Socioeconomic value of orthopedic devices: evidence and methodological challenges

Corinna Sorenson1,2
Michael Drummond2,3
1LSE Health, London School of Economics, London, UK; 2European Health Technology Institute for Socioeconomic Research, Brussels, Belgium; 3Centre for Health Economics, University of York, York, UK

Abstract: With continued technological advances in orthopedic devices and increasingly limited health care resources, greater attention will be placed on substantiating the socioeconomic value of these devices. Therefore, this study focused on a systematic review of available economic evaluations of selected orthopedic devices (n = 33 studies) to assess their impact on different clinical and economic outcomes. The existing evidence suggests that they have important benefits to patients, including reduced risk of fractures, increased mobility and functioning, and enhanced quality of life, and do so cost effectively or with cost savings. However, we have identified several methodological obstacles to sufficient ascertainment of value, such as a lack of robust information on health economic outcomes and long-term evidence. We also identify areas where additional research is needed to assess more fully the value of orthopedic devices.

Keywords: medical devices, orthopedics, health economic evaluation

Introduction
With advances in health care and improvements in broader socioeconomic conditions, there is a growing elderly population in most countries. In the UK, for example, estimates suggest that the number of people aged 65 years and over could rise by more than 40 percent in the next 30 years.1 This will have a significant impact on health and social services, and create greater demand for prevention and treatment in orthopedics. New technological innovations in orthopedic care, such as hip implants, will therefore assume an important role in ensuring that the aging population remains active and independent and that younger populations maintain their mobility and productivity over time.

As demand for these treatments proliferates, there will be an increased focus on substantiating their impact on costs and health outcomes. Indeed, there is growing interest in the economic evaluation of health technologies to demonstrate that their benefits outweigh the costs. In some countries, economic evaluations are formally required as part of the reimbursement process, but this is largely confined to pharmaceuticals.2 However, increasingly, the evaluation of medical technologies is garnering interest among policy makers and established health technology assessment agencies. For example, the technology appraisal program of the National Institute for Health and Clinical Excellence (NICE) in England and Wales considers all health technologies, including devices.

Consequently, there is a growing need for medical devices in the orthopedic sector to demonstrate socioeconomic value, because health care decision-makers are...
keen to maximize the benefits from the use of their budgets. However, in general, there tends to be minimal existing evidence examining the value of medical devices, as well as limited discussion around how such value is (or can be) best ascertained. This is especially true of orthopedic devices, where, for example, there are numerous studies on hip replacement, but few studies available examining their actual use.

Although the case for strengthening the available evidence base on orthopedic devices is fairly strong, there are several well documented challenges in undertaking these evaluations.\(^3\)\(^4\) In particular, it has been suggested that devices pose significantly greater challenges than the evaluation of drugs, which, as previously noted, are more often included in formal requirements.

To address some of these issues, this paper focuses on assessment of the socioeconomic value of medical devices in the field of orthopedics, based on a systematic review of available relevant economic evaluations. The review had four principal aims: to review the evidence on how orthopedic devices impact on various dimensions of value (eg, health outcomes, cost); to identify the main methodological challenges faced by those undertaking assessments; to identify the situations where orthopedic devices are likely to demonstrate good value and where they are not; and to understand the primary ways in which devices differ from drugs and how such variations might be taken into account when assessing value. We begin by outlining the research methods, followed by a discussion of the findings across the aforementioned aims.

**Materials and methods**

Economic evaluations were identified by searching the NHS Economic Evaluation Database (NHS EED). The NHS EED conducts comprehensive literature searches of health and social science databases (eg, Medline, Embase, Scopus), identifying studies which explore the economic aspects of health care treatments and programs. Studies that are considered to be full economic evaluations, ie, those comparing the costs and consequences of alternative health care programs and treatments, are reviewed and a structured abstract is produced.\(^3\) It currently contains around 7000 quality-assessed abstracts of full economic evaluations. Given its systematic review of the literature and thorough classification of studies, NHS EED was considered to be a comprehensive and authoritative source of economic evaluations and a reliable and efficient way of identifying studies for the review.

