Osteoporosis International (2020) 31:2461–2471
https://doi.org/10.1007/s00198-020-05513-x

ORIGINAL ARTICLE

Cost-effectiveness of balloon kyphoplasty and vertebroplasty versus conservative medical management in the USA

T. J. Hopkins[1] · S. Eggington[2] · M. Quinn[3] · C. I. Nichols-Ricker[3]

Received: 4 March 2020 / Accepted: 19 June 2020 / Published online: 12 July 2020

Abstract

Summary The cost-effectiveness of surgical versus conservative medical management of vertebral compression fractures in the US was analyzed in the context of inpatient versus outpatient treatment. Surgical intervention (balloon kyphoplasty and vertebroplasty) was found to be cost-effective relative to conservative medical management at a US willingness-to-pay threshold.

Introduction To date, only one published study has evaluated the cost-effectiveness (C/E) of balloon kyphoplasty (BKP) or vertebroplasty (VP) in US Medicare patients with osteoporotic vertebral compression fractures. This study further evaluates the C/E of surgical treatment vs. conservative medical management (CMM), expanding on prior modeling by accounting for quality-adjusted life-years gained.

Methods A Markov microsimulation model of 1000 patients was constructed. Cost data were based on an analysis of Medicare claims payments, with propensity-score matching performed for BKP and VP vs. controls (CMM). Mortality inputs were based on US life tables, modified to account for age at initial fracture, presence of subsequent fracture(s), and relative risk of mortality by treatment. Separate incremental cost-effectiveness ratios (ICERs) were calculated for BKP and VP in inpatient and outpatient surgical treatment locations to account for individual clinical profiles presenting to each.

Results The discounted ICER for inpatient BKP vs. CMM was $43,455 per QALY gained; for outpatient BKP vs. CMM, $10,922; for inpatient VP vs. CMM, $39,774; and for outpatient VP vs. CMM, $12,293. Probabilistic sensitivity analysis confirmed that both BKP and VP would be considered C/E vs. CMM at a US willingness-to-pay (WTP) threshold of $50,000/QALY in 80% and 100% of 500 model simulations, respectively. The most sensitive parameters included quality of life estimates and hazard ratios for mortality.

Conclusion While VP and BKP are more expensive treatment options than CMM in the short term, model results suggest interventional treatment is cost-effective, among patients eligible for surgery, at a US WTP threshold. This conclusion supports those from economic analyses conducted in EU-member countries.

Keywords Balloon kyphoplasty · Cost-effectiveness · Markov model · Osteoporosis · Vertebral compression fracture · Vertebroplasty

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00198-020-05513-x) contains supplementary material, which is available to authorized users.

✉ T. J. Hopkins
thomas.hopkins@duke.edu

[1] Department of Anesthesiology and Pain Medicine, Duke Health, Duke University, 201 Trent Drive, 4313 Duke South, Orange Zone, DUMC, Box 3094, Durham, NC 27710, USA

[2] Medtronic Global Health Economics & Reimbursement, Tolochenaz, Switzerland

[3] Medtronic Pain Therapies Health Economics & Reimbursement, Fridley, MN, USA

Abbreviations

BKP	Balloon kyphoplasty
BMD	Bone mineral density
CCI	Charlson comorbidity index
CEAC	Cost-effectiveness acceptability curve
CMM	Conventional medical management
CMS	Centers for Medicare and Medicaid Services
ICER	Incremental cost-effectiveness ratio
LY	Life-year
PSA	Probabilistic sensitivity analysis
QALY	Quality-adjusted life-year
QoL	Quality of life
VCF	Vertebral compression fracture

VP Vertebroplasty
WTP Willingness to pay

Introduction

By 2025, the incidence of osteoporotic fracture in the USA is predicted to grow to affect 3 million patients per year, with a lifetime risk between 40 and 50% in women and 13 and 22% for men with a history of osteoporosis and an associated $25 billion in healthcare costs [1, 2]. Of those, nearly a quarter will be vertebral compression fractures (VCF), which have been estimated to affect as many as 1 in 4 women and 1 in 5 men over the age of 50, and more than 1 in 3 women over the age of 70 [3].

As a known sequela of osteoporosis, VCFs can be either asymptomatic or debilitatingly painful. Regardless of pain presentation, primary VCF is associated with 5-fold elevated risk of subsequent VCF and a doubling in the risk of hip fractures [4]. These and other osteoporotic fractures can set in motion a cycle of progressive spinal deformity, musculoskeletal disability, morbid injury, and mortality, along with significant demand on associated healthcare resources [1, 5–8].

Conservative medical management (CMM)—inclusive of pharmaceutical pain management, bed rest, bracing, and physical therapy—has long been the standard of care for VCFs because the risks associated with open surgery on this frail and comorbid population were seen as too great [5, 9]. However, this strategy is not optimal when considering the risk of subsequent functional outcomes and mortality [8]. In recent decades, minimally invasive techniques for reducing and stabilizing VCFs, such as vertebroplasty (VP) and balloon kyphoplasty (BKP), have changed the course of osteoporotic spinal degeneration, safely relieving pain, supporting functional recovery, and dramatically reducing mortality risk, even among very elderly patients [8, 10, 11]. A recent analysis of mortality in more than 2 million patients with VCF, expressed in epidemiological terms of number needed to treat (NNT), showed that approximately 15 patients need to be treated with BKP instead of CMM to potentially avoid one VCF-related death within 1 year of fracture [12]. NNT was similarly low for vertebroplasty relative to CMM (22.8 treated to avoid one death within 1 year of fracture).

Despite these improvements, the direct, short-term costs of these surgical interventions are significantly higher than those of CMM. Separate from cost considerations, literature has shown improvement in patient quality of life and mortality risk, resulting in cost-effective conclusions in ex-US geographies [13–17]. To our knowledge, only one US-based cost-effectiveness model exists that evaluates treatment with BKP or VP relative to CMM; however, this model evaluated the cost-effectiveness of life-years gained and not the impact on patient quality of life, an important consideration [13]. Furthermore, a recent shift in procedure setting in the USA from inpatient to outpatient warrants more detailed evaluation of cost-effectiveness in various clinical care settings, which prior studies do not address.

