

CLINICAL BENEFIT AND COST EFFECTIVENESS OF TOTAL KNEE ARTHROPLASTY IN THE OLDER PATIENT

F. Krummenauer¹, C. Wolf¹, K.-P. Günther², S. Kirschner²

¹Clinical Epidemiology and Health Economy Unit (Head of the unit: Prof. Dr. F. Krummenauer),

²Department of Orthopedic Surgery (Head of the department: Prof. Dr. K.-P. Günther),
University Hospital Carl Gustav Carus, Medical Faculty of the Technical University of Dresden, Germany

Abstract

Purpose: Total knee arthroplasty (TKA) is an effective, but also cost-intensive health care procedure for the elderly. Furthermore, bearing demographic changes in Western Europe in mind, TKA-associated financial investment for health care insurers will increase notably and thereby catalyze discussions on resource allocation to Orthopedic surgery. To derive a quantitative rationale for such discussions within Western Europe's health care systems, a prospective assessment of both the benefit of TKA from a patient's perspective as well as its cost effectiveness from a health care insurer's perspective was implemented.

Methods: A prospective cost effectiveness trial recruited a total of 65 patients (60% females), who underwent TKA in 2006; median age of patients was 66 years (interquartile range 61 – 74 years). Before and three months after surgery patients were interviewed by means of the EuroQol-5D and the WOMAC questionnaires to assess their individual benefit due to TKA and the subsequent inpatient rehabilitation. Both questionnaires' benefit estimates were transformed into the number of gained quality adjusted life years [QALYs]. Total direct cost estimates for the overall care were based on German DRG and rehabilitation cost rates [€]. The primary clinical endpoint of the investigation was the individual number of QALYs gained by TKA based on the WOMAC interview; the primary health economic endpoint was the marginal cost effectiveness ratio (MCER) relating the costs to the associated gain in quality of life [€/QALY].

Results: Total direct costs for the overall procedure were estimated 9549 € in median. The WOMAC based interview revealed an overall gain of 4.59 QALYs (interquartile range 2.39 – 6.21 QALYs), resulting in marginal costs of 1795 €/QALY (1488 – 3288 €/QALY). The corresponding EuroQol based estimates were 2.93 QALYs (1.75 – 5.59 QALYs) and 3063 €/QALY (1613 – 5291 €/QALY). Logistic regression modelling identified the patients' age as the primary determinant of cost effectiveness (Likelihood Ratio $p = 0.006$): patients younger than 60 years showed a median gain of 6.45 QALYs and median marginal costs of 1463 €/QALY, patients between 60 – 70 years 5.47 QALYs and 1744 €/QALY, patients older than 70 years 2.76 QALYs and 3186 €/QALY.

Conclusion: TKA was proven to be cost effective from a health care insurer's perspective, although its marginal costs notably increased with increasing age. Note, however, that this age-related gradient in marginal cost effectiveness is of comparable order as the changes in cost effectiveness due to variation of the underlying assessment instrument.

Key words: Total Knee Arthroplasty (TKA), benefit outcome, quality of life, cost effectiveness

INTRODUCTION

Recent demographic changes in Western Europe imply an increasing demand for health care procedures in the elderly and thereby increasing resource allocation by health care insurers. However, bearing limitations of the insurers' total budgets in mind, resource allocation decisions cannot be exclusively based on clinical rationales, but must also involve economic aspects [19]. As a compromising consequence, both the clinical effectiveness and the financial investment of a medical treatment must be considered simultaneously to derive a funding decision. The marginal and incremental cost effectiveness ratio concept [14] has been proven to provide quantitative allocation rationales in this setting, which allow for both easy interpretation and direct comparison with treatment alternatives: The cost effectiveness ratio relates the costs of a treatment to its benefit from a patient's perspective, mostly estimated in terms of monetary units per gained quality adjusted life year (QALY). Estimation of the treatment's effectiveness in terms of QALYs allows for a patient-related benefit interpretation as well as for comparison of its cost effectiveness estimate with the corresponding health economic characteristic of alternative treatments. In particular, the estimation of a treatment's cost effectiveness enables health care insurers to evaluate its patient-related benefit in comparison to other treatments, which already underwent this decision process for refunding. Note, in addition, that such health economic treatment characteristics cannot only be derived for a treatment procedure in general, but also allow for patient-specific funding recommendations: estimation of the treatment's cost effectiveness among patient sub samples (for example by strati-

fication of the cost effectiveness estimate for age or working status): These sub sample estimates will characterize the treatment's cost / benefit relation in more detail and thereby may identify patient groups with a more promising (and therefore fundable) cost / benefit relation, as it would have been assumed for the overall treatment. Therefore the concept of marginal and incremental cost effectiveness evaluation finds increasing acceptance and application among both clinicians and health economists. By providing a quantitative and therefore transparent rationale in resource allocation discussions, its particular value in the need for health economic evaluation of treatment concepts for older patients is obvious.

