Eur Spine J (2009) 18 (Suppl 3):S338–S347
DOI 10.1007/s00586-009-0939-3

ORIGINAL ARTICLE

The impact of economic evaluation on quality management in spine surgery

Norbert Boos

Received: 22 December 2008 / Accepted: 5 March 2009 / Published online: 1 April 2009

Abstract Health care expenditures are substantially increasing within the last two decades prompting the imperative need for economic evaluations in health care. Historically, economic evaluations in health care have been carried out by four approaches: (1) the human-capital approach (HCA), (2) cost-effectiveness analysis (CEA), (3) cost-utility analysis (CUA) and (4) cost-benefit analysis (CBA). While the HCA cannot be recommended because of methodological shortcomings, CEA and CUA have been used frequently in healthcare. In CEA, costs are measured in monetary terms and health effects are measured in a non-monetary unit, e.g. number of successfully treated patients. In an attempt to develop an effectiveness measure that incorporates effects on both quantity and quality of life, so-called Quality Adjusted Life Years (QUALYs) were introduced. Contingent valuation surveys are used in cost-benefit analyses (CBA) to elicit the consumer's monetary valuations for program benefits by applying the willingness-to-pay approach. A distinguished feature of CBA is that costs and benefits are expressed in the same units of value, i.e. money. Only recently, economic evaluations have started to explore various spinal interventions particularly the very expensive fusion operations. While most of the studies used CEA or CUA approaches, CBAs are still rare. Most studies fail to show that sophisticated spinal interventions are more cost-effective than conventional treatments. In spite of the lack of therapeutic or cost-effectiveness for most spinal surgeries, there is rapidly growing spinal implant market demonstrating market imperfection and information asymmetry. A change can only be anticipated when physicians start to focus on the improvement of health care quality as documented by outcome research and economic evaluations of cost-effectiveness and net benefits.

Keywords Costs · Cost analysis · Cost-effectiveness · Cost-benefit analysis · Economic evaluations · Low back pain · Lumbar spinal fusion · Value-based competition

Introduction

Recent epidemiological studies have reported a life time prevalence of low back pain (LBP) of up to 84% [11]. Even though the rate of those individuals who develop chronic LBP (cLBP) leading to disability is estimated to be only about 1%, costs caused by restricted activity and bed days due to chronic LBP (cLBP) as well as its treatment costs have become a substantial socioeconomic burden in western industrial countries [10, 27, 61]. Related to gross domestic product (GDP), the total average health expenditure share across OECD countries increased from 6.6% in 1980 to 9.0% [range 15.3% (USA) to 6% (South Korea)] in 2005 (http://www.oecd.org/health/healthdata).

A cost-of-illness study of back pain in the Netherlands estimated the total direct medical costs of back pain at USD 367.6 million which is about 1.7% of the GDP or nearly one-sixth of the 1991 health care expenses of the Netherlands [86]. The major part of direct medical costs, i.e. USD 200 million, was caused by hospital costs. However, the total indirect costs were found to be USD 4.6 billion of which USD 3.1 billion were due to absenteeism of work and USD 1.5 billion due to disablement indicating that 93% of the costs due to back pain are

N. Boos (✉)
Centre for Spinal Surgery, University of Zurich,
University Hospital Balgrist, Forchstrasse 340,
8008 Zurich, Switzerland
e-mail: norbert.boos@mysmx.ch

indirect costs. Coyte et al. [13] found a sum of USD 8.1 billion for Canada in 1994 corresponding to 1.07% of the GNP. Similar figures were reported in a 'cost-of-illness' study from the UK exploring the socio-economic costs of back pain [61]. Maniadakis et al. [61] estimates the direct health care cost of back pain in 1998 to be £ 1,632 million. Approximately 35% of this cost relates to services provided in the private sector and therefore is most likely paid directly by patients and their families. However, the direct cost of back pain was found to be insignificant compared to the cost of informal care and the production losses related to it, which total £ 10,668 million. With these figures, back pain imposes a greater economic burden than other diseases such as coronary heart disease, Alzheimer's disease, stroke, insulin-dependent diabetes mellitus, epilepsy, benign prostatic hyperplasia, diabetes, multiple sclerosis, lower respiratory tract infections, deep vein thrombosis and pulmonary embolism, depression, critical limb ischaemia, and migraine [61]. Even though these results cannot be directly compared to other countries, they may give an idea of the economic impact of back pain in western industrial countries.

Spinal surgery for persistent LBP is one of the most rapidly growing surgical disciplines in medicine substantially increasing health care costs for one of the most common medical problems. The most frequently performed spinal interventions are lumbar discectomy for a disc herniation, spinal decompression for spinal stenosis and spinal fusion for degenerative disc disease [85]. A variety of factors contributed to the rising frequency of spinal surgery over the last decade. Changes in the population (increasing super-annuation), technological advances (improved anaesthetic techniques, pedicle screw fixation devices, fusion cages), and uncertainty regarding indications, as well as the financial incentives for surgeons, hospitals and the implant industry may have synergistic effects [16].

In spite of the frequent use of these interventions, scientific evidence for the most frequently performed spinal is sparse [85]. Debate is still continuing on the therapeutic efficacy of these interventions to cure back problems compared to natural history and non-operative treatment. Particularly, no convincing evidence can be found to support the use of spinal fusion over a non-operative cognitive behavioural treatment approach for degenerative disc disease [9, 21, 49]. Nevertheless, this intervention is one of the most frequently performed. The lack of scientific evidence for spinal fusion in degenerative disc disease has recently prompted the discussion whether spinal fusion should become a case for restraint [16]. In an area of limited financial health care resources, there is an increasing demand not only to analyse the therapeutic efficacy of the treatment modalities but also to consider related economic and societal costs.

Theoretical background of economic assessments

To better understand economic evaluations it is necessary to review two of the most important principles of economics. Scarcity is a fundamental fact of life. Economics can therefore be defined as the study of how people make choices under conditions of scarcity and of the result of those choices for society [22]. The scarce resources in healthcare prompt the urgent need for comprehensive economic evaluations to efficiently allocate sparse resources. Fundamental in economics is the cost-benefit principle indicating that an individual (or a firm or a society) should take an action if the extra benefit from taking the action is at least as great as the extra benefit [22].