In the present study, we adapted and extended on an existing cost-effectiveness model originally built for a UK perspective [17]. Our goal was to evaluate the cost-effectiveness of active surgical intervention (with BKP or VP, excluding other spinal augmentation procedures) compared with CMM from the US Medicare payer perspective, while accounting for differences among patients undergoing treatment on an inpatient versus outpatient basis.

Methods

Model overview

The model was developed to assess the cost-effectiveness, from a US Medicare perspective, of BKP and VP, compared with CMM in patients with osteoporotic VCF. The model was designed to build on an existing UK-based model, reported by Svedbom et al. [17], while adapting demographic, clinical, and cost inputs to reflect a US Medicare population.

Modeling strategy and structure

Svedbom and colleagues developed their model in TreeAge Pro (Williamstown, MA) with a front-end in Microsoft Excel (Redmond, WA) [17]. For the present analysis, the model was re-built in Excel.

A Markov microsimulation approach used a 6-month cycle with a starting population of 1000 simulated patients. Given that VCFs are more common in older populations, a lifetime horizon was used in the base-case analysis to allow sufficient time for the benefits of each intervention to accumulate. Health states included treatment status, repeat or subsequent fracture (at any level), and death (absorbing state), as shown in Fig. 1. Subsequent fractures at any level were assumed to have been managed using the same intervention as for the initial fracture.

As individual patients were simulated over a lifetime perspective, the model calculated patient utility, costs, and risk of mortality over time. Microsimulation was required to accurately capture the timing of specific events such as subsequent fractures, and to model their impact on cost, quality of life (QoL), and mortality. Health outcomes were modeled in terms of life-years (LYs) and quality-adjusted life-years (QALYs) to allow a full cost-utility analysis to be undertaken. Costs, life-years and QALYs were discounted at 3% per year as per recommendations for economic analyses undertaken for the USA [18].

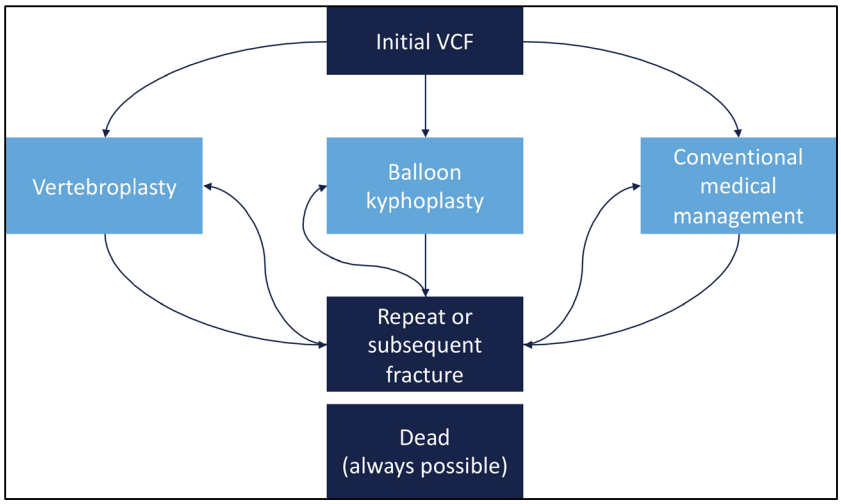

Fig. 1 Markov model state-transition diagram. Repeat or subsequent fractures were assumed to be treated in the same way as the initial fracture[a]

[a]VCF, vertebral compression fracture

The model was designed to calculate separate estimates of cost-effectiveness for each intervention versus CMM, by initial treatment setting (inpatient and outpatient) given the different patient demographics, comorbidity status, and subsequent post-acute care use and mortality rates observed by treatment setting in an exploratory analysis of Centers for Medicare and Medicaid Services (CMS) claims data.

Model inputs

Clinical data

Patient population

Base-case demographic characteristics for patients in the model were based on analysis of CMS 100% standard analytical file claims data (2014–2016). Patients were selected based on the date of their first VCF diagnosis, then followed for up to 2 years after discharge from the index encounter. Follow-up ended at the earliest of the time of death, receipt of a spinal fusion procedure (with follow-up censored to the day prior to spinal fusion), or the end of 2 years of follow-up data. Patients were excluded if they had previous history of BKP or VP, if they were a CMM patient who underwent BKP or VP anytime during follow-up, and if they had any diagnosis of cancer. Propensity-score matching was used to select similar CMM patients versus surgically treated patients, with matching run separately by surgical treatment setting to reflect different patient-level clinical profiles correlated with inpatient versus outpatient treatment (Table 1). Further details on the retrospective analysis of CMS claims data for model inputs can be found in the Supplemental content, Supplemental methods.

Subsequent fracture risk

Patients in the model were at risk of subsequent fracture(s) at any spinal level. This risk was applied by taking a baseline fracture risk for the population in the USA and applying relative risks for prior fracture, reduced bone mineral density (BMD), and the use of bisphosphonates. Patients could experience more than one subsequent fracture during the model time horizon.

Age- and gender-specific VCF risks in the general population were taken from a study by Ettinger et al., who used US hospital discharge data to derive annual sex-specific rates of osteoporotic fractures (hip, vertebra, humerus, and forearm) in five-year age bands [19]. The rates reported for VCF were converted into six-month probabilities to match the cycle length used in the model and were then weighted according to the male/female split for the selected treatment comparison (Supplemental Content Table S1).

Patients were assigned an additional risk factor modifier to reflect higher incidence of fracture among individuals with a previous VCF (Supplemental Content, Table S2). This was applied in the form of an age-specific relative risk compared with an individual of the same age but with no prior VCF. Data were taken from the UK cost-effectiveness analysis reported by Svedbom et al. who used results from Klotzbuecher et al. to calculate age-specific relative versus those without prior fracture. The relative risks were adjusted for the prevalence of vertebral fracture in the general population and were assumed to be constant for the remainder of the model horizon (Supplemental Content Table S2) [17, 20].

