

Determinants of cost-effectiveness in lumbar spinal fusion using the net benefit framework: a 2-year follow-up study among 695 patients

Rikke Soegaard · Cody E. Bünger · Terkel Christiansen · Finn B. Christensen

Received: 3 January 2007 / Revised: 29 March 2007 / Accepted: 6 April 2007 / Published online: 23 May 2007
© Springer-Verlag 2007

Abstract Up to one third of patients undergoing lumbar spinal fusion show no improvement after the procedure and thus, despite evidence from RCTs, there might be a rationale for observational studies clarifying indications. Similarly, selection of the right patients for the right procedure could have significant impact on cost-effectiveness, which in some countries, in turn, affects whether procedures are to be available through the National Health Service. The aim of this study was to investigate determinants of cost-effectiveness in lumbar spinal fusion. An observational cohort study with 2-year follow-up was conducted: 695 patients who underwent lumbar spinal fusion from 1996 to 2002 were included and followed for 2 years. Patients had a localized segmental pathology and were diagnosed with MRI-verified isthmic spondylolisthesis (26%) or disc degeneration (74%). The surgical techniques were non-instrumented posterolateral fusion (14%), instrumented posterolateral fusion (54%), and circumferential fusion (32%). Societal costs and improvement in functional disability (Dallas Pain Questionnaire) were transformed into a net benefit measure. Classical linear regression of the net benefit was conducted using predictors of age, sex, diagnosis, duration of

pain, smoking habits, occupational status, severity of disability, emotional distress, surgical technique, and number of levels fused. The main results were that two determinants were found to negatively influence net benefit: smoking and diagnosis, whereas two others were found to be positively associated with the net benefit: severe disability and emotional distress. In conclusion, predicting net benefit reverses the picture usually seen in studies predicting clinical outcomes, because the response variable is based on improvement over time rather than end-point measures alone. Smoking habits, diagnosis, pre-operative disability, and pre-operative emotional distress were found to be significantly associated with the net benefit of spinal fusion.

Keywords Lumbar spinal fusion · Chronic low back pain · Economics · Net benefit · Determinants

Introduction

Since the 1930s, lumbar spinal fusion has been a controversial intervention for the treatment of chronic low back pain [5]. The level of supporting evidence has increased during the last decade but similarly, so has the rate of surgery, at a pace not seen in other major orthopedic procedures. Deyo et al. [9] investigated potential explanations for this increase and found no association with reports of clarified indications or improved efficacy, but with the approval of new implants.

The increasing rate of surgeries constitute an immediate burden on health care budgets; however, this burden can very quickly become inferior to the counterfactual scenario, namely that costs of care are adding up for the remaining life-expectancy of patients should they not

R. Soegaard (✉) · C. E. Bünger · F. B. Christensen
Spine Unit, Orthopaedic Research Laboratory,
University Hospital of Aarhus, Norrebrogade 44,
Building 1 A, 8000 Aarhus C, Denmark
e-mail: rikke.sogaard@ki.au.dk

T. Christiansen
Health Economics Unit, Institute of Public Health,
University of Southern Denmark, Odense, Denmark

receive the optimal treatment. Lumbar spinal fusion has been found to be the optimal treatment in selected patients; however, a proportion of 24–30% of patients do not benefit [6, 11, 22, 27]. For that reason, despite evidence from RCTs, there might be a persisting rationale for observational studies that elucidate indications for this procedure.

A definition of the magnitude of a clinically relevant improvement is not straightforward; hence the proportion of patients not having benefited from surgery is somehow arbitrary. Nowadays, cost-effectiveness studies enter the scene in a similar manner as do clinical RCTs and rejects procedures on grounds of poor cost-effectiveness for, perhaps, only a fraction of patients. Therefore, in parallel to the persisting rationale for observational clinical studies, there may also be a rationale for cost-effectiveness studies stepping back and seeking to understand the mechanisms underlying the ratio of costs and effects. For example, in a similar manner as for clinical observational studies, if a procedure is associated with very high costs for particular patient characteristics it may be unfair to pool over these characteristics when investigating cost-effectiveness.

Importantly, this perspective is not to suggest selection of patients on grounds of their expected cost-effectiveness ratio, which is conflicting with ethics of equity. Rather is the rationale understood from a counterfactual scenario of no understanding of determinants for cost-effectiveness in which procedures may fail to prove cost-effective simply because they are tested in a heterogeneous population in terms of cost-effectiveness.

A recent study by Soegaard et al. [23] investigated determinants of net benefit among consecutive patients who had undergone lumbar spinal fusion. Pre-operative psychosocial distress and smoking were found to significantly influence net benefit while type of surgical procedure was not, hence, the heterogeneity of the study population may hold even greater impact to cost-effectiveness than the surgical technique itself. This study, the first to utilize such an approach, is, however, limited by virtue of its narrow perspective and limited sample size.

