

Cost effectiveness of disc prosthesis versus lumbar fusion in patients with chronic low back pain: randomized controlled trial with 2-year follow-up

Peter Fritzell · Svante Berg · Fredrik Borgström · Tycho Tullberg · Hans Tropp

Received: 23 December 2009 / Revised: 12 September 2010 / Accepted: 17 October 2010 / Published online: 5 November 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract This randomized controlled health economic study assesses the cost-effectiveness of the concept of total disc replacement (TDR) (Charité/Prodisc/Maverick) when compared with the concept of instrumented lumbar fusion (FUS) [posterior lumbar fusion (PLF) /posterior lumbar interbody fusion (PLIF)]. Social and healthcare perspectives after 2 years are reported. In all, 152 patients were randomized to either TDR ($n = 80$) or lumbar FUS ($n = 72$). Cost to society (total mean cost/patient, Swedish kronor = SEK, standard deviation) for TDR was SEK 599,560 (400,272), and for lumbar FUS SEK 685,919 (422,903) (ns). The difference was not significant: SEK 86,359 (−45,605 to 214,332). TDR was significantly less

costly from a healthcare perspective, SEK 22,996 (1,202 to 43,055). Number of days on sick leave among those who returned to work was 185 (146) in the TDR group, and 252 (189) in the FUS group (ns). Using EQ-5D, the total gain in quality adjusted life years (QALYs) over 2 years was 0.41 units for TDR and 0.40 units for FUS (ns). Based on EQ-5D, the incremental cost-effectiveness ratio (ICER) of using TDR instead of FUS was difficult to analyze due to the “non-difference” in treatment outcome, which is why cost/QALY was not meaningful to define. Using cost-effectiveness probabilistic analysis, the net benefit (with CI) was found to be SEK 91,359 (−73,643 to 249,114) (ns). We used the currency of 2006 where 1 EURO = 9.26 SEK and 1 USD = 7.38 SEK. It was not possible to state whether TDR or FUS is more cost-effective after 2 years. Since disc replacement and lumbar fusion are based on different conceptual approaches, it is important to follow these results over time.

P. Fritzell (✉)
Department of Orthopedic Surgery, Center for Clinical Research,
Falun Hospital, Dalarna, Nissers Väg 3,
SE-79182 Falun, Sweden
e-mail: peter.fritzell@ltdalarna.se

S. Berg · T. Tullberg
Stockholm Spine Center, Löwenströmska Hospital,
Upplands Väsby, SE-19489 Stockholm, Sweden
e-mail: svante.berg@spinecenter.se

T. Tullberg
e-mail: tycho.tullberg@spinecenter.se

F. Borgström
i3 Innovus, Klarabergsviadukten 90 D,
SE-111 64 Stockholm, Sweden
e-mail: fredrik.borgstrom@i3innovus.com

F. Borgström
LIME/MMC, Karolinska Institutet, Stockholm, Sweden

H. Tropp
Department of Orthopedic Surgery, University Hospital,
SE-58183 Linköping, Sweden
e-mail: hans.tropp@lio.se

Keywords Disc prosthesis · Lumbar fusion · Cost-effectiveness · Cost-utility · Health economic evaluation

Introduction

In Sweden, as in many other European countries, total healthcare costs for society are reported to be approximately 8% of gross national product [1], and the total costs for low back pain and also chronic low back pain have been reported to be substantial [2, 3]. Approximately 80% of costs attributable to back pain are reportedly indirect costs while 10–20% can be attributed to direct medical care, and up to 75% of all costs have been ascribed to those still on sick leave after 3 months chronic low back pain (CLBP)

due to loss of productivity [4]. Inpatient treatment for back pain in Sweden has been reported to account for 17% of total healthcare costs to society [5]. Surgery for back pain accounts for a relatively small but rising part of total hospital costs, and thereby also health care costs, for back pain, which may be attributed to expanded indications for surgery and new surgical techniques.

The gold standard for selected patients today is lumbar fusion (FUS), aimed at reducing pain, decreasing disability (including returning to work), and increasing quality of life. A variety of techniques can be used to perform fusion surgery; so far no specific technique has proven to be superior in a randomly assigned patient population. It has been reported in several reviews, and also in the comprehensive Swedish spine register (“SweSpine”: <http://www.4s.nu>), that average patient satisfaction after fusion surgery is approximately 70% [6–8]. It is therefore not surprising that new surgical procedures are introduced almost yearly, and that patient selection prior to undertaking any type of lumbar surgery is stressed in the literature [9].

Total disc replacement has gained in popularity in recent years, although the technique has been known since the early 1980s [10]. One potential advantage of TDR may be the removal of a possibly painful disc. Also, restoring and preserving mobility in the motion segment(s) could theoretically minimize risk of future discomfort caused by progressive degenerative changes in adjacent motion segments, which in theory may result from unfavorable altered biomechanics due to fusion [11–14].

When new technologies like TDR are introduced to the market, it should be in the interest of all parties to evaluate whether they are cost-effective—do they provide more value compared with today’s gold standard? Cost differences and clinical differences should be compared, as well as cost-effectiveness/utility, which is often expressed as cost/quality adjusted life years (QALYs) gained [15].

This health economic RCT with 2-year follow-up was conducted at Study Center, the most productive spine surgery department in Sweden with more than 1,200 elective procedures performed annually for various degenerative spine disorders. The economic perspectives addressed were those of society and the healthcare sector. Clinical results were presented in a recent article in the ESJ [16], see “Discussion”. Follow-up was 2 years. Neither the study secretary (CB) nor the first author of this paper (PF), were associated with the study department, Stockholm Spine Center.

The aims of the study were to:

1. Define and compare costs associated with TDR in the lumbar spine, and instrumented lumbar FUS.
2. Compare cost-effectiveness/utility using the quality of life instrument EQ-5D.