The effect of osteoporosis on risk of fracture was incorporated using a relative risk adjustment of 2.3 for patients with reduced BMD, based on a report by Marshall and colleagues [21]. Because bisphosphonate therapy is known to reduce the

Table 1 Demographic inputs by treatment comparison[a]

	BKP	CMM	*P* Value	VP	CMM	*P* Value
Inpatient						
N	2,071	2,071		710	710	
Age, mean	81.6	82.2	0.028	81.4	81.8	0.287
Female, %	82	83	0.387	82	83	0.780
CCI score ≥ 2, %	42	44	0.630	47	48	0.494
Outpatient						
N	3,708	3,708		1,042	1,042	
Age, mean	78.9	79.3	0.015	79.5	80.4	0.006
Female, %	78	81	0.013	79	80	0.704
CCI score ≥ 2, %	22	22	0.701	24	24	0.985

[a] Beta distributions were used to represent uncertainty for both age and gender split. Values based on propensity-score-matched cohorts in an analysis of CMS claims data

BKP, balloon kyphoplasty; *CMM*, conventional medical management; *VP*, vertebroplasty; *CCI*, Charlson Comorbidity Index

risk of VCF in patients with osteoporosis, it was assumed that all patients in the model would be receiving this therapy. Accordingly, a relative risk adjustment of 0.6 was applied in the first 5 years of the model to mirror the approach used by Svedbom et al. [17]

The following equation summarizes the fracture risk calculation in each cycle:

$$\text{p\_fracture} = \text{p\_frac\_pop} \times \text{rr\_prior\_frac} \times \text{rr\_low\_bmd} \times \text{rr\_bisphos}$$

where p_fracture = risk of fracture in current 6-month model cycle, p_frac_pop = age- and gender- specific baseline risk of fracture in the general population, rr_prior_frac = age-specific relative risk of subsequent fracture due to prior fracture, rr_low_bmd = relative risk of fracture due to low BMD (2.3), and rr_bisphos = relative risk of fracture due to use of bisphosphonates (0.6).

It was not possible to reliably quantify the risk of subsequent fracture by treatment group using CMS claims, as there is no unique diagnosis code for a subsequent (or repeat) VCF versus the initial event. Therefore, due to a lack of suitable data, the risk of subsequent fracture was assumed to be independent of the treatment received for the initial fracture for the purposes of this analysis.

Mortality risk

Patients' risk of death during each six-month model cycle was composed of a baseline risk combined with additional risks for treatment and treatment setting. Baseline mortality risk was based upon US life tables from the Centers for Disease Control and Prevention [22]. These risks were converted to a six-month cycle length and weighted according to the male-female split observed in CMS claims data for each treatment comparison.

Mortality risk among patients with VCF is higher relative to the general population following a fracture, but the risk declines over time as other age-associated risk factors begin to outweigh the VCF-specific effect. Additionally, patients with subsequent fractures are at higher risk of death than those with no subsequent fracture. Therefore, time-dependent relative risks were applied, based on an analysis linking Swedish inpatients and death registers and used in the analysis by Ström et al. [16] These time-dependent risks were applied to the baseline mortality rate according to each patient's age, time since last fracture, and subsequent fracture status [17].

Analysis of patient-level data from Medicare claims data for each treatment comparison included an assessment of survival outcomes across propensity-score-matched treatment groups over the 2 years post-fracture in Cox proportional hazard models. In each case, a hazard ratio was estimated for risk of mortality (within 2 years of initial VCF) for the surgical treatment versus CMM groups (Supplemental Content Table S3).

To adjust for the differential clinical profile between patients treated in the inpatient and outpatient settings, the incremental mortality risk by treatment setting was implemented in the model using the hazard ratio for mortality from a Cox proportional hazards model using all matched patients (surgical treatment and controls) in the CMS claims analysis (hazard ratio = 2.27 for mortality for inpatient vs. outpatient, standard deviation = 0.05). This incremental mortality risk was applied in all treatment comparisons in the model and applied to all treatment groups, including CMM. This could be seen as counter-intuitive, as CMM patients do not undergo surgery, making treatment setting irrelevant. However, due to the propensity-score matching approach used, patients receiving

CMM had demographic and clinical profiles closely matched to those undergoing surgical intervention, and thus, the additional risk applied to these patients is meant to represent the underlying patient clinical severity and comorbidity burden.

Taken in summary, the following equation summarizes the mortality risk calculation in each cycle:

$$\text{p\_death} = \text{p\_death\_pop} \times \text{rr\_subseq\_frac} \times \text{hr\_treatment} \times \text{hr\_setting}$$

where p_death = risk of death in current 6-month model cycle, p_death_pop = age- and gender-specific baseline risk of death in the general population (per 6 months), rr_subseq_frac = relative risk of death among patients with prior fracture, hr_treatment = hazard ratio of death for each treatment comparison (surgical versus CMM), (Table S3), and hr_setting = hazard ratio of death by initial treatment setting (inpatient vs. outpatient).

Cost data

The cost parameters used in the model were based on the CMS claims analysis described above, with costs evaluated for each of the 4 propensity-score-matched cohorts. In this study, the term "costs" refers to Medicare reimbursement amounts. All costs were inflated to 2016 USD using the medical care component of the Consumer Price Index.

Procedure

For each treatment comparison, the cost of the initial procedure was averaged across all patients undergoing BKP or VP (Supplemental Content Table S4). No surgical procedure cost was applied to the CMM treatment group since no procedure was performed for these patients.

Post-acute care

Various elements of post-acute care were used to calculate a mean 2-year follow-up cost per patient, by treatment group and treatment setting (Supplemental Content Table S5). Post-acute costs evaluated included home health, skilled nursing facility, hospice, and inpatient rehabilitation. Conditional mean costs were then multiplied by the proportion of patients with any utilization by care setting to obtain a total mean post-acute cost per patient.

Outpatient follow-up

Mean Medicare costs of outpatient follow-up care in the 2 years following the initial fracture were also included in the model (Supplemental Content Table S6). Costs were extracted for each of the 4 treatment groups and divided by 4 in order to spread the aggregate 2-year follow-up cost over the first four 6-month model cycles. Unlike post-acute care use, 100% of patients were assumed to have outpatient visits. No further outpatient costs were applied beyond 2 years, except in the case of subsequent fracture.