Orthopedic surgery is pretty concerned with the treatment of the older patient, since the demand for total hip and knee arthroplasty is catalyzed by both demographic changes and availability of effective and safe treatment procedures. In particular, total knee arthroplasty (TKA) presents a treatment offer to the older patient, who often suffers from a severe loss in quality of life due to gonarthrosis, but also from prognostically relevant comorbidity. Therefore a direct cost assessment of TKA based on diagnose related groups (DRGs) will result in a rather cost-intensive reimbursement for health care insurers. Accordingly, the German DRG rate for unilateral TKA amounts to 7500 € plus the direct costs for postoperative rehabilitation, which at least amount to additional 2000 €. On the other hand, TKA has been proven to be effective from the patients' perspective concerning regain of both function and quality of life. Whether the amount of this patient benefit is associated with the individual patients' sociodemographic characteristics, however, is a matter of sensitive discussion [4, 16].

Bearing recent demographic changes in Western Europe and the increasing limitations of financial resources in Western health care systems in mind, this investigation was implemented to investigate the order and determinants of the cost effectiveness of TKA in the elderly.

PATIENTS AND METHODS

The primary intention of this investigation was to derive an estimate for the cost effectiveness of TKA under stratification of putative (multivariate) outcome determinants such as age, gender or occupational state. Therefore a prospective patient-based investigation with the aim of estimating the cost effectiveness among the patients under observation was implemented instead of performing a Markov decision analysis based on meta analysis estimates.

STUDY DESIGN AND PATIENT CHARACTERISTICS

This prospective cost effectiveness investigation comprised individual data of patients, who underwent unilateral TKA at the authors' University Hospital in 2006. A total of 65 patients with the clinical indication for TKA were enrolled consecutively and asked for participation in this prospective health economic trial. After written informed consent patients underwent a written interview by means of the WOMAC and the

EuroQ(u)ol-5D questionnaire one week before surgery. The interviews were coordinated by a study nurse, who offered assistance to the interview if required. Furthermore the study subjects were invited for a three months recall to undergo the same interview. This study design was positively rated by the local Independent Ethics Committee by June 25th 2005.

The 65 patients (60% females) spent a median duration of 9 days in hospital. They showed a median age of 66 years (interquartile range 61 – 74 years, range 45 – 78 years); 31% of patients were older than 70 years, 23% younger than 60 years. Female patients (39% older than 70 years at the day of surgery) showed a median age of 67 years versus 65 years among male patients (19% older than 70 years). Among the 65 study patients, 25% reported to live without a partner or family members; 12% reported an academic degree. Only 8% reported their occupational status as "working" or reported to "restart working" after rehabilitation.

OPERATIVE PROCEDURE

The study population consisted of patients with primary or secondary osteoarthritis of the knee joint stage 3 or 4 according to Kellgren. The patients were routinely operated under tourniquet after admission of perioperative antibiotic prophylaxis in a laminar airflow room. Through a straight skin incision and a medial parapatellar approach a cemented Natural Knee II (Fa. Zimmer, Germany) total knee arthroplasty without resurfacing of the patella was implanted. Postoperatively all patients received a risk adjusted anticoagulation treatment, for example, high risk patients received arixtra 1x1 s.c. injection over six weeks. Postoperatively all patients were allowed for full weightbearing. After discharge from the hospital the patients were transferred to a inpatient rehabilitation unit for additional three weeks.

CLINICAL BENEFIT EVALUATION AND MODEL ASSUMPTIONS

The primary clinical endpoint of this investigation was the individual gain in quality of life as assessed by means of the WOMAC questionnaire: Before and three months after surgery patients answered the 24 WOMAC items, which were documented in terms of a five-staged ordinal scale. The 24 items were averaged and transformed into a utility scale of 0 – 100%, where the scale maximum 100% indicates the optimum well-being among each of the 24 items. The intraindividual difference post – pre of this transformed WOMAC index was then considered as a surrogate for the patients' clinical benefit achieved by TKA. This surrogate was then extrapolated alongside the patients' theoretical rest life expectancy, where a Gaussian life expectancy distribution was assumed with mean expectancy of 85 years for women and 80 years for men (both underlying assumed standard deviations of 10 years). Based on this model assumption the individual (rest) life expectancy of a patient was simulated. The WOMAC based clinical benefit was then assumed to persist over this patient's simulated rest life period and proportionally extrapolated over time. To account for time-depen-

dent loss in the primary clinical benefit as assessed shortly after surgery, the overall benefit estimate was discounted at an annual discounting rate of 3%. If, for example, a patient had a simulated rest life expectancy of 10 years after surgery and reported a 25% increase in terms of the WOMAC based utility scale, his individual crude benefit was estimated as $10 \text{ years} * 25\% = 2.5$ quality adjusted life years (QALYs). After discounting at the annual 3% loss rate a net TKA benefit of 2.18 QALYs was then estimated for this patient.

However, since a 10 years revision rate of at least 5 – 10% must be assumed after TKA [10], the above simulation of rest life expectancy also introduced a 10 years censoring for 5% of the patients by random selection. This model assumption allowed to also simulate the fact of loss in quality of life when a revision becomes necessary (and thereby imposes new financial investment).

PRIMARY ENDPOINTS AND TARGET PARAMETERS

The primary clinical endpoint of this investigation was the individual gain in quality of life as illustrated above [QALYs], when assessed by means of the WOMAC questionnaire. The most important secondary endpoints were the underlying WOMAC-based benefit estimate [%], the corresponding EuroQuol-based benefit estimate [%] and its quality of life derivate [QALYs], which was estimated in the same manner as the WOMAC-based analogue.