Patients and methods

Patients had suffered at least 12 months from what was understood to be discogenic low back pain in one or two motion segments between L3 and S1 [16]. Additional nonspecific leg pain was allowed. Nonspecific conservative treatment had been tried and failed. Diagnosis was mainly based on medical history, clinical examination, radiographs, and MRI, though some patients had also undergone preoperative discography and facet blocks in an effort to identify pain-generating segment(s) [17].

Between 2003 and 2005, a total of 152 selected patients (age 21–55) suffering from therapy-resistant CLBP were randomized by means of closed envelope technique to either TDR ($n = 80$), or to instrumented lumbar FUS ($n = 72$) [16]. See Flow chart Fig. 1, and inclusion criteria Table 1. For demographics and baseline status, see Table 2.

In the disc replacement group, patients were randomized to receive one of three prostheses: Charité ($n = 26$, DePuySpine, Raynham, MA, USA), Prodisc ($n = 28$, Synthes Spine, West Chester, PA, USA), or Maverick ($n = 26$, Medtronic, Memphis, TE, USA). When TDR was performed, a vascular surgeon prepared the anterior retroperitoneal approach, and the spine surgeon removed the disc, mobilized the segment, and inserted the disc prosthesis. Of the 72 fusion procedures, 44 were performed as an instrumented posterior lumbar fusion (PLF) and 28 as an instrumented posterior lumbar interbody fusion (PLIF),

Cost Data,

- Databases, Patient records, Questionnaires (Cost diary)
1. Hospital care Stockholm Spine Center, including:
 - Physician visits
 - Physician other interventions
 - Physical therapist visits
 - X-ray, CT, MRI
 - Disc provocation
 - Facet blocks
 - Corset
 - Surgical procedure
 - Hospital stay
 - Reoperation
 - Rehabilitation
 2. Primary care, including visits:
 - Physician
 - Physical therapist
 - Chiropractor/Naprapath
 3. Back-related drug use
 4. Family support

Effect Data in this study,

1. Quality of life, EQ-5D (0–1 max)
EQ-5D was used for cost-effectiveness analysis.
2. Back pain, VAS (0–100 max)
3. Disability, ODI (0–100 max)
4. Self-rating of success (ordinal)

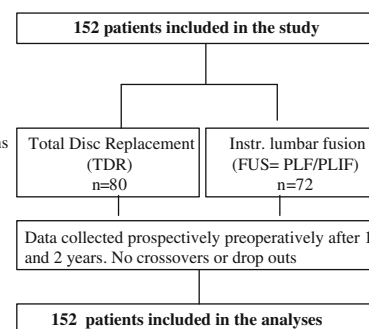


Fig. 1 Comparative health economic analyses were performed between TDR using either Charité, Prodisc or Maverick and instrumented lumbar FUS, performed either as PLF (posterior lumbar fusion), or as a PLIF (posterior lumbar interbody fusion). Analyses of statistical significance were performed as Intention to treat, which was the same as per protocol, as there were no crossovers in this study

Table 1 Study population

Inclusion criteria	Exclusion criteria
Age 20–55 years	Spinal stenosis with a need for decompression
Low back pain with or without leg pain for more than 1 year	Facet joint arthritis more than mild
If leg pain, LBP should dominate	Three or more painful levels on clinical examination
Patients not better after a rehab program scheduled for >3 months	No obvious painful level(s) at discography—if performed
Radiographic confirmation of disc degeneration (X-ray, CT, MRI)	Isthmic spondylolysis/olisthesis
Oswestry disability index of at least 30 or back pain (VAS) over 50/100 the week before inclusion	Deg. of spondylolisthesis (>3 mm)
Patients open to the two treatment options	Major deformity
Signed informed consent	Clinically important osteoporosis
	Previous relevant spine procedures
	Previous disc/vertebral infections
	Spinal tumor
	Inability to understand information
	Non-Swedish residents (possible problems with follow-up)
	Pregnancy
	Other medical contraindications

Table 2 Demographics and status at index episode

	TDR <i>n</i> = 80	FUS <i>n</i> = 72	<i>p</i>
Women (%)	60	58	ns
Age (SD)	40.2 (8.1)	38.5 (7.8)	ns
On sick leave (%)	70	70	ns
Blue collar workers (%)	46	37	ns
Smokers (%)	10	11	ns
Previous spinal surgery (%)	12	11	ns
LBP \geq 2 years duration (%)	79	87	ns
VAS back pain (SD)	62.3 (20.8)	58.5 (21.7)	ns
VAS leg pain (SD)	32.8 (26.4)	43.7 (28.2)	0.02
EQ-5D (SD)	0.42 (0.31)	0.36 (0.33)	ns
ODI (SD)	41.8 (11.8)	41.2 (14.6)	ns

Index episode = TDR or FUS plus postop indoor stay
TDR total disc replacement, *FUS* fusion

according to the preference of the surgeon. Instruments used were the Monarch for PLF, supplemented with Brantigan cages in case of a PLIF (both implants DePuy-Spine, Raynham, MA, USA). Approximately 80% (118/152) of all procedures were performed by one surgeon (SB). In the current study we compared the “concepts” of prosthesis and fusion, why we did not compare costs or cost-effectiveness with regard to the specific implants/procedures used. This was because we did not consider it probable that any procedure would produce relevantly different results in a 2-year period.

Directly after the “index procedure” (TDR or FUS), patients in both groups were instructed to be as active and mobile as possible. A soft orthosis was used for 6 weeks in

the TDR group as it was generally recommended in the literature at the time of the study. An orthosis was not considered relevant in the FUS group since all patients were stabilized using instrumentation. Discharge from the hospital was determined based on the patient’s ability to ambulate independently, under oral pain management. After this “index episode” (index procedure plus inpatient stay), the operating surgeon standardized the sick leave period to 2 months for all patients, with 50% leave for an additional month. After this 3-month period, the patients had the option to contact either the operating surgeon or the primary care doctor to assess need for additional sick leave.