Subsequent fracture

Any patient experiencing a subsequent fracture, whether at the same level or at a different level, was assumed to be managed in the same manner as for the initial fracture.

Inpatient costs for subsequent fracture events were summarized using Medicare claims (Supplemental Content Table S7). In the CMM group, the cost of subsequent fracture was based on the mean cost of a readmission in the inpatient setting with a primary diagnosis of VCF, with no BKP or VP procedure performed.

In addition to the procedure costs of a subsequent fracture event, the costs of post-acute and outpatient care were included for each subsequent fracture, using the same post-acute cost data as for the initial fracture (Supplemental Content Tables S5 and S6**).**

Quality of life

Patients' QoL was estimated using inputs from several sources. The FREE-2 trial, the results of which were first reported by Wardlaw et al., compared BKP with CMM, with QoL data collected using the Euroqol 5-dimension (EQ-5D) questionnaire, administered at 1, 3, 6, 12, and 24 months follow-up [23, 24]. We obtained patient-level data from the FREE-2 trial and recalculated patient utility values by applying the US-specific preference weights from Fryback et al. to the patient-level EQ-5D domain scores (Supplemental Content Table S8) [25]. EQ-5D data were not collected at 18 months in the FREE-2 study, so the utility weight for each treatment at this time point was estimated by interpolation between the utilities at 12 and 24 months. Because the FREE-2 study did not include a VP treatment arm, utility weights for VP patients were assumed to be identical to those observed for BKP.

Beyond 24 months, utility in the CMM group was based on the 24-month utility calculated for CMM patients at the 24-month visit of the FREE-2 trial. For the BKP and VP groups, patients' utility between 24 months and 36 months was assumed to decrease linearly from the value at the 24-month visit to equal the 36-month value in the CMM group (0.668). Beyond 36 months, utility for VP and BKP patients was set equal to CMM patients. If a patient suffered a subsequent fracture, utility values were re-set. This assumption of no additional utility benefit by the end of 3 years is consistent with the assumptions made by Svedbom et al. [17], namely

that any effects of treatment on QoL would be confounded at time points further from the intervention, at which point the older age of the patient cohort plays a more central role in utility.

The Medicare-aged population used for the base case of this model (mean 79–82 years) was older than patients in the FREE-2 trial (mean 72–74 years). To incorporate the effect of age, all utility weights used in the model were adjusted by multiplying by age-specific general-population utilities from the USA (Supplemental Content Table S9), as reported in 10-year age bands by Szende et al. [26]

Data analyses and model validation

For each treatment comparison, the incremental cost-effectiveness ratio (ICER) was calculated according to the formula:

$$\mathrm{ICER} = \frac{\Delta(\text{Cost surgical treatment} - \text{Cost CMM})}{\Delta(\text{QALY surgical treatment} - \text{QALY CMM})}$$

We then undertook a series of scenario analyses to explore the uncertainty related to specific model inputs. The base-case model assumed that the utility weights for patients undergoing VP were identical to those applied to patients receiving BKP. To test the impact of changing this assumption, we varied the VP utilities within the ranges of their 95% confidence intervals (i.e., setting all utilities to the lower limit and again to the upper limit of each distribution). For completeness, we conducted an equivalent analysis for the BKP treatment comparisons. Furthermore, we used a similar approach to vary the age-specific relative risks of subsequent fracture within their confidence limits, running the model after setting all these parameters to their lower 95% confidence and again using the upper limit of each distribution.

Probabilistic sensitivity analysis (PSA) was performed to explore the impact of the joint uncertainty in all inputs upon the cost-effectiveness results. To obtain meaningful results from a PSA, appropriate distributions were selected to represent the uncertainty in each input parameter. To reflect the skew often found in cost data and to avoid sampling negative cost values, Gamma distributions were applied to all cost parameters using the mean and standard deviation of each cost input from the Medicare analysis. Variability in utility parameters was represented via beta distributions, again using the mean and standard deviation reported for each utility, to ensure sampled values bounded by 0 and 1. Beta distributions were also used for proportions such as those for gender, patients receiving each type of post-acute care, and the baseline fracture probabilities. Uncertainty in relative risks and hazard ratios was represented using lognormal distributions, and the mean age of patients in the model was sampled from a normal distribution. For some inputs, no information was reported in the literature regarding the level of uncertainty (e.g., a confidence interval or standard deviation). For such parameters, distributions were approximated by assuming the 95% confidence interval around the mean was represented by values 20% either side of the mean.

The PSA for each treatment comparison was run by sampling 500 sets of the input parameters (outer loop), within which 1000 patients were simulated (inner loop) each time. Mean costs and QALYs were recorded for each treatment group for each outer loop simulation (i.e., averaging across the 1000 patients in the inner loop), to give 500 sets of costs and QALYs for each group. This information was then used to construct a cost-effectiveness plane and cost-effectiveness acceptability curve (CEAC) for each treatment comparison to determine the level of uncertainty in the ICER.

Finally, we reviewed our model reporting against the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist to ensure complete reporting on all components of our model [27].

Results

Between 2014 and 2016, a total of 75,524 patients in the CMS 100% standard analytical file claims data were diagnosed with VCF. After all inclusion and exclusion criteria were applied, the final dataset was represented by 2071 patients in the inpatient BKP group, 3708 in the outpatient BKP group, 720 in the inpatient VP group, and 1042 in the outpatient VP group, with an equivalent number of propensity-matched CMM patients for each treatment group. Table 1 summarizes demographics for each treatment group.

Table 2 shows the deterministic cost-effectiveness results for each treatment comparison. Results are presented separately for discounted and undiscounted scenarios. The discounted ICER for inpatient BKP versus CMM was $43,455 per QALY gained (Table 2). For outpatient BKP versus CMM, the ICER was $10,922 per QALY. The ICER for the inpatient VP versus CMM was $39,774 per QALY gained, and the ICER for outpatient VP versus CMM was $12,293 per QALY gained.