We used a broad initial search strategy. The NHS EED database was canvassed using the following search terms: “orthopedic”, “hip”, “knee”, “shoulder”, “ankle”, “elbow”, “arthroplasty”, and “joint”. The search was not restricted by publication date; however, all relevant available studies were published between 1990 and 2009 at the time of the search. All available abstracts were reviewed, with duplicate abstracts identified and eliminated. In addition to the initial and primary search of NHS EED, a subsequent search was conducted, because NHS EED is known to incorporate a time lag between date of publication and review and inclusion of studies in the database. In light of this potential limitation, we rereviewed the database six months following our initial search to ensure inclusion of all relevant articles.

After the abstracts had been identified, the full published papers were obtained and reviewed. We used a standardized data extraction form, developed in Microsoft Access, to review each article, facilitate data extraction, and subsequent analyses. The form was developed to meet the research aims and was based on a variety of sources, including the systematic review guidelines produced by the organization that oversees NHS EED.

Two trained reviewers read each article, extracted the relevant data, and then convened a consensus review to resolve any discrepancies; the reviewers were in agreement in 95% of cases. Reviewers were not masked to the identity of the authors or the journal where the study was published. For each economic evaluation, the descriptive characteristics collected included year of publication, journal, country of origin, intervention type, comparator interventions, and study funding source. Information on the methods used, including type of analysis, health outcome measures and inclusion of costs, and results obtained (eg, effectiveness, costs, cost-effectiveness) were also gathered. Finally, any methodological challenges identified in the evaluation were noted.

Several analyses of the dataset were performed. These included identification of the types of devices evaluated, the types of economic evaluation conducted, the cost-effectiveness of devices over time, the relationship between the source of study funding and the level of cost-effectiveness reported, and the methodological problems identified.

**Results**

**Basic study details**

A total of 33 relevant economic evaluations published from 1996 to 2008 were identified and reviewed (Appendix 1). Of the total evaluations, 19 related to hip-specific devices (eg, hip implants, hip protectors), six concerned the knee...
(eg, knee implants), two addressed both hip and knee implants, and the remainder assessed other relevant devices, such as ankle and shoulder replacements. Within these broad categories, there was substantial variation in the technology evaluated and the comparator(s) used across studies. Total hip replacement was compared with a range of alternative treatment options or strategies, including total hip replacement with a waiting period (as compared with immediate replacement), one-stage replacement (versus two-stage replacement), nonoperative management, and cemented and hydroxyapatite-coated hip implants (as compared with cementless and noncoated implants, respectively). Different types of implants were also evaluated. However, with regard to hip protectors, these were compared with no treatment or intervention in all studies. The range of comparators for knee implants was also diverse, including implantation versus nursing home placement, revision total knee replacement (versus primary total knee replacement), manual replacement (versus computer-aided), and minimally invasive techniques and unicompartmental knee arthroplasty compared with total knee replacement, respectively.

Table 1 provides a summary of study characteristics for the reviewed evaluations. The rate at which economic evaluations of orthopedic devices were published gradually increased over time at around 2–3 studies per year until the mid 2000s, when there was almost a doubling in annual publications (Figure 1). Although it appears that there were no evaluations published in 2009, this may be due to the lag time between publication and inclusion in the NHS EED database, as discussed earlier. Based on the analysis, there were no identifiable trends in terms of which orthopedic devices were being evaluated over time.

Study methods

The vast majority of studies adopted a payer or hospital perspective, focusing on direct health care costs only. Only one study adopted a broader societal perspective, examining indirect costs and benefits. Five additional studies purported that a societal perspective was used, but upon review, only direct costs and benefits were included and examined in the analyses.