Cost data

Direct and indirect costs were identified, measured, and valued. In Table 3 costs “per item” are presented. Non-surgical cost items specified (injections and blocks included) were added as patients utilized them after the index episode (surgical procedure). These treatments/diagnostic procedures were valued using a local “cost per patient” reimbursement system used in Study Center, and were used also for cost estimations regarding a few surgical reinterventions performed at other centers.

All costs, with the exception for three preoperative radiographic examinations (one X-ray, one CT-scan, and one MRI), are included from the index episode and onwards. Direct cost was defined as costs incurred through utilization of healthcare resources thus specified in Table 3, by each individual patient. Indirect cost was defined as patient reimbursement to cover work absenteeism in Sweden (sick leave).

Table 3 Costs per item—2006 currency

Hospital costs ^a	SEK	EURO	USD
Physician per visit	2,003	216	271
Other actions by physician ^b	501	54	68
X-ray per investigation	600	65	81
CT per investigation	1,500	162	203
MRI per investigation	2,500	270	339
Diagnostic tests ^a			
Disc injection/episode	11,000	1,188	1,491
Facet injection/episode	11,000	1,188	1,491
Corset	1,000	108	136
Index episode (TDR or FUS) ^a			
Total disc replacement (TDR)	105,613	11,405	14,311
Instrumented fusion (PLF or PLIF)	107,111	11,567	14,514
Complications/adverse events ^a			
Debridement/postoperative infection ^c	21,029	2,271	2,849
Extraction of implant ^c	26,608	2,873	3,605
Refusion due to pain ^c	109,317	11,805	14,813
Reoperation abdominal hernia ^c	20,185	2,180	2,735
Indoor rehabilitation period ^c	41,000	4,428	5,556
Primary care ^c			
Physician	1,447	156	196
Physical therapist	723	78	98
Physician private	724	78	98
Physical therapist private	1,447	156	196
Chiropractors and naprapaths private	724	78	98
Relatives, cost/hour			
Travel, shopping, house cleaning	280	30	38
Medical drug costs according to “FASS” ^d			
Reimbursement 1 day on sick leave			
Male	2,397	259	325
Female	2,059	222	279

1 EURO = SEK 9.26 SEK and 1 USD = SEK 7.38

^a Costs according to “cost per patient”, Stockholm Spine Center. Index episode = Index procedure (TDR or FUS) plus additional inpatient stay

^b Estimated cost 1/4 of cost/visit

^c National board of health and welfare (social styrelsen)

^d FASS, a publication with information of drug prizes

^e Sensitivity analysis performed. Reoperations merged. Also 3% discount per year

We obtained information from Statistics Sweden (SCB. <http://www.scb.se>), the Study Center and from the cost diary [18]. The mailed cost diary, used to assess utilization of different services, was submitted by patients to the study secretary (CB) after 1, 3, 6, 12, 18, and 24 months. All information about part-time or full-time sick leave, as well as part-time or full-time work (25–50–75–100%), was converted into full-day equivalents for purposes of analysis. Patients who failed to respond at any of the six specified follow-up periods were contacted by phone (CB).

Clinical data

Clinical outcome was determined based on data from the Swedish Spine Register (“SweSpine”), <http://www.4s.nu>, which is gathered by questionnaires mailed to all surgically treated spine patients in Sweden. All patients were also seen for clinical follow-up by one of the authors (SR), although no data were recorded at these visits. Health economic evaluations commonly use the EQ-5D as the main clinical outcome measure (0–1-graded scale, all minus values were transferred to zero, where 0 = “death”, and 1 = perfect quality of life). EQ-5D is a general quality of life instrument that assesses health status based on five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression [19]. In addition, clinical outcome was reflected as improvement in back pain (VAS, 0–100) and disability (ODI, 0–100) [20]. Patient-reported outcome was also presented using a five-point ordinal scale: total relief of pain—much better—better—unchanged—worse.

Analyses

The cost-effectiveness (CEA), in this context the same as cost-utility (CUA) [21], was estimated using the EQ-5D instrument. The accumulated QALYs for TDR and FUS over 24 months were calculated using area under the curve based on point estimates of QoL at baseline, 12 months, and 24 months, shown in Fig. 2. The incremental cost-effectiveness ratio (ICER), i.e., the cost per QALY gained, was estimated by dividing the difference between all relevant direct and indirect costs associated with the two procedures by the difference in effects (QALYs). The ICER indicates the difference in costs per additional effect unit gained using TDR instead of FUS.

The ICER necessarily relates to “willingness to pay” (WTP) for a QALY in order to determine whether a treatment strategy is cost-effective, in this case TDR when compared with FUS [15, 22, 23]. WTP is associated with the value a purchaser (e.g., the healthcare system) puts on a QALY. The threshold for WTP for a QALY is not established and differs from country to country. In Sweden, a usually referred range for the value of a QALY is SEK 500,000–700,000 (LFN <http://www.tlv.se>). This WTP for a QALY can be derived from the value of a statistical life as measured in a contingent valuation on road safety in Sweden [24]. The tables and figures are denominated in Swedish kronor (SEK), but costs in EURO and USD are presented as well in Table 3. 1 EURO = SEK 9.26 and 1 USD = SEK 7.38 (2006 annual exchange rates).

A probabilistic analysis using bootstrapping methods was undertaken to assess the uncertainty of the data and results were presented as cost-effectiveness planes [25, 26].

