osteoporotic vertebral compression fractures is difficult to estimate because not

all fractures come to the attention of clinicians and they are sometimes overlooked on X-rays. Estimates presented in the manufacturers' submissions ranged from 7073 people per year in England identified as potential candidates for vertebral augmentation to 27,051 people per year in England and Wales admitted to hospital for osteoporotic vertebral compression fractures, vertebral fatigue or collapsed fractures. The Assessment Group commented that the figure of 7073 (estimated using data from Dr Foster Intelligence) appeared to be more relevant to the decision problem because it excludes diagnoses other than osteoporosis, for example malignancy or trauma. Clinically evident osteoporotic vertebral compression fractures are associated with an increase in mortality.

- 1.3 The general aim of treatment is to restore mobility, reduce pain, and minimise the incidence of new vertebral compression fractures. Non-invasive treatment (such as medication for pain relief, bed rest and the use of back braces) for vertebral compression fractures is focused on the alleviation of symptoms and spinal support. Surgery is rarely indicated, but may be considered in people whose condition is refractory to medical therapy and in whom there is continued vertebral collapse and severe pain. NICE interventional procedure guidance supports the use of PVP (IPG 12) and BKP (IPG 166) as options for the treatment of vertebral fractures. The guidance notes that these procedures should only be carried out after prior discussion with a specialist multidisciplinary team and in an appropriately resourced facility, which has access to a spinal surgery service. For PVP, the guidance also states that the procedure should be limited to people whose pain is refractory to more conservative treatment.

## 2 The technologies

### **Percutaneous vertebroplasty**

- 2.1 PVP involves the injection of bone cement, typically polymethylmethacrylate (PMMA), into the vertebral body (the large, cylindrical part of the vertebra). It can be performed with the patient under sedation (usually a local anaesthetic) and with an analgesic. PVP may be used to provide pain relief for people with painful osteoporotic vertebral compression fractures and to strengthen the bone to prevent future fractures. The procedure does not directly restore vertebral body height.
- 2.2 Several bone cements are available for carrying out PVP. The high-viscosity Confidence Spinal Cement System (Johnson and Johnson) is indicated for the fixation of pathological fractures of the vertebral body during PVP or BKP procedures, and different costing options are available based on the number of vertebral levels to be treated. The weighted average cost of the kit is £1472. Low-viscosity cements are available and, based on list prices provided to NICE by 3 manufacturers (Cook, Orthovita and Stryker), the Assessment Group estimated a cost of £800 per low-viscosity cement PVP procedure.

### **Percutaneous balloon kyphoplasty**

- 2.3 BKP is a variation of PVP. It involves the insertion of a balloon-like device (inflatable bone tamps) into the vertebral body. The balloon is then slowly inflated until the normal height of the vertebral body is restored or the balloon reaches its highest achievable volume. When the balloon is deflated, the space is filled with bone cement. Vertebral body stents can also be inserted before the cement is added. During the procedure, which can potentially restore vertebral body height and reduce curvature of the spine, the patient is anaesthetised (either by local or general anaesthetic).

- 2.4 The KYPHON BKP kit (Medtronic) is available in the UK for BKP. KYPHON BKP is a CE-marked, single-use sterile pack with a list price of £2600.50. The kit includes 2 Kyphon Xpander inflatable bone tamps. The bone cement included in the kit, Kyphon ActivOs bone cement with hydroxyapatite, is a PMMA cement to which hydroxyapatite (a calcium compound believed to promote osseointegration) has been added. Two alternative cements for use in kyphoplasty are Kyphon KyphOs FS calcium phosphate bone substitute and Kyphon HV-R bone cement.

**Percutaneous balloon kyphoplasty with stenting (stentoplasty)**

- 2.5 BKP with stenting involves the insertion of a small balloon catheter surrounded by a metal stent into the vertebral body using radiographic guidance and either local or general anaesthetic. The balloon catheter is inflated with liquid under pressure to create a space in which the stent is expanded. The balloon catheter is then deflated and withdrawn, but the stent is left in position within the vertebra and maintains the height of the cavity into which high-viscosity PMMA bone cement is then injected. It is intended to prevent loss of height after the balloon is deflated but before the cement is injected, which can occur with percutaneous balloon kyphoplasty without stenting. The use of a vertebral body balloon, an optional site preparation device, enables the operator to identify how much space can be created for stent expansion.
- 2.6 A vertebral body stenting system (Synthes) is available, consisting of a vertebral body stent catheter, an inflation system, a vertebral body stenting access kit, and an optional vertebral body balloon catheter. Commercial-in-confidence prices were provided by Synthes: 1 level total cost of [REDACTED] and 2 level total cost of [REDACTED]. The manufacturer stated that vertebral body stenting is relatively new on the market and therefore there is limited clinical evidence

available in the form of 5 case series. The Assessment Group did not consider stentoplasty because of a lack of robust evidence.

### **Adverse reactions**

- 2.7 For both PVP and BKP, adverse reactions relate to: insertion of a needle (such as infection, venous bleeding and damage to neural or other structures); complications related to the leakage of bone cement or the displacement of bone marrow and other material by the cement; systemic reactions to the bone cement (such as hypotension and death); complications related to patient positioning and anaesthesia (such as fracture of the rib or sternum in patients with severe osteoporosis); and systemic infection. In addition, BKP can be associated with balloon rupture, and stentoplasty with a greater risk of procedure-generated adjacent fractures.

## **3 Remit and decision problem**

- 3.1 The remit from the Department of Health for this appraisal was: to appraise the clinical and cost effectiveness of PVP and BKP (with or without vertebral body stenting) for the treatment of osteoporotic vertebral fractures.

	Final scope issued by NICE	Additional comments or specifications in the Assessment Group's protocol
<b>Population</b>	<p>People with painful osteoporotic vertebral fractures.</p> <p>If evidence allows, consideration will be given to subgroups defined by:</p> <ul style="list-style-type: none"> <li>• time between fracture and treatment</li> <li>• people with and without fracture-related deformity before treatment</li> <li>• people who received inpatient care before treatment and people who did not.</li> </ul> <p>People with malignancy-related vertebral fractures and people with neuropathy in the absence of osteoporotic compression fractures are outside the scope of this appraisal.</p>	No additional comments or specifications
<b>Intervention</b>	<p>Percutaneous vertebroplasty.</p> <p>Percutaneous balloon kyphoplasty with or without vertebral body stenting.</p>	<p>Percutaneous vertebroplasty.</p> <p>Percutaneous balloon kyphoplasty without vertebral body stenting.</p> <p>Percutaneous balloon kyphoplasty with vertebral body stenting was not considered due to lack of clinical evidence (see section 2.6)</p>
<b>Comparators</b>	<p>The interventions should be compared with each other.</p> <p>Non-invasive management (without the use of either intervention).</p> <p>Assessment of evidence should also include consideration of clinical trials in which the sham procedure was performed.</p>	<p>The interventions should be compared with each other.</p> <p>Non-invasive management (without the use of either intervention).</p> <p>Operative placebo with local anaesthesia (OPLA).</p>

The Assessment Group noted that non-invasive management could include 'no treatment' in people in whom the relevant active comparators were not tolerated. Injection of local anaesthesia to the affected vertebral body was also considered a relevant comparator because this has been used as a 'sham' intervention in double-blind placebo-controlled trials of PVP. The



Assessment Group's clinical specialist suggested that administration of local anaesthesia with facet joint injection was routinely offered in the UK as a minimally invasive intervention before considering patients for vertebral augmentation. The assessment report described these procedures as 'operative placebo with local anaesthesia' (OPLA) rather than sham procedures because of ongoing debate as to whether these procedures actually constitute a sham intervention.

	Final scope issued by NICE	Additional comments or specifications in the Assessment Group's protocol
<b>Outcomes</b>	<ul style="list-style-type: none"> <li>• pain</li> <li>• functional status/mobility</li> <li>• vertebral body height and angular deformity</li> <li>• progression of treated fracture</li> <li>• rate of new vertebral fractures</li> <li>• mortality</li> <li>• adverse effects of treatment</li> <li>• health-related quality of life</li> </ul>	<p><u>Primary outcomes</u></p> <ul style="list-style-type: none"> <li>• health-related quality of life</li> <li>• back-specific functional status/mobility</li> <li>• pain/analgesic use</li> <li>• vertebral body height and angular deformity</li> <li>• incidence of new vertebral fractures</li> <li>• progression of treated fracture</li> </ul> <p><u>Secondary outcomes</u></p> <ul style="list-style-type: none"> <li>• all-cause mortality</li> <li>• symptomatic and asymptomatic leakage of cement (for example, into adjacent intervertebral discs)</li> <li>• periprocedural balloon rupture</li> <li>• postoperative complications (including infection)</li> <li>• other adverse events</li> </ul>
<b>Economic evaluation</b>	<p>The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per quality-adjusted life year.</p> <p>The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.</p> <p>Costs will be considered from an NHS and personal social services perspective.</p>	No additional comments or specifications

## 4 Clinical-effectiveness evidence

4.1 The Assessment Group carried out a systematic review and identified 9 randomised controlled trials that met the inclusion

criteria. The term optimal pain management (OPM) was adopted to encompass comparators such as optimum pain medication, conservative treatment and non-surgical management included in the trials. Two studies (Buchbinder, n=78; INVEST, multinational, n=131) compared PVP with operative placebo with local anaesthesia. Five studies (Farrokhi, n=82; VERTOS, n=46; VERTOS II, n=202; Blasco, n=125; Rousing, n=50) compared PVP with OPM. One study (FREE, multinational, n=300) compared BKP with OPM and 1 study (Liu, n=100) compared PVP with BKP. Table 1 lists the trials and their outcomes.