The aim of present study was to investigate determinants of net benefit in a 2-year societal perspective among patients undergoing lumbar spinal fusion. Patient characteristics previously found capable of predicting clinical outcomes were hypothesized to also be able to predict cost-effectiveness: smoking [3], diagnosis [3, 4], duration of pain [25], pre-operative occupational status [3, 28], and pre-operative psychological distress [8, 14, 25]. Surgical technique and number of levels fused were included to adjust for possible confounding in all analysis.

Materials and methods

Study design and material

The study design was an observational cohort study with a 2-year follow-up among patients undergoing lumbar spinal fusion from January 1996 to December 2002 in a single center at a university hospital. Patient data were recruited from a clinical database prospectively monitoring outcomes; criteria for inclusion were (1) severe, chronic low back pain and leg pain due to localized segmental pathology caused by a) isthmic spondylolisthesis (grade I-II), (b) primary disc degeneration (no previous surgery), or (c) secondary disc degeneration (previous spinal surgery), and (2) pre-operative data on functional disability.

Effect measure

The primary effect measure was change in functional disability from pre-operatively to 2 years post-operatively as assessed by means of the Dallas Pain Questionnaire (DPQ) [16]. In contrast to the original discontinuous visual-analogue-scale of the DPQ, our data was collected using a continuous visual-analogue-scale making data amendable to classical linear regression. Prior to calculating change over time, the four disability scales were summarized into a single-index.

Imputation was conducted to replace missing values in single items of otherwise completed questionnaires; horizontal (intra-patient) means of non-missing values within individual areas of functional disability were calculated and used for imputation. This procedure was applied in pre-operative data of 16% of the patients and in post-operative data of 14% of the patients.

Non-response in 2-year disability was imputed by means of a regression approach. Within the sample of patients having complete follow-up, a model was fitted that predicted 2-year post-operative single-item-values from the pre-operative and 1-year post-operative measurements. Imputation of non-response was conducted in 25% of patients and the impact of such was investigated by appending two parallel analyses: one in respondents only and another in the whole cohort, assuming non-respondents performed only similar to the lower quartile of respondents.

Economic methodology

The overall perspective of the present study was a societal perspective and hence, all possible activity and

resource consumption occurring as consequence of the fusion regimen contributes to overall cost estimates. In economic terms, costs are defined as total opportunity costs from a societal perspective. Unless otherwise described below, the time frame was fixed by the date of index surgery –180 days (to encompass diagnostics) to the date of index surgery +720 days (2-year follow-up). All costs are expressed in 2004-EUR and both costs and effects were discounted at an annual rate of 3%.

Index surgery, re-operations, and re-hospitalizations

Service utilization in the secondary health care sector, that is, index surgery, re-operations, re-hospitalizations, and visits to outpatient clinics, regardless of whether they took place at public or private hospitals, were recorded uniquely per patient in The National Patient Registry (run by the National Health Service). Service utilization was valued using national average unit costs provided by the Diagnosis-Related-Grouping (DRG) system.

Service utilization in the primary health care sector

Contacts to general practitioners, medical specialists, physiotherapists, psychologists, or chiropractors were recorded uniquely per patient in The National Health Insurance Service Registry (run by the National Health Service). Contacts were valued from collective agreements that are negotiated every second year between the National Health Service and the professional societies of medical specialists or therapists. Specific to contacts to general practitioners, costs were marked-up to 130% as a “charge to cost mark-up” [1]. Valuation of recumbent transportation (at the time of discharge and for the first visit to the outpatient clinic for clinical follow-up) is dependent on the geographical distance from the patient’s home to the hospital but, for practical reasons, an average value was adapted from the literature [23].

Medication

Prescribed medication was recorded uniquely per patient in the Register of Prescribed Medication (run by Danish Medicines Agency). Medication was valued using market prices.

Patient’s costs

Over the counter medication [22], visits to complementary practitioners [22], and family time and paid help [12] were

estimated from the literature and applied as a fixed base-cost. Transportation for hospital treatment was included and valued by half the market price of recumbent transportation that was supplied by the primary health care sector.

Productivity costs

Productivity costs incurred by the patient’s absenteeism and work disability were recorded from the Social Science Research Register (run by Statistics Denmark), which records number of days at work, part-time work, sick leave, retirement, pension etc. uniquely per patient. The Friction Cost method was used for valuation of productivity costs [15]. This method takes into account that an employer can replace an employee that has been work disabled within a certain friction time, therefore, society suffers productivity costs in this period of friction only (in contrast to the Human Capital method [21] that fully accumulates absent time as production loss). Also included in productivity costs are the costs of hiring and training an employee for replacement in the event of long-term sick leave (>14 days) or early retirement due to illness. The cost of replacing an employee was estimated at EUR 1,341. Calculations were based on a three-month friction period and 2004 national average gross incomes in age- and gender-matched groups.