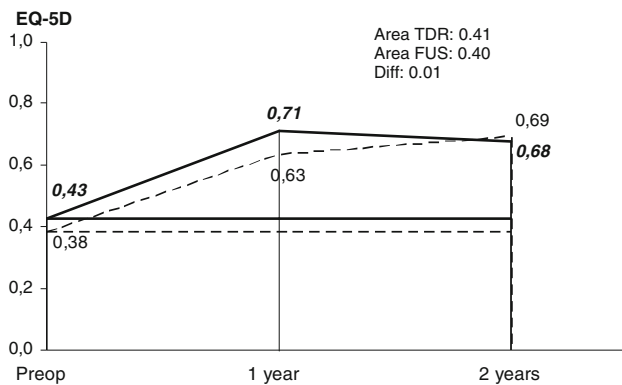


Fig. 2 Area under the curve, demonstrating total improvement after 2 years. EQ-5D: 0 = "death", 1 = "perfect quality of life". *Larger area* indicates better gain in quality of life. ***Bold and italic*** total disc replacement; *hatched* instrumented fusion

Because of potential problems with interpreting cost-effectiveness ratios in probabilistic analysis, net benefits (NBs) were used to analyze the results. For example, if both incremental values are negative, the ICER is positive, but the interpretation of results would be the opposite if incremental values were positive. Using NBs avoids this ambiguity. Net benefit is defined as $NB = \lambda Q - C$, where λ is the WTP for a QALY, Q is the incremental QALYs, and C is the incremental cost. A positive NB suggests treatment is cost-effective (depending on uncertainty, here confidence intervals), while a negative NB suggests the opposite. To assess NB, a value must be placed on a QALY; and a WTP for a QALY of SEK 500,000 (about EURO 54,000; USD 67,750) was assumed for this study.

The "intention-to-treat principle" was used to guide statistical analysis, but analysis in this study, with no crossovers, was equivalent to "per protocol" analysis.

Results after 2 years were reported from the perspective of cost to society, including both direct costs (healthcare sector) and indirect costs (loss of productivity due to sick leave). In addition, costs from a healthcare perspective were reported separately.

Sensitivity analyses were performed on excluding reoperations in both groups, also on costs for inpatient rehabilitation postop, and on discounting.

Statistics

Power calculation

We used the results presented in the Swedish Lumbar Spine Study [27], and estimated the total cost to society for lumbar fusion at SEK 700,000 (EURO 75,594; USD 94,851), and 20% less for disc replacement, or SEK 575,000 (EURO 62,095; USD 77,913). Standard deviation

(SD) was estimated from that same study at SEK 250,000 (EURO 26,998; USD 33,875). To achieve 80% power and a 5% level of significance, a total of 64 patients were required in each group. It was decided to expand the study groups to 72 patients each to allow for potential dropout. Power calculation regarding clinical outcome was reported in a separate study [16], and was in the same range as the power calculation for this health economic study.

Since improvement and return to work rate are dichotomous variables, we used the McNemar exact test; for continuous variables, we used the Wilcoxon signed test. For testing differences between the two groups regarding costs and other non-normal clinical variables, we used the Mann–Whitney U test. All baseline data were compared between the study groups using a significance level of 0.05. To analyze confidence intervals for cost and effect differences and for ratios, we used the bootstrapping technique (resampling 10,000 times) [28]. Data were entered into the SPSS statistical program (version 17.0).

Approval was obtained from the ethics committee at Karolinska Hospital in Stockholm, and all patients provided written informed consent.

Results

Follow-up of cost diaries was 100% at 1, 3, and 6 months, 95% after 12 months, 96% at 18 months, and 99% at 24 months. Follow-up of clinical outcome after 1 and 2 years was 100% for returned questionnaires. In both cases, the protocols may have been missing some answers, but missing answers were few and we did not detect any specific patterns.

Public economic assistance for patients on sick leave (part- or full-time) or other unrelated public assistance (part- or full-time) at the time of the index episode involve various types of reimbursement systems, which made detailed estimates difficult, but overall amounted to about 70% of patients in both groups, i.e., approximately 30% in both groups were working part or full time at baseline. After 2 years, according to the cost diaries, 76% in the TDR group and 72% in the FUS group were working part- or full-time (ns). Among those returning to full- or part-time work, about 20% in both groups had changed to lighter duties after 2 years. The number of full days of sick leave (with SD) following the index episode in the TDR group and the FUS group among those returning to work full- or part-time was 185 (146) and 252 (189), respectively ($p = 0.129$). Number of sick leave days among those who were working full-time after 2 years in the two groups was 139 (108) and 166 (132), respectively ($p = 0.740$), while sick leave days among those returning to part-time work was 336 (159) and 419 (173), respectively ($p = 0.211$).

Table 4 Hospital data (SD)

Index episode	TDR <i>n</i> = 80	FUS <i>n</i> = 72	<i>p</i> ^a
Intraoperative blood loss, ml ^b	560 (400)	444	ns
Operating time, hours	2.3 (0.8)	2.7 (0.6)	<0.05
Length of hospital stay, days	4.4 (1.6)	5.9 (1.2)	<0.05

^a Mann–Whitney *U* test

^b A few patients in the TDR group had an intraoperative bleeding due to vein injury, with no sequele in either patient

Table 4 presents data associated with the index episode and indoor stay after this episode. Reoperation in the TDR group was 10/80 patients (10%), and in the FUS group 26/72 patients (38%). Implant removal was carried out on 20/72 patients (28%) in the FUS group.

Costs

Societal perspective

The mean cost/patient (with SD) in the TDR group was SEK 599,560 (400,272), and in the FUS group, SEK 685,919 (422,903). The difference (with CI) expressed as TDR minus FUS was not significant: 86,359 (−214,332 to 45,605).