In a one-way sensitivity analysis, the mortality benefit associated with surgery was "turned off," i.e., the hazard ratio was set to 1. Resulting ICERs were as follows: inpatient BKP vs. CMM was $283,579 per QALY gained; outpatient BKP vs. CMM was $55,485 per QALY gained; inpatient VP vs. CMM was $314,958 per QALY gained; and outpatient VP vs. CMM was $53,077 per QALY gained. These results confirmed that mortality was a key driver of model results. Varying the BKP and VP utility weights within their 95% confidence limits led to ICER ranges as follows: $37,152 to $53,321 (BKP inpatient); $8,698 to $15,022 (BKP outpatient); $34,673 to $47,357 (VP inpatient); and $9,599 to $17,566 (VP outpatient). Varying the age-specific relative

Table 2 Deterministic results

Treatment	Discounted results			Undiscounted results		
	Costs, USD	Life-years	QALYs	Costs, USD	Life-years	QALYs
BKP versus CMM; inpatient ($N = 2,071$)						
CMM	32,324	2.95	1.47	32,901	3.17	1.58
BKP inpatient	58,986	3.98	2.08	60,463	4.39	2.29
Incremental	26,662	1.03	0.61	27,561	1.22	0.71
ICER per LY	25,813			22,564		
ICER per QALY	43,455			38,661		
BKP versus CMM; outpatient ($N = 3,708$)						
CMM	24,234	6.16	3.08	25,923	7.17	3.59
BKP outpatient	32,972	7.50	3.88	35,829	8.96	4.63
Incremental	8,738	1.34	0.80	9,906	1.80	1.05
ICER per LY	6,544			5,508		
ICER per QALY	10,922			9,466		
VP versus CMM; inpatient ($N = 720$)						
CMM	31,005	2.95	1.47	31,569	3.17	1.58
VP inpatient	61,342	4.27	2.23	63,041	4.74	2.48
Incremental	30,338	1.32	0.76	31,472	1.57	0.89
ICER per LY	22,918			20,008		
ICER per QALY	39,774			35,240		
VP versus CMM; outpatient (1,042)						
CMM	23,789	6.03	3.02	25,437	6.99	3.50
VP outpatient	32,301	7.15	3.71	35,399	8.49	4.39
Incremental	8,512	1.12	0.69	9,962	1.50	0.89
ICER per LY	7,578			6,654		
ICER per QALY	12,293			11,151		

BKP, balloon kyphoplasty; *CMM*, conservative medical management; *ICER*, incremental cost-effectiveness ratio; *LY*, life-year; *QALY*, quality-adjusted life-year; *VP*, vertebroplasty

risks of subsequent fracture within their 95% confidence limits yielded the following ICER ranges: $42,375 to $45,400 (BKP inpatient); $10,004 to $11,970 (BKP outpatient); $38,516 to $41,494 (VP inpatient); and $10,712 to $13,915 (VP outpatient).

A full PSA was undertaken to evaluate the cumulative uncertainty for all model inputs. Cost-effectiveness planes confirm all model results, for each treatment comparison, fell into the upper-right quadrant with costs greater for the surgical cohorts but with higher QALY values relative to CMM (Fig. 2a–d). Probabilistic results were clustered around the deterministic result, with greater variability in the QALY outcomes relative to costs.

Figure 3a–d summarizes probabilistic results in CEAC format. At willingness-to-pay (WTP) thresholds greater than $45,000, BKP in the inpatient setting had a higher probability of being cost-effective relative to CMM. Using the typical US WTP threshold of $50,000 per QALY, BKP had an ~ 80% chance of being considered cost-effective across all model runs (Fig. 3a). For BKP in the outpatient setting at the US threshold of WTP, outpatient BKP had an ~ 100% probability of being considered cost-effective compared with CMM across all model runs (Fig. 3b). Results for inpatient and outpatient VP versus CMM were similar (Fig. 3c–d).

In further model validation, we compared the two-year Kaplan-Meier curves for survival derived using the CMS claims analysis vs. model predictions. Results were highly concurrent, with absolute differences between predicted and observed survival proportions ranging from 0 to a maximum of 2% for all comparisons. These results suggest that the model accurately captures the mortality effects associated with treatment.

Discussion

This analysis focused on a comparison of surgical intervention (BKP or VP) versus CMM in patients with osteoporotic VCF, in both inpatient and outpatient settings. In all 4 of the treatment scenarios tested, surgical intervention was predicted to