**Table 1 Outcomes reported in the included randomised controlled trials**

Study	Outcomes reported
Blasco 2012	Quality of life (QALEFFO) at baseline, 2 weeks, and 2, 6 and 12 months Pain at baseline, 2 weeks, and 2, 6 and 12 months Analgesic use at baseline, 2 weeks, and 2, 6 and 12 months Symptomatic vertebral fractures
Buchbinder 2009	Average pain during 24-hour period; pain at rest, and pain in bed at night (11-point VAS at 1 week, 1, 3, and 6 months) Quality of life (AQoL, QUALEFFO, EQ-5D) at 1 week, and 1, 3, and 6 months Back pain-related disability (modified RDQ score at 1 week, 1, 1, 3, and 6 months) Patients' perception of pain at 1 week, 1, 3, and 6 months Opioid use at 1 week, and 1, 3, and 6 months Adverse events at 1 week, and 1, 3, and 6 months
Farrokhi 2011	Average pain during 24-hour period (Huskisson's 10-point scale at 1 week and 2, 6, 12, 24, and 36 months) Functional quality of life (non-validated Persian translation of Oswestry Disability Index at 1 week and 2, 6, 12, 24, and 36 months) Vertebral body height (measured radiographically at 2, 6, 12, 24, and 36 months) Sagittal index (measured radiographically at 2, 6, 12, 24, and 36 months) Mobility on day 1 after start of intervention Cement leakage Adverse events
FREE	Quality of life (SF-36 PCS and EQ-5D at baseline, 1, 3, 6, 12 and 24 months) Function (RDQ score at baseline, 1, 3, 6, 12 and 24 months) Non-pharmacological therapies at baseline, 1 and 12 months Pain (11-point scale and analgesic use at baseline, 1, 12 and 24 months) Changes in spinal deformity (postoperatively and at 24 months) Patient satisfaction at 24 months Days of restricted activity at 1, 12, and 24 months Incident fractures at 12 and 24 months Procedural safety and other adverse events at 12

	and 24 months
INVEST	Back pain-related disability (modified RDQ) Average pain during 24-hour period (11-point scale, modified Deyo-Patrick Pain Frequency and Bothersomeness Scale) Opioid use Quality of life (SF-36 PCS and MCS, EQ-5D) Functional status (SOF-ADL)
Liu 2010	Pain on a 10-point scale at 3 days and 6 months Postoperative vertebral body height Postoperative kyphotic wedge angle Adjacent fractures
Rousing 2009	Pain on a 10 cm VAS at 12–24 hours, 3 and 12 months Quality of life (SF-36 PCS and MCS at 3 and 12 months; also EQ-5D, Barthel Index, and MMSE in subgroup only) Effect of pain on daily life (Dallas Pain Questionnaire) Function (objective functionality tests – tandem test, timed “Up & Go” test, repeated chair test – at 3 and 12 months, in subgroup only) Incident fractures at 3 and 12 months Intraoperative cement leakage
VERTOS	Back pain recorded on an 11-point scale 1 day and 2 weeks after vertebroplasty or initiation of optimal pain medication Analgesic use score 1 day and 2 weeks after vertebroplasty or initiation of optimal pain medication Quality of life (QUALEFFO completed 2 weeks after vertebroplasty or initiation of optimal pain medication) Back pain-related disability (RDQ completed 2 weeks after vertebroplasty or initiation of optimal pain medication)
VERTOS II	Pain (11-point scale at 1 day, 1 week, and 1, 3, 6 and 12 months; analgesic use at 1 day, 1 week and 1 month) Quality of life (QUALEFFO and EQ-5D) Back pain-related disability (RDQ) Secondary fractures (X ray at 1, 3 and 12 months) Vertebral body height loss ‘during follow-up’
AQoL, Assessment of Quality of Life; MCS, mental component summary; MMSE, mini-mental state examination; PCS, physical component summary; QUALEFFO, Quality of life questionnaire of the European Foundation for Osteoporosis; RDQ, Roland-Morris Disability Questionnaire; SF-36, Short-Form Health Survey with 36 questions; SOF-ADL, Study of Osteoporotic Fractures and Activities of Daily Living; VAS, visual analogue score	

- 4.2 The Assessment Group highlighted that only the Buchbinder and INVEST studies were double blind, which could result in risk of bias in the remaining studies. In addition, the FREE study was at risk of bias because of inclusion of less than 80% of patients in the final analysis, an unexpected imbalance in drop-outs, and selective reporting of outcomes. The studies comparing PVP with OPM (Blasco, Farrokhi, Rousing, VERTOS, and VERTOS II) varied in quality, with the Farrokhi study being least at risk of bias. The only study to compare PVP with BKP (Liu et al.) was poorly reported and potentially at risk of bias from a number of sources. It also appeared to be underpowered to identify statistically significant differences in effectiveness between the 2 interventions.
- 4.3 Only the Blasco, FREE, and INVEST studies appeared to be adequately powered for their primary outcomes. The Assessment Group stated that, because most studies were underpowered for most outcomes, the absence of a statistically significant treatment effect should not necessarily be taken as evidence that no such difference exists. In addition, the generalisability of the studies was questioned because only the Buchbinder and INVEST studies provided adequate information about the operating clinicians' training and experience, making it difficult to assess the extent to which study results may be replicable elsewhere. There was also some potential for bias because of crossover in the Blasco, FREE and VERTOS II studies. The results for the primary and secondary outcomes are discussed below.

#### **Percutaneous vertebroplasty compared with operative placebo with local anaesthesia (OPLA)**

- 4.4 The Buchbinder and INVEST studies reported pain measured on either a numeric rating scale or VAS, with higher scores indicating more severe pain. The Buchbinder study reported no statistically

significant differences in the change from baseline between groups at 1 week, or at 1, 3 or 6 months in the primary outcome of overall pain, with an adjusted mean difference of 0.7 (95% confidence interval [CI] -0.4 to 1.8) at 1 week, -0.5 (95% CI -1.7 to 0.8) at 1 month, -0.6 (95%CI -1.8 to 0.7) at 3 months and -0.1 (95% CI -1.4 to 1.2) at 6 months. The INVEST study reported no statistically significant differences in the change from baseline between groups for overall pain at 3 days, 1 week and 1 months, with an adjusted mean difference of 0.4 (95% CI -0.5 to 1.5,  $p=0.37$ ) at 3 days, 0.1 (95% CI -0.8 to 1.1,  $p=0.77$ ) at 1 week, and -0.7 (95%CI -1.7 to 0.3,  $p=0.19$ ) at 1 month. The Assessment Group stated that in the INVEST study, 64% of patients randomised to PVP and 48% of patients randomised to OPLA reported a clinically meaningful improvement in pain (that is, a decrease of 30% or more) at 1 month, but these results were not statistically significant ( $p=0.06$ ). In addition, a meta-analysis of the individual patient data from both studies also found no statistically significant improvement in change from baseline between groups at 1 month, with an adjusted mean difference of -0.6 (95% CI -1.4 to 0.2). Also, the number of patients taking opioids for pain decreased over time in both the PVP and OPLA groups in both studies. In the meta-analysis of individual patient data from the Buchbinder and INVEST studies, after adjusting for baseline opioid use, patients randomised to PVP were statistically significantly more likely to be taking opioids at 1 month than patients randomised to placebo (relative risk [RR] 1.25, 95% CI 1.14 to 1.36,  $p<0.001$ ). Therefore the trend towards a higher proportion of patients in the PVP group achieving an improvement of 30% or more in pain scores at 1 month, may have been influenced by the fact that the PVP group was more likely than the placebo group to be using opioids at that point.

4.5 The Buchbinder study also reported pain outcomes in terms of QUALEFFO (Quality of Life Questionnaire of the European Foundation for Osteoporosis) pain scores and found no statistically significant differences between groups. Data were also collected on perceived pain: this was classified as 'better' if pain was moderately or substantially better than before the intervention, and 'worse' if it was moderately or substantially worse. There were no statistically significant differences in the proportion of patients in these categories at any time point. The INVEST study also reported on the frequency with which patients experienced pain, and the impact of pain on their daily lives. For PVP and OPLA, both pain frequency and pain 'bothersomeness' decreased between baseline and 1 month with point estimates favouring PVP. However, these results were not statistically significant.

4.6 The Buchbinder study presented health-related quality-of-life results based on AQoL (Assessment of Quality of Life), EQ-5D and QUALEFFO measures. The INVEST study presented health-related quality-of-life results based on EQ-5D and SF-36. No difference was found between the PVP and OPLA groups based on AQoL scores. EQ-5D scores in the Buchbinder study were only available for 79% of patients in the PVP group and 73% in the OPLA group. Both studies found no statistically significant difference between PVP and OPLA in terms of short- or medium-term outcomes. This result was supported by the meta-analysis of individual patient data at 1 month, which indicated that the result was not statistically significant (adjusted mean difference 0.03, 95% CI -0.02 to 0.08). The Assessment Group highlighted that, because the minimum clinically important difference for back pain on the EQ-5D scale is 0.08, the confidence interval for the pooled data only just included the possibility of a clinically important difference favouring PVP. Based on QUALEFFO scores in the Buchbinder study, the only statistically significant result was at



1 week with an adjusted mean difference of  $-4.0$  (95% CI  $-7.8$  to  $-0.2$ ). The Assessment Group stated that, because no minimum clinically important difference had been proposed for the QUALEFFO, the clinical significance of this result was not clear. The INVEST study found no statistically significant differences between treatment groups at any point using SF-36 scores.