Healthcare perspective

The mean cost per patient in the TDR group was SEK 147,750 (73,408), and in the FUS group, SEK 170,746 (58,290). The difference was significant: 22,995 (1,202 to 43,055) Table 5. The costs per patient associated with the reinterventions in the two groups are presented in Table 6.

Table 5 Mean cost/patient in the TDR and FUSION group 2 years following treatment start (SEK)

Total disc replacement versus fusion	TDR		FUS		Diff. ^a	95% CI Bootstrapping
	(<i>n</i> = 80)	Std	(<i>n</i> = 72)	Std		
1. Hospital costs index procedure (cost per patient)	105,613	25,688	107,111	15,589	−1,498	(−8,093 to 5,338)
2. Hospital costs after index procedure ^b	15,502	24,074	33,347	42,353	−17,845	(−29,111 to −7,078) ^b
3. Total hospital costs (1 + 2) ^b	121,115	36,897	140,458	42,253	−19,343	(−32,355 to −6,747) ^b
4. Primary/Private Care	24,146	64,793	26,726	28,597	−2,580	(−16,167 to 14,795)
5. Back-related drugs	2,489	3,461	3,562	4,819	−1,073	(−2,458 to 238)
6. Healthcare perspective (direct costs) (3 + 4 + 5) ^b	147,750	73,408	170,746	58,290	−22,996	(−43,055 to −1,202) ^b
7. Family support, housekeeping	12,879	16,114	16,685	21,444	−3,806	(−10,213 to 2,095)
8. Productivity loss (indirect costs)	438,931	371,542	498,488	389,313	−59,557	(−179,503 to 62,816)
9. Societal perspective (direct + indirect costs) (6 + 7 + 8)	599,560	400,272	685,919	422,903	−86,359	(−214,332 to 45,605)

^a Minus in the Diff column means that fusion was more costly, however not significantly so with exception for ^b

^b Hospital costs and also Health care costs differed significantly between groups due to more reoperations (mostly removal of implants) in the FUS group

Table 6 Reoperations and costs after 2 years

	<i>n</i>	%	Costs/pat ^b
TDR (<i>n</i> = 80)	8	10	7,015
FUS (<i>n</i> = 72)	26 ^a	36	22,399
^b SEK diff (<i>p</i> < 0.05)			15,384

^a 20/72 (28%) patients in the FUS group had their implants removed, explaining the high reoperation rate in this group

^b SEK

Clinical outcome

Quality of life significantly improved in both surgical groups when comparing preoperative status with the situation at one and 2 years, TDR showing 0.41 EQ-5D units and FUS showing 0.40 EQ-5D units, which translates to a nonsignificant QALY gain of 0.01 units in favour of the TDR group over 2 years, Fig. 2. Other clinical results were described in detail in a previous study [16], where after 2 years, if accumulated difference was measured as in this article, no statistical difference was found between TDR and FUS with regard to back pain (VAS) or disability (ODI). On an ordinal scale (total relief–much better–better–unchanged–worse), more patients in the TDR group experienced total relief from back pain after 2 years, 30 versus 15% (*p* = 0.031), while there was no difference between groups if total relief and much better were combined Table 7.

Cost-effectiveness and net benefit

The ICER for EQ-5D using TDR instead of instrumented fusion was SEK 1,863,590 and is illustrated together with uncertainty in a cost-effectiveness plane, Fig. 3. The mean ICER was located in the southeast quadrant, indicating that

Table 7 Clinical results with standard deviations (SD)

	Baseline data		FU 2 years		TDR–FUS 2 years diff ^c	<i>p</i> ^c
	TDR ^a <i>n</i> = 80	FUS ^b <i>n</i> = 72	TDR ^a <i>n</i> = 80	FUS ^b <i>n</i> = 72		
Back pain VAS (SD)	62.3 (20.8)	58.5 (21.7)	25.4 (29.8)	29.2 (24.6)	7.6	ns
Function ODI (SD)	41.8 (11.8)	41.2 (14.6)	20.0 (19.6)	23.1 (17.0)	3.7	ns
Total relief of pain (%) ^d			24/80 (30)	11/72 (15)	15	0.031
Much better (%) ^d			32/80 (40)	35/72 (49)	–9	ns
Total relief/much better ^d			56/80 (70)	46/72 (64)	6	ns

^a Total disc replacement

^b Lumbar fusion

^c MW *U* test

^d Ordinal scale: totally pain free, much better, better, unchanged, worse

^e minus in the 2 years diff column, indicates that the difference was in favor of FUS

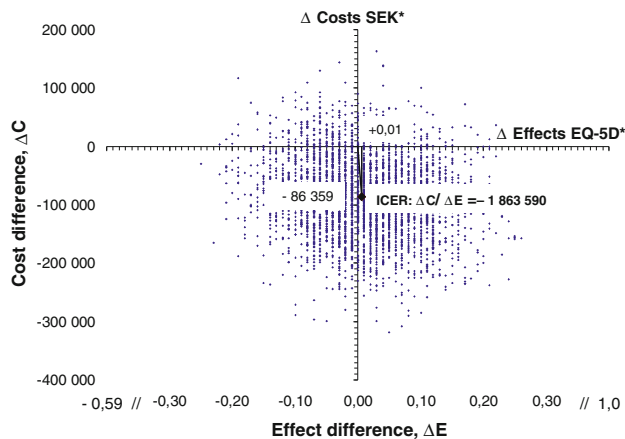


Fig. 3 Cost-effectiveness plane illustrating ICER for TDR compared with FUS. ICER is located in the south-east quadrant, indicating that TDR was less costly and slightly more effective. However not significantly so. *Difference (Δ) in costs and effects between TDR and FUS. Minus in costs and plus in effects favors TDR. Statistics: Bootstrapping was used. The dotted area represents the uncertainty (“uncertainty box”), with representations in all four quadrants, illustrating the uncertainty in the calculations

TDR was cost-saving (i.e., less costly) and associated with a small improvement in QoL, albeit not significant compared with FUS. In order to address uncertainty the net benefit approach was used. This showed that TDR could not be demonstrated as a cost-effective strategy when compared with FUS (SEK 91,359; CI –73,643 to 249,114).