Statistical analysis

A net benefit framework was applied to convert cost-effectiveness ratios into a parameter of net benefit [24]. The net-health-benefit (NHB) for patient i is defined by

$$\text{NHB}_{i\lambda} = (E_i^0 - E_i^2) - (C_i/\lambda)$$

where E denotes effect (disability), superscripts of 0 and 2 refer to baseline and 2-year post-operative status, C denotes total costs, and λ denotes decision-makers’ threshold for willingness to pay per effect unit (hereafter referred to as threshold).

Possible determinants were investigated by means of classical ordinary-least-squares regression with 95% bias-corrected, bootstrapped confidence intervals for the coefficients. Bootstrapping refers to a non-parametric simulation technique where repeated drawings from the sample are taken, assuming the sample is representative of the population [10]. To investigate whether assumptions of classical linear regression are satisfied, the extent of interaction among coefficients was evaluated by means of pair-wise correlation and scatter diagrams. It was found necessary to adjust for the interaction between pre-oper-

ative occupational statuses and age (receiving a pension is associated with older age). Model validity was further investigated by evaluating whether residuals had the same distributions, could be characterized by a normal distribution, and whether any observations deviated. The following plots were evaluated and demonstrated with reasonable validity of the model: (1) residuals versus predicted values, (2) residuals versus explanatory variables, (3) standardized residuals versus explanatory variables, (4) leverage versus explanatory variables, and (5) leverage versus standardized residuals. Analysis was conducted using statistical software STATA Intercooled (ver 8.0, StataCorp, USA).

Ethical issues

The study was completely observational and the collection of register data was approved by the national data protection agency. The purpose of this study was blinded to patients and doctors.

Results

Complete follow up was attained in 100% of patients with respect to costs and in 75% of patients with respect to effects; however, on the basis of imputation of the 2-year follow-up questionnaire in non-respondents, all patients were included for analysis. The reasons for patients not filling in the 2-year follow-up questionnaire are unknown, but a large proportion is probably attributable to logistical problems because data were collected on a routine basis and not under the enthusiastic responsibility of a researcher. Table 1 shows baseline characteristics of the full study base ($n = 695$) and the subgroup of non-respondents to the 2-year follow-up questionnaire ($n = 177$).

Figure 1 illustrates patients' functional disability pre-operatively and 2 years post-operatively. There was a highly significant improvement ($P < 0.0001$) in all dimensions of functional disability.

Table 2 illustrates patients' service utilization and its related costs in different perspectives. The sub-perspective of productivity costs represents the costs of patient's absenteeism or early retirement from the labor market valued by two different methods (Human Capital and Friction Cost). Evidently, the magnitude of productivity costs is dependent on the method chosen for valuation, and for the grand total, the conservative Friction Cost method was chosen. The grand total of the regimen of lumbar spinal fusion was an average cost of EUR 38,937 (37,942; 40,015) by a 2-year societal perspective.

The following range of determinants were tested in regression models: age, gender, diagnosis, duration of pain, smoking habits, occupational status, severity of disability, and emotional distress. In addition, analysis was controlled by including surgical technique and number of levels fused. For a threshold of decision-makers' willingness to pay \geq EUR 250 per effect unit, only four determinants were found to significantly influence average net benefit: diagnosis, smoking, severity of disability, and emotional distress as illustrated in Fig. 2. Being diagnosed with secondary degeneration (degeneration + previous spinal surgery in comparison to degeneration without previous spinal surgery) significantly reduces the net benefit for a threshold \leq EUR 375, this being due both to a poorer outcome and increased costs. The influence of patients' smoking habits is significant regardless of the threshold as smoking is associated with a poorer clinical outcome and, at the same time, increased costs. The influence of the pre-operative level of disability is significant for thresholds at, or above, the level accepted today.

Severe disability in comparison with moderate or intermediate disability increases the net benefit. The explanation behind this more or less surprising finding can be linked to the effect parameter of improvement over time; a high value of pre-operative disability presents a relatively great potential for improvement whereas a moderate value has no such potential. It is primarily clinical improvement (not cost savings) that contributes to the higher net benefit among patients with severe disability pre-operatively. The fourth determinant, expressing significant influence to the net benefit, is pre-operative emotional distress (in comparison with no emotional distress). The course of this curve is similar to that of disability: pre-operative emotional distress is associated with increasing net benefit. This, again, is due to the fact that pre-operative emotional distress leaves a relatively greater potential for improvement in comparison with no emotional distress; however, the magnitude of this relationship is relatively high because emotional distress is also associated with decreased costs.