In terms of the approach used to assess value, about one quarter (27%) were cost-consequence analyses, where differences in cost between the use of the device and its comparator were compared and presented alongside comparisons of other outcomes. Another 15% were cost-effectiveness analyses, where differences in costs and the major outcome, measured in terms of clinical effects, were assessed (eg, cost per hip fracture avoided). Multiple endpoints were used to assess outcomes, with treatment effects expressed most frequently in mortality, preoperative and postoperative complication and infection rates, clinical knee or hip rating scores, health-related quality of life (eg, pain, mobility, sleep), and the incidence of fractures or dislocations. In the case of hip protectors, in particular, patient compliance with wearing the protector, patient satisfaction, and the efficacy of the protector were also considered. As can be inferred from the limited number of studies assuming a societal perspective, direct costs were predominantly assessed, which most frequently included procedural costs (including any repair or revision surgical costs), hospital, rehabilitation and follow-up costs, and the

| Table 1 Characteristics of orthopedic economic evaluation literature |
|---------------------------------|----------------|----------------|
| Characteristic                  | Studies (n = 33) | Percentage of studies |
| **Device type**                 |                |                  |
| Hip implants                    | 11             | 34              |
| External hip protectors         | 8              | 24              |
| Knee implants                   | 6              | 18              |
| Hip and knee implants           | 2              | 6               |
| Other                           | 6              | 18              |
| **Country of study**            |                |                  |
| United States                   | 15             | 46              |
| United Kingdom                  | 7              | 21              |
| Finland                         | 2              | 6               |
| Germany                         | 2              | 6               |
| Other                           | 7              | 21              |
| **Sponsorship**                 |                |                  |
| Professional associations       | 6              | 18              |
| Government                      | 5              | 15              |
| Nonprofit organizations         | 2              | 6               |
| Industry                        | 2              | 6               |
| Academia                        | 1              | 2               |
| Multiple sources of funding     | 4              | 12              |
| No funding                      | 2              | 6               |
| Did not specify                 | 11             | 34              |
| **Study theme**                 |                |                  |
| Elderly                         | 11             | 34              |
| Nursing home residents          | 5              | 15              |
| Children                        | 1              | 2               |
| None/not stated                 | 20             | 61              |
| **Journal**                     |                |                  |
| Journal of Bone and Joint Surgery | 7        | 21              |
| Journal of the American         | 3              | 9               |
| Geriatrics Society              |                |                  |
| Seminars in Arthroplasty        | 3              | 9               |
| International Journal of Technology | 5      | 15              |
| Assessment in Health Care       |                |                  |
| Orthopaedics                    | 2              | 6               |
| Other                           | 13             | 40              |

Note: *Nonexclusive.
cost of the device itself. Only one study actually included indirect costs, expressed in the wages foregone by employed patients receiving a total hip replacement. There were no studies that examined long-term or lifetime costs.

In the largest proportion of studies (49%), the various outcomes were combined in a single generic measure of health gain, ie, the quality-adjusted life-year (QALY) in a cost-utility analysis. This approach is favored by several of the government health technology assessment agencies, including NICE, because it allows comparisons of cost-effectiveness (expressed in terms of the incremental cost per QALY gained) across different areas of care. In an additional 6% of studies, QALYs gained were assessed in combination with another summary measure of benefit, most commonly the number of hip fracture-related deaths and hip fractures avoided. The remaining studies used a cost-minimization (3%) approach.

Health outcomes and costs

Given the range and diversity of the studies reviewed, it is difficult to provide meaningful analyses of the therapeutic benefit and costs across these technologies. However, based on the evidence, a number of general findings can be discussed. In the case of hip protectors, the evidence suggests that their use reduces the mean fracture risk by around 4% (2.3%–6.5%) and protects against hip fracture-related deaths. Decreased risk of ill health and improved mobility resulted in a mean QALY gain of 0.02 (range 0.0074–0.0406).