The primary economical endpoint were the total direct costs [€], based on German DRG rates for the initial TKA and the subsequent indoor rehabilitation. The latter was deterministically modelled by assumption of a “flat rate” of 2100 € (referring to average daily costs of 100 € during a three weeks inpatient rehabilitation).

The main target parameter of the cost effectiveness evaluation was the marginal cost effectiveness ratio (MCER), which relates the direct costs as assessed by the primary economical endpoint to the associated patient benefit as assessed by the primary clinical endpoint [€/QALY]. Both cost and cost effectiveness estimation were performed from the German compulsory health care insurer’s perspective.

To briefly illustrate the estimation and interpretation of the MCER endpoint [14], the following numerical example will be sketched out: If, for example, a patient’s net benefit was estimated 2.18 QALYs and the overall direct DRG costs for TKA and subsequent rehabilitation were calculated 9549 €, this patient’s marginal costs were estimated as $\text{MCER} = 9549 \text{ €} / 2.18 \text{ QALYs} = 4380 \text{ €} / \text{QALY}$. Note, that the above study design allows for estimation of this MCER parameter for each of the study subjects; the MCER’s distribution therefore illustrates heterogeneity in the cost effectiveness of TKA among sub samples of patients. If a second patient reports a clinical benefit of 4.50 QALY and underwent the same clinical procedure at an identical cost level, his individual MCER is estimated $2122 \text{ €} / \text{QALY}$ and therefore obviously indicates a better individual cost / benefit relation as obtained for the previous patient. For the sake of illustration of these MCER estimates the marginal costs of monofocal cataract surgery may be mentioned, which were estimated $1916 \text{ €} / \text{QALY}$ [14]. At least the second of

the above model patients would therefore illustrate a comparable cost effectiveness of cataract and TKA surgery (and thereby motivate corresponding funding recommendation).

The following will introduce an exploratory cut-point of 2000 € / QALY such as motivated by the marginal cost effectiveness order of monofocal cataract surgery [14], which is undoubted to be fully refundable by health care insure insurers.

STATISTICAL ANALYSIS

The distributions of continuous endpoints such as the primary clinical endpoint and the MCER were described by medians and quartiles (graphically on non-parametric box plots, accordingly) to take account for possible statistical outliers. The description of categorical endpoints was based on absolute and appropriate relative frequencies. These methods were applied to the overall sample as well as to sub samples. Intraindividual comparisons were based on the description of difference distributions for continuous endpoints and on total frequencies in contingency tables for categorical endpoints.

The significance evaluation of intraindividual changes in continuous endpoints was based on pairwise sign tests, in categorical endpoints on pairwise McNemar tests. Sub sample comparisons were based on pairwise Wilcoxon two sample tests and Kruskal/Wallis multi sample tests for continuous and on pairwise Fisher tests for categorical endpoints. Results of these tests were summarized in terms of p-values. Due to the rather exploratory character of the sub sample comparisons, these p-values were not formally adjusted for multiplicity. A p-value < 0.05 therefore indicates locally significant sub sample differences.

A multivariate re-analysis of the cost effectiveness data was based on categorization of the primary endpoint: TKA treatment was considered to be individually cost effective, if the underlying MCER was estimated $\leq 2000 \text{ €} / \text{QALY}$ (see section 2.4). This exploratory endpoint allowed for multiple logistic regression modelling of simultaneous cost effectiveness determinants like age, gender, working and family state. Logistic regression modelling was performed by means of Likelihood Ratio tests (forward selection at a local 5% significance level).

All numerical and graphical evaluations were based on the software SPSS® (release 12.0 for Windows®).

SENSITIVITY ANALYSIS

The results of a cost effectiveness evaluation crucially depend on the validity and precision of the underlying input data: Note that, for example, the model assumptions on rest life expectancy of patients (85 years for female and 80 years for male patients) or on survival of prosthesis (loss after ten years in 5% of the patients) can imply bias into the overall MCER estimation. Therefore a deterministic variation of such input parameters and model assumptions was introduced to simulate their possible imprecision and the resulting impact on the MCER estimate: In the above setting the model parameter on patient life expectancy, which was

applied to the QALY transformation of the WOMAC and EuroQuol results, was varied deterministically over ± 5 years according to recent age-stratified age tables provided by German authorities: The latter indicated a total life expectancy of 75.59 years for male and 81.34 years for female newborns in Germany. Furthermore they demonstrate a rest life expectancy of 19.51 years (males) versus 23.72 years (females) for people aged between 60 – 65 years and of 15.79 years (males) versus 19.44 years (females) for people aged between 65 – 70 years. The resulting imprecision in the rest life expectancy estimates was therefore summarized by the above deterministic assumptions.

The input parameter on prosthesis survival was varied [10] over $\pm 3\%$ (thereby ranging from 2 – 8%) and the “flat rate” assumed for the direct costs of postoperative rehabilitation over ± 200 € per patient (therefore ranging between 1900 – 2300 €).

