

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Effect of Health-Related Quality-Of-Life Instrument and Quality-Adjusted Life Year Calculation Method on the Number of Life Years Gained in the Critical Care Setting

Tarja Vainiola, MNSc^{1,*}, Risto P. Roine, MD, PhD², Ville Pettilä, MD, PhD³, Taru Kantola, MD, PhD³, Pirjo Räsänen, PhD², Harri Sintonen, Prof⁴

¹National Institute for Health and Welfare, Helsinki, Finland; ²Helsinki and Uusimaa Hospital District, Helsinki, Finland; ³Helsinki University Central Hospital, Helsinki, Finland; ⁴University of Helsinki, Helsinki, Finland

ABSTRACT

Objectives: Health-related quality of life (HRQoL) and quality-adjusted life-years (QALYs) gained are basic elements in the cost-utility evaluations of health care. Different HRQoL instruments produce different scores for the same patient, and thus also a different number of QALYs. We examined the effect of these factors on the number of QALYs gained and the cost per QALY in the critical care setting. **Methods:** In 937 patients having been treated in the critical care setting in the Helsinki University Central Hospital the HRQoL scores were measured by the EQ-5D and 15D 6 and 12 months after start of treatment, and QALYs were calculated using four different sets of assumptions regarding recovery from disease. **Results:** The mean number of QALYs gained during the first year after treatment ranged from 0.178 ± 0.206 to $0.550 \pm$

0.508 and the consequent cost per QALY from €38,405 to €118,668 depending on HRQoL instrument and assumptions used in the calculations regarding recovery from disease. **Conclusions:** The HRQoL instrument and the assumptions employed regarding recovery from disease have a great influence on the results of cost-utility analyses and should, therefore, be explicitly described in studies reporting QALYs. Furthermore, a common consensus on which calculation method should be used within critical care would be urgently needed. **Keywords:** HRQoL, QALY, calculation method, EQ-5D, 15D.

Copyright © 2011, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

Introduction

Recently, attention has been paid to differences in the methodology of estimating quality-adjusted life years (QALYs) for patients undergoing an intervention, and how these differences affect the results of cost-utility analyses [1,2]. The calculation of QALYs is based on quality of life scores (utilities) that are obtained from patients filling in a generic utility-generating health-related quality of life (HRQoL) questionnaire at baseline, and at one or several time points during the study period (e.g., at 6 and 12 months). Within critical care, however, it is not usually possible for a patient to fill in questionnaires and, therefore, assumptions about the baseline HRQoL have to be made. One commonly used approach is to assume that without treatment a patient would die and use a baseline HRQoL score of zero [3–5]. Another possibility is to ask health care professionals to estimate the baseline HRQoL based on patient records [6].

Another problem in the calculation of QALYs is the fact that different instruments produce different HRQoL scores for the same patient [7–9]. Furthermore, even when using the same instrument, differences in the number of QALYs gained may arise from at least two sources. The first of them is related to assumptions concerning the evolution of HRQoL between measurement points because continuous measurement is usually impossible. Three assumptions have been presented: 1) HRQoL changes linearly between measurements

(most frequently used in the literature); 2) HRQoL remains constant from one measurement to the next and then changes overnight; and 3) HRQoL changes at the midpoint between measurements [10]. In the critical care setting the latter two alternatives are problematic when the calculation of QALYs is based on the assumption that without treatment patients would die. Consequently, some critical care QALY calculations have used a fourth alternative, namely the assumption that the change in HRQoL for better or worse takes place immediately [11]. The second source for possible variance in the number of QALYs is related to how the commonly used area-under-the-curve (AUC) method is applied because it makes a difference if the whole AUC is calculated or only the area above/below the baseline utility value [1,2].

Because QALYs, and the cost per QALY, play an increasingly important role in the allocation of health care resources, we set out to study, and using real life data, what the magnitude of the above mentioned factors is on the results of cost-utility analyses in the critical care setting.