- 4.7 Both the INVEST and Buchbinder studies assessed back-specific functional status using the modified 23-point version of the Roland-Morris Disability Questionnaire (RDQ). However, neither study found any statistically significant differences for short-term (3 days to 2 weeks) or medium-term (1 month to 6 months) outcomes. In addition, the meta-analysis of individual patient data from both studies indicated no statistically significant difference between treatment groups at 1 month in terms of mean RDQ scores with an adjusted mean difference of  $-0.8$  (95% CI  $-0.9$  to  $2.4$ ). The INVEST study included a post-hoc analysis to identify the proportion of patients who achieved a clinically meaningful improvement in physical disability related to back pain at 1 month. This improvement was not defined, but was presumably measured in terms of a reduction in the RDQ score. There was no statistically significant difference between the proportion of patients in each group who achieved a clinically meaningful improvement (40% of the PVP group and 41% of the OPLA group,  $p=0.99$ ). Meta-analysis of the individual patient data from both studies also found no statistically significant difference in the proportion showing an improvement of at least 3 units or of at least 30% in RDQ scores. The INVEST study also reported mean SOF-ADL scores at baseline and 1 month, with no statistically significant difference between treatment groups in change from baseline with an adjusted mean difference of  $0.4$  (95% CI  $-0.8$  to  $1.6$ ,  $p=0.51$ ).

### **Percutaneous vertebroplasty compared with optimal pain management**

**(OPM)**

4.8 All studies comparing PVP with OPM reported pain measured on a numeric rating scale or VAS and the results are presented in table 2. The Farrokhi, Rousing and VERTOS II studies found statistically significant improvements between groups in short- and medium-term changes from baseline in pain following PVP. However, the Assessment Group highlighted that the favourable result reported by Rousing at 1 month was unreliable because of the high risk of recall bias because these data were collected almost a year after the event. The VERTOS II and Farrokhi studies also found statistically significant long-term improvements in the change from baseline between groups. However, the Assessment Group noted that in the VERTOS II study, at a minimum clinically important improvement of 2 or more points, the 95% CI included the possibility that the results were not clinically meaningful. In the study by Blasco, statistical significance in change from baseline was only reported at 2 months, when the result favoured PVP.

4.9 Results for the total number of patients taking opioids indicated that there were no statistically significant between-group differences in the relative risks of taking opioids, other than at baseline. However, the Assessment Group highlighted that the results from this study were difficult to interpret. This was because even though a statistically significantly higher proportion of patients in the PVP group needed opioids at baseline, the proportion of patients needing opioids fell noticeably from baseline to 2 weeks, and then gradually thereafter. However, in the control group this proportion rose steeply at 2 weeks and remained elevated for 6 months, then fell substantially at 12 months. In addition, in both treatment groups, the number of patients needing opioid analgesia at 12 months was smaller than the number for whom data were missing. In the VERTOS study there was no statistically significant between-group difference in the use of pain medications at

baseline. However, at both 1 day and 2 weeks, the mean analgesic use score had reduced in the PVP group and increased in the control group, resulting in statistically significant differences that favoured PVP. In VERTOS II, analgesic use was said to be statistically significantly reduced in the PVP group compared with the control group at 1 day, 1 week, and 1 month ( $p < 0.0001$ ,  $p < 0.001$ , and  $p = 0.033$  respectively), but not at later stages of follow-up. However, the Assessment Group highlighted that the actual figures were not presented.

**Table 2. Adjusted mean difference in overall pain scores after percutaneous vertebroplasty for treating osteoporotic vertebral fractures. Statistically significant figures shown in bold.**

Study	Time (change from baseline)	Adjusted Mean difference in pain scores, percutaneous vertebroplasty compared with optimal pain management (95% CI). Negative values favour PVP.
Blasco 2012	2 weeks	0.18 (-0.95 to 1.31)
	2 months	<b>-1.48 (-2.94 to -0.02)</b>
	6 months	-0.48 (-1.84 to 0.88)
	12 months	-0.73 (-2.10 to 0.64)
Farrokhi 2011	1 week	<b>-4.3 (-5.11 to -3.49)</b>
	2 months	<b>-4.1 (-5.28 to -2.92)</b>
	6 months	<b>-3.1 (-4.17 to -2.03)</b>
	36 months	<b>-2.1 (-3.36 to -0.87)</b>
Rousing 2009	12–24 hours	NR
	1 month	<b>-1.7 (-2.21 to -1.19)</b>
	3 months	<b>-0.5 (-0.05 to -0.95)</b>
	12 months	0.4 (-0.03 to 0.83)
VERTOS	2 weeks	<b>-1.0 (-0.5 to -2.5)</b>
VERTOS II	1 week	<b>-2.4 (-3.11 to -1.70)</b>
	1 month	<b>-2.6 (-3.37 to -1.74);</b> $p < 0.0001$
	6 months	<b>-1.9 (-2.84 to -0.97)</b>
	12 months	<b>-2.0 (-2.80 to -1.13);</b> $p < 0.0001$

- 4.10 The Farrokhi and Blasco studies reported changes in vertebral body height and angular deformity. However, their results are not comparable because it is not clear whether they used the same methods of measuring vertebral height. The Blasco study found no statistically significant difference between treatment groups in change in vertebral body height from baseline at 12 months. In contrast, the Farrokhi study found that PVP was associated with statistically significant improvements in mean vertebral body height that were sustained throughout the first year but not thereafter, and with statistically significant improvements in angular deformity that were sustained throughout the 36-month follow-up period. It was suggested that these results may be related either to patients being in the prone position used during PVP or to the high pressure produced within the vertebra by the injected cement, both of which can expand the vertebra and correct kyphotic deformity to some extent. The VERTOS II study reported data relating to the progression of treated fractures during follow-up. At the last follow-up (mean 11.4 months, median 12.0 months, range 1–24 months), statistically significant moderate or severe height loss was seen in 11 vertebrae in 12% of patients in the PVP group, compared with 39 vertebrae in 41% of patients in the control group ( $p < 0.001$ ).
- 4.11 The Rousing study assessed health-related quality of life using the Dallas Pain Questionnaire, which is designed to evaluate the impact of chronic pain on a patient's life. Only the score for work and leisure at 3 months reached statistical significance, favouring PVP compared with OPM. However the Assessment Group noted that this was the unadjusted score. In each domain, baseline scores were noticeably lower in the PVP group than in the OPM group. When this was adjusted for by comparing changes from baseline in each group rather than crude scores, all point estimates favoured conservative management whereas previously all except those for social interest at 12 months had favoured PVP. These

differences were statistically significant for all outcomes except work and leisure at 3 months and anxiety and depression at both 3 and 12 months.

- 4.12 The Rousing and VERTOS II studies assessed health-related quality of life using EQ-5D. In the Rousing study scores were only available for 58% in the PVP group and 71% in the OPM group. Results indicated an adjusted mean difference of  $-0.085$  (95% CI  $-0.15$  to  $-0.02$ ) at 3 months and  $-0.169$  (95% CI  $-0.23$  to  $-0.11$ ) at 12 months. The VERTOS II collected EQ-5D data throughout the study but only reported baseline values. Therefore follow-up values were not available.
- 4.13 The Blasco, VERTOS and VERTOS II studies assessed health-related quality of life using the QUALEFFO. Results from the Blasco study indicated that there was a non-statistically significant improvement in scores with PVP compared with OPM at all time points in the short- and medium-term. The VERTOS study found that PVP was associated with better short-term total QUALEFFO score than OPM. In VERTOS II, after adjusting for baseline differences, there was a statistically significant difference in QUALEFFO scores at 1 year that favoured PVP ( $p < 0.0001$ ); however, the actual figures were not reported. The Rousing study assessed health-related quality of life using SF-36 and reported no statistically significant differences between treatment groups at any point.
- 4.14 The VERTOS and VERTOS II studies assessed back-specific functional status using the original 24-point version of the RDQ. The VERTOS study reported that the between-group difference in change from baseline at 2 weeks favoured PVP compared with OPM, but the statistical significance could not be calculated because neither standard deviation nor standard error was reported. The VERTOS II study reported a statistically significant

difference that favoured PVP at 1 year compared with OPM ( $p < 0.0001$ ); however, the actual numbers were not available and its clinical importance was not indicated.

4.15 The Farrokhi study used a modified Oswestry Disability Index (ODI) and reported that PVP was associated with a statistically significant improvement in change from baseline at all times from 1 week to 36 months compared with OPM. The Assessment Group noted that, because the minimum clinically important difference for back pain on the ODI was 4 points, these differences were clinically meaningful throughout. The Farrokhi study also noted that all 40 patients in the PVP group could walk 1 day after PVP, but only 2% of patients in the OPM group could walk 1 day after OPM, indicating a relative risk of 28.32 (95% CI 5.88 to 136.45,  $p < 0.0001$ ).

4.16 The Rousing study reported functional outcomes using the Barthel Index, using the version scored from 0 to 20, with lower scores indicating greater disability. Data were only available for a subset of the study population. At 12 months, the absolute score was statistically significantly better in the PVP group than in the OPM group. However, once the difference in baseline scores was taken into account, the difference between groups was no longer statistically significant. The Assessment Group stated that, because the baseline measurement is relatively high, this result may indicate that there is a ceiling effect whereby there is little scope for PVP to improve functional outcome more than OPM. The Rousing study also reported 3 observer-assessed tests of physical function for a subset of the population: the tandem test, timed up and go test, and repeated chair test. No statistically significant differences between groups were noted at 3 or 12 months but, because baseline values were not reported, the clinical meaningfulness of this result in terms of change from baseline was not clear.