Twenty-five per cent of patients ($n = 177$) did not report endpoint disability and their 2-year status was predicted from their baseline and 1-year status. The uncertainty associated with such methodological choice relates to whether non-respondents were systematically different from respondents. Assuming they were not different, a complete parallel analysis was undertaken for respondents only; the results were not significantly deviating from those of base-case analysis. Assuming non-respondents were different, and would have expressed a value of disability only similar to the lower quartile of the disability level of respondents, another complete parallel analysis was

Table 1 Baseline characteristics at trial entry of 695 patients having lumbar spinal fusion

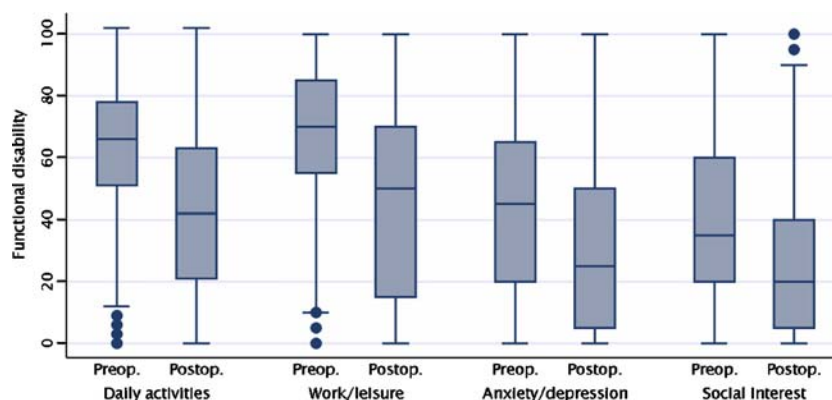
Baseline characteristics	Non-respondents 2-year follow-up (<i>n</i> = 177)	All patients (<i>n</i> = 695)
Age, mean (SD)	52 (14)	50 (13)
Females	103 (58)	423 (61)
Diagnosis ^a		
Spondylolisthesis	33 (18)	180 (26)
Primary degeneration	73 (41)	252 (36)
Secondary degeneration	72 (41)	262 (38)
Duration of pain >2 years	116 (72)	482 (73)
Smokers	79 (55)	290 (48)
Occupational status		
Working	44 (25)	206 (30)
Sick-listed	41 (23)	212 (31)
Retired or pensioned due to sickness	52 (29)	184 (26)
Retired or pensioned due to age	40 (23)	93 (13)
Classification of disability ^b		
Slight or intermediate disability	25 (14)	101 (15)
Major disability	49 (28)	210 (31)
Major disability + emotional distress	100 (57)	376 (55)
Type of spinal fusion		
Non-instrumented posterolateral	36 (20)	95 (14)
Instrumented posterolateral	91 (52)	373 (54)
Circumferential	49 (28)	225 (32)
Vertebral levels fused at index surgery >1	111 (63)	375 (54)

Values are number of patients (%) unless stated otherwise

^a Primary degeneration = no previous surgery and secondary degeneration = degeneration after previous surgery

^b According to classification by Ozguler et al. [20]

Fig. 1 Effectiveness of lumbar spinal fusion in a cohort of 695 patients; functional disability at baseline and 2-year follow-up in four dimensions of life measured by the Dallas Pain Questionnaire. *Boxes* represent medians, 25- and 75-percentiles, and whiskers are upper and lower adjacent values to the percentiles. All changes over time are significant ($P < 0.0001$)



undertaken; again, results were not significantly deviating from base-case analysis.

Discussion

The influence of potential determinants on the net benefit of lumbar spinal fusion was investigated in a cohort of 695 patients. Two determinants were found to be negatively associated with the net benefit: smoking (in comparison with no smoking) and secondary degenera-

tion (degeneration + previous spinal surgery in comparison with degeneration without previous spinal surgery), whereas two others were found to be positively associated with the net benefit: pre-operative severe disability (in comparison with moderate or intermediate disability) and pre-operative emotional distress (in comparison with no emotional distress). Whereas several other studies have investigated predictors of clinical outcomes in lumbar spinal fusion, this study is to the best of our knowledge the first to use a net benefit approach.

The findings of smoking and secondary degeneration being negatively associated with the net benefit were expected due to the hypotheses set forward a priori to the study. Although the hypothesized relationship between net benefit and clinical predictors was not defined by sign, we implicitly expected the relationship to point in the same direction as seen in the clinical studies, for example, that smoking is a negative predictor of outcomes [2, 3, 26]. Therefore, at first sight, it was unexpected to identify severe disability and emotional distress as being associated with an increased net benefit. There are high-quality clinical studies investigating the influence of psychological factors on outcome. Recently, Andersen et al. [2] reported a large-scale ($n = 566$) prospective cohort study investigating predictors of functional disability, using the same instrument as that of present study, and concluded

that pre-operative severe disability and emotional distress are potent predictors of a poor outcome. Also recently, Trief et al. [26] reported a prospective cohort study among 160 patients concluding that pre-operative emotional distress predicts a poor outcome. These findings are supported by findings of the Swedish Lumbar Spine Study reporting that the presence of neurosis predicts a poor outcome [14].