Methods

The data were collected prospectively in the Helsinki University Hospital between January 1, 2003, and December 31, 2004. The study was approved by the local Ethics Committee. Two HRQoL

* Address correspondence to: Tarja Vainiola, National Institute for Health and Welfare, P.O. Box 30, FI-00271 Helsinki, Finland.

E-mail: tvainiol@mappi.helsinki.fi.

1098-3015/\$36.00 – see front matter Copyright © 2011, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

doi:10.1016/j.jval.2011.05.047

questionnaires were mailed to patients alive and with a known address 6 and 12 months after start of intensive care unit (ICU) or high-dependency unit (HDU) treatment together with a one-page letter, informed consent form, and a prepaid envelope for returning the signed consent form and the questionnaires. In case of nonresponse one reminder was sent. The data used for analysis consist of patients who returned the completely filled HRQoL questionnaires both at 6 and 12 months, or had died during that period (because their HRQoL score at 6 or 12 months was known; that is, 0) and on whom diagnostic information was available. For those who died before the end of the follow-up, QALYs were calculated using their observed survival time. Furthermore, only patients receiving care on an emergency basis (i.e., those whose admission to both the hospital and ICU or HDU occurred on the same day) were included and patients undergoing routine follow-up in an ICU after an elective procedure were excluded. One year direct secondary care costs (in- and outpatient hospital treatment) were obtained from the Ecomed clinical patient administration system (Datawell Ltd., Finland), where all costs of treatment of individual patients in the hospital are routinely stored.

The HRQoL Instruments

The EQ-5D consists of five dimensions with one item on each: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension is divided into three levels: no problems, some problems, and severe problems. For the calculation of EQ-5D scores we used the UK time-trade-off tariff, which is the most commonly used valuation system for the EQ-5D internationally. According to it the HRQoL scores range from -0.59 to 1, where one means full health and zero stands for death. No health state can obtain a score between 0.88 and 0.99, and negative scores indicate health states worse than death (WTD) [12].

The 15D consists of 15 dimensions with one item on each: breathing, mental function, speech, vision, mobility, usual activity, vitality, hearing, eating, elimination, sleeping, distress, discomfort and symptoms, depression, and sexual activity. Each dimension is divided into five levels from no problems to extreme problems. The HRQoL scores of the 15D range from 0 to 1, with one being equivalent to full health and zero equivalent to death [13].