Economic valuation of public goods

An economic evaluation is a comparison of alternative actions in terms of their costs and benefits, and the purpose of economic evaluation is to answer about allocation of resources [47]. Environmental goods, such as air quality, water quality and bio-diversity have much in common with health, in that market failures indicate that government intervention in the market is often desirable. The fact that government intervention occurs causes economists to wish for the benefits of such intervention to be compared with the costs [37]. Early forms of such cost-benefit analyses were introduced in the USA during the "Great Depression" of the 1930s. Later on, it has been used more and more in environmental economics as well as in medical care.

Today these methods have gained increased acceptance among academic economists as well as policy-makers. They are widely considered to be versatile and powerful methodologies for valuing non-traded goods and services [37]. Particularly for environmental field and public transport guidelines form the National Oceanic and Atmospheric Administration (NOAA) have been proposed on how to conduct such analyses [37, 76].

Economic evaluations in health care

Economic evaluations are made at different levels in the health care system, i.e. by patients, by doctors, by hospitals and by third party-payers [47]. The purpose of performing health economic evaluations must be to provide decision-makers—be they physicians, health administrators or politicians—with information that can support their allocation of scarce resources [2]. Quality of life is an important economic benefit, something we are willing to pay for, individually and/or collectively [47]. Since resources are limited, however, we have to choose among different interventions, within or outside the healthcare system, that all have the potential for improvement in quality of life

[47]. For such evaluations to be useful in the eyes of decision-makers, the studies must be adapted to the decision-making context of readers and not only to methodological demands [2].

Historically, several strategies have been applied in analysing programs with non-monetary effects [82].

Human-capital approach

In 1676, Sir William Petty calculated that better medicine in England would save 200,000 lives, lives he wished to value monetarily to argue that the requisite funds would be well spent [82]. This approach to economic evaluation is usually referred to as the human-capital approach [6]. In the standard human-capital approach, it is assumed that the value of society of an individual's life is measured by future production potential, usually calculated as the present discounted value of expected labour earnings [56]. The human-capital approach cannot be recommended as a basis for measuring the value of improved health in economic evaluations, because (1) it is not rooted in the theoretical foundations of welfare economics, (2) it discriminates people not in the labour force, since increase in production are measured as wages earned, and (3) ignores the intrinsic value of good health and quality of life [6, 7].

Shadow pricing

In shadow pricing, non-monetary effects are valued as the most nearly comparable prices, e.g. relief from arthritis pain might be valued as the cost of analgesics [82]. The money spent on analgesics depends not just on the amount of arthritic pain but also on the efficacy and production costs of aspirin and other products. The problem with shadow pricing is that, too often the most comparable prices do not adequately reflect the value of the public good [82].

Cost-effectiveness analyses

The limitations of the human-capital approach led to the development of cost-effectiveness analyses first published by Klarman in 1968 [6]. In CEA, costs and effects are measured in monetary and physical units, e.g. life-years gained, number of successfully treated patients. The decision rule in CEA is to maximize the effectiveness for a given budget and best suited to comparison of alternative treatments that have the same one-dimensional goal [45]. However, Johannesson [43] argues that cost-effectiveness analysis is best viewed as a subset of cost-benefit analysis, where the aim of the analysis is to estimate the cost function of producing health effects. He also concluded that to interpret and use cost-effectiveness analysis as a tool to maximize the health effects for one specified real-world budget, will be inconsistent with a societal perspective and is likely to lead to major problems of suboptimization. However, it is important to note that this method does not escape the valuation problem (e.g. price per life-year that society is willing to pay).

Cost minimization analyses

This approach is a special form of cost-effectiveness analysis and involves a given and agreed outcome [53]. Alternative treatment modalities can then be ranked according to their costs [74, 75].

Cost-utility analyses

In CUA, costs in monetary terms are related to the non-monetary benefits of a program. This may be used to build league tables of medical interventions. These tables show the different amounts of money that have to be paid to gain a comparable increment of quality of life by different medical programs. Thus, CUA lead to a ranking of measures but is not conclusive about the threshold upon which a measure still should be performed. Quality adjusted life years (QALY) are mostly used in CUA as the effect parameter. In the first investigation, utilities for different health states are assessed. The utility of perfect health is rated as 1 whereas the utility of death is rated as 0. The estimated duration of each health state is then multiplied by the corresponding utility. By transforming these products into products of the same amount but assuming a perfect health state (=1) one receives the QALY for each health state, which can then be compared. However, several restrictions of this method have to be considered: (1) utilities of the different health states have to remain stable during the whole life, and (2) the number of living-years a subject is willing to renounce in order to reach a higher health state must be independent of the living-time left of this subject (constant proportional trade-off). Furthermore, studies using QALYs or life-years gained often lack a relevant discussion of society's willingness-to-pay per QALY or life-years gained [2].

The methodological approach of CEA/CUA analyses as well as cost assessments was covered in a preceding article in the *European Spine Journal* and the interested reader is referred to this article [84].

Cost-benefit analysis

The primary difference between CBA and CEA/CUA is the way in which health benefits or outcomes are measured. In CBAs both costs and health benefits are measured in monetary units, while CEA/CUA measures health in

non-monetary units such as life-years gained or quality adjusted life-years (QUALs) [3]. So far, CBA is not yet widely used in health-related economic evaluations but provide theoretical and practical advantages which deserve more attention.

Methodology of cost-benefit analyses

There are two main approaches for CBA in the health care sector. The revealed-preference approach investigates the actual choice that individuals make between health and wealth (e.g. salary compensation for a more risky job) [6]. The second approach is referred to as contingent valuation (CV) method. The CV is a survey in which respondents are asked how much they are hypothetically prepared to pay (willingness-to-pay, WTP) for different programs [6].

According to Olsen and Smith [65], there are three theoretical advantages for the application of WTP in health care: (1) WTP is theoretically founded in welfare economics, (2) WTP enables a more comprehensive valuation of benefits than QALYs' and (3) the CBA allows to improve allocative efficiency. In spite of the theoretical and methodological advantages for WTP surveys, these benefits have not been used effectively so far in public health policy [65]. Because of the aforementioned advantages many researchers favour CBA for evaluating healthcare interventions [3].