### **Percutaneous vertebroplasty compared with percutaneous balloon kyphoplasty**

4.17 The Liu study was a prospective, comparative randomised study that assessed pain before and after treatment, and vertebral body height and angular deformity. No other outcomes, including health-related quality of life, were included. It assessed pain measured on a VAS and reported no statistically significant differences between PVP and BKP in the short or medium term with an adjusted mean difference of  $-0.2$  (95% CI  $-0.43$  to  $0.03$ ) at 3 days and  $0.1$  (95% CI  $-0.28$  to  $0.48$ ) at 6 months. However, the Assessment Group highlighted that it did not appear to have been powered to do so. The study also assessed changes in vertebral body height and angular deformity, reporting that BKP was associated with statistically significant greater improvements in both postoperative vertebral body height and angular deformity than PVP. However, it was unclear whether it measured postoperative vertebral body height and angular deformity at 3 days or at 6 months.

### **Percutaneous balloon kyphoplasty compared with optimal pain management**

4.18 The FREE study reported pain using the recommended measure of global pain severity, the bodily pain subscale of the SF-36, in which higher scores represent better health. Results indicated that the difference between treatment groups in average improvement over a period of 12 months was 9.2 points, statistically significantly greater with BKP than with OPM (95% CI 3.9 to 14.6,  $p=0.0008$ ). The FREE study also reported pain measured on a numeric rating scale and reported statistically significant long-term differences between groups. However, the Assessment Group noted that these differences did not appear to be clinically important because a difference between groups of 2 or more points indicated a clinically meaningful difference. The results are presented in table 3.

**Table 3: Adjusted mean difference in overall pain scores after percutaneous percutaneous balloon kyphoplasty for treating osteoporotic vertebral fractures. Statistically significant figures shown in bold.**

Study	Time (change from baseline)	Adjusted mean difference, balloon kyphoplasty compared with optimal pain management (95% CI)
FREE	1 week	<b>-2.2 (-1.6 to -2.8)</b>
	1 month	<b>-1.9 (-2.5 to -1.3)</b> , p<0.0001
	3 months	-1.57
	6 months	-1.61
	12 months	<b>-0.9 (-1.5 to -0.3)</b> , p=0.0034
	24 months	<b>-0.80 (-1.39 to -0.20)</b> , p=0.009

4.19 The FREE study reported analgesic use. It was found that BKP was associated with a statistically significantly reduced risk of needing opioid medication at 1 month and 6 months, but not at 12 or 24 months. The Assessment Group highlighted that the FREE study did not report changes in vertebral body height even though maintenance of vertebral body height was one of its secondary outcome measures and is 1 outcome for which BKP might be expected to provide additional benefit compared with PVP. The study protocol stated that vertebral body height was only to be measured in patients having BKP, making comparison with OPM impossible. The study reported a statistically significant improvement from baseline with BKP in the kyphotic angle of the index fracture at 24 months. The Assessment Group noted however that the clinical significance of this result is not clear.

4.20 The FREE study assessed health-related quality of life using EQ-5D, and SF-36. Using EQ-5D, statistically significant differences in outcomes favouring BKP over OPM were reported at 1, 12, and



24 months. However, the Assessment Group highlighted that, at a minimum clinically important difference for back pain of 0.08, the confidence intervals at 3, 6, 12, and 24 months included the possibility of clinically unimportant effects. Using the SF-36 physical component summary score, the study reported a statistically significant mean difference of 5.2 (95% CI 2.9 to 7.4,  $p < 0.0001$ ) at 1 month. While the results remained statistically significant at 3 months and 6 months, the confidence intervals included the possibility of achieving a clinically unimportant result, and after 6 months there was no statistically significant difference between treatment groups. The FREE study also reported psychological wellbeing, which was assessed by the SF-36 mental component summary score, and identified no statistically significant differences between treatment groups, although the confidence intervals included the possibility of potential clinically important treatment effects favouring BKP compared with OPM at time points up to 12 months.

- 4.21 The FREE study assessed back-specific functional status using the original 24-point version of the RDQ. It reported that BKP was associated with statistically significantly better outcomes compared with OPM at 1 and 12 months, but not at 24 months; moreover, at 12 months the confidence intervals included the possibility of failing to achieve clinical importance. The FREE study also reported that BKP was associated with a statistically significant reduction in the risk of needing walking aids at 1 month, but not at 12 months. However, the Assessment Group noted the data were not robust because, in the control group, the number of patients needing walking aids at 12 months was smaller than the number for whom data were missing. The FREE study also recorded the number of patients who reported 1 or more days of bed rest because of back pain in the previous 14 days. At 1 month, patients in the BKP group reported on average 2.9 fewer days of restricted activity because of

back pain in the previous 14 days than patients in the OPM group (95% CI 1.3 to 4.6,  $p < 0.001$ ), but at 12 months the difference was no longer statistically significant (1.6 days, 95% CI  $-0.1$  to 3.3,  $p = 0.0678$ ). The actual numbers of days of restricted activity in each group were not reported. However, the 12-month data in both groups were not robust because patients for whom data were missing outnumbered patients for whom there were data.

### Adverse events

- 4.22 The Assessment Group presented adverse events reported in the trials. Large observational studies and individual case reports were used to supplement randomised controlled trial data on adverse events.
- 4.23 Based on trial data, it was noted that none of the studies found any statistically significant differences in overall mortality between treatment groups. However, the Assessment Group noted that they were not powered for this outcome. The Assessment Group also combined data from Blasco, Rousing and VERTOS II studies comparing PVP with OPM because they reported overall mortality at the same time point (12 months). However, statistical significance was still not achieved, although the point estimate favoured PVP, with a risk ratio of 0.68 (95% CI 0.30 to 1.57,  $p = 0.37$ ).
- 4.24 All but the Liu and INVEST studies reported cement leakages identified using imaging equipment. All stated that they used a PMMA cement. The Assessment Group stated that, because none referred specifically to high-viscosity cement, it was assumed that low-viscosity cement was used in all studies. The Blasco study found that, although the cement leaks that they reported were not associated with immediate clinical complications, cement leakage into the inferior disk was associated with an increased risk of incident vertebral fracture (odds ratio [OR] 7.17, 95% CI 1.69 to

69.30),  $p=0.0008$ ). The Farrokhi study reported 13 asymptomatic leaks and 1 symptomatic leakage into the epidural space. The symptomatic leakage caused severe right lower-extremity pain and weakness but following immediate decompression through a bilateral laminectomy and evacuation of bone cement, the patient could walk unassisted with no radicular pain after 2 months. The Rousing study stated that none of the cement leaks caused neurological symptoms. In the VERTOS II study, most leakages were discal or into segmental veins, none were into the spinal canal and all patients remained asymptomatic. In this study, an asymptomatic cement deposition in a segmental pulmonary artery was also reported and 54 patients who had had PVP subsequently had a computed tomography (CT) scan after a mean follow-up of 22 months (median 21 months, range 6–42 months). Although during the procedure the operators had not reported fluoroscopically-visible cement migration towards the lungs in any of these patients, at follow-up 26% (95% CI 16% to 39%) had pulmonary cement embolism visible on CT scan. However, all the affected patients were asymptomatic. In the FREE study, most leaks were endplate or discal leakages, with 1 foraminal leakage, no leakages to the spinal canal, and no cement emboli.

- 4.25 All but the Liu and Blasco studies provided some information relating to peri- or postoperative complications. In the INVEST study, 1 patient had an injury to the thecal sac during PVP that resulted in hospitalisation. In addition, 1 patient who had received OPLA was hospitalised overnight after the procedure with tachycardia and rigors of unknown cause. In the VERTOS study, in a patient originally randomised to OPM who requested PVP after 2 weeks, an intrapedicular cement spur broke on manipulation by the bone biopsy needle and caused a small cortical chip fracture at the medial border of the pedicle. The patient recorded an increase in pain score at 1 day but the pain was relieved using analgesics

and local anaesthetic infiltration of the involved pedicle; there were no neurological sequelae. In VERTOS II, the patients needed additional intravenous analgesia in 30% of procedures; 2 patients needed atropine because of pain-induced vasovagal reactions and, in 1 case, the procedure had to be stopped because the patient developed an acute asthma exacerbation during PVP; however, the procedure was performed successfully a week later.

- 4.26 The Buchbinder, FREE and VERTOS II studies reported postoperative infections that were potentially related to treatment. In the Buchbinder study, prophylactic cephalothin was usually administered intravenously immediately after cement injection. Osteomyelitis developed in a patient who did not receive such prophylaxis because of multiple drug allergies, and surgical drainage and antibiotic treatment were needed approximately 2 weeks after randomisation, leading to full recovery. In the FREE study, a recurrent urinary tract infection was exacerbated by catheterisation; this patient also developed spondylitis near the cement in the vertebral body 376 days after surgery and the inflammation had not resolved by 24 months despite antibiotic therapy. Sepsis/septic shock was reported in 1 patient in the BKP group, but also in 3 patients in the OPM group. The Assessment Group also noted that 3 patients who underwent BKP subsequently had pulmonary embolisms and the earliest of these was at 46 days postoperatively. However, the significance of these embolisms is not discussed. In VERTOS II, 1 patient developed a urinary tract infection after PVP. The Farrokhi study stated that no infections occurred, while the Rousing study stated that there were no adverse events other than cement leaks. In the remaining 4 studies (Blasco, INVEST, Liu, VERTOS), no postoperative infections were mentioned, suggesting that none may have occurred.