A second direction of sensitivity analysis considered the measurement instrument, which was applied to assess the primary clinical endpoint: whereas the WOMAC questionnaire rather estimates a semi-functional outcome, which is highly associated with health-related quality of life, the EuroQuol questionnaire rather estimates general health-related quality of life. Both estimates can be considered as patient-related surrogate endpoints and therefore characterize the clinical benefit of TKA from the patient’s perspective. To quantify the assessment instrument’s impact on the

MCER estimate, the EuroQuol based MCER’s distribution was considered in terms of a sensitivity analysis for the WOMAC based results.

RESULTS

Both the clinical and the economical primary endpoint could be evaluated for each of the 65 patients under investigation. One week before surgery, the overall sample reported a median WOMAC utility score of 41% (34 – 47%), which hardly varied among age decades: whereas patients younger than 60 years at the day of surgery reported a median score of 38%, patients between 60 – 70 years and patients older than 70 years reported a median score 41%, respectively. Three months after surgery the overall sample reported a WOMAC score of 81% (64 – 88%), which varied among the age decades from 72% over 82 and 81%, respectively. The corresponding median EuroQuol score rose from 40% (32 – 54%) before surgery to 70% (65 – 78%) after surgery with a similar age-related gradient in favour of the older patients as observed for the WOMAC based score.

Table 1 summarizes the intraindividual changes in the underlying 24 WOMAC items and illustrates substantial and statistically significant improvement for each aspect under consideration (McNemar $p < 0.001$ for each of the items): the fraction of patients reporting subjective improvement three months after surgery

Table 1. Relative frequencies for self-reported improvement or worsening in the 24 WOMAC questionnaire items three months after versus one week before TKA.

	improvement	worsening
pain		
when walking on flat ground	78 %	3 %
when walking up / down stairs	82 %	3 %
which discontinuous night sleep	57 %	8 %
in sitting or lying	68 %	10 %
in straight standing	87 %	1 %
stiffness		
after first morning wake up	64 %	13 %
after over-day rest	78 %	6 %
difficulties		
when stepping down stairs	76 %	5 %
when stepping up stairs	78 %	6 %
in standing up when sitting	79 %	3 %
in straight standing	89 %	3 %
when picking things from floor	76 %	8 %
when walking on flat ground	78 %	2 %
when entering a car or bus	76 %	8 %
when going shopping	76 %	7 %
when putting on socks.etc.	74 %	6 %
when putting off socks etc.	84 %	5 %
in getting out of bed	76 %	8 %
when lying in bed	68 %	3 %
when entering or leaving a bath tub	62 %	0 %
in sitting for a longer time	68 %	7 %
when sitting on the toilet	76 %	8 %
in doing less easy homework	72 %	3 %
in doing easy homework	76 %	5 %

exceeds 70% for most items, where maximum gain was reported for items associated with pain symptoms.

CLINICAL BENEFIT EVALUATION

The median intraindividual increase in the WOMAC score was 38% (25 – 45%, sign test $p < 0.001$), the corresponding EuroQuol score increase was estimated 30% (15 – 40%, $p = 0.008$) in median (Table 2). The transformation of these score increases into quality adjusted life years revealed a median benefit of 4.59 QALYs (2.39 – 6.21 QALYs) for the WOMAC score and of 2.93 QALYs (1.75 – 5.59 QALYs) based on the EuroQuol questionnaire.

According to the model-associated impact of life expectancy into QALY calculation, but in contradiction to the above age-related gradient in the postoperative score estimates in favour of the older patients, the age decades notably varied in both QALY benefit

estimates: Whereas the intraindividual WOMAC score increase did neither statistically nor clinically significantly differ between the three decades under consideration (median increases 34%, 41% and 38%, respectively, pairwise Wilcoxon tests' p-values between 0.095 and 0.260), a statistically significant age difference was observed for the QALY estimates: patients older than 70 years reported a median clinical benefit of 2.76 QALYs, whereas patients between 60 – 70 years and patients younger than 60 years reported median benefits of 5.47 and 6.48 QALYs, respectively (pairwise Wilcoxon test p-values versus patients older than 70 years < 0.001 for both age groups, respectively). The same tendency holds for the EuroQuol based benefit estimates with medians of 1.75 QALYs versus 4.25 and 4.78 QALYs (Wilcoxon $p = 0.004$ and $p < 0.001$, respectively).

Table 3 summarizes the results of clinical benefit estimation under stratification for age, gender and ac-

Table 2. Medians and quartiles for the distribution of the pre- and postoperative WOMAC index and EuroQuol-5D utility index [%] in patients undergoing total knee arthroplasty, their respective intraindividual change [%] with p-values derived from sign tests, the resulting gain in patient quality of life [QALYs] and the marginal cost effectiveness ratio [€/QALY] from the health care insurers' perspective.

	Median	Q ₁ – Q ₃	p(sign test)
WOMAC pre	41 %	34 – 47 %	
WOMAC post	81 %	64 – 88 %	
WOMAC change	38 %	25 – 45 %	$p < 0.001$
EQ-5D pre	40 %	32 – 54 %	
EQ-5D post	70 %	65 – 78 %	
EQ-5D change	30 %	15 – 40 %	$p = 0.008$
WOMAC based gain	4.59 QALYs	2.39 – 6.21 QALYs	
EQ-5D based gain	2.93 QALYs	1.75 – 5.59 QALYs	
WOMAC based marginal costs	1795 €/ QALY	1588 – 3288 €/ QALY	
EQ-5D based marginal costs	3064 €/ QALY	1613 – 5291 €/ QALY	

Table 3. Medians and quartiles for the distribution of the gain in patient quality of life [QALYs] after total knee arthroplasty and the corresponding marginal cost effectiveness ratio [€/QALY] from the health insurers' perspective, stratified for gender, age and occupational status of patients at the day of surgery.

