Eur Spine J (2006) 15: 1165–1173
DOI 10.1007/s00586-005-0031-6

REVIEW ARTICLE

Rikke Soegaard
Finn B. Christensen

Health economic evaluation in lumbar spinal fusion: a systematic literature review anno 2005

Accepted: 15 November 2005
Published online: 21 December 2005

R. Soegaard (✉) · F. B. Christensen
Orthopaedic Research Laboratory, Spine Unit, University Hospital of Aarhus, Nørrebrogade 44, Bldg. 1, 8000 Aarhus C, Denmark
E-mail: rikke.sogaard@ki.au.dk
Tel.: +45-8949-4102
Fax: +45-8949-4150

Abstract The goal of this systematic literature review was to assess the evidence for cost-effectiveness of various surgical techniques in lumbar spinal fusion in conformity with the guidelines provided by the Cochrane Back Review Group. As new technology continuously emerges and divergent directions in clinical practice are present, economic evaluation is needed in order to facilitate the decision-makers' budget allocations. NHS Economic Evaluation Database, MEDLINE, EMBASE and Cochrane Library were searched. Two independent reviewers (one clinical content expert and one economic content expert) applied the eligibility criteria. A list of criteria for methodological quality assessment was established by merging the criteria recommended by leading health economists with the criteria recommended by the Cochrane Back Review Group. The two reviewers independently scored the selected literature and the disagreement was resolved by means of consensus following discussion. Key data were extracted and the level of evidence concluded. Seven studies were eligible; these studies reflected the diversified choices of economic methodology, study populations (diagnosis), outcome measures and comparators. At the conclusion of quality assessment, the methodological quality of three studies was judged credible. Two studies investigated posteolateral fusion (PLF) ± instrumentation in different populations: one investigated non-specific low back pain and one investigated degenerative stenosis + spondylolisthesis. Both studies reflected that cost-effectiveness of instrumentation in PLF is not convincing. The third study concerned the question of circumferential vs anterior lumbar interbody fusion and found a non-significant difference between the techniques. In conclusion, the literature is limited and, in view of the fact that the clinical effects are statistically synonymous, it does not support the use of high-cost techniques. There is a great potential for improvement of methodological quality in economic evaluations of lumbar spinal fusion and further research is imperative.

Keywords Economics · Cost-effectiveness · Spinal fusion · Systematic review · Lowback pain

Introduction

In addition to the demands for evidence-based regimens, today's orthopedic societies are faced with the challenge of treating the patients by means of cost-effective regimes. Recent randomized controlled trials (RCTs) have investigated the effect of lumbar spinal fusion in patients with chronic low back pain [5, 6, 13, 15, 16, 29, 30, 41].

In comparison to non-invasive treatments, at least two RCTs demonstrate the evidence of improved clinical outcome [15, 29] whereas two recent trials question the superiority of surgery over time [4, 9]. On the background of the evidence favoring surgery, Fischgrund et al. [13] investigated the effect of transpedicular instrumentation as a support for posteolateral fusion (PLF) in patients suffering from degenerative spondylolisthesis and spinal stenosis. The addition of instrumentation produced no reduction in pain. These findings were replicated by Moller et al. [30] who investigated the effect of instrumentation in isthmic spondylolisthesis (any grade). The first long-term study, published by Christensen et al. [6] concluded bidirectional findings; in isthmic spondylolisthesis (grades 1 and 2). the outcome in non-instrumented cases was superior to those in which the instrumentation was used and in primary degeneration, the reverse situation was found. Other studies have investigated the surgical technique for instrumented lumbar spinal fusion by different choices of intervention and control groups [5, 16, 20, 37]. None of these studies were able to conclude a significant difference between the different approaches to spinal fusion. Thus, the extent to which these approaches differ in terms of cost-effectiveness is highly interesting.

The literature of economic evaluations in chronic low back pain continues to grow as interdisciplinary barriers and methodological weaknesses are overcome [18, 20, 22–24, 27, 34, 37]. On the other hand, the quality of much of the literature on economic evaluations has been criticized as being inferior [17, 19] and moreover, the literature is characterized by broad methodological differences, making it difficult for decision-makers to navigate in the search of value for money. The objective of the present study was to conduct a systematic literature review to assess evidence for the cost-effectiveness of various surgical techniques in lumbar spinal fusion.