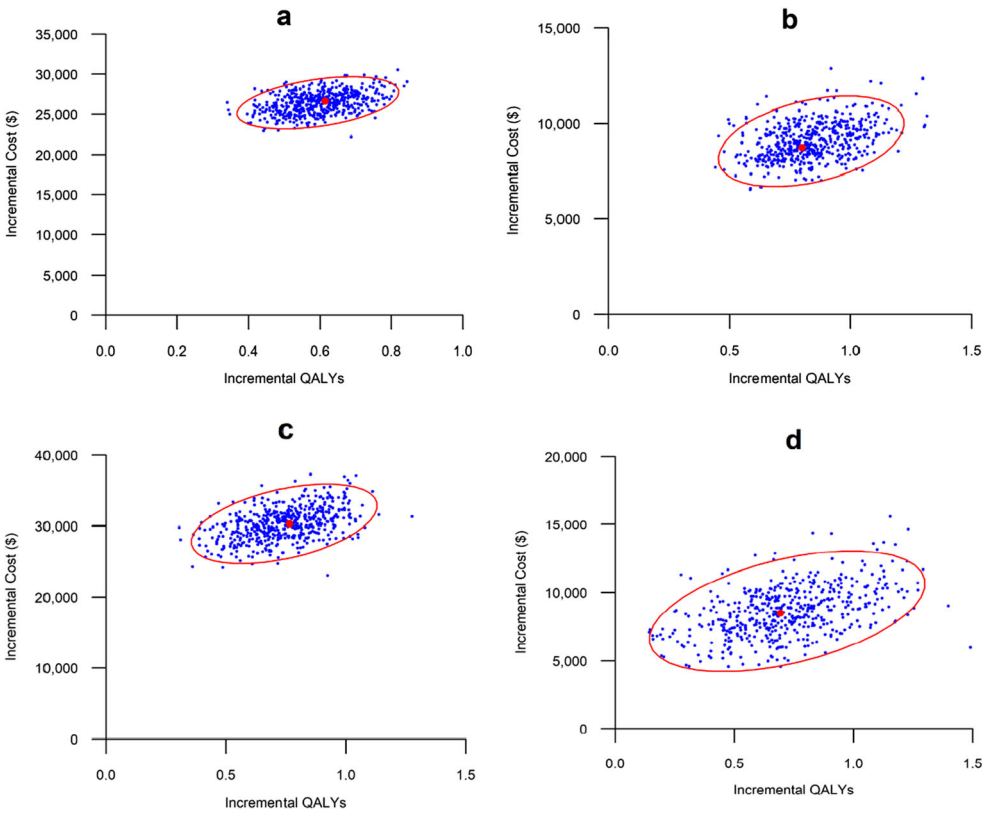

Fig. 2 Cost-effectiveness scatter plots. **a** BKP vs CMM (inpatient). **b** BKP vs CMM (outpatient). **c** VP vs CMM (inpatient). **d** VP vs CMM (outpatient)

be cost-effective compared to CMM at current WTP thresholds, with ICERs ranging from $11,000 to $43,000 per QALY gained (Table 2). BKP and VP had similar cost-effectiveness profiles compared with CMM over a lifetime horizon. This may be due to the lack of utility inputs specific to VP at two-year follow-up; our inputs relied on the BKP arm of FREE-2 trial data [24]. The results suggested that both BKP and VP are more cost-effective in the outpatient setting than in the inpatient setting. This effect was not solely due to the surgical location itself and the associated cost, but more importantly

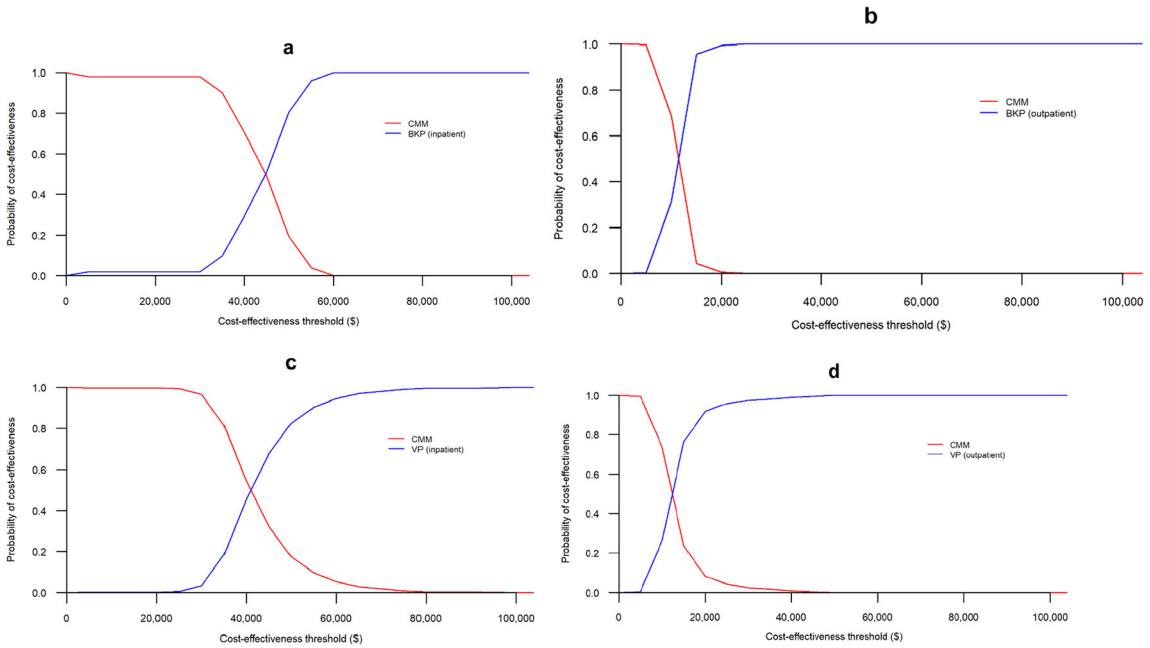

Fig. 3 Cost-effectiveness acceptability curves

the different patient-level clinical profiles who underwent treatment in each setting. Scenario analyses suggested that the model results were robust to changes in the utility weights (and their application to VP patients) and the relative risk of subsequent fracture.

The probabilistic analyses suggested a high probability of cost-effectiveness for both BKP and VP, compared with CMM (Figs. 2 and 3). Survival outcomes were validated by comparing 2-year survival projections from the model with empirical survival data from the CMS analysis, which indicated that the model could accurately replicate short-term mortality outcomes. The model assumption of a survival benefit associated with surgery is not only supported by other analyses of Medicare claims data [10, 12, 13] but also an analysis of claims from a German health insurance fund covering 2.4 million persons [28].

Our model was based largely on a similar analysis from the UK [17], and our findings that BKP is more cost-effective than CMM are consistent with the UK model. In that study, Svedbom and colleagues found that the ICER for BKP was estimated at GBP 2,706 per QALY compared with that of CMM. They performed a sensitivity analysis varying mortality and QoL benefits from BKP and confirmed that the cost-effectiveness of BKP versus CMM was robust.

Similar cost-effectiveness findings have been found in other studies of BKP. In a report that served as an earlier precursor to Svedbom and colleagues' model, Ström et al. found in 2010 that BKP was cost-effective compared with CMM in the UK, with an ICER of 8,840 GBP [16]. In 2011, Fritzell and colleagues performed a cost-effectiveness analysis using FREE trial data combined with Swedish societal costs (from the index procedure through 24 months follow-up), finding an ICER of SEK 884,682 (€92,154 or $134,043) [14]. While authors concluded that cost-effectiveness could not be demonstrated, cost data collected were limited, based on a total of 63 patients and reliant on hospital billing systems, hospital price lists, and patient diaries.