	WOMAC based gain [QALYs]	WOMAC based marginal costs [€/QALY]	EQ-5D based gain [QALYs]	EQ-5D based marginal costs [€/QALY]
total sample	4.59	1795	2.93	3064
gender				
female	3.79	2142	2.93	3260
male	5.61	1660	3.08	2847
age				
< 60 years	6.48	1463	4.78	1894
60 – 70 years	5.47	1744	4.25	2093
> 70 years	2.76	3186	1.75	5455
working status				
working	6.48	1474	2.45	3908
not working	4.46	1829	2.93	2961

tual working state. According to Table 3, a statistically significant difference in TKA benefit was observed between patients under recent employment and those out of employment (work-seeking, drawing a pension etc.) with a median difference of 6.48 versus 4.46 QALYs (Wilcoxon $p = 0.004$). Note that the median difference between these sub samples is of comparable order as already observed for the age sub samples, when benefit estimation is based on the WOMAC interview. Regarding the fact, that none of the patients older than 70 years was still under employment, the working state gradient must in fact be interpreted as being confounded by the age-related gradient already mentioned above. If, however, QALY calculation is based on the EuroQuol scores, neither a clinically relevant nor a statistically significant association with the patients' working state is observed (median benefit estimates 2.45 versus 2.89 QALYs, Wilcoxon $p = 0.522$). Note, however, the smaller median benefit estimate for patients under employment, which is in contradiction to the WOMAC based results.

The same tendencies hold for a gender specific evaluation: Table 3 illustrates a higher TKA benefit for male patients (5.61 QALYs in median) than for females (3.79 QALYs in median, Wilcoxon $p = 0.002$). Note, however, that the gender samples also differ in age distribution (see section 2.1) with fractions of 39% (females) versus 19% (males) of patients older than 70 years at the day of surgery. Therefore the gender-associated difference in benefit estimation must also be considered as a confounding effect of the dominating age-related gradient. Again only a slight sample difference was observed in the EuroQuol based benefit medians (2.93 QALYs for females versus 3.08 QALYs for males).

COST EFFECTIVENESS EVALUATION

The overall sample demonstrated a median cost effectiveness ratio of 1795 € / QALY (1588 – 3288 € / QALY) when MCER estimation used the WOMAC utility scores; based on the EuroQuol scores median marginal costs of 3064 € / QALY (1613 – 5291 € / QALY) were found. A fraction of 53% patients showed individual marginal costs < 2000 € / QALYs based on the WOMAC interview versus 33% based on the EuroQuol questionnaire.

According to the observations in QALY estimates, a statistically significant and economically relevant association between the MCER and the patients' age was observed (Fig.1): based on the WOMAC data, the marginal costs for patients younger than 60 years were estimated 1463 € / QALY, for patients between 60 – 70 years and patients > 70 years 1744 and 3186 € / QALY, respectively (Kruskal/Wallis $p = 0.001$). A similar tendency was observed in the EuroQuol based MCERs (Fig. 2) with respective medians of 1894, 2093 and 5455 € / QALY (Kruskal/Wallis $p < 0.001$). Table 3 summarizes the MCER estimates among gender, age and working state sub samples: It illustrates a slightly better cost / benefit relation for male and for employed patients than for females and for patients out of employment, respectively.

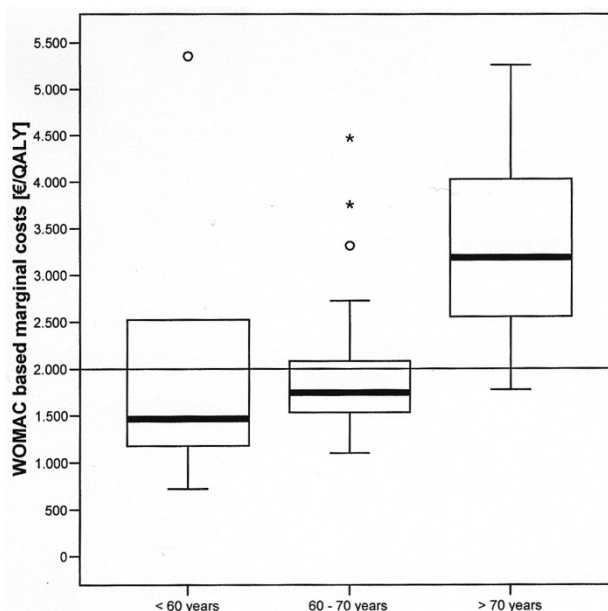


Fig. 1. Box whisker plots for the distribution of the marginal costs [€/QALY] for total knee arthroplasty, based on quality of life gain assessment by means of the WOMAC arthrosis index and stratified for the patients' age decade at the day of surgery