Methods

A systematic literature review was conducted in conformity with the method guidelines provided by the Back Review Group in the framework of the Cochrane Collaboration [42, 43]. The review team for the present investigation consisted of two reviewers, each with backgrounds in methodology; one reviewer is a clinical content expert (FBC) and one reviewer an economic content expert (RS). Any disagreement was solved by consensus following discussion.

Criteria for inclusion

1. *Economic evaluation*. The definition by Drummond et al. [8] was adapted: (1) there should be a comparison of two or more alternatives and (2) both costs (inputs) and consequences (outputs) of the alternatives should be examined.
2. *Study design*. Randomized and non-randomized, controlled trials were eligible. Economic modeling studies based upon these selected designs were eligible as well.
3. *Participants*. Adult patients suffering disabling, chronic low back pain. No criteria for diagnosis were established.
4. *Interventions*. Lumbar spinal fusion approached surgically by PLF ± instrumentation, circumferential fusion (360°), interbody fusion via a posterior approach (PLIF), interbody fusion via an anterior approach (ALIF) or interbody fusion via a transforaminal approach (TLIF). Relevant comparators, in addition to the listed techniques, comprise non-invasive (conservative) treatment or laminectomy alone.
5. *Cost measures*. Costs should be reported in a monetary unit and comprise at least the (incremental) costs of the hospital admission and the surgical procedure (so called direct costs) from a hospital perspective.
6. *Effect measures*. Physical units and validated instruments were accepted.
7. *Language*. English, German and French literature was considered eligible for inclusion.

Search strategy

The search strategy comprised the following databases:

1. The NHS Economic Evaluation Database (NHS EED) provided by the CRD (Center for Review and Dissemination) was searched by text words and subject headings "arthrodesis OR spinal fusion OR laminectomy" and limited to record types "economic evaluations".
2. MEDLINE (PubMed format) was searched by the highly sensitive search strategy suggested by Robinson and Dickersin [36] in addition to the guidelines made by van Tulder et al. [42]. The intervention at issue, lumbar spinal fusion, was defined by "Laminectomy" [MeSH] OR "Spinal Fusion" [MeSH] OR ("Arthrodesis" [MeSH] AND ("Lumbosacral Region" [MeSH] OR "Low Back Pain" [MeSH])). To identify studies investigating costs and consequences, the MeSH term "Costs and Cost Analysis" was used. MeSH is Medical Subject Headings.
3. EMBASE was searched by means of EMTREE thesaurus terms "spine fusion OR spine stabilization OR laminectomy" (each with parameters/exp/mj) combined by "AND economic evaluation" (added parameter/exp).
4. The Cochrane Library's database was searched by MeSH terms "arthrodesis OR spinal fusion OR

laminectomy'' (exploded trees) combined with ''Cost–Benefit Analysis'' (exploded trees).To the best of our knowledge, there are no systematic reviews that could enable us to double-check our search by means of screening reference lists. Nor have we been able to identify content experts in economic evaluation (except for those being reviewed here) that could have been consulted for the insurance of completeness of the search results. Searches 1 and 2 were conducted on January 21, 2005 and searches 3 and 4 were conducted on January 26, 2005.

Selection of papers

One reviewer performed the electronic literature searches (RS). Two reviewers (RS and FBC) independently examined the titles and abstracts in relation to the criteria for eligibility; if title and abstract did not provide the information necessary for determining eligibility, the complete text was acquired for perusal. After a series of discussions of the inclusion criteria, agreement between reviewers was 100%.

Data extraction

Author, publication year, country, number of participants, diagnosis, interventions, hospital costs of interventions, economic evaluation classification, study design, length of follow-up, effect parameter or instrument, rating of clinical outcome, rating of economic outcome and the author's overall conclusions about ratio of cost-consequence were extracted. All costs were converted into 1999 fiscal year by annual rate of 5% and into EUR by the Purchasing Power Parity Index of the Organization for Economic Co-operation and Development (OECD) [31].