Theoretical background

The first fundamental value judgment that is made in welfare economics is known as the Pareto principle, which states that a change is desirable if it makes some individual(s) better off without making some other individual(s) worse off. If price and quantity take anything other than their equilibrium values, a transaction that will make at least some people better off without harming others can always been found [22]. This Pareto efficiency relies on three conditions: (1) efficient exchange, (2) efficient allocation of factors, and (3) efficient output choice. It is important to note that the Pareto principal proposes nothing about the distribution of goods [44]. Since markets will not always lead to Pareto efficient outcomes due to market failure (e.g. monopolies, presence of externalities, imperfect information, public goods, etc.), there may be a role for public intervention apart from pure redistribution of income.

Characteristics

In spite of its use in other areas of public policy, CBA has not been widely applied in the health and social sectors [17, 52, 63], largely because of the difficulties associated with placing monetary values on the so-called "intangible benefits of health and social care provisions" [38]. In a contingent valuation survey, consumers are asked to consider a hypothetical scenario where a market exists for the benefits of the public good evaluated (e.g. a CV survey might ask WTP questions for the health benefits of cleaner air due to some programme) [17]. This scenario proceeds to the hypothetical contingency that such a market exists to determine what consumers would be willing to pay [17].

There is an increasing interest in WTP as a measure of health benefits in recent literature on economic evaluation in health care [6, 64, 65]. There are several reasons why WTP is preferred over other methods as mentioned above. There is a large variation of types of questions being asked in health care contingent valuation method studies [17]. This and the fact that reporting of the applied methods is poor and often not transparent make it difficult to classify and appraise the literature. Many studies focus on methodological contents of CBA and WTP. Since no gold standard has been established to handle difficulties in interpreting data as for example due to income effects even more research in the methodological field will be necessary [15, 18, 19, 30, 42, 52, 73].

Nevertheless, CBA and WTP have been found to be valuable approaches to assess patients' benefits in health care and it has also been shown that CBA is not inferior in terms of applicability, comprehensiveness to responders or test–retest reliability to other techniques for eliciting public preferences for health care such as CEA/CUA [4–6, 28, 36, 54, 58, 59, 64, 71].

Response formats

Four survey techniques can be used to estimate WTP [24, 70, 77]: (1) open-ended questions (OE), (2) bidding-games (BG), (3) payment card (PC) and (4) closed-ended questions (CE). In OE, subjects are asked directly how much they would be willing to pay for a commodity. In BG, individuals are asked whether or not they are willing to pay a certain amount. If the answer is yes the bid is increased until the respondent is no longer willing to pay. If the answer is no then the bid is decreased until the respondent says yes to a certain bid. In PC, respondents are presented with a choice of cards showing different ranges of amounts of money. Then they are asked to decide which card represents the most they would be willing to pay. In CE, a certain amount to be paid is offered to the respondent who has to answer whether or not he would be willing to pay this amount by only saying yes or no. The only information to be obtained from each individual is whether his or her maximum WTP is above or below the bid offered.

In about 50 studies systematically reviewed by Ryan et al. [70], the response rates of 22–99% were found indicating

varying acceptability. The response rate depended on the technique used for data assessment. Blumenschein [8] concluded that the dichotomous choice contingent valuation method overestimates willingness to pay, but it may be possible to correct for this overestimation by sorting out "definitely sure" yes responses. Green et al. [32] identified strong anchoring effects in single referendum questions in contingent valuation surveys on WTP for public goods that lead to systematically higher estimated mean responses from yes/no referendum responses than from open-ended responses. The application of the WTP technique to health care has focused on using the PC and CE approaches [70].

Reliability

This is defined as repeatability of results over a given time and usually assessed as test–retest reliability whereby a sample of respondents repeats the same exercise after a short period of time [70]. Ryan et al. [70] noted that implicit in the measure of reliability is an assumption that preferences exist (are complete) and are stable over time which may not necessarily be the case in health care. Based on an extensive review of CV surveys in health care, Klose [52] and Ryan et al. [70] have found that reproducibility is rarely investigated and reveal only mediocre correlations.

Validity

Three types can be differentiated: (1) content validity refers to the extent to which a measure takes account of all things deemed important in the construct's domain. (2) Criterion validity or external validity is concerned with whether the measure adopted measures what the researcher is trying to measure. (3) Construct validity can be separated in two types. The convergent validity measures the extent to which results are consistent with other measures that are held to measure the same construct. The theoretical internal validity assesses the extent to which the results are consistent with a priori expectations [70].

An important aspect of content validity is that the hypothetical CV scenario presents valuation tasks and choices realistic, e.g. by choosing payment vehicles that are usual in the health task and choices. The positive influence of health gain and income on WTP demonstrating theoretical validity is strongly indicated by several papers [52].

Sources of bias

Several sources of possible bias are known and discussed controversially. As contingent valuation methods and therefore WTP usually work with hypothetical questions, it is not easy to build a bridge between the hypothetical WTP and the real WTP [8]. A number of laboratory experiments have studied this relationship by using dichotomous choice (yes/no) questions and in most cases hypothetical WTP exceeded real WTP [8, 46]. The influence of household income has been highlighted in several papers. A higher household income has found to be associated with a higher WTP in some investigations while no association could be found in others [18, 64].

Another source of disturbance is the starting point bias: Respondents are influenced by the first number presented. Literature again allows no definite conclusion on this problem. While some investigations showed starting point bias others did not [52]. There is also some evidence suggesting that patients prefer established procedures over new ones even if they are equal or better [87].

CBA in health economics

Contingent valuation studies have been used so far in three scenarios: (1) valuing prevention, (2) valuing treatment and services, and (3) valuing health states [62]. Most studies remain in a purely hypothetical scenario which might be difficult to imagine for majority of patients. Patients are frequently asked to value a hypothetical treatment with a given outcome [12, 20, 30, 81, 82, 88].

Although the interest in contingent valuation studies has substantially increased in many different areas in healthcare [17, 63], only a few studies have been performed in the field of musculo-skeletal disorders, i.e. cervical spondylotic myelopathy [50], rheumatoid arthritis [74, 75] and osteoarthritis [14, 20, 82]. Most of them explored methodological aspects, and only one study has empirically assessed the benefits of a surgical intervention, i.e. joint arthroplasty for knee and hip osteoarthritis in a group of individuals operated on because of osteoarthritis [14].