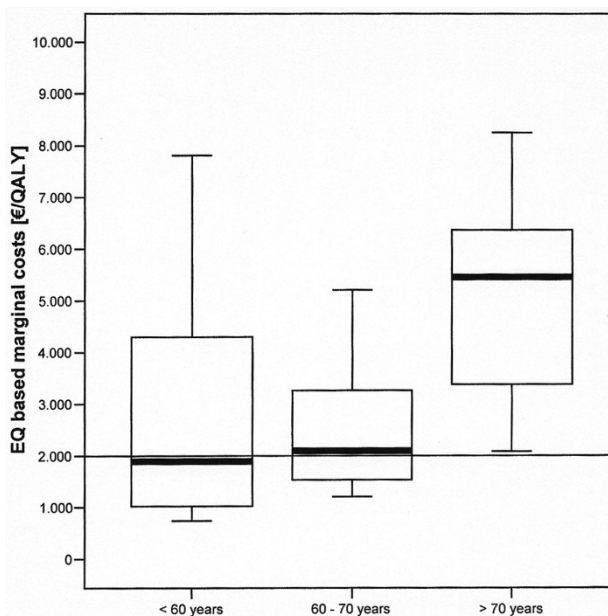


Fig. 2. Box whisker plots for the distribution of the marginal costs [€/QALY] for total knee arthroplasty, based on quality of life gain assessment by means of the EuroQuol-5D health utility index and stratified for the patients' age decade at the day of surgery

MULTIVARIATE ANALYSIS

Logistic regression modelling of the exploratory endpoint "individual MCER < 2000 € / QALY" confirmed the dominating age impact on the cost effectiveness outcome: Whereas neither gender nor working nor family state were found to be significantly as-

sociated with the WOMAC based endpoint (Likelihood Ratio $p = 0.201$, $p = 0.475$ and $p = 0.364$, respectively), the statistical significance of the age-related gradient was strictly reproduced (Likelihood Ratio $p = 0.006$). As already observed by univariate analysis, patients between 60 – 70 years and patients younger than 60 years significantly contrasted from patients older than 70 years (Likelihood Ratio $p = 0.012$ and $p = 0.002$, respectively).

SENSITIVITY ANALYSIS

Sections 3.1 and 3.2 already indicated the assessment instruments' substantial impact on the MCER outcome: In general clinical benefit and cost effectiveness estimates turned out relevantly and significantly more conservative when being based on the EuroQuol instead of the WOMAC questionnaire (McNemar $p = 0.017$). In terms of the exploratory endpoint "individual MCER < 2000 € / QALY", only 60% of the overall sample are concordantly rated "cost effective" or "not cost effective" by the respective individual WOMAC and EuroQuol based MCER estimates. Bearing the MCER distribution's medians 1795 and 3064 € / QALY in mind, an interquartile sensitivity range of 1588 – 5291 € / QALY must be assumed due to changes in the assessment instrument. Note that this imprecision is in fact due to the instrument at hand and not an epi-phenomenon of the chosen strategy for QALY calculation (section 2.1): The primary basis of the QALY estimates are the WOMAC and EuroQuol based utility scores (as estimated in terms of the respective intraindividual score changes). The latter differed significantly in the overall sample (sign test $p = 0.006$) with a median difference of 9% (11% for females and 6% for males in median).

In comparison to the numerical impact of the underlying assessment instrument, the deterministic variation of the model parameters mentioned in section 2.6 can be regarded negligible: variation of the expectable prosthesis failure rate ($5\% \pm 3\%$) implied a variation of 1755 – 1812 € / QALY for the initial WOMAC based MCER estimate 1795 € / QALY, the variation of the patients' individual life expectancy (85 ± 5 and 80 ± 5 years) a corresponding range of 1580 – 1905 € / QALY. If the individual life expectancies' distribution was simulated according to the age-stratified rest life expectancy estimates outlined in section 2.5, a rather similar range of 1618 – 1822 € / QALY was observed for the MCER estimate. However, the initially observed statistical significance of the age-related MCER gradient was strictly reproduced (Likelihood Ratio $p = 0.015$), where patients between 60 – 70 years and patients younger than 60 years significantly contrasted from patients older than 70 years.

Finally the variation of the assumed cost rate for postoperative rehabilitation ($2100 \text{ €} \pm 200 \text{ €}$) resulted in a median MCER range of 1712 – 1878 € / QALY. None of these deterministic sensitivity analyses revealed a median MCER larger than 2000 € / QALY, which implies encouraging robustness of the MCER estimation strategy chosen for the primary analysis.

DISCUSSION

This investigation intended to quantify the clinical benefit (in terms of quality of life improvement) and the individual cost / benefit relation of total knee arthroplasty from a health insurer's perspective. Primary intention of the cost effectiveness evaluation was to derive quantitative cost effectiveness characteristics for TKA in Western European patients and thereby to provide rationales for recent discussions on resource allocation to Orthopedic surgery in Western European health care systems. The above data revealed a marginal cost effectiveness ratio of about 1800 € / QALY and therefore demonstrates unilateral TKA to be as cost effective as for example monofocal cataract surgery with an MCER of about 1900 € / QALY. Nevertheless, an economically relevant and statistically significant association of the TKA cost effectiveness outcome with the patients' age at surgery was observed: Whereas TKA in patients < 60 years was found to be thoroughly cost effective, the procedure must be considered rather more expensive for patients older than 70 years regarding a median cost effectiveness ratio of over 3000 € / QALY. It must be emphasized that this age-related gradient in cost effectiveness is strictly a surrogate of the assumption of smaller rest life expectancy for those patients (which means a smaller benefit time after surgery) as compared to younger patients: The crude clinical benefit was found merely comparable among these decades with median WOMAC score increases of about 34 – 41% and even a slightly higher benefit reported by older patients. After correction of this benefit estimate for the expectable life time, however, a statistically significant age gradient was observed in the QALY and MCER estimates.