4.27 Three studies (Blasco, FREE, VERTOS II) reported the number of patients who had new radiographic vertebral fractures during the study period. None of these studies found a statistically significant difference between treatment groups. However, the Assessment Group noted that the fracture incidence data may be biased in the FREE study because loss to follow-up was higher in the OPM group than in the BKP group (23% compared with 37%) and the drop-out rate outnumbered the event rate in the OPM group. The Assessment Group also performed an exploratory meta-analysis combining data from these 3 studies and stated that although the point estimate favoured OPM, statistical significance was not achieved with a risk ratio of 1.16 (95% CI 0.85 to 1.59,  $p=0.47$ ). The Assessment Group also stated that fractures in adjacent vertebrae are more likely to be associated with therapy than fractures in more distant vertebrae. It was noted that the Blasco study found that 82% of new fractures in the PVP group were adjacent to the index vertebra, compared with 27% in the OPM group (OR 16.00, 95% CI 1.03 to 835.12,  $p=0.0101$ ). The FREE study reported that 23.7% patients in the BKP group and 16.7% patients in the OPM group had a radiographic fracture adjacent to the index fracture; however, the difference was not statistically significant (RR 1.42, 95% CI 0.83 to 2.45,  $p=0.20$ ). Similarly, in VERTOS II, the risk of adjacent rather than distant fracture was not statistically significantly different in the PVP and OPM groups ( $p=0.23$ ), nor did such fractures occur statistically significantly sooner in the PVP group than in the conservative therapy group ( $4.6\pm 5.4$  compared with  $6.1\pm 5.9$  months,  $p=0.48$ ). The only risk factor for either the occurrence or the number of new fractures was the number of vertebral fractures at study entry, which is itself an indicator of the severity of osteoporosis. The Rousing study found that over 12 months there were more radiographic fractures in the PVP group than in the OPM group (7 compared with 4, statistical

significance not reported). The relevant results from the Buchbinder and INVEST trials have not yet been published.

- 4.28 The Assessment Group noted that the most meaningful fracture outcome measure was the proportion of patients who experienced at least 1 clinically important fracture in an adjacent vertebra. Five studies (Buchbinder, Farrokhi, FREE, Rousing and VERTOS) reported the overall incidence of new clinical vertebral fractures, but the VERTOS study did so only for the PVP group. None of the studies that reported this outcome in both treatment groups identified a statistically significant difference between treatment groups. The Blasco study stated that 71% of the radiographic fractures in the PVP group were clinical but, compared with 9% in the OPM group (OR 25.67, 95% CI 3.04 to 216.8,  $p=0.029$ ), the number of patients who had clinical vertebral fractures were not reported.
- 4.29 The Assessment Group stated that randomised controlled trials may not detect long-term or rare adverse events, so large case series ( $n>200$ ) and individual case reports were also examined to gain an estimate of the incidence of more common adverse events from large cohorts, as well as rarer but serious events, which are published as individual case reports. The Assessment Group did not identify any publications of registry data that were specific to patients with osteoporotic vertebral compression fractures, but the Medtronic submission included academic-in-confidence claims-based data relating to the US Medicare population from 2005–09 and to subscribers to a major German health insurance fund from 2005–10. These reports compared mortality and complication risks in patients with osteoporotic vertebral compression fractures who had operations and those who did not. Further details are available on page 123 of the assessment report. The Assessment Group stated that it is unclear how generalisable these results are to

patients treated in England and Wales. In addition, 14 large case series were identified, of which 10 reported data on adverse reactions associated with PVP and BKP. Detailed results are available on page 130 of the assessment report.

4.30 Based on evidence from the trials, case series and case reports, the Assessment Group stated that treatment-related deaths appeared to be rare, but cement leakage was common, particularly with PVP. Many cement leaks were not associated with immediate clinical complications, but others were associated with serious problems such as pulmonary embolism, radiculopathy, and temporary or permanent motor deficits. Several procedure-related deaths were noted. Moreover, it was noted that there was no good evidence to prove that leaks which were asymptomatic in the short term did not have long-term implications. Peri- and postoperative complications other than cement leak appeared to be rare, though potentially serious. In particular, infectious complications were potentially fatal, and frequently needed treatment with further surgical intervention. To reduce the risk of such complications, it has been recommended in the literature (Lee et al. 2007) that PVP or BKP should not proceed until the patient has made a complete recovery from any existing infections, and that, in cases of recent infection, either antibiotics should be prescribed on a long-term basis to avoid deep infection, or a cement-antibiotic mixture should be used. It was noted that intraoperative balloon perforation during kyphoplasty seems unlikely to lead to any serious complications. In addition, it is likely that PVP and BKP may be associated with increased rates of new vertebral fractures, and in particular adjacent fractures, but the quality of the evidence was not very good.

4.31 The Assessment Group also noted that it was unclear whether PVP or BKP was safer because direct comparisons were unavailable.

However, Yang et al. (2010) conducted a review that found rates of specific complications (cement leakage, new compression fractures, pulmonary embolism, and radiculopathy) were all statistically significantly higher with PVP than with kyphoplasty (all  $p < 0.05$ ). They also found that cement leakage rates were lower in procedures carried out in neurosurgery departments (20.6%) and orthopaedic departments (24.7%) than in radiology departments (52.9%). The Assessment Group noted that this could be confounding if the vertebroplasties were carried out by the radiologists. None of the included studies referred to the radiation risks to patients associated with PVP and BKP. The Assessment Group highlighted that though these risks were low, they were not trivial.

- 4.32 Finally, the Assessment Group highlighted the potentially serious complications resulting from the comparators to PVP and BKP. Bed rest could result in muscle wasting and deconditioning, and these effects have been associated with deep vein thrombosis, pulmonary emboli, reduced muscle blood flow, red cell volume, capillarisation, and oxidative enzymes. Narcotic analgesics are associated with several undesirable side effects, including cognitive impairment and nausea, and non-steroidal anti-inflammatory drugs are associated with gastrointestinal problems. Both registry studies (US Medicare and a German health insurance fund) indicated that, even after adjusting for comorbidities, survival is longer in patients with osteoporotic vertebral compression fractures who have vertebral augmentation than in patients who do not. The Assessment Group stated that the reasons for this pattern are not clear, but it may have to do with avoiding the problems associated with conservative treatment.



## Subgroups

- 4.33 The Assessment Group stated that no data were identified relating to subgroups with and without fracture-related deformity before treatment or to subgroups relating to patients who were inpatients at the time of randomisation. However, some data were available for subgroups based on baseline pain severity and time from fracture to intervention. No subgroup data were available for BKP.
- 4.34 It was noted that the Staples et al. analyses of individual patient data from the Buchbinder and INVEST studies included patients grouped by baseline pain severity. The p-values were not reported, but no statistically significant differences in RDQ scores, EQ-5D scores, or pain scores were identified between patients with severe pain (score >8 on a 0–10 rating scale) or mild-to-moderate pain (score <8) at baseline. Please see tables 22, 23 and 24 in the assessment report for further details. In both treatment groups, the decrease in pain was greater in the subgroup that had more severe pain at baseline than in the subgroup with less severe baseline pain. The Assessment Group stated that this could simply reflect a greater potential for improvement. Based on this, the Assessment Group also stated there was no reason to suppose that outcomes would differ between patients who were inpatients before treatment and patients who were not because receipt of inpatient care may be influenced by factors other than clinical factors such as pain severity. For example, patients who are bedridden with severe pain may not be hospitalised, if they have adequate support networks in terms of both family and friends, and community services.
- 4.35 The INVEST study reported data by baseline pain duration and a post hoc subgroup analysis of the effect of treatment on pain at 1 month by baseline pain-duration categories found no statistically significant difference between PVP and OPLA. It was noted that the INVEST study was underpowered for this analysis. Therefore, a

meta-analysis of individual patient data from the INVEST and Buchbinder studies to assess the effectiveness of PVP in patients with fracture pain of recent onset (less than 6 weeks) compared with pain of longer duration was carried out. Because the INVEST study allowed crossover after 1 month, outcomes were only compared up to that time point. No statistically significant differences in RDQ scores, EQ-5D scores, or pain scores were identified between patients whose pain was of recent onset and patients whose pain duration exceeded 6 weeks. Please see tables 19, 20 and 21 in the assessment report for further details.

### **Summary**

4.36 The Assessment Group considered that the literature available was of variable quality and the most noteworthy methodological issue was lack of blinding for both patients and outcome assessors in most of the trials. It was noted that the literature suggested that both PVP and BKP provided substantially greater benefits than OPM in open label trials but, in double-blind, trials, PVP was shown to have no more benefit than local anaesthetic, and no trials of BKP compared with local anaesthesia have been conducted. Quality of life was most often assessed with the EQ-5D or QUALEFFO and findings indicated greater improvements with both these measures in the open label trials of PVP (Blasco, Rousing, Farrokhi, VERTOS, VERTOS II). However, no differences in quality of life were observed in either of the placebo-controlled, double-blind trials (Buchbinder, INVEST). Four open-label studies (Farrokhi, FREE, Rousing, VERTOS II), found statistically significantly greater improvements in pain among patients who had operations, and the double-blind trials found no or a small non-significant benefit. Although there was a trend toward greater pain reduction in the PVP group in the INVEST study, the Assessment Group noted that this may have been confounded by a higher level of opioid use among the PVP group.

- 4.37 A meta-analysis of mortality suggested that PVP might be associated with reductions in mortality. However, it was noted that this effect failed to reach statistical significance and the included trials were not designed to detect this outcome. The Assessment Group noted that it was possible that there was a causal difference in mortality between patients treated using OPM and patients receiving BKP or PVP given the size of the effect, and that appropriately taking into account the potential endogeneity of the treatment would tend to reduce the point estimate of the effect size but may or may not eliminate it completely. The Assessment Group concluded that it was not possible to say with certainty whether there was a difference in mortality between patients having BKP and PVP caused by the treatment. In addition it was noted that, if BKP and PVP were assumed to have a mortality benefit, there would be uncertainty around whether OPLA would also produce a mortality benefit.
- 4.38 The Assessment Group noted that complications of PVP and BKP include pulmonary embolism, periprocedural hypotension, radiculopathy, damage to surrounding tissue, paraparesia, paraplegia, rib fracture, and postoperative infection. It was noted that most of these complications were associated with leakage of bone cement outside the treated vertebra. Although intradiscal leakage is unlikely to lead to complications, epidural leakage can have serious consequences, and several procedure-related deaths have been reported. It was noted that the incidence of serious complications is rare, but the long-term implications of clinically silent cement leakages and pulmonary emboli were not clear.