In all cases, women gained QALYs by use of hip protectors, but in a couple of studies there was some loss in QALYs in men due to the inconvenience of wearing the protector. Total hip and knee replacement brought about improvements in clinical rating scores, ranging from an additional 25–35 points in hip scores and 15–56 points in knee scores. In addition, improvements in functional health and well-being scores were found across several studies, as well as in other health-related quality of life measures, eg, pain, sleep, and mobility. Such improvements resulted in an average gain in QALYs of 2.01 (−0.02 to 6.88).

Similar to estimates of therapeutic benefit, there was significant variability in cost estimates, depending on the technology under review, comparator(s) used, patient population, study time horizon, and cost methodology used. However, in the case of hip protectors, use almost always resulted in cost-savings ($68–$230 per person). This was also true for studies that examined total hip replacement and total knee replacement against no intervention or nonoperative management, but there were too few studies (n = 2) to arrive at any substantive conclusions.

In a few studies, the cost of the device itself was deemed to be a primary cost driver. However, device prices generally change over time, due to arrival of new products on the market, iterative developments, or ways in which they are procured in different health systems, and the short time horizons used in the majority of studies would not reflect price reductions over time. This differs from the case of
Evidence of value for money

Evidence of value for money is most easily interpreted in the case of studies calculating the incremental cost per life-year or QALY gained, because standards for judging this type of cost-effectiveness analyses are available. For example, NICE uses a cost-effectiveness threshold of £20,000 per QALY gained, with a justifiable range rising to £30,000 per QALY, depending on specific circumstances. Only rarely does it approve a technology for use in the UK National Health Service with an incremental cost-effectiveness ratio (ICER) exceeding £30,000 per QALY. Similar benchmarks are available for other jurisdictions; in many European countries, an ICER of less than €50,000 per QALY is typically viewed as cost-effective, as is an ICER of less than $100,000 per QALY in the US.28

To examine value for money, we focused on hip protectors and cementless hip implants, which allowed more consistent comparisons given that there was standardization in the comparator used (ie, no intervention and cemented implants, respectively). The ICERS under both scenarios are well below international thresholds. In the case of hip protectors, where cost per QALY was measured (n = 5 studies), protectors were superior to no intervention across all studies, meaning their use resulted in cost savings, while preventing hip fractures and providing gains in QALYs.9-13 For the two studies comparing cementless and cemented hip implants, the ICER ranged from $91 to $1815 (estimates converted to 2010 US dollars).17,29

However, these summary data often disguise a much more complex situation. In particular, most economic evaluations are subject to considerable uncertainty in the input parameters. For example, whereas a trial-based estimate of the relative effectiveness of the device is typically known with a given confidence interval, extrapolation of effectiveness into the future normally involves considerable uncertainty. The same may be true of other factors, including the incidence of adverse events or complications. The standard approach for dealing with parameter uncertainty is to undertake a sensitivity analysis, altering key parameter values in order to assess how they impact study results. Overall, this review highlighted a number of parameters that could impact the mean cost-effectiveness ratio. For hip protectors, these included baseline incidence of fractures, cost of protectors, and utility values.9-13,24 While the cost-effectiveness ratio was most sensitive to estimated revision rates, implant costs, utility values, and patient characteristics, namely age and gender, in studies of hip implants.22,29-31 In those evaluations assessing other devices (eg, ankle and knee implants), the durability of the implant and the estimated utility values were most influential on cost-effectiveness.20,21,32,33 However, no sensitivity analysis was undertaken in more than one third of the total studies, which is not in line with current methodological standards for economic evaluation.

Another factor potentially influencing the results in the case of hip protectors is the use of drugs for osteoporosis. However, none of the studies specified the extent to which patients were also receiving such treatments, resulting in uncertainty regarding possible impacts on health outcomes and costs.