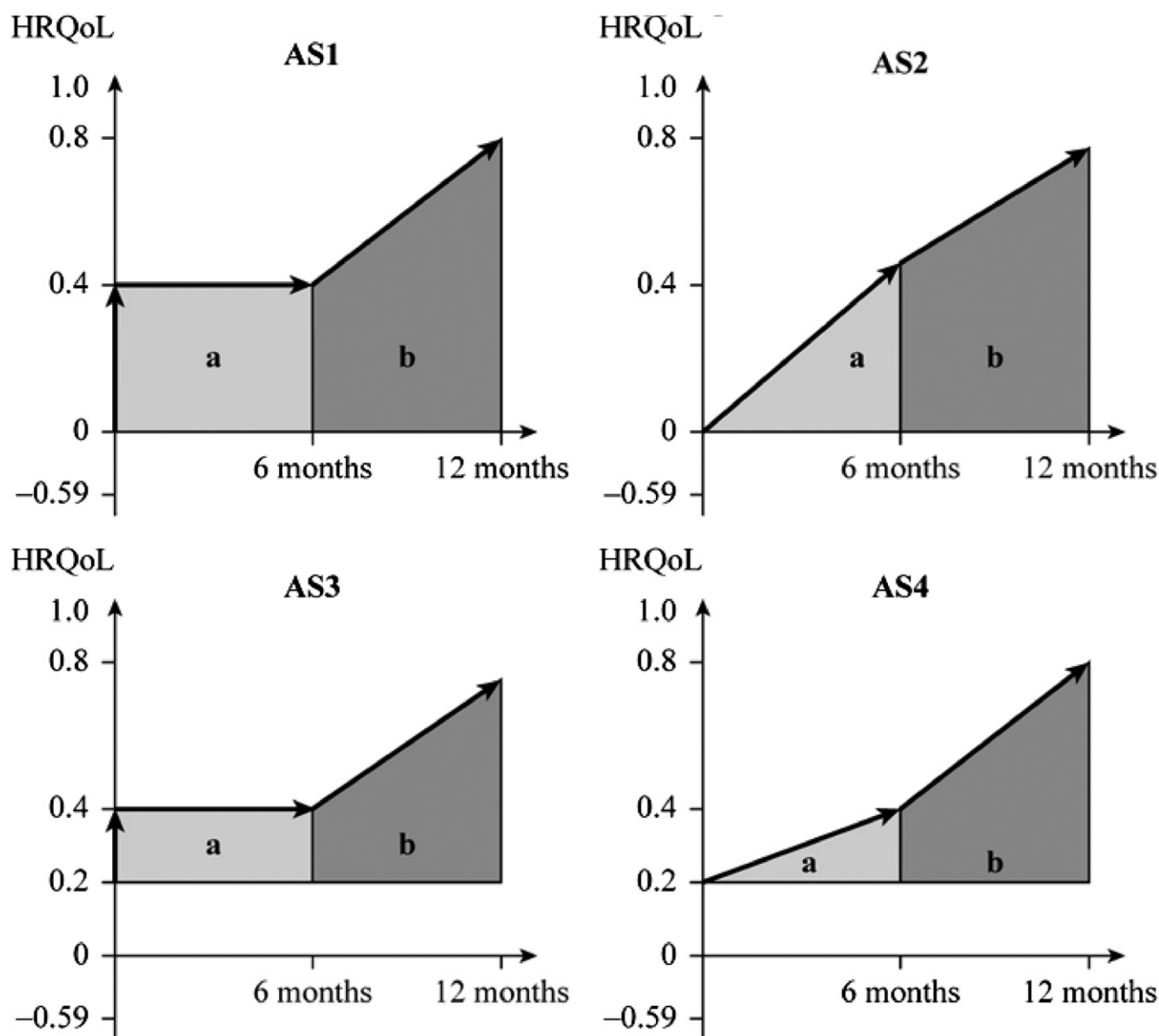


Fig. 1 - Quality-adjusted life year calculation methods based on Assumption Sets 1 to 4.

Assumptions used for Calculation of QALYs

The number of QALYs produced by ICU or HDU treatment within the year following the admission to ICU or HDU care was calculated on the basis of four different sets of assumptions and using two different baseline HRQoL scores (either zero assuming that without treatment patients would die, or a HRQoL score assessed by health care professionals based on patient records).

Assumption Set 1 (AS1). Without treatment patients would die; that is, their HRQoL score would be zero and thus no QALYs would be gained. With treatment the HRQoL changes immediately to the level observed at the first follow-up point (in our case at 6 months) after which it changes linearly to the HRQoL observed at the second follow-up point (in our case at 12 months). This implies that for those who die before the first follow-up, the QALY gain is zero. Theoretically the maximum number of QALYs gained is one. The AS1 panel of Figure 1 illustrates these assumptions and how the QALYs gained by treatment (shaded area under the curve = area a + area b = $0.4 \times T1 + [(0.8 + 0.4)/2] \times T2 = 0.5$) are estimated as the whole AUC. T1 is time alive in years between admission and the first follow-up and T2 is time alive in years between the first and second follow-up. With EQ-5D all these changes can also take place in the negative area (i.e., WTD) between -0.59 and 0 (this applies also to all other assumption sets).

Assumption Set 2 (AS2). The assumption that without treatment the patients would die immediately is the same as in AS1, but with treatment the change in HRQoL from baseline to the first follow-up is assumed to take place linearly. Theoretically the maximum number of QALYs gained is 0.750. The AS2 panel of Figure 1 illustrates these assumptions and how the QALYs gained by treatment (shaded AUC = area a + area b = $(0.4/2) \times T1 + [(0.4 + 0.8)/2] \times T2 = 0.4$) are estimated as the whole AUC.