Altogether, the literature seems consistent in reporting that negative preoperative psychological factors are potent predictors of a poor outcome whereas, in contrast, the present study found a positive association between such psychological factors and net benefit. The explanation for this differential is intuitive considering the nature of the net benefit measure: the potential for improvement is greatest for patients with the poorest

Table 2 Service utilization and its costs in 695 patients followed 2 years after lumbar spinal fusion

	Service utilization	Costs
<i>Primary health care sector</i>		
Contacts to		
GPs	20 (17;23)	234 (200;276)
Specialist doctors	2 (1;2)	48 (36;62)
Physiotherapists	11 (8;14)	86 (59;118)
Psychologists	0 (0;0)	4 (1;8)
Chiropractors	0 (0;0)	2 (1;3)
Medication	64 (59;70)	1,037 (924;1,167)
Transportation	3 (3;3)	502 (488;517)
Total primary health care		1,895 (1,760;2,047)
<i>Secondary health care sector (hospitals and outpatient clinics)</i>		
Related to back pain		
Index surgery	1	9,240 (9,113;9,363)
Revision surgery	0.1 (0.1;0.1)	797 (609;1,004)
Other admissions	0.5 (0.4;0.6)	2,246 (1,925;2,587)
Outpatient clinic	5.3 (5.0;5.6)	1,183 (1,101;1,276)
Related to comorbidity		
Surgery	0.4 (0.4;0.5)	1,972 (1,669;2,308)
Other admissions	0.4 (0.3;0.5)	2,226 (1,639;2,968)
Outpatient clinic	5.3 (5.0;5.6)	648 (555;760)
Total secondary health care		18,312 (17,408;19,291)
<i>Patients' costs</i>		
Over the counter medication ^a	N/A	58
Complementary medicine ^a	11	325
Transportation	42 (40;44)	4,004 (3,824;4,212)
Family time/paid help ^b	484	15,991
Total patients' costs		16,416 (16,405;16,428)
<i>Productivity costs</i>		
Friction cost method	20 (17;24)	2,314 (1,953;2,709)
Human capital method	569 (535;602)	55,341 (51,967;58,872)
Total costs of spinal fusion ^c		38,937 (37,942;40,015)

Values are number of services (95% CI) and mean discounted 2004-EUR (95% CI) [Bootstrapped, bias-corrected confidence intervals (10,000 replications)]

^a Estimated from the literature, Riviero-Arias et al. [22]

^b Estimated from the literature, Fritzell et al. [12]

^c With inclusion of productivity costs estimated by the Friction cost valuation method