The only available US-based cost per life-year gained analysis was reported in 2012 by Edidin and colleagues [13]. Authors analyzed the cost-versus-mortality benefit alone (not incorporating quality-adjusted life-years), finding that the cost per life-year gained ranged from $1,863 to $6,687 for BKP and from $2,452 to $13,543 for VP compared with CMM.

In the present study, we re-built the model in Microsoft Excel to adapt it to the US setting. The model was populated with demographic data from a CMS claims analysis, US-specific VCF risks, and health state utilities from the FREE-2 study which were recalculated to reflect US preference weights. An analysis of CMS claims data provided all the model cost inputs, and in addition was used to estimate the treatment effect on patient mortality risk. To our knowledge, this is the first analysis to evaluate patient subgroups by treatment setting.

By stratifying patients into 4 groups based on treatment and inpatient versus outpatient setting, we were able to show that BKP and VP are cost-effective at a willingness-to-pay threshold of < $50,000 per QALY. Generally speaking, ICER values < $50 k are considered cost-effective treatments in the USA [29].

There are some limitations to the model which warrant discussion. Long-term patient outcomes were modeled using a combination of data from clinical studies (QoL data from the FREE-2 study) [24], observational studies (baseline fracture risk inputs and general-population utilities), and data from previous economic analyses (e.g., fracture risk adjustments). However, the benefits of surgery are expected to be obtained over a longer period of time than the two-year follow-up of the FREE-2 study, and thus, alternative data and assumptions were required to extrapolate the data to a lifetime horizon.

A propensity-score-matched analysis of CMS claims data was performed to identify similar patients for each pairwise treatment comparison. The matching exercise was performed using large numbers of patients to ensure comparability of the groups and generalizability to the patient population being studied. The patients in the CMS analysis were older than those in the FREE-2 study; however, the utility weights obtained from the FREE-2 study were adjusted for the age of patients in the CMS analysis, as were the ongoing risks of mortality and subsequent fracture. Thus, this difference in demographics has been addressed to the greatest extent possible.

Another limitation identified in this analysis concerns the mortality benefit associated with patients receiving BKP or VP. This effect was incorporated using hazard ratios from a Cox proportional hazards model from the CMS claims analysis. The model results indicate that the use of BKP and VP in the outpatient setting is cost-effective even if no mortality benefit is assumed; however, in the inpatient setting, the ICERs rise above accepted WTP thresholds for both BKP and VP versus CMM. The mortality benefit is therefore a key driver of cost-effectiveness, which, dependent on a patient's expected lifespan, will vary conclusions on cost-effectiveness at the individual level, rather than at the population level.

The population evaluated in this model included only the subset of patients receiving CMM that had demographic and clinical characteristics similar to those undergoing interventional treatment. Therefore, our conclusions around the cost-effectiveness of surgical interventions relative to CMM are limited to the subset of patients with similar profiles as patients selected for surgery. In an analysis of Medicare claims data by Ong et al., prevalence of VCF was 209,337 in 2014, while incidence of BKP and VP was 8,419 and 29,679 procedures respectively [30]. Therefore, there is a large population treated with conservative management (for a variety of

reasons) that were not evaluated in our present cost-effectiveness analysis.

In conclusion, the results of our model suggest that, from a US Medicare payer perspective, BKP and VP are both more costly than CMM in the near term, but are associated with improved overall survival and quality-adjusted survival, thereby making BKP and VP cost-effective compared with CMM. These results are consistent with previous, non-US cost-effectiveness analyses of BKP and VP. The observed mortality benefit associated with BKP and VP is a key driver of cost-effectiveness. It is noteworthy that the simulated patient population represents a subset of the total Medicare population. Therefore, the cost-effectiveness findings are generalizable to patients who were historically eligible for BKP or VP, including patients treated with CMM. While the results may not extend to patients who would have been considered poor candidates for surgery, this information is critically important to consider when creating policy and considering value-based payment models.

Acknowledgments The authors thank Jeanne McAdara PhD for her professional assistance with manuscript preparation.

Funding information Medtronic paid for CMS data used for model inputs and manuscript preparation.

Compliance with ethical standards

Conflicts of interest TJH received no funding from Medtronic for this study. SE, MQ, and CNR are full-time employees of Medtronic.