Methodological quality assessment

Two reviewers independently applied a list of criteria for quality assessment. This list was developed by merging an instrument focused on economic evaluations by Drummond et al. [8] with an instrument focused on clinical trials and used by the Cochrane Back Review Group [42]. The merged list was subsequently reduced according to priority of practicality. The list of criteria is shown in Table 1. Each criterion was answered simply by ''positive'', ''negative'' or ''ambiguous'' and the horizontal sum of ''positives'' was considered to reflect the overall methodological quality of a study. Further, vertical criteria sums of ''negatives'' were considered to reflect the general problems in the literature.

Data analysis

Due to methodological differences—and especially the diversity of effect measures and intervention groups—no quantitative analysis was possible. Following quality assessment and data extraction, conclusions were drawn within each specific surgical approach encompassed by the present work and the five levels of evidence suggested by the Cochrane Back Review Group were applied [42].

Results

Literature search

The literature searches revealed a total of 130 references: 20 references indexed by NHS EED, 52 by PubMed, 45 by EMBASE and 13 by the Cochrane Library. Seven references were found eligible for review [14, 20, 22–24, 34, 37]. Except for the article by Hacker and that by Ray, all the articles are joint publications but henceforth will be referred to by the first author only.

Description of included studies

Table 2 presents an overview of the literature by means of data extraction from the included studies. The studies included were published from 1997 to 2004 and—except for one Swedish study—they have all been conducted in the USA. One study was a modeling study built on data extractions from original literature, while the others comprised sample sizes ranging from 46 through 294 patients. In total, the literature comprises 676 patients plus the implicit inclusions summarized in the modeling study.

The studies represent diversified patient populations: Fritzell and Hacker defined the patient population by chronic, non-specific low back pain, Klara and Ray defined their populations by degenerative disc diseases, the studies by Katz and Kuntz were targeted within degenerative stenosis and finally, Schöfferman reported a mixed population of spondylolisthesis, disc degeneration, stenosis and scoliosis.

Three studies investigated the cost-effectiveness of instrumentation in PLF (non-instrumented PLF compared to instrumented PLF) and found incremental costs of the instrumentation to lie between 4,298 and 5,537 EUR. Incremental costs of anterior support in PLF were found to amount to 2,383 EUR in one study. Two studies contrasted circumferential fusion against PLIF and found incremental costs of the circumferential technique between 10,789 and 14,886 EUR. The surgical technique of PLIF alone, however, is not very common today and these incremental costs were measured

Table 1 Criteria for the quality assessment of economic evaluations in lumbar spinal fusion, 2005

Aim and perspective[a]	
1a	Was a well-defined question posed in an answerable form?
1b	Was a viewpoint for the analysis stated and was the study placed in any particular decision-making context?
Patient selection[b]	
2a	Were the eligibility criteria specified?
2b	Was a method of randomization used to allocate the patients?
Intervention[b]	
3a	Were index and control interventions explicitly described?
3b	Was the compliance acceptable in all groups?
Outcome measurement—costs[a]	
4a	Were all the important and relevant costs and consequences identified?
4b	Were costs and consequences measured accurately in appropriate physical units?
4c	Were costs and consequences valued credibly?
Outcome measurement—effect[b]	
5a	Was the outcome assessor blinded to the intervention?
5b	Were the outcome measures relevant?
5c	Was the withdrawal/dropout rate described and acceptable?
Economic methodology[a]	
6a	Were costs and consequences adjusted for differential timing?
6b	Was an incremental analysis of costs and consequences of alternatives performed?
6c	Was allowance made for uncertainty in the estimates of costs and consequences?
6d	Did the presentation and discussion of the study results include all issues of concern to users?
Statistics[b,c]	
7a	Was the sample size for each group described and discussed in relation to power considerations (if not calculations)?
7b	Did the analysis include the intention-to-treat analysis?
7c	Were the synthesis of costs and benefits reported and if so, were measures of variability presented accordingly for the primary outcome measures?