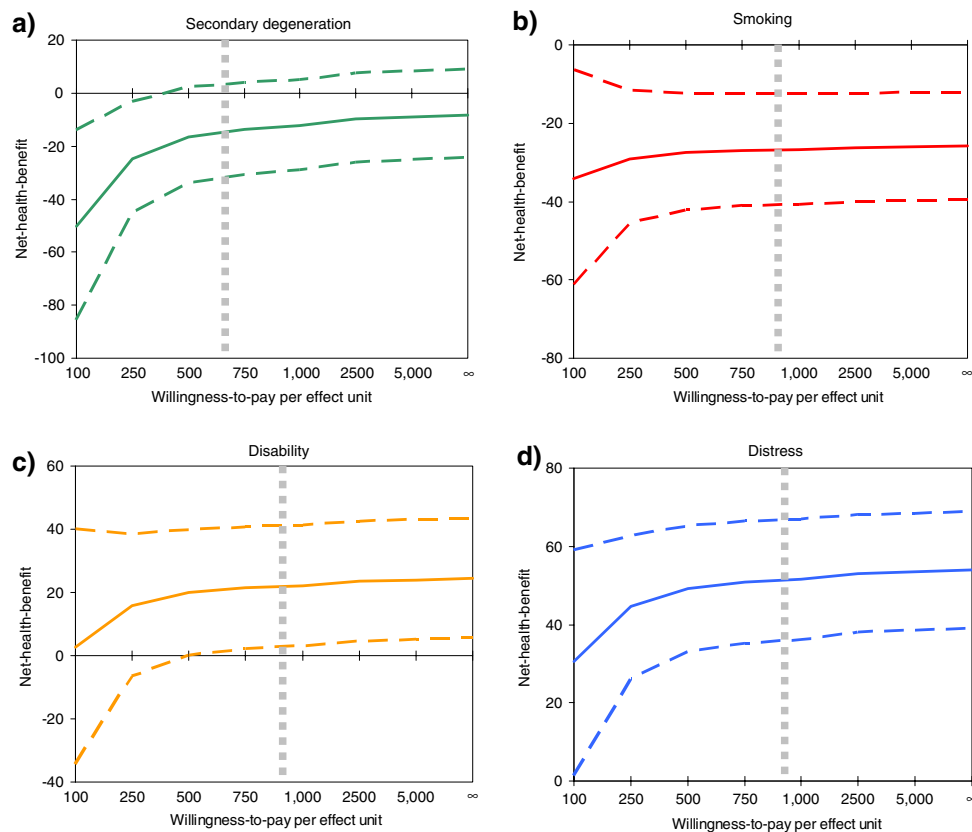


Fig. 2 Significant determinants of the net-health-benefit (NHB) of lumbar spinal fusion (adjusted for age, gender, diagnosis, duration of pain, smoking habits, severity of disability, emotional distress, surgical technique, and number of levels fused where relevant). Curves are average NHB as a function of willingness-to-pay with 95% bias-corrected, bootstrapped CI (*dotted curves*). The *vertical dotted grey line* is the willingness-to-pay accepted today. **a** The influence of secondary degeneration is associated with significantly lower NHB compared with primary degeneration, however, this influence becomes insignificant with increasing values of willingness-

to-pay. **b** The influence of smoking is relatively constant over willingness-to-pay; smoking significantly lowers the NHB (in comparison with no smoking). **c** The influence of pre-operative disability is slightly increasing for increasing values of willingness-to-pay; severe disability is associated with higher NHB (in comparison with moderate or intermediate disability). **d** The influence of pre-operative emotional distress is slightly increasing with increasing values of willingness-to-pay; emotional distress is associated with higher NHB (in comparison with no emotional distress)

baseline values. In contrast, studies predicting clinical outcomes use baseline health states to predict end-point outcomes; all things being equal, the patients expressing the poorest baseline health status have the poorest odds for a high end-point outcome. A secondary analysis of the present material using clinical end-points as the response variable confirmed findings of the literature as described above.

In relation to the internal validity of present study, effects were prospectively measured using a validated instrument (DPQ) while validity of costs relies upon validity of health-service-registers recording service utilization. Their validity has been investigated and found to be solid for the present purpose, that is, an agreement of 75–90% with respect to diagnoses and, most likely, considerably higher validity with respect to service utilization and

its costs because payment to health care providers can not be effected until reporting to registries is accomplished [13, 17, 19].

The net benefit approach is a simple transformation of the cost-effectiveness ratio in order to avoid the problems associated with a ratio founded on both a monetary and an effectiveness scale [7, 24]. Accordingly, the properties of the two measures are basically the same and, in particular, investigating average net benefit (or average cost-effectiveness) holds academic interest alone. There is absolutely no theoretical justification for average net benefit measures in relation to policy decisions.

We chose to impute endpoint disability due to a 25% non-response in one parameter. Whether imputation is justified or not, it was found not to alter findings and the reader may argue it is therefore redundant. One can,

however, also argue that although imputation holds no intrinsic interest, the fact that it appeared not to alter findings is a necessary condition for this study being valid at all. A sufficient condition would then be that the method used to impute is also valid; we argue for a non-linear method because recent long-term trials agree on the greatest improvement occurring between baseline and 1-year follow up, whereas hereafter the curve kinks [6, 11, 22, 27]. Our model was based on a regression fitted in respondents which was, in turn, applied in non-respondents to predict their endpoint disability from their baseline and 1-year disability levels. This implies one assumption, namely that non-respondents follow the same course over time as do respondents but, in fact, it otherwise allows non-respondents to systematically differ from respondents on all other parameters of analysis. We argue that this is sounder than simply ignoring non-respondents in base-case analysis.

In terms of external validity, it remains uncertain how close the material is to a consecutive series, although in principle, all elective patients in our center are included in the clinical database from which inclusions were drawn. We recommend careful comparison of settings prior to generalization of results to Western, specialized centers of spine surgery. However, not only should clinical settings be considered as, among others, informal care may be valued differently in different countries and hence item costs are not necessarily externally valid.

The main objective of this study has been to test the net benefit framework, thereby stimulating the discussion about cost-effectiveness and the methodology for investi-

gating such factors in spine surgery. It may raise ethical questions about whether patients should be selected on grounds of their expected cost-effectiveness ratio and that is definitely not intended as it conflicts with ethics on equity. Rather, should the findings serve as markers when designing future stochastic evaluations or, perhaps, even rejecting stochastic evaluations in favor of modeling studies (allowing for heterogeneity among other things) because this work has demonstrated that certain determinants have more influence to cost-effectiveness than the intervention itself. The approach of modeling alternative strategies for the investigation of cost-effectiveness is a recent trend in UK where the National Institute for Health and Clinical Excellence in fact recommends modeling studies over stochastic trials [18].