Sensitivity analyses

1. Exclusion of all reoperation costs (36% in the FUS and 10% in the TDR group).
2. Calculating in a 3% discount per FU year. Exclusion of reoperations did not alter the conclusions from a

societal perspective but from the healthcare perspective, the significant cost difference between the study groups was eliminated, illustrating the impact of reoperation costs. Discounting resulted in a small but significant cost difference to the advantage of TDR, see Table 8.

Discussion

We could not demonstrate any difference with regard to societal costs or cost-effectiveness when comparing TDR with instrumented lumbar fusion (FUS) in patients suffering from CLBP since at least 1 year. In a previous study by Berg et al. (the current study uses the clinical data reported in that study), the authors could not demonstrate clinical superiority with regard to any of the two procedures using pain (VAS), Function (ODI and SF36), and patient satisfaction, although significantly more patients in the TDR group reported to be totally pain free after 2 years (30 vs. 15%) [16].

From the Societal perspective we found no significant differences in costs after 2 years. From the healthcare perspective there was a significant cost difference in favor of TDR, mainly due to a higher reoperation rate in the FUS group (see below). No concept could be demonstrated to be cost-effective on behalf of the other.

Several studies report that TDR and FUS provide approximately the same clinical outcome over 2 years [11, 16, 29–31]. However, some authors express serious concerns about using this procedure except in scientific trials [11], but others are more optimistic [32]. The “non-inferiority trial” comparing Charité with anterior standalone BAK instrumentation, reports superior outcome both clinically and economically in the disc prosthesis group [33],

Table 8 Sensitivity analyses after 2 years

	TDR (<i>n</i> = 80)	Std	FUS (<i>n</i> = 72)	Std	Diff	95% CI ^a
Healthcare perspective: reoperation costs excluded	140,736	69,650	148,348	42,784	7,611	(−11,992 to 24,783)
Healthcare perspective: discounting 3% over 2 years	147,751	73,408	170,746	58,289	21,636	(1,407 to 40,530)
Societal perspective: reoperation costs excluded	592,546	397,790	663,520	416,132	70,975	(−57,728 to 199,534)
Societal perspective: discounting 3% over 2 years	599,560	400,272	685,919	422,903	81,254	(−40,905 to 204,762)

There was a small but significant cost difference in the perspective of the health care sector when discounting with 3% per year

^a Bootstrapping

but this comparison could be questioned because of the use in the control group of stand-alone BAK cages, now recognized as an inferior fusion technique.

We used the EQ-5D instrument to assess QALYs. This instrument has been shown to be satisfactory for analyzing health-related quality of life in the field of spine surgery [34] and to be a useful tool for analyzing cost-effectiveness. In the current study, the minimal gain of 0.01 EQ-5D units (on a one-point scale) after 2 years make it practically impossible to conclude that TDR was associated with a higher gain in QALYs (diff. in costs/diff. in effects gained × time) compared with FUS, even though there was also a nonsignificant cost difference in favor of TDR. The point estimate of the ICER was located in the southeast quadrant of the cost-effectiveness plane Fig. 3 indicating a possible advantage for TDR, but the probabilistic analysis using bootstrapping showed widespread distribution of the ICERs in all four quadrants why it therefore not significantly favored either procedure.

Cost diaries were used to estimate resource utilization, including number of sick leave days after the index episode. One reason we did not use information from the Swedish Social Insurance Agency was a controversy over principals among the authorities during the study period as to how the central registers documenting length of sick leave and other reimbursement systems could be made available to researchers. However, this was a randomly controlled trial for which high internal validity can be assumed, in which we used patient based “cost diaries” on six different occasions with secretarial back up when needed, and in addition interviews by a study nurse also on six different occasions over the 2-year follow-up period, for which reason we consider the cost comparison between the two study groups to be valid.

The threshold of SEK 500,000 (EURO 54,000; USD 67,750) for WTP for an extra QALY gained used in the current study can be debated. Most commonly used WTP thresholds for a QALY vary between EURO 40,000 and 80,000 [35]. The UK has often used a threshold of GBP 30,000 (EURO 39,000) in health technology appraisals (here synonymous with health economic evaluations) conducted by NICE [36]. However, the WTP for a QALY

was not crucial for interpretation of results in this study since no difference in QALY gain was found between TDR and FUS.

According to the scientific literature, the cost-effectiveness of TDR when compared with FUS after 2 years is not yet established. Guyer et al. [37] tried to assess this issue using an economic model developed by DePuy Spine, where lumbar disc prosthesis (Charité, DePuy Spine) on one level was compared with instrumented fusion at unspecified levels. They conclude that the disc prosthesis procedure was probably cost-effective from a hospital perspective. These results could possibly be considered in light of the current study, where hospital costs were significantly higher in the fusion group, however, mainly due to more reoperations in the FUS group, which was not the case in the Guyer study.

Our pre-study power calculation estimates the 2-year cost to society for FUS at SEK 700,000 (EURO 75,600; USD 94,850), and 20% less for TDR, or SEK 575,000 (EURO 62,100; USD 77,900). The discrepancy was primarily based on the assumption that patients receiving disc prosthesis would return to work earlier than fusion patients. The power calculation proved to be fairly accurate with a total cost to society of approximately SEK 686,000 for the FUS group and SEK 600,000 for the TDR group, a difference of about 13% (ns). In addition, number of sick leave days (according to cost diaries) was approximately 60 days less in the TDR group (ns). The uncertainty reflected by wide confidence intervals and large standard deviations, hampered conclusions regarding cost-effectiveness in favor of either treatment strategy from a societal perspective. Also, this study was not powered to detect differences in number of sick leave days (post-study analysis power was found to be 0.54, compared with the pre assumed and relevant 0.8).