Assumption Set 3 (AS3). Without treatment patients are not assumed to die, but to stay at the baseline self-reported or proxy-assessed HRQoL level for the whole follow-up period. With treatment the HRQoL changes immediately to the level measured at 6 months after which it changes linearly to the HRQoL score observed at 12 months. The maximum number of QALYs gained depends on the baseline HRQoL score. The AS3 panel of Figure 1 illustrates these assumptions and how the QALYs gained by treatment (shaded AUC = area a + area b = $0.4 \times T1 + [(0.4 + 0.8)/2] \times T2 - 0.2 \times T3 = 0.3$) are estimated as the area above/below the baseline HRQoL score. T3 is the length of the follow-up in years (T1+T2).

Assumption Set 4 (AS4). The assumptions are otherwise the same as in the AS3, but with treatment the change in HRQoL from baseline to the first follow-up and between follow-ups is assumed to take place linearly. The maximum number of QALYs gained depends on the baseline HRQoL score. The AS4 panel of Figure 1

illustrates these assumptions and how the QALYs gained by treatment (shaded AUC = area a + area b = $[(0.2+0.4)/2] \times T1 + [(0.4 + 0.8)/2] \times T2 - 0.2 \times T3 = 0.25$) are estimated as the area above/below the baseline HRQoL score.

The Baseline HRQoL

For the AS3 and AS4 the HRQoL of patients at admission was assessed retrospectively based on information obtained from patients' medical and nursing records. All baseline assessments were performed by two persons: a physician (TK or RPR) and a registered nurse with a long history of ICU work (TV). The information was mapped onto the 15D and EQ-5D questionnaires and the corresponding scores were derived. The patients were classified into ten diagnostic groups according to the ICD10 and every 10th patient was selected from each diagnostic group for baseline evaluation. The target was to evaluate 18 patients in each diagnostic group but because all patient records could not be retrieved, and 19 patients were excluded because they received routine follow-up care after an elective procedure, the final number of patients evaluated was 112. The HRQoL proxy assessment was based on the status of the patient at admission to the ICU or HDU. The baseline HRQoL used for the analyses was the average of all assessments in each diagnostic group (Table 1).

Statistical Methods

The results of QALY calculations are presented using descriptive statistics (i.e., mean, standard deviation, range, and proportions). Differences in the mean number of QALYs obtained with the two HRQoL instruments were compared with paired samples t test, and differences in medians and distributions with the Wilcoxon Signed Ranks Test. Chi-square test was used to find out if the distribution of proxy-assessed patients across diagnostic groups deviates from that of all ICU patients across these groups. $P < 0.05$ was considered statistically significant.

Results

There were 3600 patients treated at HCU or HDU during the study period. Of them 1990 patients fulfilled the inclusion criteria. Because 451 patients died before the first follow-up there were 1539 patients alive at 6 months and of them 486 (31.6%) returned both questionnaires. Between 6 and 12 months another 17 patients died and 469 patients returned the 12-month questionnaire. One-year mortality was thus 23.5%. Those who had died before the 6-month follow-up were included in the analyses (because their HRQoL score could be set at zero) as were also those who had returned the

Table 1 – Mean proxy-assessed (by two health care professionals) baseline health-related quality of life scores according to the EQ-5D and the 15D based on information obtained from patient records.

Diagnostic group	Mean EQ-5D score	Mean 15D score	No. of assessed patients	Percentage of patients in proxy-assessed baseline group/percentage of patients in intensive care unit population
Resuscitated patients	-0.59	0.11	15	13.4/8.5
Neurological diseases	-0.51	0.15	10	8.9/9.6
Respiratory organ diseases	-0.41	0.25	11	9.8/5.6
Intoxication	-0.42	0.21	16	14.3/13.7
Infectious diseases	-0.46	0.24	14	12.5/13.2
Gastrointestinal diseases	-0.38	0.39	12	13/13.8
Other diseases	-0.25	0.41	19	17.0/14.1
Vascular diseases	-0.34	0.53	7	6.3/8.1
Heart diseases	0.09	0.53	8	7.1/13.4
On average	-0.39	0.29	112	

Table 2 – The number of QALYs gained using the 15D and the EQ-5D for the assessment of health-related quality of life and four different assumption sets for calculation of quality-adjusted life years (QALYs) (N = 937).