## **5 Comments from other consultees**

- 5.1 Patient groups stated that the main issues for people with osteoporotic vertebral fractures were pain, posture, and the ability to work and care for themselves and, consequently, emotional

wellbeing and quality of life. The painful and debilitating nature of the condition was highlighted. Patient groups stated that current treatment options included analgesia, limited bed rest, physiotherapy, bracing and facet joint injection. In very severe cases, a course of subcutaneous or nasal calcitonin, or an infusion of intravenous bisphosphonates may be given. The main advantage of PVP and BKP would be to help with short- and medium-term pain relief, which would lead to an improvement in quality of life. Another important advantage would be relief from physical symptoms and levels of disability such as height loss, which has a direct impact on swallowing, digestion, self care, ability to work and emotional wellbeing. However, comments also indicated concerns around the risk of complications from the procedure, including adverse effects caused by implantation of a balloon or leakage of cement and the increased risk of new vertebral fractures in the adjacent vertebra. Comments also highlighted concerns around any unknown adverse effects, as well as worry about the condition getting worse following treatment.

- 5.2 Patient groups noted that people with severe symptoms who are not responding to standard therapy and in whom quality of life may be greatly affected were more likely to benefit, and that it was less clear whether people with persistent or chronic back pain caused by fracture were likely to benefit. It was also noted that people who cannot have local or general anaesthesia and people living in areas where there is no local service may find it difficult to use PVP and BKP.

## **6 Cost-effectiveness evidence**

- 6.1 The literature review conducted by the Assessment Group identified 1 Markov cohort model that assessed the cost effectiveness of BKP compared with OPM in patients who were hospitalised with vertebral compression fracture (Strom et al.

2010). The model simulated the experiences of patients until death or age 100, with EQ-5D scores taken directly from the FREE study. It was assumed that the EQ-5D scores would be independent of the intervention 3 years post BKP or OPM with a linear decline between 12 months and 36 months. The risks of future vertebral fracture and the risks of mortality after vertebral fracture were incorporated. The base case assumed a cohort of 70 year-old women and men with a T-score of  $-2.5SD$  (T-score is defined as the number of standard deviations from the average bone mineral density of healthy young women) and estimated that BKP would be associated with an additional cost of £1494 to obtain 0.169 quality-adjusted life years (QALYs), resulting in an ICER of £8840 per QALY gained for BKP compared with OPM. The Assessment Group highlighted that this model was updated by Medtronic to include PVP as an intervention and to incorporate the potential beneficial effect of BKP and PVP on mortality, and therefore these results have been superseded.

### **Medtronic model**

6.2 The objective of the Medtronic model was to determine the cost effectiveness of BKP compared with PVP and OPM in patients who were hospitalised with vertebral compression fracture. It was a Markov tunnel model with a lifetime time horizon and an NHS perspective was adopted. Costs and utilities were discounted at 3.5% and a time cycle of 6 months was used. In the base case, it was assumed that patients were 70 years old with a T-score of  $-3.0SD$ , in line with data from the FREE and VERTOS II trials. People remained in their initial treatment health state (progressing through the substates) until death or an additional vertebral fracture occurred. For all patients, a subsequent vertebral fracture was assumed to be treated using OPM. The transition probabilities for further vertebral fractures were calculated from equations that were a function of the patient's bone mineral density compared with that

of a young woman, age, previous fracture status and the imputed ratio between hip and vertebral fractures at each age, assuming that the Swedish ratio (from Strom et al. 2010) was applicable to the UK. The transition probabilities to death used data from the Human Mortality Database for patients in the UK and the relative risks of mortality reported in Strom et al. for people with a prior vertebral fracture.

6.3 The utility values for BKP and OPM were taken directly from the FREE trial. The utility values for PVP were estimated assuming that the difference between PVP and OPM reported in the VERTOS II trial could be directly added to the OPM scores in the FREE trial. Because the QALY data for VERTOS II were presented only at baseline, 1 month and 12 months, the manufacturer inferred the average utility across the 1 year time horizon. It was assumed that the difference in utility between BKP and OPM would linearly decline across 1 year so that there was no difference 3 years after the intervention. For PVP, it was assumed that the utility after the first year (which was not recorded) would progress similarly to that for BKP. It was assumed that the utility would decline after 2 years in accordance with population normalised data. The model assumes that both BKP and PVP are associated with a mortality benefit compared with OPM. The hazard ratio for death for BKP was set at ■■■ and for PVP was set at ■■■ based on the US registry data. No adverse events were included in the model except recurrent fracture, with lack of data being the reported reason for the omission. The submission stated that consequences of adverse events may be substantial.

6.4 The list price of a BKP kit is £2600.50, and the submission also noted an average selling price of £1900. The cost of PVP was assumed to be ■■■. The costs of the preliminary phase, the operating phase and the postoperative phase were updated from

those previously been reported in Strom et al. (2010). The length of stay following each intervention was reported to be taken from Hospital Episode Statistics 2010/11 data, and the assumed cost per day in hospital was taken from NHS Reference costs 2009/10/11. The deterministic analyses of BKP compared with OPM and of PVP compared with OPM gave ICERs of £2167 and £2053 per QALY gained respectively. The deterministic analysis of BKP compared with PVP gave an ICER of £2510 per QALY gained. Probabilistic analyses of BKP compared with OPM, PVP compared with OPM and BKP compared with PVP gave ICERs of £2118, £2100 and £2174 per QALY gained respectively.

- 6.5 Sensitivity analyses were conducted to study the impact of changes to the time horizon, the discount rate for costs and QALYs, the proportion of health utility benefit from the pivotal trial, the health utility offset time, post fracture mortality, the price of PVP compared with BKP, the unit costs per bed day, the assumed T-score of the cohort, the age of the cohort, the removal of bisphosphonate treatment, and the assumption that all patients were male. The ICER for BKP compared with OPM and BKP compared with PVP remained below £15,000 per QALY gained in all instances. The assumed mortality effect (which was more favourable to BKP than PVP) was a key driver of the cost-effectiveness results and, when it was assumed that there was no mortality benefit associated with either PVP or BKP, then the ICER of BKP compared with PVP was £27,340 per QALY gained. It was noted that the ICERs for both PVP and BKP compared with OPM remained low, with the key change being the ICER for BKP compared with PVP. The sensitivity analysis on the assumed length of hospital stay following BKP also increased the ICER of BKP compared with PVP to over £20,000 per QALY gained.

**Johnson and Johnson model**

- 6.6 The objective of the Johnson and Johnson model was to determine the cost effectiveness of PVP, BKP and OPM. In addition, a comparison with OPLA was carried out as a scenario analysis. A 1-year treatment-state model with costs and benefits discounted at 3.5% was developed, and an NHS perspective was adopted. The base-case analysis considered the treatment of all patients with vertebral compression fractures, regardless of the nature of these fractures. This all-patient analysis uses clinical evidence derived from all studies in the network meta-analysis and therefore could include patients who would not be expected to benefit from surgical intervention. An analysis based on a target population was also included to reflect the patient population that is expected to benefit most from surgical interventions. This analysis made use of clinical data from studies that considered patients with fractures 3 months old or less. The Assessment Group noted, however, that no differential effects of PVP compared with placebo were observed when the duration of pain was divided into the categories '6 weeks or less' and 'more than 6 weeks'.
- 6.7 Within the model, patients were assigned a VAS score at baseline and then at 2 weeks depending on the intervention received. The treatment-dependent VAS was updated at 1 month, 6 months and 12 months. A regression analysis was used to determine the relationship between VAS and EQ-5D, based on data for both outcomes derived from the network meta analysis. This relationship then allowed quality-of-life changes to be estimated in the model, based on the VAS scores reported at multiple time points. The Assessment Group stated that, in the network meta-analysis conducted by the manufacturer, no attempt was made to extrapolate or interpolate data from the randomised controlled trials if they did not report VAS scores at the designated time intervals, and this could cause discrepancies within the longitudinal data.



Secondly, a further trial, Blasco et al., was published after completion of the manufacturer's systematic review. This trial had similar VAS scores for both PVP and OPM, with both values being relatively high. If the manufacturer had included this study, the VAS scores in all arms would have increased and the relative difference between OPM and both PVP and BKP would have been reduced. It was assumed that there were no procedure-related adverse effects.

- 6.8 A bottom-up costing approach based on published data from Strom et al. was used for PVP and BKP. Costs for OPM were not specifically modelled because all patients, including patients receiving PVP or BKP, would receive OPM. The costs of the preliminary phase, the operating phase and the postoperative phase were taken from Strom et al. and inflated to 2009/10 prices, and were assumed to be applicable to both BKP and PVP. The costs for PVP varied according to the number of levels that need to be treated, and were reported to be £1358 for 1 level, £1784 for 2 levels and £1848 for 3 levels. Based on the estimated distribution of surgical procedures between 1, 2 and 3 levels by Dr Foster, the average weighted cost was estimated to be £1472. The Assessment Group noted that, in the text, the manufacturer stated that 11cm<sup>3</sup> of cement was needed for the 2-level procedure, the calculations included 7cm<sup>3</sup> of cement. If 11cm<sup>3</sup> was used, the average weighted cost would increase to £1546. The cost of the BKP kit reported in Strom et al. was inflated to a 2009/2010 cost of £2842.
- 6.9 The base-case results indicated that PVP was both more effective and less costly than BKP and therefore dominated BKP. The analysis of PVP compared with OPM gave an ICER of £4392 per QALY gained and BKP compared with OPM gave an ICER of £14,643 per QALY gained. The target population results also

indicated that PVP dominated BKP. The analysis of PVP compared with OPM gave an ICER of £4755 per QALY gained and BKP compared with OPM gave an ICER of £15,006 per QALY gained. When the scenario analysis including OPLA was considered, results indicated that PVP dominated OPLA, but that OPLA dominated BKP. The comparison of OPM with OPLA gave an ICER of £4853 per QALY gained.