Conclusion

In conclusion, predicting net benefit reverses the picture usually seen in studies predicting clinical outcomes, because the response variable is based on improvement over time rather than end-point measures alone. Smoking habits, diagnosis, pre-operative disability, and pre-operative emotional distress were found to be significantly associated with the net benefit of spinal fusion.

Appendix

Absolute regression equations on NHB determinants are given in Table 3.

Table 3 Absolute regression equations on determinants of net-health-benefit (NHB)

Determinant	NHB ₁₀₀	NHB ₂₅₀	NHB ₅₀₀	NHB ₇₅₀	NHB _{1,000}	NHB _{2,500}	NHB _{5,000}	NHB _∞
Age	-1 (-3;1)	-1 (-2;1)	-1 (-2;1)	0 (-2;1)	0 (-2;1)	0 (-1;1)	0 (-1;1)	0 (-1;1)
Female	-26 (-55;2)	-15 (-32;2)	-12 (-26;3)	-11 (-25;4)	-10 (-24;4)	-9 (-23;5)	-9 (-22;5)	-8 (-21;6)
Diagnosis								
Primary degeneration (vs. spondylolisthesis)	-37 (-4; -71)	-20 (-42;4)	-14 (-34;7)	-14 (-31;4)	-12 (-29;5)	-10 (-26;7)	-9 (-25;8)	-7 (-26;13)
Secondary degeneration (vs. primary deg.)	-50 (-86; -14)	-25 (-45; -3)	-17 (-34;2)	-12 (-31;9)	-11 (-30;9)	-9 (-28;11)	-9 (-28;11)	-11 (-27;6)
Duration of pain >2 years.	-23 (-52;6)	-16 (-35;3)	-14 (-30;2)	-13 (-29;2)	-13 (-28;3)	-12 (-27;3)	-12 (-27;3)	-12 (-27;3)
Smoking	-34 (-61; -6)	-29 (-46; -12)	-28 (-42; -12)	-27 (-41; -12)	-27 (-41; -12)	-26 (-40; -12)	-26 (-40; -12)	-26 (-39; -12)
Occupational status								
On sick leave (vs. working)	-123 (-320;60)	-31 (-136;79)	0 (-87;93)	9 (-74;99)	14 (-68;102)	24 (-56;108)	27 (-53;111)	35 (-44;119)
Retired due to sickness (vs. working)	71 (-126;273)	-26 (-146;98)	-58 (-161;41)	-69 (-168;33)	-75 (-171;26)	-84 (-178;13)	-87 (-180;9)	-83 (-175;13)
Pension due to age (vs. working)	-103 (-434;241)	-10 (-216;213)	20 (-164;215)	31 (-151;219)	36 (-134;221)	45 (-131;230)	48 (-128;232)	44 (-135;227)
Disability/emotional distress								
Severe disability (vs. moderate-intermediate)	-3 (-40;34)	16 (-7;38)	20 (0;40)	21 (2;40)	22 (3;41)	24 (4;42)	24 (5;43)	24 (5;43)
Emotional distress (vs. severe disability)	31 (-1;59)	45 (26;52)	49 (33;55)	51 (35;56)	52 (35;67)	53 (38;68)	53 (38;68)	54 (39;69)
Surgical technique								
Instrumented posterolateral (vs. non-inst.)	-25 (-67;16)	-16 (-34;3)	-6 (-29;18)	-4 (-27;19)	-3 (-26;20)	-2 (-24;21)	-1 (-23;21)	-1 (-23;21)
Circumferential (vs. inst. posterolateral)	31 (2;60)	11 (-16;38)	10 (-6;26)	9 (-7;24)	8 (-8;23)	6 (-9;21)	6 (-9;21)	6 (-10;21)
Number of vertebral levels fused >1	-5 (-36;27)	-1 (-21;18)	-0 (-17;16)	0 (-16;16)	0 (-15;16)	1 (-15;16)	1 (-15;16)	1 (-14;17)
<i>n</i>	566	566	566	566	566	566	566	566
Probability >F	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
<i>R</i> ²	0.1027	0.1256	0.1449	0.1523	0.1559	0.1622	0.1642	0.1660
Adj. <i>R</i> ²	0.0749	0.0985	0.1184	0.1260	0.1298	0.1362	0.1383	0.1402
Root MSE	163.31	100.92	87.31	84.24	83.03	81.43	81.08	80.81