Situations in which devices deliver greater economic value

Discussion of the cost-effectiveness results above shows that the main question is not one of whether medical devices deliver economic value, but the circumstances in which they do so. To answer this question precisely, it would be necessary to examine each application of a given device in detail. However, it is possible to make several general conclusions for the purposes of this review. Some of these are fairly obvious. For instance, the greater the relative treatment effect of the device compared with the alternative, the lower the ICER. Similarly, the higher the incremental cost of the initial procedure compared with the alternative, the higher the ICER.

In addition, the baseline risk in the patient population treated is often an important factor. This is because economic evaluations compare the absolute improvement in clinical outcome with the increased cost. Therefore, even if the device has the same relative treatment effect in a low-risk population, it will not be as cost-effective as in a high-risk population. Indeed, across studies it was generally found that device use was more cost-effective in high-risk patients, which is mainly due to reductions in mortality and greater improvements in mobility or functioning and health-related quality of life. Studies defined high-risk differently, but commonly used indicators including greater baseline incidence of fracture and postoperative relative risk of fracture, as well as prior lack of success or response to medical therapy.8,10,21-23 In some cases, high-risk related to the age and gender of patients, where cost-effectiveness was greatest in younger and female patient
populations. For example, given that women are more at risk for hip fracture than men and become so at an earlier age, the use of hip protectors was determined to be of most benefit if initiated earlier (ie, 75–80 years) in women and later in men (ie, 85 years).

Finally, there are some important factors relating to specific devices. With particular regard to hip protectors, the rate of patient compliance with wearing a protector had an influence on value for money. Protectors were found to be most cost-effective when patient adherence was not lower than 50%.

To this end, one study demonstrated greater cost-effectiveness in institutionalized populations, including patients residing in nursing homes, where staff can encourage appropriate use. Of note, the effectiveness of hip protectors depends, in part, on whether they are properly fitted. Moreover, for total hip and knee replacement, maximum benefit was attained when the need for revision was minimized (ie, less than 5% or the device maintained durability for at least ten years).

### Challenges in economic evaluation of orthopedic medical devices

One of the main motivations for undertaking this review was to document the main methodological challenges in evaluating medical devices. This issue is particularly pertinent in some countries, such as the UK, where the same type and quality of economic evidence that is required for pharmaceuticals is being suggested for devices.

It can be seen from Figure 2 that the predominance of non-randomized studies was one of the most frequent challenges. As a result, several studies noted that use of observational data or a retrospective design may diminish the robustness of the analysis. The other most frequently mentioned challenges included uncertainty around the data parameters used and the concern about the representativeness of the patient population or study setting, which would impact the generalizability of the findings. In relation to the issue around data uncertainty, several researchers mentioned the fact that some of the data needed for the analyses were either not available or were of too poor quality to use with any confidence. Short study time horizons and small sample sizes in the clinical studies in this field were also highlighted. Moreover, for many of the studies, important outcomes, including quality of life, a full range of relevant costs, and long-term outcomes, were not collected, which may limit their ability to capture fully the value of these devices. While few analysts gave specific reasons why such outcomes were not considered, it was noted across several studies that the summary benefit measure used (eg, fractures avoided) did not capture quality of life and, in some cases, there were inherent deficiencies in some of the measures used, such as clinical scoring and health-related quality of life instruments. In both cases, these challenges may limit robust assessment of the costs and benefits of a device.

At first glance, most of these challenges mirror those found in economic evaluations of pharmaceuticals. However, one of the challenges noted in the review - the rapid evolution of devices - is specific to this field. This was considered to

---

**Figure 2** Key methodological challenges to economic evaluation of orthopedic devices.

**Abbreviation:** HRQoL, health-related quality of life.
pose difficulty in capturing the value of technology adequately. Indeed, iterative development often means there is no substantial “steady-state” period during which the device can be evaluated in a controlled clinical study and, thus, no foundation of clinical evidence can be accumulated. For many devices, the lack of randomized controlled trials would indeed present a major challenge. In addition, while not explicitly outlined as a challenge, one of the key issues in evaluating the value of devices is the fact that the performance and cost of devices often depends not only on the device itself, but also on how it is used, which often entails a learning curve for the user, such as the physician or surgeon and, in the case of hip protectors, the patient. For example, in the case of hip protectors, it was noted that the importance of patient adherence was discussed as important to ensure optimal user performance and utility derived from protectors.