- 6.10 Several other scenario analyses were conducted incorporating data from the OPLA trials but assuming that these could be pooled with OPM; extending the time horizon to beyond 1 year for both the base case and target populations; using an alternative bottom-up costing methodology and payment-by-results tariff for both the base case and target populations; using direct EQ-5D values for both the base case and target populations. PVP dominated BKP in all scenarios. The ICER for PVP compared with OPM ranged from £568–£13,595 per QALY gained in the base case and from £2550–£16,497 per QALY gained in the target population.
- 6.11 Univariate sensitivity analysis was conducted comparing PVP with OPM and PVP with BKP. It was found that the main drivers of the analysis were the efficacy of the treatment, that is, the VAS score at various time points, and costs (driven by the length of stay, cost per bed day and surgical equipment costs) for both base-case as well as target-population analyses. Probabilistic sensitivity analyses results were broadly similar to the deterministic results in the base-case analysis; PVP was still estimated to dominate BKP for both base-case and target-population analyses. In addition, in the base-case analysis, the probabilistic ICER for PVP compared with OPM was estimated to be £4388 per QALY gained in the base-case analysis and £4711 in the target population analysis. The probabilistic ICER for BKP compared with OPM was estimated to be £14,718 per QALY gained in the base-case analysis and

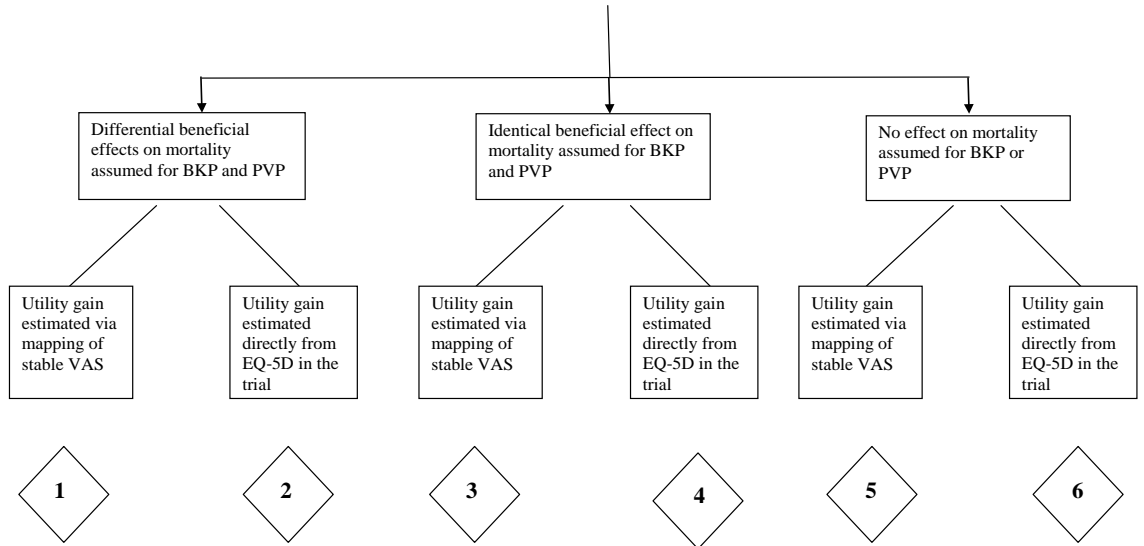
£15,010 in the target population analysis. In addition, the Assessment Group stated that there appeared to be a typographical error in the manufacturer's mathematical model in which only 10% of patients receiving BKP were assumed to consume operating-room resources. The Assessment Group assumed that this value was intended to be 100%. As such, the overall cost-effectiveness results are likely to favour to BKP.

### **Assessment Group model**

6.12 The objective of the Assessment Group model was to determine the cost effectiveness of PVP, BKP and OPM. In addition, because of the uncertainty around whether OLPA should be considered, analyses were presented both including and excluding OPLA. Although the Assessment Group presented 2 foundation analyses, they stated that these represented 2 of many plausible scenarios rather than a base case. The Assessment Group stated that, given the uncertainty around whether there were mortality benefits, the results were split into 3 categories based on the underlying assumption of: a differential effect assumed for BKP and PVP, in which both were better than OPM; a pooled analysis assuming identical effects with BKP and PVP, in which both were better than OPM; and 1 in which no mortality benefit of BKP or PVP was assumed. The effect of OPLA was varied in sensitivity analyses for the first 2 categories and assumed to be equal to OPM in the third. The Assessment Group also stated that there was a difference in results based on whether the EQ-5D was taken directly from the trials (INVEST, FREE, Buchbinder and Rousing) or whether the mapping of stable VAS scores from the network meta-analyses to EQ-5D was adopted. Please see pages 192–5 of the assessment report for details of mapping of VAS to EQ-5D. Analyses were conducted based on the Buchbinder, Free and INVEST studies. The Rousing study was excluded because of the imbalance in EQ-5D at the start of the study between the PVP and OPM arm. Six

plausible scenarios were presented as outlined in figure 1 (figure 1 from the assessment report).

**Figure 1** Derivation of the Assessment Group's six scenarios.



Sensitivity Analyses will be undertaken for each of the six scenarios.

Scenarios 2, 4 and 6 will also be subdivided into the results from Buchbinder, the FREE trial, and INVEST

An exploratory analysis was also presented assuming that all patients were provided with a facet joint injection before PVP, based on clinical advice that this is common practice. In addition, exploratory analyses using high-viscosity cement and accounting for additional patient education were also presented.

6.13 The model consisted of 5 health states: the starting health state of post-osteoporotic vertebral compression fractures for which BKP, PVP, OPLA or OPM has been carried out; a subsequent additional vertebral fracture; a subsequent hip fracture; both a subsequent vertebral and a hip fracture; and death. For simplicity, only 1 further vertebral fracture and 1 hip fracture were permitted. A time horizon of 50 years was assumed to represent patients' lifetimes and the model employed 36 monthly time cycles followed by 47 yearly time cycles. The rationale for the different cycle length was that there may be a utility difference between interventions in the initial period

following a procedure, which was more easily incorporated into monthly time cycles. A life table methodology, rather than a half-cycle correction, was employed to take into consideration that all transitions did not take place at the end of the time cycle. Both costs and benefits were discounted at 3.5% per annum. The model did not include the potential disutility associated with anxiety about the prospect of future fractures, or the potential reduction in bone mineral density associated with prolonged bed rest.

6.14 The transition probabilities were taken from the literature. To take into consideration that a patient's bone density is likely to deteriorate over time, a decrease of 0.255SD per 5-year age group was incorporated in accordance with data from the literature, and it was assumed that women and men with the same T-score would have the same risks of fracture. If a patient was assumed to be taking a bisphosphonate, the assumed effect on vertebral fractures was based on relative risks reported in the literature. This effect was assumed to last for 5 years, with a linear decline in effect over a 5-year period, so that the relative risk was 1 after 10 years. The risk of hip fracture or vertebral fracture for patients was assumed to be independent of whether the patient was simulated to have a subsequent vertebral fracture or hip fracture.

6.15 The mortality associated with hip fracture was taken from Stevenson et al. (2009) and the mortality associated with vertebral fracture was taken from a UK study (Jalava et al. 2003). The model allowed an increased risk of mortality in the year of subsequent fracture. It was also assumed that mortality following hip fracture could not be lower than either mortality associated with a vertebral fracture or lower than that of general mortality in the underlying age- and gender-matched population. The underlying mortality through other causes than fracture was taken from interim life tables from the Office for National Statistics, and it was assumed

that all patients would die before they were 101 years. When BKP, PVP or OPLA were assumed to have positive mortality effects compared with OPM, these were incorporated in the model for a user-defined period (set to 5 years in the base case). It was assumed that mortality benefit would cease immediately after the user-defined period. The relative risks associated with treatment were assumed to apply to all-cause mortality, and to the increase associated with vertebral fractures, but not to hip fractures. Please see pages 180–5 of the assessment report for further details around the calculation of transition probabilities.

**6.16** The hazard ratios within the 3 scenarios used to explore the effects of mortality were calculated from academic-in-confidence data provided by Medtronic and are presented in table 4. Data about the effect of OPLA on mortality were not available. It was assumed that the effect was half of that observed for PVP because this was observed with VAS data. This equated to a hazard ratio of [REDACTED] when a differential effect was assumed, [REDACTED] when a pooled effect was assumed and 1 when no effect was assumed. The effect of OPLA on mortality was adjusted in sensitivity analyses.

**Table 4. The hazard ratios within the 3 scenarios used to explore the effects of mortality associated with BKP, PVP and OPM. All values compared with OPM. A lower number indicates that the intervention is associated with a longer life expectancy.**

Scenario	Hazard ratio (95% CI)	
	Balloon kyphoplasty	Percutaneous vertebroplasty
Differential effects	[REDACTED]	[REDACTED]
Pooled effects	[REDACTED]	[REDACTED]
No effect	1	1

**6.17** Utilities for all health states were assumed to be a function of: gender; age; procedure undertaken; time since procedure; time at

which patients treated with OPM were assumed to have the same utility as patients treated with an active intervention; the disutility associated with vertebral fractures that occurred more than 1 year ago; and the mapping of VAS scores onto the EQ-5D. In addition, in the health state in which a patient sustained an additional vertebral fracture and remained alive or an additional hip fracture and remained alive, the disutility associated with a vertebral fracture or hip fracture in the year of occurrence was considered relevant. In the cycle of the subsequent vertebral or hip fracture, a QALY decrement was automatically applied to account for the associated pain. For the health state in which patients sustained both a vertebral fracture and a hip fracture and remained alive, the disutility associated with a vertebral fracture and a hip fracture in the year of occurrence as well as in subsequent years was taken into account. By definition, the utility within the death state was zero. It was assumed that there were no cost or QALY implications associated with adverse events. However, a sensitivity analysis was conducted assuming that the QALY losses associated with BKP and PVP were 0.02.