	AS1		AS2		AS3		AS4	
	15D	EQ-5D	15D	EQ-5D	15D	EQ-5D	15D	EQ-5D
Mean	0.42	0.37	0.31	0.27	0.23	0.55	0.18	0.41
SD	0.41	0.40	0.31	0.30	0.28	0.54	0.21	0.40
Median	0.19	0.08	0.39	0.07	0.13	0.42	0.10	0.32
Min	0	-0.30	0	-0.22	-0.21	-0.07	-0.11	-0.05
Max	1	1	0.75	0.75	0.88	1.59	0.66	1.20
Negative QALY (%)	0	1.4	0	1.4	16.5	2.0	16.5	2.0
Cost/QALY	50,412	57,713	67,271	76,811	90,657	38,405	118,668	51,269

6-month questionnaires but had then died before the second follow-up at 12 months. Altogether the study group consists thus of 937 patients, 47.1% of the basic population.

The respondents and nonrespondents differed statistically significantly regarding length of stay at ICU or HDU (5.5 vs. 4.3 days; $P < 0.001$), costs of treatment (€21,103 vs. €18,129; $P 0.006$) and age (61 vs. 54; $P < 0.001$) indicating that respondents had a slightly more severe disease than the nonrespondents. Sixty-two percent of respondents were men and the mean length of hospital stay was 14.0 days. The mean utility score at 6 months for patients alive was 0.820 (CI 95% 0.808–0.832) for the 15D and 0.714 (CI 95% 0.690–0.738) for the EQ-5D, respectively. At 12 months the mean scores were 0.830 (CI 95% 0.818–0.842) and 0.727 (CI 95% 0.702–0.752), respectively. The distribution of proxy-assessed patients across diagnostic groups did not deviate statistically significantly from that of all ICU patients across these groups ($P = 0.151$).

The mean number of QALYs gained calculated using the 15D or the EQ-5D differed significantly in all assumption sets; both between instruments and within each instrument ($P < 0.001$) (Table 2). With the AS1 and AS2 the 15D produced on average slightly more QALYs than the EQ-5D. By contrast, with the AS3 and AS4 the EQ-5D produced on average two times more QALYs than the 15D. With the AS3 and AS4 the 15D produced considerably more often negative QALYs than the EQ-5D, but the latter produced negative QALYs with all assumption sets. The variation in the number of QALYs gained (in percentages of standard based on AS1 and EQ-5D and set at 100) using different sets and HRQoL instruments, is shown in Figure 2. During the 1-year follow-up the 15D and EQ-5D produced one QALY with AS1 (which is the maximum number of QALYs that can be gained when using HRQoL instruments with scores ranging between zero and one) in 1.9% and 10.1% of cases, respectively. The EQ-5D produced more than one QALY with AS3 and AS4 in 30.5% and 7.4% of cases, respectively. The maximum

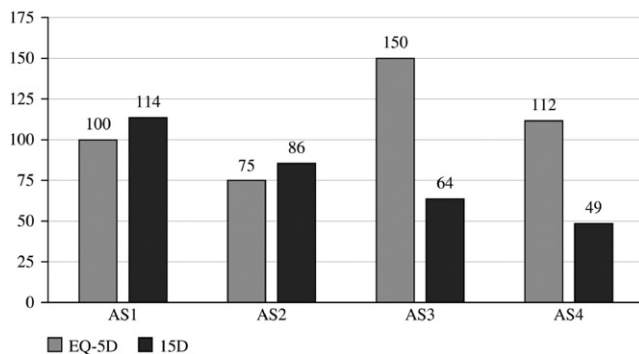


Fig. 2 – The variation in the number of QALYs gained (as percentage of standard based on assumption set AS1 and EQ-5D and set at 100) using different sets and health-related quality of life instruments (N = 937).

number of QALYs within a year was 1.594 with the EQ-5D. For the 15D the maximum QALY gain was one QALY (Table 2).