The authors wish to emphasize that their personal interpretation of this result is the confirmation of funding TKA for patients older than 70 years! Note that in this investigation patients of this decade showed an age range between 70 and 78 years – even under maximum formal interpretation of the model assumptions on rest life expectancy (80 – 85 years, see section 2.1) each patient in this investigation is therefore complimented with a minimum rest life expectancy of at least two years. Bearing the fact in mind that the indication and decision for TKA is rather an "ultima ratio" offer to patients suffering from immobility and continuous pain, to prolong this painful state for over a minimum of two life years cannot be justified from an ethical perspective.

METHODOLOGICAL ASPECTS

The cost effectiveness of TKA has yet been widely discussed and quantified [3, 15, 16] in patient samples recruited in the U.S.A. and the U.K. as well as in Scandinavia. Its overall performance is accepted to be cost effective and may be slightly superior to that of total hip arthroplasty [1, 2, 3, 7, 8, 11, 16]. In particular the observation of inferior cost effectiveness profiles in patients older than 70 years is a well-accepted fact [4, 16]. Recent investigations rather concentrate on possibilities of improvement in the cost effectiveness of TKA by, for example, introducing the advantages of

clinical pathway applications [12] in internal resource optimization, by making use of computer assisted surgery tools [9] or less invasive surgery [17, 18], or by improving patient compliance by providing oral and written information as well as physiotherapeutic exercises before TKA.

However, a lot of this recently published information is – according to the actual research hypothesis at hand or due to the cost reimbursement strategies on site – rather based on the health care provider's perspective instead of the health care insurer's. Furthermore, some authors do not provide appropriate outcome measures in terms of utility scores or quality adjusted life years, which enable the comparison of TKA cost effectiveness characteristics with those of concurring treatment offers (and thereby allow health care insurers to apply the cost effectiveness information in local resource allocation discussions). As a consequence, only few published data allows for direct and quantitative comparison of its cost effectiveness characteristics with the above MCER estimates, which were primarily derived for interpretation in the setting of Western European health care systems.

Rissanen et al. report a standardized utility score increase of 9% in patients younger than 60 years, of 5% in patients aged between 60 – 70 years and 2% for older patients. If median ages of 55, 65 and 75 years are assumed for these sub samples, a crude extrapolation for a median life expectancy of 85 years would result in respective crude benefit estimates of 2.70 QALYs, 1.00 QALYs and 0.2 QALYs respectively. Whereas the age-related gradient observed by Rissanen et al. is strictly reproduced by the German patient data presented in the above, the absolute order of clinical benefit estimates based on the latter crucially differs from the Scandinavian benefit characteristics. Since Rissanen et al. reported overall treatment costs of about 10.000 € per patient with a 15% cost increase for patients older than 70 years, the resulting MCER estimates would dramatically exceed the sensitivity range < 5.500 € / QALY as obtained here.

Two main putative reasons may have induced these deviations: Whereas the WOMAC index used in the above investigation is rather a surrogate of clinical function in daily life [13], and therefore its postoperative increase only presents a proxy estimate of TKA utility, the 15D questionnaire results reported by Rissanen et al. more directly estimate changes in health-related quality of life. To obtain comparability with this approach the EuroQuol-5D [5] questionnaire was used in the recent setting, which in fact resulted in significantly smaller clinical benefit estimates. The latter, however, were still found much larger than the ones reported by Rissanen et al. [16]. A second direction of explanation for the clinically relevant difference in the benefit estimates may draw attention to differences in the underlying study designs: Whereas the above German data was acquired three months after surgery, the postoperative Scandinavian outcome assessment took place 24 months after TKA. It can only be hypothesized, whether a three months recall period may have resulted in somewhat optimistically biased results due to a persisting benefit from inpatient rehabilitation care, but which would soon have turned more conser-

vative at an additional six months recall. Since the process of recovering and continuous improvement after TKA must be assumed to last for at least one year, these several concurring possibly conservative or liberal biases complicate the direct comparison of its three months and two years outcome.

In particular, Rissanen's two years follow-up design directly allows for the observation of short-term complications and the resulting re-treatment costs and outcome: In the above cohort, no revision to any reasons was observed within three months after initial TKA, which would have notably reduced the respective patients' individual cost effectiveness estimate [6]. Two directly postoperative complications due to comorbidity were observed, but no additional costs were observed for their immediate re-treatment (as induced by the German DRG system, which does not consider additional real costs for complication treatment within 30 days after the initial TKA).

Both design deviations from the Scandinavian register based investigation may have imposed a liberal bias on the WOMAC and EuroQuol outcome estimates. Therefore further recalls of the above patient cohort are necessary to reproduce or at least robustify the TKA outcome estimates found here.