[a]Adopted from Drummond [8]
[b]Adopted from van Tulder et al. [42, 43]
[c]7a and 7c are own modifications

approximately 10 years ago. Finally, one study investigated the potential cost-effectiveness by approaching 360° without the PLF (270° fusion) and found a cost reduction of 2,795 EUR.

With the exception of one study, all studies were conducted as cost-effectiveness analyses. One study (Kuntz) was a cost–utility analysis and conducted as a modeling study building on data from 15 original papers. In all but the modeling study, the epidemiological designs adapted were longitudinal cohort studies. Fritzell performed a random allocation and Schöfferman a quasi-random allocation whereas the others were observational studies. Except for the Ray study, all studies had at least 2 years of follow-up.

In relation to the measurement of effects, the literature is truly diversified. Hacker and Ray base their effect measurement solely on such physical facts as operating time and blood loss. No validated instrument, targeted functional ability or pain, was applied to measure the outcome. The other studies implemented such validated instruments as, e.g., the Oswestry Disability Index [11] and the Prolo Scale [33].

Very few studies were capable of detecting significant differences in effects among interventions. With respect to hospital costs, four studies concluded significant differences among techniques, with the least invasive technique being the least costly and the most invasive technique being the most costly. An explicit synthesis of incremental cost-effectiveness was reported in only two studies: Fritzell reported that the additional costs of an additional effect unit (Oswestry Disability Index) amounts to approximately 1,000 EUR when treating the patients surgically (vs conservative treatment) whereas Kuntz reported that the cost–utility ratio for instrumented PLF (vs non-instrumented PLF) amounts to approximately 2 million EUR per quality-adjusted-life-year.

Quality assessment of included studies

The reviewers independently applied the criteria for quality assessment. The list of criteria is presented in Table 1 and the scorings are presented in Table 3. Overall, the quality of only three studies was found to be methodologically credible; the remainder of the studies did not fulfill more than 50% of the applied criteria. The agreement between the reviewers was good since only

Table 2 Economic evaluations in lumbar spinal fusion (January, 2005)

References	Country	*N*	Diagnosis	Surgical approach	Cost (EUR)[a]	Economic evaluation	Design	Length of follow-up	Effect parameter	Rating clinical outcome[b]	Rating economic outcome[c]	Authors conclusion: ratio of cost-consequence
Fritzell et al. [14]	Sweden	294	Unspecific LBP > 2 years	Non-surgical PLF non-instrumented PLF instrumented 360° fusion	4,408 6,510 10,808 13,191	CEA	RCT	2 years	Global improvement, pain (VAS)[d], disability (ODI)[e], return to work	NS	PLF PLF + inst. 360°	Cost-effectiveness depends on the value of effect units
Hacker [20]	USA	75	Disabling LBP > 6 months	360° fusion PLIF (BAK device)	(46,566) (31,680)	CEA	Cohort	2 years	Fusion, blood loss, operating time, graft pain, return to work, satisfaction, complications	PLIF 360°	NS	PLIF cost-effective
Katz et al. [22]	USA	163	Degenerative stenosis	Laminectomy PLF non-instrumented PLF instrumented	8,025 11,765 16,485	CEA	Cohort	2 years	Health status (SIP)[f], walking capacity, pain, satisfaction	Fusion Laminec.	Laminec. PLF PLF + inst.	Non-instrumented PLF cost-effective
Klara et al. [23]	USA	46	Degenerative disc disease	ALIF (threaded cage) PLIF (threaded cage) PLIF (plate)	(7,659) (8,688) (9,817)	CEA (CMA)	Cohort	2 years	Function (Prolo)[g], fusion, bed days, complications	NS	ALIF Others	Concomitant posterior stabilization not cost-effective
Kuntz et al. [24]	USA	–	Degenerative stenosis + spondylolisthesis	Laminectomy PLF non-instrumented PLF instrumented	13,375 17,154 22,691	CUA	Markov model (review)	10 years	QALY	Fusion Laminec.	Laminec. PLF PLF + inst.	Non-instrumented PLF cost-effective
Ray [34]	USA	50	Disabling LBP from disc degeneration	360° fusion PLIF (treaded cage)	(31,263) (20,474)	CEA (CMA)	Cohort	3–29 months	Fusion, blood loss, operating time, re-opeations	NS	NS	PLIF (treaded cage) cost-effective
Schöfferman et al. [37]	USA	48	Spondylolisthesis Disc degeneration Stenosis Scoliosis	360° fusion 270° fusion	(23,052) (20,258)	CEA (CMA)	Quasi-RCT	24–45 months	Pain (NRS)[h], disability (ODI)[e], operating time, blood loss, bed days	NS	NS	270° fusion may be cost-effective