There is some evidence showing that the performance of users improves over time, so it is important to assess the value of a device after an average performance level has been reached. In addition, costs may reduce over time, provided that the relevant hospital staff secure and maintain proficient use of the new device or procedure and resources are organized to support the introduction of the technology effectively. Given that these challenges are somewhat unique to devices and the short duration of the majority of the studies, it was surprising that only two analyses mentioned such issues.

**Discussion**

While there is evidence available for the effectiveness of various medical technologies used in orthopedics, there is limited evidence on their socioeconomic value, especially across multiple devices. Our study aimed to meet this gap through a systematic review of available economic evaluations of orthopedic devices.

The existing evidence for select orthopedic devices suggests that they improve the functional status and health-related quality of life of patients and, in the context of hip protectors, reduce the risk of fracture. Moreover, they achieve these benefits at good value for money, and for hip protectors, with cost savings. Value for money continued to be demonstrated with variations in the data parameters used for evaluation across all technology types. However, maximum value was attained when orthopedic devices were used in populations with a greater baseline incidence of fracture, in women, and in older patients. As one might expect, hip protectors were most cost-effective when patients adhered to regular use and, in the case of prosthetic implants, when a low rate of revision was achieved.

In addition to examining the socioeconomic value of orthopedic devices, we identified potential methodological challenges in assessing the costs and benefits of such technologies. Understanding such issues will be of increasing importance, because decision-makers are starting to require formal evaluation of medical devices for decisions regarding resource allocation. One of the most important challenges was the lack of robust randomized controlled trials, partly because of the nature of devices, where blinding of patients is not feasible, as well as the fact that too few studies address all relevant outcomes, particularly health economic endpoints, such as, quality of life and indirect costs. Furthermore, the studies available are characterized by short time horizons and small sample sizes, which may not capture the full costs and benefits of a device to patients and the health care system. Some of the other challenges identified were unique to medical devices, where the value of a device depends somewhat on the skills, knowledge, and abilities of the user (ie, performance “in use”) and are frequently undergoing modifications or iterative development to improve performance. This makes it difficult for devices to be evaluated adequately in clinical studies or at least until an average performance level has been attained.

It is important to note that there are a number of limitations to our study. First, while the review included quality-assessed economic evaluations, the NHS EED database is not an entirely conclusive source of all available economic evaluations on orthopedic devices. We did not include, for example, health technology assessment reports published by NICE and other similar agencies, although these studies would be included if they were also published in the peer-reviewed journals searched by the NHS EED. Furthermore, although we strived to address the potential time lag between study publication and inclusion in the NHS EED, we may not have captured all of the most recent evaluations. Second, our broad approach of considering all orthopedic devices in the review introduced some challenges. There was considerable diversity in the studies in terms of technologies examined, comparators, and summary benefit measures (eg, QALYs gained versus number of hip fractures avoided in the case of hip protectors) used, and patient groups evaluated. Therefore, deriving robust conclusions regarding the costs and benefits across medical technology used in the sector proved difficult. In addition, we did not evaluate the merits of clinical or modeling assumptions included nor were we able to assess the quality of the data used in the evaluations. However, all studies included in the NHS EED are assessed for quality. Finally, some of the results presented, namely on cost-effectiveness, are not static, and
the costs of the devices and their associated benefits may have changed since publication of the studies. However, if changes have occurred over time, it is most likely that costs have declined and/or that performance has improved with iterative development of technology. Therefore, conclusions around value for money of orthopedic devices will presumably stay the same, if not improve.