Economic evaluations in spinal surgery

Increasing data are gathered on the societal costs and burden of back pain [61] whereas economic evaluations of spinal surgery are still sparse [79]. Only recently the concept of cost-effectiveness and cost-benefit was applied to spinal surgery.

Cost assessment studies

Earlier studies [1, 34, 48, 51, 60, 68] compared the intervention costs of alternative surgical procedure (mainly spinal fusion) without appropriate treatment control group. However, many of these studies exhibited methodological limitations related to study design and assessment of

economic parameters [79]. Based on a review by Soegaard [79], three early studies were found to be methodologically credible [26, 55, 72] but only two provided a formal synthesis of costs and effects [26, 55]. Since the review of Soegaard [79], additional studies were published which predominantly focussed on the cost side. These studies dealt with a comparison of one-level lumbar total disc arthroplasty (TDA) versus lumbar fusion (i.e. anterior lumbar interbody fusion with iliac bone crest or BMP, instrumented posterolateral fusion) [33], one- or two-level TDA versus circumferential fusion [57], or explored patients' demands to the primary health sector based on a comparison of three post-operative rehabilitation protocols [79].

An interesting analysis was provided by Polly et al. [66] who compared lumbar spinal fusion to other surgical procedures i.e. total knee replacement, total hip replacement, and coronary artery bypass surgery based on the average reimbursement costs per SF-36 PCS (physical component summary). Although this study can be criticized for methodological limitations and flaws, the calculations indicate that lumbar fusion compares well to other well-accepted medical interventions.

Cost-effectiveness studies

With regard to the health care scarce resource formal CEA and CUA relate to our ambition to select the best available treatment for the money available and could therefore serve as a bridge of understanding among clinicians, managers, healthcare policy-makers and politicians [25]. The first CEA with an appropriate methodology according to current standards was provided by Kuntz et al. [55] who compared the cost and benefits of laminectomy alone and laminectomy with concomitant lumbar fusion (instrumented or non-instrumented) for patients with degenerative lumbar spondylolisthesis and spinal stenosis. The costs of the surgery were $14,700 for laminectomy without fusion, $21,500 for decompression with non-instrumented fusion, and $30,200 for decompression with instrumented fusion (based on data from Boston, 1990–1993). According to this analysis, laminectomy with non-instrumented fusion resulted in an additional 42 quality-adjusted days per person over a 10-year period at an incremental cost of $5,900, yielding an incremental cost-effectiveness ratio of $56,000 per QUALY when compared to laminectomy without fusion. Laminectomy with instrumented fusion cost an additional $8,700 (over a 10-year period) per person as compared with non-instrumented fusion and resulted in one additional quality-adjusted day with a cost-effectiveness ratio of $3,112,800 per QUALY. The authors concluded that non-instrumented fusion not only enhances pain relief, but also increases costs and complications. Instrumented fusion is regarded as very expensive compared to the incremental gain in health outcome.

A further study on lumbar fusion techniques considered the cost-effectiveness of titanium cages versus femoral ring allografts for anterior lumbar interbody fusion [23], indicating that titanium cages are not cost-effective.

From a health economic perspective, the evaluation may not only encompass diagnosis and treatment costs, but must also consider the societal dimension, i.e. the costs of production loss from absenteeism and disability [79]. These costs by far exceed those of diagnosing and treating the target disease [31]. More importantly, the CEA should include a non-surgical control group to add to our understanding of the societal impact of the target procedure.

The first full CEA fulfilling theses requirements compared three alternative surgical treatment options with a non-operative control group ($n = 284$ patients) for degenerative disc disease. Fritzell et al. [26] reported that the surgical group required significantly higher societal costs than in the nonsurgical group (SEK 704,000 vs. SEK 636,000). The direct cost per patient was significantly higher for the surgical group (SEK 123,000 vs. SEK 65,200) for the control group. The incremental cost-effectiveness ratio (ICER, for explanation see [84]), illustrating the extra cost per extra effect unit gained by using fusion instead of nonsurgical treatment were for improvement (i.e. much better, better, unchanged, worse) SEK 2,600, for back pain: SEK 5,200, for disability (ODI): SEK 11,300, and for return to work: SEK 4,100. The authors concluded that both direct and indirect 2-year costs were significantly higher for lumbar fusion compared with non-surgical treatment but all treatment effects were significantly in favour of surgery. The probability of lumbar fusion being cost-effective increased with the value put on extra effect units gained by using surgery.

A similar study was conducted by Rivero-Arias et al. [69] assessing the cost-effectiveness of spinal fusion over an intensive cognitive behavioural rehabilitation programme in patients with chronic low back pain. At 2 years, a significant difference in mean total cost of £3,300 per patient was observed (£7,830 for the surgery group, and £4,526 for the intensive rehabilitation group). There was no significant difference in mean QALYs over the trial period. The incremental cost-effectiveness ratio was estimated to be £48,588 pounds per QALY gained. These data indicated that surgical spinal stabilization may not be a cost-effective use of scarce healthcare resources.

When we want to allocate resources with regards to the cost-effectiveness of treatment modalities it is mandatory to better understand the confounding variables on the cost and the effect side. In this context, Soegaard et al. [80] were the first to explore factors influencing cost-effectiveness in a study comparing posterolateral instrumentation and anterior

intervertebral support as adjuncts to posterolateral lumbar fusion. The costs of non-instrumented posterolateral fusion, instrumented posterolateral fusion and instrumented posterolateral and anterior lumbar fusion averaged DKK 88,285, DKK 94,396, DKK 120,759, respectively. The authors found that the regimes' net benefit was significantly affected by smoking, functional disability in psychosocial life areas, multi-level fusion and surgical technique but no correlation was found between treatment costs and treatment effects. An incremental analysis suggested that a limited probability of posterior instrumentation being cost-effective, whereas the probability of anterior intervertebral support being cost-effective increased as willingness-to-pay per effect unit increases. This study highlighted that patient characteristics have more influence on cost-effectiveness than the surgical technique itself.