NS denotes a statistically non-significant difference at $P = 0.05$

PLF posteolateral lumbar fusion; *PLIF* posterior lumbar interbody fusion; *ALIF* anterior lumbar interbody fusion; *CEA* cost-effectiveness analysis, *CMA* cost-minimization analysis, *CUA* cost–utility analysis; *RCT* randomized, controlled trial

[a] Average hospital costs converted to 1999 price index by rate 5%. If information on price year is missing, the publication year is adapted as the price year. Currency conversions were made by means of OECD Purchasing Parity Index [31]. Costs listed in parentheses are charges from a third-party perspective

[b] Rating by the best outcome first

[c] Rating by the least costly first

[d] Visual analogue scale

[e] Oswestry Low Back Disability Index

[f] Sickness Impact Profile

[g] Prolo Scale

[h] Numerical rating scale

three papers required discussion after the first iteration. Seven criteria (of 19 in total) produced a score ≥4 "negatives", which may be considered to reflect the general methodological problems of the literature. The first emergent problem concerns the statement of an explicit perspective for the evaluation. Only the studies by Fritzell and Kuntz stated the perspective of their analysis. The next problem concerns the issue of random allocation, which is of even greater importance in economic evaluation analyzed by means of incremental ratio statistics. Fritzell and Schöfferman satisfied this criterion. The measurement of costs should build upon relevant and conscious methodology parallel to the measurement of clinical effects. Physical units should always be reported prior to cost estimates; Hacker, Katz and Schöfferman demonstrated this. Costing by means of group estimates should contain implicit analysis of sensitivity with respect to the findings' sensitivity to under- or overestimation of both central activity measurements and valuations. The criteria of allowance for uncertainty and the criteria of discussing relevant issues of concern to the users should be fulfilled by completion and discussion of an analysis of sensitivity; only Fritzell and Kuntz carried this out. The combination of "no significant difference" and "no explicit power considerations" provides very little information—especially since economic evaluations comprise at least four stochastic estimates (respectively, costs and effects in both intervention- and control-group) of which the costs have been found to reflect greater variation than effects [26, 45]. This problem is severe in as many as five of seven studies. Finally, only Fritzell and Kuntz reported the explicit synthesis between costs and effects, which should be the paramount analysis in a cost-effectiveness evaluation.

Discussion

The present systematic review was aimed at identifying the evidence concerning the cost-effectiveness of lumbar spinal fusion. Ultimately, we found that the volume of references with credible methodological quality to be extremely limited. There is no other systematic review in the literature that has reviewed this field, which is otherwise a highly emergent aspect of health-services management.

Economic evaluations require, in addition to awareness of possible bias from underlying clinical trials, the important considerations in terms of perspective for analysis, the method for costing, incremental analysis and reporting of uncertainty surrounding the incremental cost-effectiveness ratio (ICER). Problems in these areas were widely apparent in the reviewed literature.

From an economic theoretical point of view, the relevant perspective for economic evaluation in spinal fusion is a societal perspective: the costs of production loss from absenteeism and disability constitute an economic impact which greatly exceeds those of the diagnostic, therapeutic and treatment regimens [26, 28]. It is also evident that the cost-effectiveness of an in-hospital treatment cannot be determined without considering the extra-hospital service utilization, as demonstrated in a recent study of service utilization after lumbar spinal fusion and in-hospital rehabilitation [38]. However, when budgets are already earmarked for spinal fusion and the goal is to maximize the gain of health effects within an allocated budget, an administrator's (hospital management's) perspective is justified. Because only two studies in the reviewed literature performed an analysis from the full-scale, societal perspective, we chose to limit data extractions and quality assessment to an administrator perspective.