CONCLUSION

A three months recall after TKA revealed an overall cost effectiveness ratio of 1795 € / QALY, which showed a statistically significant and economically relevant association with the patients' age at the time of surgery. Patients older than 70 years showed a similar WOMAC based clinical benefit as found for younger patients, but a more pessimistic QALY based cost effectiveness characteristic. The data at hand do not contradict TKA in patients older than 70 years.

Comment: The content of this manuscript was presented during the 2007 annual meeting of the German Society for Orthopedics and Orthopedic Surgery (DGOOC) in Berlin, as well as during the 2007 annual meeting of the German Society for Medical Informatics, Biometry and Epidemiology (GMDS) in Augsburg.

Acknowledgement: The authors are grateful to Ms Karen Faulkner, MD, for a native speaker revision of this manuscript.

Conflict of Interests statement: The authors have no commercial or political interests in the methodological or medical aspects presented in this paper. The investigation was granted by Zimmer® GmbH Deutschland; the study nurse position of Claudia Wolf was sponsored by this grant to allow for personal patient interviews and data management.

REFERENCES

1. Antoniou J, Martineau PA, Filion KB, Haider S, Zukor DJ, Huk OL, Pilote OL, Eisenberg JL. In-hospital cost of total hip arthroplasty in Canada and the United States. *J Bone Joint Surg Am* 2004; 86A: 2435-2439
2. Bozic KJ, Saleh KJ, Rosenberg AG, Rubash HE. Economic evaluation in total hip arthroplasty: analysis and review of the literature. *J Arthroplasty* 2004; 19: 180-189
3. Brauer CA, Rosen AB, Olchanski NV, Neumann PJ. Cost-utility analysis in orthopedic surgery. *J Bone Joint Surg Am* 2005; 87: 1253-1259

4. Brander VA, Malhotra S, Jet J, Heinemann AW, Stulberg SD. Outcome of hip and knee arthroplasty in persons aged 80 years and older. *Clin Orthop Relat Res* 1997; 345: 67-78
5. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37: 53-72
6. Burns RW, Bourne RB, Chesworth BM, MacDonald SJ, Rorabeck CH. Cost effectiveness of revision total knee arthroplasty. *Clin Orthop Rel Res* 2006; 446: 29-33
7. Chang RW, Pellisier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. *JAMA* 1996; 275: 858-865
8. Chapman RH, Stone PW, Sandberg EA, Bell C, Neumann PJ. A comprehensive league table of cost-utility ratios and a sub-table of "panel-worthy" studies. *Med Decis Making* 2000; 20: 451-467
9. Dong H, Buxton M. Early assessment of the likely cost-effectiveness of a new technology: a Markov model with probabilistic sensitivity analysis of computer-assisted total knee replacement. *Int J Technol Assess Health Care* 2006; 22: 191-202
10. Ethgen O, Bryere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty: a qualitative and systematic review of the literature. *J Bone Joint Surg Am* 2004; 86: 963-974
11. Gillespie WJ, Pekarsky B, O'Connell DL. Evaluation of new technologies for total hip replacement – economic modelling and clinical trials. *J Bone Joint Surg Br* 1995; 77: 528-533
12. Healy WJ, Iorio R, Ko J, Appleby D, Lemos DW. Impact of cost reduction programs on short-term patient outcome and hospital cost of total knee arthroplasty. *J Bone Joint Surg Am* 2002; 84: 348-353
13. Kane RL, Saleh KJ, Wilt TJ, Bershadsky B. The functional outcome of total knee arthroplasty. *J Bone Joint Am* 2005; 87: 1719-1724
14. Krummenauer F, Landwehr I. Incremental cost effectiveness evaluation in clinical research. *Eur J Med Res* 10: 18-23
15. Lavernia CJ, Guzman JF, Gachupin-Garcia A. Cost-effectiveness and quality of life in knee arthroplasty. *Clin Orthop Relat Res* 1997; 345: 134-139
16. Rissanen P, Aro S, Sintonen H, Asikainen K, Slati P, Paavolainen P. Costs and cost-effectiveness in hip and knee replacements – a prospective study. *Int J Technol Assess Health Care* 1997; 13: 575-588
17. Slover J, Espehaug B, Havelin LI, Engesaeter LB, Fumes O, Tomek I, Tosteson A. Cost-effectiveness of unicompartmental and total knee arthroplasty in elderly low-demand patients: a Markov decision analysis. *J Bone Joint Surg Am* 2006; 88: 2348-2355
18. Soohoo NF, Sharifi H, Kominski G, Liebermann JR. Cost-effectiveness analysis of unicompartmental knee arthroplasty as an alternative to total knee arthroplasty for unicompartmental osteoarthritis. *J Bone Joint Surg Am* 2006; 1975-1982
19. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA* 1996; 276: 1253-1258

Received: May 22, 2008 / Accepted: June 24, 2008

Address for correspondence:

Prof. Dr. Frank Krummenauer
Clinical Epidemiology and Health Economy Unit
Department of Orthopedic Surgery
University Hospital Carl Gustav Carus
Medical Faculty of the Technical University of Dresden
Fetscherstr. 74 (bd 29)
D-01307 Dresden (Germany)
Phone: +49-351-458 3747
Fax: +49-351-458 4344
Email: Frank.Krummenauer@uniklinikum-dresden.de