When we interpret cost-effectiveness data on surgical procedures we must also consider the follow-up period which can have a substantial impact on the result as outlined by Rivero-Arias et al. [69]. The conclusions could even be reverted when patients undergoing rehabilitation instead of surgery require secondary surgery. Similarly, we should consider the cost-effectiveness of the type of post-surgical rehabilitation as investigated by Soegaard et al. [78].

In other areas of spinal surgery, economic evaluations are still very sparse. Economic data from the Spine Patient Outcomes Research Trial [83] comparing non-operative ($n = 416$) versus surgical treatment ($n = 775$) for lumbar disc herniation indicate that the mean difference in QALYs over 2 years was 0.21 in favour of surgery. Surgery was more costly than non-operative treatment with regard to the total direct ($20,237 vs. $5,804) and total indirect costs ($7,089 vs. $3,321). The cost per QALY gained for surgery relative to non-operative care was $69,403 using general adult surgery costs and $34,355 using Medicare population surgery costs. Tosteson et al. [83] concluded that surgery for lumbar disc herniation was moderately cost-effective when evaluated over 2 years but the estimated economic value of surgery varied considerably according to the method used for assigning surgical costs.

Cost-benefit studies

While the cost-effectiveness and cost-utility of several spinal interventions has been explored [23, 26, 69, 80, 83], CBA using the CV approach with WTP is very rare in the field of spinal surgery. So far, only one study [35] attempted to clarify the feasibility of a cost-benefit approach in spinal surgery. In this pilot study, the authors demonstrated the feasibility of the CV approach with ex post willingness-to-pay/willingness-to-accept (WTP/WTA) questions in 115 patients who underwent lumbar fusion, discectomy, or decompression. The patients were asked to respond to an ex post questionnaire on their WTP/WTA for their respective intervention. Additional questions addressed socio-demographics, household income, and clinical outcome. WTP/WTA was related to the actual intervention costs as well as clinical outcome and the data were then combined within a formal CBA framework [35]. Almost 90% of the respondents (105/115) were satisfied or very satisfied with the treatment, 76.2% considered the surgical result as good or excellent, and 75.7% would choose the operation for a given hypothetical treatment cost. The average costs of lumbar fusion (13,800 €) were almost twice the costs for lumbar decompression (7,000 €). Discectomy was least cost intensive (5,200 €). The main components of cost were wages for personnel and medical services, while implants accounted for 24.0% of the total costs in the fusion group. In the fusion group, maximum WTP was one-fifth lower than the actual procedure costs (not known to the participants), while WTP exceeded costs in the discectomy and the decompression group by 37 and 10%, respectively. The individuals' financial situation was the strongest predictor for WTP. Pain improvement, present pain, duration of hospitalization, and estimated intervention costs were significant independent predictors in the expected direction for the WTP. Calculation of net benefits showed that spinal decompression and discectomy are both within the realms of being cost-beneficial with positive net benefits while spinal fusion gave rise to a net welfare loss. Given a hypothetically average wealth and monthly household income of Switzerland's population, maximum WTP was substantially higher than the actual intervention costs (fusion: +55%; decompression: +47%; discectomy: +227%).

This pilot study demonstrated that a CV survey using a WTP approach is feasible and permits the application of CBA in spinal surgery. Although the majority of patients answered the questions in a reasonable fashion, refinement of the methodological approach are mandatory to improve reliability and validity. If the preliminary results are reproduced in a full economic study that also includes indirect costs, CBA may offer an additional tool to help decision-makers in the, often tense, relationship between economics and healthcare [35].

Conclusions and recommendations

The costs for spinal surgery in particular spinal stabilization are continuously rising despite a lack of clinical evidence for the effectiveness of these interventions compared to non-operative treatment [9, 21]. This trend is perpetuated due to a variety of factors such as demographic changes, advances in technology, unclear indications, and

financial incentives for the involved parties may have had synergistic effects [16]. Despite the increasing use of spinal interventions, scientific evidence for their therapeutic efficacy compared to natural history and non-operative treatment is sparse. This is particularly true for instrumented fusion for degenerative disc disease, one of the most costly spinal intervention [29]. The subjective outcome of patients is highly variable prompting the notion that market growth cannot be reasonably based on consumer's demand despite the urge to empower and engage consumers of health care (consumer-driven health care) [39–41]. The lack of scientific evidence on one hand is contrasted by the implant market on the other hand, indicating the presence of substantial market imperfections and asymmetric information. Porter and Teisberg [67] recently stressed that the current preoccupation with cost shifting and cost reduction undermines physicians and patients. These authors [67] propose a strategy for health care reform that is market based but physician led focusing on a value-based competition. With such a competition to improve results, patients will receive better care, physicians will be rewarded for excellence, and costs will be contained. According to Porter and Teisberg [67], three principles should guide this change:

- creating value for the patients, i.e. improve the quality of care
- organizing medical practice around medical conditions and care cycles rather than around specialities or procedure
- measuring results i.e. risk-adjusted outcomes and costs.

Following these principles, Porter and Teisberg feel that professional satisfaction will increase and current pressures on physicians will decrease. Unless physicians improve health and health care value for patients, they will inevitably face ever-increasing administrative control of medicine [67].

Limited healthcare resources increasingly demand that evidence is obtained not only on therapeutic efficacy of treatment modalities but also on costs. It is the task for all of us to convincingly demonstrate the therapeutic and cost-effectiveness as well as the net benefit of surgical treatments to improve the quality of spinal health care. More importantly, physician must take a lead because only medical teams can improve the value of care which is the only solution to ailing health systems [67].

Conflict of interest statement None of the authors has any potential conflict of interest.