Our results do not provide any information concerning the cost-effectiveness of spinal surgery as a procedure per se, i.e., comparison of cost-effectiveness of surgery versus no surgery in this patient population, using the procedures addressed in this article. However, the collective data allow us to estimate potential QALY gain from the data derived in this study. Assuming that patients remain at the same

QoL level without surgery, the QALY gain over 2 years after surgery can be estimated at approximately 0.45 (on a one-point EQ-5D scale). This gain is similar to the QALY gain that has been estimated for total hip arthroplasty in Sweden [38]. Cost-effectiveness can be roughly estimated by conservatively assuming that initial hospitalization costs are the only extra cost of surgery (about SEK 140,000). The ICER of surgery when compared with no surgery would then be about SEK 340,000. If WTP for a QALY is assumed to be approximately SEK 500,000 (as in this study), this estimate can be considered cost-effective and possibly represents an underestimate of real cost-effectiveness since it does not consider issues such as fewer days of sick leave. However, if patients who do not undergo surgery improve due to natural course of disease, this assumption may be erroneous.

The current study focused on evaluating the cost-effectiveness of TDR versus FUS. Another interesting issue to consider is the economic value of spinal surgery as a procedure per se. One way to address this issue, based on the information collected for this study, may be to elaborate on potential cost reduction from patients returning to work after surgery. At baseline about 70% of patients were not working, while 2 years post-surgery 40% (TDR and FUS patients combined) of the patients had returned to work (full- or part-time). Assuming that the patient sample contains an equal proportion of part- and full-time workers, average productivity loss for these patients, had they not been operated, can be calculated as follows: 500 working days during 2 years $\times 0.7 \times \text{SEK } 1,700 = \text{SEK } 595,000$. The average cost over 2 years related to sick leave for the operated patients in this study was SEK 467,142. Potential productivity gain from operating would then be SEK 127,858 per operated patient. By multiplying this figure with the number of patients in the study (152), the total estimated value of the productivity gain would be about SEK 19 million. From this perspective spine surgery could provide considerable value in terms of productivity gain. This is a rough calculation based on 2-year data and a number of assumptions, but we find it interesting and believe it could be further addressed in a separate study that would also include reimbursement data from the social insurance agency covering both surgical patients and nonsurgical patients suffering from the same disorder. Moreover, others have come up with similar estimates for back pain patients awaiting surgery and reached the same interesting hypothesis.

Potential problems with the current study

This was a single center study conducted at the most productive spine unit in Sweden with over 1,200 procedures

for degenerative disc disorders performed annually. One surgeon (SB) performed the index operation in 80% of the patients, and it is possible that patient selection, although there were strict inclusion and exclusion criteria to balance this, and surgical skills play a role, for which reasons the results achieved in this study therefore may not be generally replicable elsewhere. However, when comparing patient demographics, baseline values and clinical outcome with other studies and also with national registers, this seems to be less of a concern.

Although the index episode was equally costly for both TDR and FUS, total hospital costs after 2 years, as well as total costs to the healthcare sector, were significantly greater in the FUS group. This difference can be ascribed to a high reoperation rate, 36% in the FUS group (of which 28% was attributed to implant removal as the implant was diagnosed by the surgeon as a pain generator), when compared with 10% in the TDR group. In most cases the indication for reoperation was determined by one surgeon (SB), who also performed 90% of these procedures. Obviously, a certain reoperation rate must be expected after any surgical procedure but considering the single center perspective, meaning that the external validity of the decision to reoperate may be called into question; we performed a sensitivity analysis where all costs for reoperation were removed from both groups. The result was that the cost difference from the perspective of both the hospital and the healthcare sector disappeared. It is possible that there is, in general, a higher reoperation rate after instrumented lumbar FUS when compared with after total disc prosthesis, but this cannot be finally concluded from this study.

The current study compares the conceptual approach using disc prosthesis (Charité, Prodisc and Maverick) with that of instrumented fusion (PLF and PLIF). The same costs were assumed for the three prostheses, and a mean cost was calculated for the two fusion techniques. It is possible that from a longer perspective, one of these procedures, and indeed one of these approaches (motion preservation vs. fusion), will turn out to be superior from the perspective of cost-effectiveness. This remains to be seen and is an important reason to continue to follow these patients for 5, 10, and 15 years, since 2 years may be considered a short time frame. The two most important variables with future potential to influence cost-effectiveness, as outlined in this study, may be differences in work capacity, and the need within any group or subgroup for additional back surgery.

Health economic comparisons between fusion and non-surgical treatment in patients suffering from CLBP have been published with somewhat different results [39–42], and it should be important to evaluate whether TDR could have advantages/disadvantages from a cost-effectiveness

perspective in this patient population. We therefore consider this RCT-based health economic evaluation with relevant follow up and no cross-over, comparing the concepts of TDR and FUS, as important. Although cost-effectiveness could not be demonstrated with regard to any concept, after 2 years we found no economic argument against using TDR in selected patients suffering from CLBP.

Conclusion

1. Societal costs showed no significant difference when compared with TDR and instrumented lumbar FUS after 2 years. From a healthcare perspective, FUS was significantly more costly, mainly due to a high reoperation rate in this group.
2. It was not possible in this study to determine whether TDR or FUS was more cost-effective for society within the 2-year time frame, although TDR was associated with less costs and a very small gain in quality of life.

Acknowledgments The study was sponsored by the following companies: DePuySpine, Synthes, Medtronic.