Values are regression coefficients (95% bias-corrected, bootstrapped CI) over hypothetical values of decision-makers' threshold for willingness-to-pay per effect unit (λ in 2004-EUR). NHB = health-effect - (societal costs/ λ); hence, the influence of costs decreases with increasing values of λ and NHB_∞ is equivalent to the equation of health effects alone. Health effect was reduction of functional disability from pre-operative to 2 years post-operative measured by the Dallas Pain Questionnaire and summarized into a single index value with range [-400;400]. Interaction variables between occupational status and age were included in equations although not listed, thus, the listed variables of occupational status and age are main-effect dummy variables. Significant determinants ($P < 0.05$) are in italics

References

1. Almen lægepraksis i Danmark. Almen praksis i sundhedstjenesten (2005) Månedsskrift for praktisk lægegerning, København
2. Andersen T, Christensen FB, Bunger C (2006) Evaluation of a Dallas Pain Questionnaire classification in relation to outcome in lumbar spinal fusion. *Eur Spine J* 15(11):1671–1685
3. Andersen T, Christensen FB, Laursen M, Hoy K, Hansen ES, Bunger C (2001) Smoking as a predictor of negative outcome in lumbar spinal fusion. *Spine* 26:2623–2628
4. Buttermann GR, Garvey TA, Hunt AF, Transfeldt EE, Bradford DS, Boachie-Adjei O et al (1998) Lumbar fusion results related to diagnosis. *Spine* 23:116–127
5. Christensen FB (2004) Lumbar spinal fusion. Outcome in relation to surgical methods, choice of implant and postoperative rehabilitation. *Acta Orthop Scand Suppl* 75:2–43
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. Claxton K, Posnett J (1996) An economic approach to clinical trial design and research priority-setting. *Health Econ* 5:513–524
8. DeBerard MS, Masters KS, Colledge AL, Schleusener RL, Schlegel JD (2001) Outcomes of posterolateral lumbar fusion in Utah patients receiving workers' compensation: a retrospective cohort study. *Spine* 26:738–746
9. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI (2005) United States trends in lumbar fusion surgery for degenerative conditions. *Spine* 30:1441–1445
10. Efron B, Tibshirani RJ (1993) An introduction to bootstrap. Chapman and Hall, New York
11. 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. *BMJ* 330:1233–1239
12. 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
13. Gaist D, Sorensen HT, Hallas J (1997) The Danish prescription registries. *Dan Med Bull* 44:445–448
14. Hagg O, Fritzell P, Ekselius L, Nordwall A (2003) Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish lumbar spine study. *Eur Spine J* 12:22–33
15. Koopmanschap MA, Rutten FF, van Ineveld BM, van Roijen L (1995) The friction cost method for measuring indirect costs of disease. *J Health Econ* 14:171–189
16. Lawlis GF, Cuenca R, Selby D, McCoy CE (1989) The development of the Dallas Pain Questionnaire. An assessment of the impact of spinal pain on behavior. *Spine* 14:511–516
17. Mosbech J, Jorgensen J, Madsen M, Rostgaard K, Thornberg K, Poulsen TD (1995) The national patient registry. Evaluation of data quality. *Ugeskr Laeger* 157:3741–3745
18. National Institute for Clinical Excellence (2004) Guide to the Methods of Technology Appraisal. NICE, London
19. Olivarius ND, Hollnagel H, Krasnik A, Pedersen PA, Thorsen H (1998) The Danish National Health Service Register. A tool for primary health care research. *Dan Med Bull* 44:449–453
20. Ozguler A, Gueguen A, Leclerc A, Landre MF, Piciotti M, Le GS et al. (2002) Using the dallas pain questionnaire to classify individuals with low back pain in a working population. *Spine* 27:1783–1789
21. Rice DP, Cooper BS (1967) The economic value of human life. *Am J Public Health Nations Health* 57:1954–1966
22. 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* 330:1239–1245
23. Soegaard R, Christensen FB, Christiansen T, Bunger C (2006) Costs and effects in lumbar spinal fusion. A follow-up study in 136 consecutive patients with chronic low back pain. *Eur Spine J* (Epub ahead of print) Jul 27, 2006
24. 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
25. Trief PM, Grant W, Fredrickson B (2000) A prospective study of psychological predictors of lumbar surgery outcome. *Spine* 25:2616–2621
26. Trief PM, Ploutz-Snyder R, Fredrickson BE (2006) Emotional health predicts pain and function after fusion: a prospective multicenter study. *Spine* 31:823–830
27. Videbaek TS, Christensen FB, Soegaard R, Hansen ES, Hoy K, Helmig P et al (2006) Circumferential fusion improves outcome in comparison with instrumented posterolateral fusion: long-term results of a randomized clinical trial. *Spine* 31:2875–2880
28. Wetzel FT, McCracken L, Robbins RA, Lahey DM, Carnegie M, Phillips FM (2001) Temporal stability of the minnesota multiphasic personality inventory (MMPI) in patients undergoing lumbar fusion: a poor predictor of surgical outcome. *Am J Orthop* 30:469–474