References

1. Ackerman SJ, Mafilios MS, Polly DW Jr (2002) Economic evaluation of bone morphogenetic protein versus autogenous iliac crest bone graft in single-level anterior lumbar fusion: an evidence-based modeling approach. Spine 27:S94–S99. doi:10.1097/00007632-200208151-00017
2. Anell A, Norinder A (2000) Health outcome measures used in cost-effectiveness studies: a review of original articles published between 1986 and 1996. Health Policy 51:87–99. doi:10.1016/S0168-8510(00)00058-0
3. Bala MV, Mauskopf JA, Wood LL (1999) Willingness to pay as a measure of health benefits. Pharmacoeconomics 15:9–18. doi:10.2165/00019053-199915010-00002
4. Bala MV, Zarkin GA (2000) Are QALYs an appropriate measure for valuing morbidity in acute diseases? Health Econ 9:177–180. doi:10.1002/(SICI)1099-1050(200003)9:2<177::AID-HEC497>3.0.CO;2-2
5. Birch S, Donaldson C (2003) Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? Soc Sci Med 56:1121–1133. doi:10.1016/S0277-9536(02)00101-6
6. Blumenschein K, Johannesson M (1996) Economic evaluation in healthcare. A brief history and future directions. Pharmacoeconomics 10:114–122. doi:10.2165/00019053-199610020-00003
7. Blumenschein K, Johannesson M (1996) Incorporating quality of life changes into economic evaluations of health care: an overview. Health Policy 36:155–166. doi:10.1016/0168-8510(95)00809-8
8. Blumenschein K, Johannesson M, Yokoyama KK, Freeman PR (2001) Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. J Health Econ 20:441–457. doi:10.1016/S0167-6296(01)00075-3
9. Brox JI, Sorensen R, Friis A, Nygaard O, Indahl A, Keller A, Ingebrigtsen T, Eriksen HR, Holm I, Koller AK, Riise R, Reikeras O (2003) Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 28:1913–1921. doi:10.1097/01.BRS.0000083234.62751.7A
10. Carron H, DeGood DE, Tait R (1985) A comparison of low back pain patients in the United States and New Zealand: psychosocial and economic factors affecting severity of disability. Pain 21:77–89. doi:10.1016/0304-3959(85)90079-X
11. Cassidy JD, Carroll LJ, Cote P (1998) The Saskatchewan health and back pain survey. The prevalence of low back pain and related disability in Saskatchewan adults. Spine 23:1860–1866. doi:10.1097/00007632-199809010-00012. Discussion 1867
12. Corso PS, Hammitt JK, Graham JD, Dicker RC, Goldie SJ (2002) Assessing preferences for prevention versus treatment using willingness to pay. Med Decis Making 22:S92–S101. doi:10.1177/027298902237713
13. Coyte PC, Asche CV, Croxford R, Chan B (1998) The economic cost of musculoskeletal disorders in Canada. Arthritis Care Res 11:315–325. doi:10.1002/art.1790110503
14. Cross MJ, March LM, Lapsley HM, Tribe KL, Brnabic AJ, Courtenay BG, Brooks PM (2000) Determinants of willingness to pay for hip and knee joint replacement surgery for osteoarthritis. Rheumatology 39:1242–1248. doi:10.1093/rheumatology/39.11.1242
15. Currie GR, Donaldson C, O'Bbrien BJ, Stoddart GL, Torrance GW, Drummond MF (2002) Willingness to pay for what? A note on alternative definitions of health care program benefits for contingent valuation studies. Med Decis Making 22:493–497. doi:10.1177/0272989X02238301
16. Deyo RA, Nachemson A, Mirza SK (2004) Spinal-fusion surgery—the case for restraint. N Engl J Med 350:722–726. doi:10.1056/NEJMsb031771
17. Diener A, O'Brien B, Gafni A (1998) Health care contingent valuation studies: a review and classification of the literature. Health Econ 7:313–326. doi:10.1002/(SICI)1099-1050(199806)7:4<313::AID-HEC350>3.0.CO;2-B
18. Donaldson C, Birch S, Gafni A (2002) The distribution problem in economic evaluation: income and the valuation of costs and