Finally, the review pointed to areas where additional and more robust evidence is needed to measure fully the value of orthopedic devices. First and foremost, for an important area of medical technology, it was surprising to find such a dearth of studies evaluating the costs and benefits across a variety of orthopedic devices. In addition to limitations in the number of studies, more research is needed on long-term health outcomes and costs. Also, as noted earlier, there was insufficient inclusion of all relevant costs across studies, especially with regard to consideration of medical costs that would have occurred in the absence of replacement surgery. This coincides with the fact that relatively few studies assessed implantation of a prosthesis against no treatment, and conducting such an evaluation would raise ethical concerns. There was also a lack of studies evaluating devices from a societal perspective and, thus, there is limited evidence of their indirect impact, eg, productivity gains, costs and benefits to carers of these technologies. This is particularly important in the case of orthopedic devices given their contributions to improved quality of life, functionality, and enhanced independence in life activities. Given that particular devices (eg, hip and knee implants) tend to be used in younger and employed populations, enhanced mobility has a positive impact on broader economic outcomes, namely productivity. To that end, there were no evaluations examining use of orthopedic devices in younger populations (ie, 40–60-year-olds), which needs further inquiry to substantiate the value for money these devices can provide in such cases. Use in younger populations may lead to earlier return to work, prevention of future fractures or other adverse events, and better functioning over a greater number of years, but such benefits may come with an increased need for revisions or multiple replacements over time. Overall, there is a need to understand better which patient populations benefit most from these technologies.

There were surprisingly few studies that took patient perspectives into account, in terms of perceived benefits and risks of orthopedic devices and preferences regarding their use. Given that use of these prostheses is elective in some cases, better understanding of patient preferences and factors driving their use is important. Such information would help to improve clinical decision-making on the part of both patients and providers, and would likely have important implications for more cost-effective use of these technologies. Finally, although there tends to be a general lack of evidence available concerning medical devices in actual use, there are increasing numbers of patient registries collecting considerable amounts of data on hip and knee replacements, especially with regard to survival rates. There are potential challenges with these data around the most effective use of such information in economic evaluations. For the reasons mentioned earlier, there are a number of unique methodological challenges associated with assessing medical devices. However, in view of these limitations and evidence gaps, our study provides some initial evidence on the value of orthopedic devices, which is an important starting point given that, to date, existing reviews of health technologies have predominantly focused on pharmaceuticals.

Disclosure

The authors report no conflicts of interest in this work.

References


**Appendix table**

**Appendix 1 Studies included in review**

March et al\textsuperscript{25}
Rasanen et al\textsuperscript{22}
Keating et al\textsuperscript{10}
Bozic et al\textsuperscript{26}
Honkanen et al\textsuperscript{12}
SooHoo et al\textsuperscript{20}
Slover et al\textsuperscript{21}
Coon et al\textsuperscript{42}
Ganjour et al\textsuperscript{13}
McBryde et al\textsuperscript{24}
Duwelius et al\textsuperscript{6}
Chang et al\textsuperscript{13}
Rorabeck et al\textsuperscript{18}
Clinkscales and Peterson\textsuperscript{40}
Rissanen et al\textsuperscript{19}
Lorenze et al\textsuperscript{17}
Givon et al\textsuperscript{19}
Gartsman et al\textsuperscript{19}
Kumar and Parker\textsuperscript{24}
Segui-Gomez et al\textsuperscript{11}
O'Shea et al\textsuperscript{14}
Colón-Emeric et al\textsuperscript{10}
Singh et al\textsuperscript{17}
SooHoo and Kominski\textsuperscript{23}
McKenzie et al\textsuperscript{22}
Briggs et al\textsuperscript{31}
Fielden et al\textsuperscript{15}
Meyer et al\textsuperscript{8}
Dong and Buxton\textsuperscript{18}
Honkanen et al\textsuperscript{14}
Karuppiah et al\textsuperscript{14}
Papakonstantinou et al\textsuperscript{23}
Marinelli et al\textsuperscript{19}