References

1. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A (2007) Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. J Bone Miner Res 22(3):465–475. https://doi.org/10.1359/jbmr.061113
2. Johnell O, Kanis J (2005) Epidemiology of osteoporotic fractures. Osteoporos Int 16(Suppl 2):S3–S7. https://doi.org/10.1007/s00198-004-1702-6
3. Kendler DL, Bauer DC, Davison KS et al (2016) Vertebral fractures: clinical importance and management. Am J Med 129(2):221.e1–221.10. https://doi.org/10.1016/j.amjmed.2015.09.020
4. Hübschle L, Borgström F, Olafsson G, Röder C, Moulin P, Popp AW, Külling F, Aghayev E, SWISSspine Registry Group (2014) Real-life results of balloon kyphoplasty for vertebral compression fractures from the SWISSspine registry. Spine J 14(9):2063–2077. https://doi.org/10.1016/j.spinee.2013.12.019
5. Musbahi O, Ali AM, Hassany H, Mobasheri R (2018) Vertebral compression fractures. Br J Hosp Med (Lond) 79(1):36–40. https://doi.org/10.12968/hmed.2018.79.1.36
6. Bliuc D, Nguyen ND, Milch VE, Nguyen TV, Eisman JA, Center JR (2009) Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. JAMA. 301(5):513–521. https://doi.org/10.1001/jama.2009.50
7. Yuan WH, Hsu HC, Lai KL (2016) Vertebroplasty and balloon kyphoplasty versus conservative treatment for osteoporotic vertebral compression fractures: a meta-analysis. Medicine (Baltimore) 95(31):e4491. https://doi.org/10.1097/MD.0000000000004491
8. Chandra RV, Maingard J, Asadi H, Slater LA, Mazwi TL, Marcia S, Barr J, Hirsch JA (2018) Vertebroplasty and kyphoplasty for osteoporotic vertebral fractures: what are the latest data. AJNR Am J Neuroradiol 39(5):798–806. https://doi.org/10.3174/ajnr.A5458
9. Garg B, Dixit V, Batra S, Malhotra R, Sharan A (2017) Non-surgical management of acute osteoporotic vertebral compression fracture: a review. J Clin Orthop Trauma 8(2):131–138. https://doi.org/10.1016/j.jcot.2017.02.001
10. Edidin AA, Ong KL, Lau E, Kurtz SM (2015) Morbidity and mortality after vertebral fractures: comparison of vertebral augmentation and nonoperative management in the Medicare population. Spine (Phila Pa 1976) 40(15):1228–1241. https://doi.org/10.1097/BRS.0000000000000992
11. Chen AT, Cohen DB, Skolasky RL (2013) Impact of nonoperative treatment, vertebroplasty, and kyphoplasty on survival and morbidity after vertebral compression fracture in the medicare population. J Bone Joint Surg Am 95(19):1729–1736. https://doi.org/10.2106/JBJS.K.01649
12. Hirsch JA, Chandra RV, Carter NS, Beall D, Frohbergh M, Ong K (2020) Number needed to treat with vertebral augmentation to save a life. AJNR Am J Neuroradiol 41(1):178–182. https://doi.org/10.3174/ajnr.A6367
13. Edidin AA, Ong KL, Lau E, Schmier JK, Kemner JE, Kurtz SM (2012) Cost-effectiveness analysis of treatments for vertebral compression fractures. Appl Health Econ Health Policy 10(4):273–284. https://doi.org/10.2165/11633220-000000000-00000
14. Fritzell P, Ohlin A, Borgström F (2011) Cost-effectiveness of balloon kyphoplasty versus standard medical treatment in patients with osteoporotic vertebral compression fracture: a Swedish multicenter randomized controlled trial with 2-year follow-up. Spine (Phila Pa 1976) 36(26):2243–2251. https://doi.org/10.1097/BRS.0b013e3182322d0f
15. Stevenson M, Gomersall T, Lloyd Jones M, Rawdin A, Hernández M, Dias S, Wilson D, Rees A (2014) Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for the treatment of osteoporotic vertebral fractures: a systematic review and cost-effectiveness analysis. Health Technol Assess 18(17):1–290. https://doi.org/10.3310/hta18170
16. Ström O, Leonard C, Marsh D, Cooper C (2010) Cost-effectiveness of balloon kyphoplasty in patients with symptomatic vertebral compression fractures in a UK setting. Osteoporos Int 21(9):1599–1608. https://doi.org/10.1007/s00198-009-1096-6
17. Svedbom A, Alvares L, Cooper C, Marsh D, Ström O (2013) Balloon kyphoplasty compared to vertebroplasty and nonsurgical management in patients hospitalised with acute osteoporotic vertebral compression fracture: a UK cost-effectiveness analysis.

Osteoporos Int 24(1):355–367. https://doi.org/10.1007/s00198-012-2102-y
18. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, Kuntz KM, Meltzer DO, Owens DK, Prosser LA, Salomon JA, Sculpher MJ, Trikalinos TA, Russell LB, Siegel JE, Ganiats TG (2016) Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. JAMA. 316(10):1093–1103. https://doi.org/10.1001/jama.2016.12195
19. Ettinger B, Black DM, Dawson-Hughes B, Pressman AR, Melton LJ (2010) Updated fracture incidence rates for the US version of FRAX. Osteoporos Int 21(1):25–33. https://doi.org/10.1007/s00198-009-1032-9
20. Klotzbuecher CM, Ross PD, Landsman PB, Abbott TA, Berger M (2000) Patients with prior fractures have an increased risk of future fractures: a summary of the literature and statistical synthesis. J Bone Miner Res 15(4):721–739. https://doi.org/10.1359/jbmr.2000.15.4.721
21. Marshall D, Johnell O, Wedel H (1996) Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ. 312(7041):1254–1259. https://doi.org/10.1136/bmj.312.7041.1254
22. Arias E, Heron M, Xu J (2017) United States life tables, 2014. Natl Vital Stat Rep 66(4):1–64
23. Wardlaw D, Cummings SR, Van Meirhaeghe J et al (2009) Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet. 373(9668):1016–1024. https://doi.org/10.1016/S0140-6736(09)60010-6
24. Boonen S, Van Meirhaeghe J, Bastian L et al (2011) Balloon kyphoplasty for the treatment of acute vertebral compression fractures: 2-year results from a randomized trial. J Bone Miner Res 26(7):1627–1637. https://doi.org/10.1002/jbmr.364
25. Fryback DG, Dunham NC, Palta M, Hanmer J, Buechner J, Cherepanov D, Herrington SA, Hays RD, Kaplan RM, Ganiats TG, Feeny D, Kind P (2007) US norms for six generic health-related quality-of-life indexes from the National Health Measurement study. Med Care 45(12):1162–1170. https://doi.org/10.1097/MLR.0b013e31814848f1
26. Szende A, Janssen B, Cabases J (2014) Self-reported population health: an international perspective based on EQ-5D. https://doi.org/10.1007/978-94-007-7596-1
27. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, Augustovski F, Briggs AH, Mauskopf J, Loder E (2013) Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. Value Health 16(2):e1–e5. https://doi.org/10.1016/j.jval.2013.02.002
28. Lange A, Kasperk C, Alvares L, Sauermann S, Braun S (2014) Survival and cost comparison of kyphoplasty and percutaneous vertebroplasty using German claims data. Spine (Phila Pa 1976) 39(4):318–326. https://doi.org/10.1097/BRS.0000000000000135
29. Shiroiwa T, Sung YK, Fukuda T, Lang HC, Bae SC, Tsutani K (2010) International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness. Health Econ 19(4):422–437. https://doi.org/10.1002/hec.1481
30. Ong KL, Beall DP, Frohbergh M, Lau E, Hirsch JA (2018) Were VCF patients at higher risk of mortality following the 2009 publication of the vertebroplasty "sham" trials? Osteoporos Int 29(2):375–383. https://doi.org/10.1007/s00198-017-4281-z