consequences of health care programmes. Health Econ 11:55–70. doi:10.1002/hec.642
19. Donaldson C, Farrar S, Mapp T, Walker A, Macphee S (1997) Assessing community values in health care: is the 'willingness to pay' method feasible? Health Care Anal 5:7–29. doi:10.1007/BF02678452
20. Ethgen O, Tancredi A, Lejeune E, Kvasz A, Zegels B, Reginster JY (2003) Do utility values and willingness to pay suitably reflect health outcome in hip and knee osteoarthritis? A comparative analysis with the WOMAC Index. J Rheumatol 30:2452–2459
21. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R (2005) Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. Br Med J 330:1233. doi:10.1136/bmj.38441.620417.8F
22. Frank RH, Bernanke BS (2001) Principles of economics. McGraw Hill, New York
23. Freeman BJ, Steele NA, Sach TH, Hegarty J, Soegaard R (2007) ISSLS prize winner: cost-effectiveness of two forms of circumferential lumbar fusion: a prospective randomized controlled trial. Spine 32:2891–2897
24. Frew EJ, Whynes DK, Wolstenholme JL (2003) Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. Med Decis Making 23:150–159. doi:10.1177/0272989X03251245
25. Fritzell P (2006) Comments on "Health economic evaluation in lumbar spinal fusion. A systematic literature review anno 2005". Eur Spine J 15:1174–1175. doi:10.1007/s00586-005-0060-1
26. Fritzell P, Hagg O, Jonsson D, Nordwall A (2004) Cost-effectiveness of lumbar fusion and nonsurgical treatment for chronic low back pain in the Swedish Lumbar Spine Study: a multicenter, randomized, controlled trial from the Swedish Lumbar Spine Study Group. Spine 29:421–434. doi:10.1097/01.BRS.0000102681.61791.12. Discussion Z3
27. Frymoyer JW, Durett CL (1997) The economics of spinal disorders. In: Frymoyer JW (ed) The adult spine. Principles and practice. Raven Press, New York, pp 143–150
28. Gafni A (1997) Alternatives to the QALY measure for economic evaluations. Support Care Cancer 5:105–111. doi:10.1007/BF01262566
29. Gibson JN, Waddell G (2005) Surgery for degenerative lumbar spondylosis: updated Cochrane Review. Spine 30:2312–2320. doi:10.1097/01.brs.0000182315.88558.9c
30. Giesler RB, Ashton CM, Brody B, Byrne MM, Cook K, Geraci JM, Hanita M, Souchek J, Wray NP (1999) Assessing the performance of utility techniques in the absence of a gold standard. Med Care 37:580–588. doi:10.1097/00005650-199906000-00007
31. Goossens ME, Evers SM, Vlaeyen JW, Rutten-van Molken MP, van der Linden SM (1999) Principles of economic evaluation for interventions of chronic musculoskeletal pain. Eur J Pain 3:343–353. doi:10.1016/S1090-3801(99)90016-2
32. Green D, Jacowitz K, Kahneman D, McFadden D (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods. Resour Energy Econ 20:85–116. doi:10.1016/S0928-7655(97)00031-6
33. Guyer RD, Tromanhauser SG, Regan JJ (2007) An economic model of one-level lumbar arthroplasty versus fusion. Spine J 7:558–562. doi:10.1016/j.spinee.2006.09.006
34. Hacker RJ (1997) Comparison of interbody fusion approaches for disabling low back pain. Spine 22:660–665. doi:10.1097/00007632-199703150-00017. Discussion 665-666
35. Haefeli M, Elfering A, McIntosh E, Gray A, Sukthankar A, Boos N (2008) A cost-benefit analysis using contingent valuation techniques: a feasibility study in spinal surgery. Value Health 11:575–588
36. Hammit JK, Graham JD (1999) Willingness to pay for health protection: inadequate sensitivity to probability? J Risk Uncertain 8:33–62. doi:10.1023/A:1007760327375
37. Hanley N, Ryan M, Wright R (2003) Estimating the monetary value of health care: lessons from environmental economics. Health Econ 12:3–16. doi:10.1002/hec.763
38. Healey A, Chisholm D (1999) Willingness to pay as a measure of the benefits of mental health care. J Ment Health Policy Econ 2:55–58. doi:10.1002/(SICI)1099-176X(199906)2:2<55::AID-MHP42>3.0.CO;2-1
39. Herzlinger RE (2003) Consumer-driven health care: implications for providers, payers, and policy-makers. Healthplan 44:26–27 29
40. Herzlinger RE (2006) Why innovation in health care is so hard. Harv Bus Rev 84:58–66 156
41. Herzlinger RE, Parsa-Parsi R (2004) Consumer-driven health care: lessons from Switzerland. JAMA 292:1213–1220. doi:10.1001/jama.292.10.1213
42. Johannesson M (1994) The concept of cost in the economic evaluation of health care. A theoretical inquiry. Int J Technol Assess Health Care 10:675–682
43. Johannesson M (1995) The relationship between cost-effectiveness analysis and cost-benefit analysis. Soc Sci Med 41:483–489. doi:10.1016/0277-9536(94)00353-U
44. Johannesson M (1996) The willingness to pay for health changes, the human-capital approach and the external costs. Health Policy 36:231–244. doi:10.1016/0168-8510(96)00815-9
45. Johannesson M, Jonsson B (1991) Economic evaluation in health care: is there a role for cost-benefit analysis? Health Policy 17:1–23. doi:10.1016/0168-8510(91)90114-D
46. Johannesson M, Liljas B, Johansson PO (1998) An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions. Appl Econ 30:643–647. doi:10.1080/000368498325633
47. Jonsson B (1996) Quality of life and health economics: where is the link? Scand J Gastroenterol Suppl 221:33–36. doi:10.3109/00365529609095553
48. Katz JN, Lipson SJ, Lew RA, Grobler LJ, Weinstein JN, Brick GW, Fossel AH, Liang MH (1997) Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. Spine 22:1123–1131. doi:10.1097/00007632-199705150-00012
49. Keller A, Brox JI, Gunderson R, Holm I, Friis A, Reikeras O (2004) Trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain randomized to lumbar fusion or cognitive intervention and exercises. Spine 29:3–8. doi:10.1097/01.BRS.0000103946.26548.EB
50. King JT Jr, McGinnis KA, Roberts MS (2003) Quality of life assessment with the medical outcomes study short form-36 among patients with cervical spondylotic myelopathy. Neurosurgery 52:113–120. doi:10.1097/00006123-200301000-00014. Discussion 121
51. Klara PM, Freidank SA, Rezaiamiri S (2003) Comparison of lumbar interbody fusion techniques using Ray threaded fusion cages and pedicle screw fixation systems. Neurosurg Q 13:20–29. doi:10.1097/00013414-200303000-00002
52. Klose T (1999) The contingent valuation method in health care. Health Policy 47:97–123. doi:10.1016/S0168-8510(99)00010-X
53. Korthals-de Bos I, van Tulder M, van Dieten H, Bouter L (2004) Economic evaluations and randomized trials in spinal disorders: principles and methods. Spine 29:442–448. doi:10.1097/01.BRS.0000102683.61791.80
54. Krabbe PF, Essink-Bot ML, Bonsel GJ (1997) The comparability and reliability of five health-state valuation methods. Soc Sci Med 45:1641–1652. doi:10.1016/S0277-9536(97)00099-3
55. Kuntz KM, Snider RK, Weinstein JN, Pope MH, Katz JN (2000) Cost-effectiveness of fusion with and without instrumentation for