The method of costing may have crucial impact on conclusions. The methods identified in the present review were in general the average costs or prices (charged third-party payers) obtained from the accounting department of the hospital. Such estimates are rarely specific to spinal fusion or even orthopedics or spine surgery. This may cause a bias that leads to differentiated under- or overestimations of the costs, which in terms of the incremental cost's estimate becomes difficult to interpret (but nevertheless have direct effect on the

Table 3 Quality assessment of economic evaluations in lumbar spinal fusion, 2005

	1a	1b	2a	2b	3a	3b	4a	4b	4c	5a	5b	5c	6a	6b	6c	6d	7a	7b	7c	Sum
Fritzell	0	1	1	1	1	1	1	0	1	?	1	1	1	1	1	1	1	1	1	16/18
Hacker	1	0	1	0	1	1	1	1	0	0	1	1	0	0	0	0	0	?	0	8/18
Katz	1	0	1	0	0	1	1	1	1	1	1	0	?	0	0	0	0	1	0	9/18
Klara	1	0	1	0	1	1	1	0	?	1	1	1	0	0	0	0	0	1	0	9/18
Kuntz	1	1	0	0	0	?	1	0	1	1	1	?	1	1	1	0	?	?	1	10/16
Ray	0	0	1	0	1	1	1	0	0	?	0	1	1	1	0	0	0	1	0	8/18
Schöfferman	1	0	1	1	1	1	0	1	1	0	1	1	0	1	0	0	0	1	0	11/19
Negatives	2	5	1	5	2	0	1	4	2	2	1	1	3	3	5	6	5	0	5	54/126

0 negative, *1* positive and *?* ambiguous

ICER; overestimated marginal costs lower the ICER and underestimated marginal costs raise the ICER). Ideally, costing should be activity-based while remaining in balance with analysis perspective (for example, a few extra minutes of surgery is insignificant in a broad societal perspective). Overall, and in addition to the specific method for costing, it would be desirable if common guidelines, i.e., the Dutch manual for costing, were adopted in terms of measuring and valuating in a transparent manner [32].

Whether or not the clinical practice is evidence-based, an economic evaluation should compare the new intervention to the best clinical alternative (clinical practice). In other words, the right approach for an economic evaluation is an incremental approach most often concerned about the cost-effectiveness of a new and more sophisticated (costly) intervention. This is not recognized by a great majority of the literature (compares average cost-effectiveness instead of incremental), which relegates the studies to be descriptive rather than decision-analytically. One might argue that the incremental approach is immediately suitable in trial settings with random allocation of patients, but newer methodology describes inherent possibilities in trials with imperfect or missing random allocation of patients as well [21]. Another recurring fact in the reviewed literature is the different choices of comparators. Obviously, a consensus about relevant comparators is lacking and the choice may differ among countries, centers and even surgeons or trialists. Such different choices of comparators (in addition to different populations and outcome measures) hindered the mutual replication. This, of course, has a direct influence on the economic evaluation as it, by virtue of its nature, cannot go beyond the underlying clinical trials. A future possibility for overcoming the challenge of comparability in outcome measurement may be to supplement the outcome evaluation by a generic measure of health-related quality-of-life (HRQoL). However, because different instruments have been found to be non-interchangeable [7], a certain consensus about the methodology is still required and the question of sensitivity and responsiveness in chronic low back pain needs further investigation as well [2, 40].

The reporting of an explicit estimate of the synthesis between costs and effects is evidently important and the uncertainty surrounding it should by all means be reported in some way. Although clinical effect differences may verge on being insignificant, economic methodology has ways of quantifying and reporting the uncertainty surrounding the ICER [12, 25]. One common approach is the reporting by means of acceptability curves. Such reporting were seen in only one of the reviewed studies, which is problematical due to the fact that an ICER itself does not follow a rank scale if it is not significantly limited in one quadrant of the cost-effectiveness plane [3]. Also, if the clinical difference between the comparators approaches 0, the ICER approaches infinity (as effect differences constitute the denominator in the ICER ratio). Finally, one of the problems with the ratio statistics includes the fact that the ICER is not amendable to regression analysis, which is central in econometrics. These arguments are the foundation for recent trends in health economics, which suggest a transformation of the ICER into a net-benefit measure [21, 39]. It may be that net-benefit (cost-effectiveness) is more significantly influenced by for example, a range of key patient characteristics than by surgical technique itself.