- 6.18 Costs within each of the health states were largely taken from Stevenson et al. (2009) and inflated to 2010/11 prices using the Hospital and Community Health Services inflation indices. The cost of the Confidence Spinal Cement System was taken from the Johnson and Johnson submission, although it was assumed that 11cm<sup>3</sup> of cement was needed for a 2-level procedure rather than 7cc. This gave an average cost of £1546 per operation. List prices for low-viscosity cements were available that, when weighted for the proportion of operations that were 1-, 2- and 3-level procedures, equated to an estimated value of £697. The Assessment Group was advised by a clinical specialist that approximately 15% of cases are more complex and would need Cortoss cement, collation or thicker cement, while younger patients

would need bone-absorbable cement. It was assumed that the added cost of these complex cases would add slightly over £100 to the average cost of an operation, resulting in an assumed cost of £800 per low-viscosity cement PVP procedure. The results for PVP have been estimated assuming the use of low-viscosity cement. The Assessment Group highlighted that, given that the estimate includes a component for using higher-viscosity cement, the price used within the analysis could be equated to a strategy in which low-viscosity cement is used for most patients, with higher-viscosity cement used in a small proportion when the clinician believes that this is appropriate. Sensitivity analyses were carried out on these average values.

- 6.19 The list price of £2600.50 per kit for BKP was inflated to take into consideration that a proportion of patients would need BKP at more than 1 level. It was assumed that the percentages reported for PVP were also applicable to BKP and that each level would need an additional pack of Kyphon HV-R bone cement, priced at £62 per pack, with the remaining instruments being reused. This resulted in the average price per patient increasing to £2639 for BKP. The cost of OPLA was assumed to be equal to that of PVP, and this was varied in the sensitivity analysis.
- 6.20 Costs for the preliminary and postoperative phase were taken from the Johnson and Johnson submission, based on clinical opinion that these were broadly correct. These were estimated to be £540 and £243 respectively. Costs for the operating phase were taken from the bottom-up costing approach provided by Johnson and Johnson, which estimated costs of £528 and was considered more realistic than those from Strom et al. (2010). The length of hospital-stay data from Medtronic were most appropriate because they used the standard HES data source, whereas the cost per day of £232 presented by Johnson and Johnson was considered most



appropriate. However, it was noted that the £232 per day value would underestimate the total costs in each arm. It was also noted that clinical advisers to the Assessment Group stated that most procedures performed would be day cases and therefore length of stay would be shorter than suggested by HES data.

- 6.21 The results for each of the plausible scenarios are presented on pages 222–62 of the assessment report. Each scenario was subjected to sensitivity analysis exploring the impacts of changes to the following assumptions: assuming a bed day cost of £0; changing the assumed cost of equipment for OPLA and the cost of the procedure; changing the time of convergence; and including potential QALY losses associated with adverse events. The Assessment Group highlighted that combinations of these sensitivity analyses may represent more plausible central estimates of the cost effectiveness of the interventions and should be provided with equal weight.
- 6.22 The Assessment Group summarised that, in scenarios 1 and 2, in which differential mortality effects were assumed, with BKP being more effective than PVP, results indicated that BKP always provided the most QALYs and always gave an ICER below £20,000 per QALY gained. This was irrespective of whether utility gain was estimated by mapping stable VAS or directly from EQ-5D in the trials, irrespective of the trials included and irrespective of the sensitivity analysis carried out. Also, it was maintained even if the cost of BKP was increased assuming a separate kit was needed for each level.
- 6.23 In scenarios 3 and 4, in which the mortality effects of PVP and BKP were assumed to be identical, and OPLA was assumed to provide half the mortality benefit compared with PVP and BKP, results indicated that BKP was dominated by PVP. This was because it effectively provided the same QALYs at a higher cost, and the

ICER for PVP compared with OPM remained below £10,000 per QALY gained. This was consistent across all assumptions except for the combination of assumptions in which: OPLA was assumed to have an identical mortality benefit to BKP and PVP; OPLA was assumed to have a lower cost than PVP; adverse events for PVP were included; and the EQ-5D data from the randomised controlled trials were used. In this instance, PVP was dominated by OPLA. However, it was noted that, if OPLA was not seen to be an appropriate comparator, the ICER of PVP compared with OPM remained below £10,000 per QALY gained.

- 6.24 In scenarios 5 and 6, in which no mortality effects were assumed for BKP or PVP compared with OPM, the conclusions depended on the assumptions made. For example, when the utility gain was estimated by mapping, PVP typically provided the most QALYs and the ICER remained below £20,000 per QALY gained. The only exception to this was when assumptions unfavourable to PVP were adopted, such as equal hospitalisation stay costs, reduced cost of OPLA, incorporation of adverse events for PVP, and an earlier convergence of EQ-5D scores. When utility gained was estimated directly from the Buchbinder and INVEST trials, BKP was always dominated by PVP, but PVP was also dominated by OPLA in some cases and had an ICER greater than £20,000 per QALY gained in some cases. When OPLA was not considered an appropriate comparator, PVP compared with OPM had an ICER greater than £20,000 per QALY gained in some cases. The Assessment Group highlighted that reducing the BKP and PVP estimates so that they are closer to the OPLA value from the network meta-analysis, will result in an unfavourable comparison of PVP with OPM, compared with raising the OPLA value to the BKP and PVP network meta-analysis estimate.

- 6.25 The Assessment Group stressed that the decision about which intervention was most cost effective was dependent on the assumptions chosen. Given the uncertainty regarding the mortality effects of the treatments (including OPLA), a definitive conclusion could not be reached. However, the Assessment Group stated that, given that a facet joint injection was commonly used, relatively inexpensive and may have considerable benefit in up to a third of patients (Wilson 2011), it was likely that it would be considered an appropriate first measure in clinical practice.
- 6.26 The Assessment Group also conducted an exploratory analysis using high-viscosity cement for all patients. It was calculated that there would need to be an additional 0.037 QALYs for the cost per QALY gained to be equal to £20,000 per QALY gained. Given that this value was greater than the value of 0.02 discounted QALYs assumed in the sensitivity analyses, the Assessment Group stated that it was unlikely that the ICER of high-viscosity cement compared with low-viscosity cement would be lower than £20,000 per QALY gained. However, it was stated that there was a possibility that operations would need to be re-performed if there was a problem with low-viscosity cement. The costs per QALY gained of high-viscosity cement at different levels of re-operation rates were estimated to explore this impact. It was calculated that there would need to be a re-operation rate in excess of 25% for a strategy of using high-viscosity cement in all patients to be cheaper than using it in a selected 15% of patients, assuming that QALYs remained unaltered. The Assessment Group noted that the only identified estimate was less than 1.5%, and also it was not certain that high-viscosity cement would have prevented the re-operation in each case. Lastly, the Assessment Group estimated the impact of changing both the QALY gained and the level of re-operations on the cost per QALY gained using high-viscosity cement. At a re-operation rate of 5% and a QALY increase of 0.02, the ICER using

high-viscosity cement was £30,158 per QALY gained. The Assessment Group stated that it was unlikely that a strategy of using high-viscosity cement in all patients rather than a subset selected by the clinician would have an ICER less than £20,000 per QALY gained.

- 6.27 The Assessment Group carried out an exploratory analysis of the effect of an initial facet joint injection on the cost effectiveness of BKP and PVP. Assuming that, based on the literature, a third of patients would respond to a facet joint injection and that they would exhibit identical VAS or EQ-5D effects regardless of the treatment arm, then the average VAS or EQ-5D difference shown in the entire population would be estimated to have increased by 50% when considering patients who did not respond to the facet joint injection. If it was assumed that the entire QALY difference was because of VAS or EQ-5D scores (rather than adverse events), then the ICER would be reduced by a third, implying that ICERs of £30,000 may be reduced to £20,000 per QALY gained if all patients had a facet joint injection initially.

## **7 Equalities issues**

- 7.1 During the scoping consultation, consultees noted that the population for this appraisal predominantly comprises older women, many of whom are also primary caregivers to other people. However, this is not considered as an equality issue because it is not expected to lead to unfair access to treatment.
- 7.2 In their submissions, professional groups noted that people who were unable to have local or general anaesthesia and people living in areas where there was no local service would find it difficult to access the technologies. In addition, Medtronic stated that the impact of vertebral compression fractures and hyperkyphosis on both health-related quality of life and mortality has been historically

overlooked in England and Wales with respect to other disease areas.

## **8 Innovation**

8.1 No claims for innovation were presented.

## **9 Authors**

### **Raisa Sidhu**

Technical Lead

### **Pall Jonsson**

Technical Adviser

## Appendix A: Supporting evidence

### *Related NICE guidance*

#### **Published**

Balloon kyphoplasty for vertebral compression fractures. NICE interventional procedure guidance 166 (2006, updated 2008). Available from [www.nice.org.uk/guidance/IPG166](http://www.nice.org.uk/guidance/IPG166)

Percutaneous vertebroplasty. NICE interventional procedure guidance 12 (2003). Available from [www.nice.org.uk/guidance/IPG12](http://www.nice.org.uk/guidance/IPG12)