Conflict of interest The study was sponsored by the following companies: DePuySpine, Synthes, Medtronic. The companies had no influence in the planning, conduction, analysing or reporting of the study or the study results.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Norlund AI, Waddell G (2000) Cost of back pain in some OECD countries. In: Nachemson AL, Jonsson E (eds) Neck and back pain. The scientific evidence of causes, diagnosis, and treatment. Lippincott Williams & Wilkins, Philadelphia, pp 421–425
2. Juniper M, Le TK, Mladsi D (2009) The epidemiology, economic burden, and pharmacological treatment of chronic low back pain in France, Germany, Italy, Spain and the UK: a literature-based review. *Expert Opin Pharmacother* 10(16):2581–2592
3. Dagenais S, Caro J, Haldeman S (2008) A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine J* 8(1):8–20
4. Andersson GB (1999) Epidemiological features of chronic low-back pain. *Lancet* 354(9178):581–585
5. SBU—Swedish Council on Health Technology Assessment (2000) Back and Neck Pain. Report 145/1, 145/2, P 306. ISBN: 91-87890-60-7
6. Brox JI et al (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(17):1913–1921
7. Fairbank J et al (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. *BMJ* 330(7502):1233
8. Fritzell P et al (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(11):1131–1141
9. Hagg O et al (2003) Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. *Eur Spine J* 12(1):22–33
10. Link HD (2002) History, design and biomechanics of the LINK SB Charite artificial disc. *Eur Spine J* 11(Suppl 2):S98–S105
11. Rohlmann A, Zander T, Bergmann G (2005) Effect of total disc replacement with prodisc on intersegmental rotation of the lumbar spine. *Spine (Phila Pa 1976)* 30(7):738–743
12. De Kleuver M, Oner FC, Jacobs WC (2003) Total disc replacement for chronic low back pain: background and a systematic review of the literature. *Eur Spine J* 12(2):108–116
13. Rohlmann A et al (2009) Effect of an artificial disc on lumbar spine biomechanics: a probabilistic finite element study. *Eur Spine J* 18(1):89–97
14. Lee CS et al (2009) Risk factors for adjacent segment disease after lumbar fusion. *Eur Spine J* 10:314–319
15. Drummond MF et al (1997) Methods for the economic evaluation of health care programmes, 2nd edn. Oxford University Press, Oxford
16. Berg S et al (2009) Total disc replacement compared to lumbar fusion: a randomised controlled trial with 2-year follow-up. *Eur Spine J* 19:205–214
17. Ito M et al (1998) Predictive signs of discogenic lumbar pain on magnetic resonance imaging with discography correlation. *Spine* 23(11):1252–1258
18. Goossens ME et al (2000) The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol* 53(7):688–695
19. Burstrom K, Johannesson M, Diderichsen F (2001) Health-related quality of life by disease and socio-economic group in the general population in Sweden. *Health Policy* 55(1):51–69
20. Fairbank JC et al (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66(8):271–273
21. Johannesson M, O’Conor RM (1997) Cost-utility analysis from a societal perspective. *Health Policy* 39(3):241–253
22. Blumenschein K et al (2001) Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. *J Health Econ* 20(3):441–457
23. Johannesson M (1996) The willingness to pay for health changes, the human-capital approach and the external costs. *Health Policy* 36(3):231–244
24. Persson U, Hjelmgren J (2003) Health services need knowledge of how the public values health. *Lakartidningen* 100(43):3436–3437
25. Black WC (1990) The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 10(3):212–214
26. Briggs AH, O’Brien BJ, Blackhouse G (2002) THINKING OUTSIDE THE BOX: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Annu Rev Public Health* 23:377–401
27. Fritzell P et al (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(4):421–434 (discussion Z3)
28. Efron B, Tibshirani RJ (1993) An Introduction to the Bootstrap. Chapman & Hall, New York, pp 168–177

29. Freeman BJ, Davenport J (2006) Total disc replacement in the lumbar spine: a systematic review of the literature. *Eur Spine J* 15(Suppl 3):S439–S447
30. Guyer RD et al (2009) Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: 5-year follow-up. *Spine J* 9(5):374–386
31. Le Huec JC et al (2005) Clinical results of Maverick lumbar total disc replacement: 2-year prospective follow-up. *Orthop Clin North Am* 36(3):315–322
32. Zigler J et al (2007) Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine* 32(11):1155–1162 (discussion 1163)
33. Blumenthal S et al (2005) A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. *Spine* 30(14):1565–1575 (discussion E387-391)
34. Solberg TK et al (2005) Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. *Eur Spine J* 14(10):1000–1007
35. Eichler HG et al (2004) Use of cost-effectiveness analysis in health-care resource allocation decision-making: how are cost-effectiveness thresholds expected to emerge? *Value Health* 7(5):518–528
36. Raftery J (2001) NICE: faster access to modern treatments? analysis of guidance on health technologies. *Bmj* 323(7324):1300–1303
37. Guyer RD, Tromanhauser SG, Regan JJ (2007) An economic model of one-level lumbar arthroplasty versus fusion. *Spine J* 7(5):558–562
38. Malchau H et al (2005) Presidential guest address: the Swedish Hip Registry: increasing the sensitivity by patient outcome data. *Clin Orthop Relat Res* 441:19–29
39. Soegaard R, Christensen FB (2005) Health economic evaluation in lumbar spinal fusion: a systematic literature review anno. *Eur Spine J* 15:1165–1173
40. Soegaard R et al (2007) Circumferential fusion is dominant over posterolateral fusion in a long-term perspective: cost-utility evaluation of a randomized controlled trial in severe, chronic low back pain. *Spine* 32(22):2405–2414
41. Fritzell P et al (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* February 29(4):421–434
42. Rivero-Arias O et al (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(7502):1239