A weakness of the present literature review is the fact that reviewers were not blinded to the authors, institutions and journals throughout the process of selection of literature, quality assessment and extraction of data. The literature is fairly limited and blinding of authors, institutions and journals would not be sufficient, as the few targeted studies conducted all have enjoyed a high impact and visibility throughout the spine research society. Omission of blinding is, however, justified from two studies of methodology, that found such labor-intensive process of blinding insignificant to the quality of a review [1, 44].

Although we adapted a highly sensitive search strategy [36, 42] and extended our search to several databases as suggested by the Cochrane Review Group [42, 43], this review may be affected by a publication bias. A comprehensive search for abstracts may have decreased the chance of the publication bias, but since economic evaluation in truth is built on clinical studies, the problem of publication bias lies perhaps not within the economic field but rather within the original literature of the clinical effect studies.

Since the conduction of the literature search for the present review, a British randomized trial investigating the cost–utility of lumbar spinal fusion vs conservative treatment has been published [35]. This investigation applies most of the methodological issues discussed above, i.e., they applied the EQ-5D [10] for the measurement of HRQoL. The authors concluded lumbar spinal fusion to be not cost-effective within a 2-year perspective; however, they reasonably underlined the sensitivity of the findings to subsequent surgical treatment in the conservative group, which goes beyond the 2-year follow-up.

Conclusions

On the basis of a systematic literature review of economic evaluations of lumbar spinal fusion, we found that the literature is limited. In other words, the use of heavy-cost techniques is not substantiated by evidence in the literature dealing with cost-effectiveness. In addition to the methodological problems relating to the economic

methodology, the reviewed literature was affected by different choices of for example, population and/or outcome measures, which hindered the mutual replication. There is a great potential for improvement of methodological quality in economic evaluations of lumbar spinal fusion and further research is imperative.

References

1. Berlin JA (1997) Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. Lancet 350:185–186
2. Brazier J, Roberts J, Tsuchiya A, Busschbach J (2004) A comparison of the EQ-5D and SF-6D across seven patient groups. Health Econ 13:873–884
3. Briggs A, Fenn P (1998) Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ 7:723–740
4. 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
5. Christensen FB, Hansen ES, Eiskjaer SP, Hoy K, Helmig P, Neumann P, Niedermann B, Bunger CE (2002) Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. Spine 27:2674–2683
6. Christensen FB, Hansen ES, Laursen M, Thomsen K, Bunger CE (2002) Long-term functional outcome of pedicle screw instrumentation as a support for posterolateral spinal fusion: randomized clinical study with a 5-year follow-up. Spine 27:1269–1277
7. Conner-Spady B, Suarez-Almazor ME (2003) Variation in the estimation of quality-adjusted life-years by different preference-based instruments. Med Care 41:791–801
8. Drummond MF (1997) Methods for the economic evaluation of health care programmes. Oxford University Press, Oxford
9. Ekman P, Moller H, Hedlund R (2005) The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: a randomized controlled study. Spine J 5:36–44
10. EuroQol group (1990) EuroQol—a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 16:199–208
11. Fairbank J (1995) Use of Oswestry Disability Index (ODI). Spine 20:1535–1537
12. Fenwick E, O'Brien BJ, Briggs A (2004) Cost-effectiveness acceptability curves—facts, fallacies and frequently asked questions. Health Econ 13:405–415
13. Fischgrund JS, Mackay M, Herkowitz HN, Brower R, Montgomery DM, Kurz LT (1997) 1997 Volvo Award Winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. Spine 22:2807–2812
14. 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
15. Fritzell P, Hagg O, Wessberg P, Nordwall A (2001) 2001 Volvo Award Winner in clinical studies: lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. Spine 26:2521–2532
16. Fritzell P, Hagg O, Wessberg P, Nordwall A (2002) Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish Lumbar Spine Study Group. Spine 27:1131–1141
17. 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
18. Goossens ME, Rutten-Van Molken MP, Kole-Snijders AM, Vlaeyen JW, Van Breukelen G, Leidl R (1998) Health economic assessment of behavioural rehabilitation in chronic low back pain: a randomised clinical trial. Health Econ 7:39–51
19. Graves N, Walker D, Raine R, Hutchings A, Roberts JA (2002) Cost data for individual patients included in clinical studies: no amount of statistical analysis can compensate for inadequate costing methods. Health Econ 11:735–739
20. Hacker RJ (1997) Comparison of interbody fusion approaches for disabling low back pain. Spine 22:660–665
21. Hoch JS, Briggs AH, Willan AR (2002) Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ 11:415–430
22. 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
23. 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
24. 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
25. Lothgren M, Zethraeus N (2000) Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Econ 9:623–630
26. Maetzel A, Li L (2002) The economic burden of low back pain: a review of studies published between 1996 and 2001. Best Pract Res Clin Rheumatol 16:23–30
27. 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
28. Maniadakis N, Gray A (2000) The economic burden of back pain in the UK. Pain 84:95–103