patients with degenerative spondylolisthesis and spinal stenosis. Spine 25:1132–1139. doi:10.1097/00007632-200005010-00015
56. Landefeld JS, Seskin EP (1982) The economic value of life: linking theory to practice. Am J Public Health 72:555–566. doi:10.2105/AJPH.72.6.555
57. Levin DA, Bendo JA, Quirno M, Errico T, Goldstein J, Spivak J (2007) Comparative charge analysis of one- and two-level lumbar total disc arthroplasty versus circumferential lumbar fusion. Spine 32:2905–2909
58. Liljas B, Blumenschein K (2000) On hypothetical bias and calibration in cost-benefit studies. Health Policy 52:53–70. doi:10.1016/S0168-8510(00)00067-1
59. Liljas B, Lindgren B (2001) On individual preferences and aggregation in economic evaluation in healthcare. Pharmacoeconomics 19:323–335. doi:10.2165/00019053-200119040-00001
60. Malter AD, Larson EB, Urban N, Deyo RA (1996) Cost-effectiveness of lumbar discectomy for the treatment of herniated intervertebral disc. Spine 21:1048–1054. doi:10.1097/00007632-199605010-00011
61. Maniadakis N, Gray A (2000) The economic burden of back pain in the UK. Pain 84:95–103. doi:10.1016/S0304-3959(99)00187-6
62. Morrison GC, Gyldmark M (1992) Appraising the use of contingent valuation. Health Econ 1:233–243. doi:10.1002/hec.4730010405
63. O'Brien B, Gafni A (1996) When do the "dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. Med Decis Making 16:288–299. doi:10.1177/0272989X9601600314
64. O'Brien B, Viramontes JL (1994) Willingness to pay: a valid and reliable measure of health state preference? Med Decis Making 14:289–297. doi:10.1177/0272989X9401400311
65. Olsen JA, Smith RD (2001) Theory versus practice: a review of 'willingness-to-pay' in health and health care. Health Econ 10:39–52. doi:10.1002/1099-1050(200101)10:1<39::AID-HEC563>3.0.CO;2-E
66. Polly DW Jr, Glassman SD, Schwender JD, Shaffrey CI, Branch C, Burkus JK, Gornet MF (2007) SF-36 PCS benefit-cost ratio of lumbar fusion comparison to other surgical interventions: a thought experiment. Spine 32:S20–S26. doi:10.1097/BRS.0b013e318053d4e5
67. Porter ME, Teisberg EO (2007) How physicians can change the future of health care. JAMA 297:1103–1111. doi:10.1001/jama.297.10.1103
68. Ray CD (1997) Threaded fusion cages for lumbar interbody fusions. An economic comparison with 360 degrees fusions. Spine 22:681–685. doi:10.1097/00007632-199703150-00021
69. Rivero-Arias O, Campbell H, Gray A, Fairbank J, Frost H, Wilson-MacDonald J (2005) Surgical stabilisation of the spine compared with a programme of intensive rehabilitation for the management of patients with chronic low back pain: cost utility analysis based on a randomised controlled trial. Br Med J 330:1239. doi:10.1136/bmj.38441.429618.8F
70. Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen ER, Russell EM, Napper M, Robb CM (2001) Eliciting public preferences for healthcare: a systematic review of techniques. Health Technol Assess 5:1–186
71. Schiffner R, Schiffner-Rohe J, Gerstenhauer M, Hofstadter F, Landthaler M, Stolz W (2003) Willingness to pay and time trade-off: sensitive to changes of quality of life in psoriasis patients? Br J Dermatol 148:1153–1160. doi:10.1046/j.1365-2133.2003.05156.x
72. Schofferman J, Slosar P, Reynolds J, Goldthwaite N, Koestler M (2001) A prospective randomized comparison of 270 degrees fusions to 360 degrees fusions (circumferential fusions). Spine 26:E207–E212. doi:10.1097/00007632-200105150-00019
73. Sculpher MJ, O'Brien BJ (2000) Income effects of reduced health and health effects of reduced income: implications for health-state valuation. Med Decis Making 20:207–215. doi:10.1177/0272989X0002000206
74. Slothuus U, Brooks RG (2000) Willingness to pay in arthritis: a Danish contribution. Rheumatology 39:791–799. doi:10.1093/rheumatology/39.7.791
75. Slothuus U, Larsen ML, Junker P (2000) Willingness to pay for arthritis symptom alleviation. Comparison of closed-ended questions with and without follow-up. Int J Technol Assess Health Care 16:60–72. doi:10.1017/S0266462300016160
76. Smith KJ, Roberts MS (2002) Cost-effectiveness of newer treatment strategies for influenza. Am J Med 113:300–307. doi:10.1016/S0002-9343(02)01222-6
77. Smith RD (2000) The discrete-choice willingness-to-pay question format in health economics: should we adopt environmental guidelines? Med Decis Making 20:194–206. doi:10.1177/0272989X0002000205
78. Soegaard R, Bunger CE, Laurberg I, Christensen FB (2008) Cost-effectiveness evaluation of an RCT in rehabilitation after lumbar spinal fusion: a low-cost, behavioural approach is cost-effective over individual exercise therapy. Eur Spine J 17:262–271. doi:10.1007/s00586-007-0479-7
79. Soegaard R, Christensen FB (2006) Health economic evaluation in lumbar spinal fusion: a systematic literature review anno 2005. Eur Spine J 15:1165–1173. doi:10.1007/s00586-005-0031-6
80. Soegaard R, Christensen FB, Christiansen T, Bunger C (2007) Costs and effects in lumbar spinal fusion. A follow-up study in 136 consecutive patients with chronic low back pain. Eur Spine J 16:657–668. doi:10.1007/s00586-006-0179-8
81. Stavem K (1999) Willingness to pay: a feasible method for assessing treatment benefits in epilepsy? Seizure 8:14–19. doi:10.1053/seiz.1998.0242
82. Thompson MS, Read JL, Liang M (1984) Feasibility of willingness-to-pay measurement in chronic arthritis. Med Decis Making 4:195–215. doi:10.1177/0272989X8400400207
83. Tosteson AN, Skinner JS, Tosteson TD, Lurie JD, Andersson GB, Berven S, Grove MR, Hanscom B, Blood EA, Weinstein JN (2008) The cost effectiveness of surgical versus nonoperative treatment for lumbar disc herniation over two years: evidence from the Spine Patient Outcomes Research Trial (SPORT). Spine 33:2108–2115. doi:10.1097/BRS.0b013e318182e390
84. van der Roer N, Boos N, van Tulder MW (2006) Economic evaluations: a new avenue of outcome assessment in spinal disorders. Eur Spine J 15(Suppl 1):S109–S117. doi:10.1007/s00586-005-1052-x
85. van Tulder MW, Koes B, Seitsalo S, Malmivaara A (2006) Outcome of invasive treatment modalities on back pain and sciatica: an evidence-based review. Eur Spine J 15(Suppl 1):S82–S92. doi:10.1007/s00586-005-1049-5
86. van Tulder MW, Koes BW, Bouter LM (1995) A cost-of-illness study of back pain in The Netherlands. Pain 62:233–240. doi:10.1016/0304-3959(94)00272-G
87. Walton SM, Graves PE, Mueser PR, Dow JK (2002) The bias against new innovations in health care: value uncertainty and willingness to pay. Value Health 5:67–70. doi:10.1046/j.1524-4733.2002.52114.x
88. Werner P, Vered I (2002) Women's willingness to pay out-of-pocket for drug treatment for osteoporosis before and after the enactment of regulations providing public funding: evidence from a natural experiment in Israel. Osteoporos Int 13:228–234. doi:10.1007/s001980200019