29. Moller H, Hedlund R (2000) Surgery versus conservative management in adult isthmic spondylolisthesis—a prospective randomized study: part 1. Spine 25:1711–1715
30. Moller H, Hedlund R (2000) Instrumented and noninstrumented posterolateral fusion in adult spondylolisthesis—a prospective randomized study: part 2. Spine 25:1716–1721
31. OECD Organisation for Economic Co-operation, Development Purchasing Power Parities and Real Expenditures—1999 Benchmark Year (2002) OECD Report 7
32. Oostenbrink JB, Koopmanschap MA, Rutten FF (2002) Standardisation of costs: the Dutch Manual for costing in economic evaluations. Pharmacoeconomics 20:443–454
33. Prolo DJ, Oklund SA, Butcher M (1986) Toward uniformity in evaluating results of lumbar spine operations. A paradigm applied to posterior lumbar interbody fusions. Spine 11:601–606
34. Ray CD (1997) Threaded fusion cages for lumbar interbody fusions. An economic comparison with 360 degrees fusions. Spine 22:681–685
35. 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. BMJ 330:1239
36. Robinson KA, Dickersin K (2002) Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. Int J Epidemiol 31:150–153
37. 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
38. Soegaard R, Christensen FB, Lauerberg I, Bünger CE (2005) Lumbar spinal fusion patients' demands to the primary health sector: evaluation of three rehabilitation protocols—a prospective randomized study. Eur Spine J (Jun 4; Epub ahead of print)
39. Stinnett AA, Mullahy J (1998) Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 18:S68–S80
40. Suarez-Almazor ME, Kendall C, Johnson JA, Skeith K, Vincent D (2000) Use of health status measures in patients with low back pain in clinical settings. Comparison of specific, generic and preference-based instruments. Rheumatology (Oxford) 39:783–790
41. Thomsen K, Christensen FB, Eiskjaer SP, Hansen ES, Fruensgaard S, Bunger CE (1997) 1997 Volvo Award Winner in clinical studies. The effect of pedicle screw instrumentation on functional outcome and fusion rates in posterolateral lumbar spinal fusion: a prospective, randomized clinical study. Spine 22:2813–2822
42. van Tulder M, Furlan A, Bombardier C, Bouter L (2003) Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. Spine 28:1290–1299
43. van Tulder MW, Assendelft WJ, Koes BW, Bouter LM (1997) Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for spinal disorders. Spine 22:2323–2330
44. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Knipschild PG (1998) Balneotherapy and quality assessment: interobserver reliability of the Maastricht criteria list and the need for blinded quality assessment. J Clin Epidemiol 51:335–341
45. Willan AR, O'Brien BJ (1999) Sample size and power issues in estimating incremental cost-effectiveness ratios from clinical trials data. Health Econ 8:203–211