Like the number of QALYs, also the cost per QALY gained varied within a very wide range depending on the HRQoL instrument and assumption set used. With the instrument constant, the assumption set used resulted at maximum over double difference in the cost per QALY. With the assumption set fixed, the choice of the instrument resulted at maximum in more than doubling of the cost per QALY (Table 2).

Discussion

This study set out to explore the implications of using different HRQoL instruments and assumption sets for assessing the number of QALYs gained within the year following admission to ICU or HDU treatment. Both the HRQoL instrument and the assumption set used had a significant effect on the total number of QALYs gained. The calculation methods based on the estimated baseline HRQoL (i.e., AS3 and AS4) produced clearly less QALYs with the 15D than the methods based on the assumption that without treatment patients would die (i.e., AS1 and AS2). But with the EQ-5D the situation was opposite, the calculation methods based on the estimated baseline HRQoL (i.e., AS3 and AS4) produced more QALYs than the methods based on the assumption that without treatment patients would die (i.e., AS1 and AS2).

The crucial assumption of the AS1 and AS2 is that without ICU or HDU treatment all patients would have died immediately or very quickly. Thus all observed survival and HRQoL of patients is attributed to ICU or HDU treatment. This assumption is certainly valid in most ICU and in many HDU patients, but not always. Thus, when applying the calculation methods AS1 and AS2, the QALYs gained by those, who without ICU or HDU treatment would have stayed alive for different durations of time at different levels of HRQoL, are omitted. Consequently, AS1 and AS2 overestimate the incremental QALYs gained by ICU or HDU treatment. On the other hand, these approaches omit the QALYs gained by those, who with ICU or HDU treatment stayed alive at different levels of HRQoL for some time, but not until the first follow-up point, and may thus also underestimate the QALYs gained by ICU or HDU treatment.

In AS1 and AS2 the assumption that all patients have at baseline the HRQoL score of zero (dead) conceals an important difference between instruments that produce negative (worse than death) values for health states (like the EQ-5D) and those, that do not (like the 15D). In the follow-up measurements with the EQ-5D the patients can obtain negative HRQoL values and thus gain negative QALYs as can be seen from Table 2. This raises a philosophical and ethical problem of whether dying immediately produces in fact more QALYs (zero) than living in a state worse than death for some time—purely mathematically zero is more than any negative number.

AS3 and AS4 require that at admission to an ICU or HDU patients are assigned a realistic HRQoL score other than zero. In most

cases it may be impossible to derive it based on patient self-report and it has to be assessed by a proxy. It can, of course, be questioned whether the HRQoL estimates by proxies are in agreement with those of the patients themselves. There is evidence from the intensive care setting to suggest though that the estimates by close proxies [14,15] and even by professionals are in relatively good agreement with those of patients [16]. Although the proxy assessed HRQoL could be acceptable, our target is not to generalize the baseline HRQoL assessed and used in this study to all studies in the critical care setting. The baseline scores are theoretical estimates derived for this study from empiric data to enable the QALY calculation by the AS3 and AS4.

The 15D score is by definition always positive, whereas the EQ-5D score can in many cases be also negative (i.e., WTD). With AS3 and AS4 the 15D (and the EQ-5D with a positive score at baseline) can produce negative QALYs if a patient dies before 6 months or if the score at 6 months is lower than at baseline. In that case the AUC is smaller than the white area under the shaded area for AS3 and AS4 in Figure 1. Thus dying brings a loss in QALYs as can be seen in a number of cases in Table 2. In addition, the EQ-5D can produce negative QALYs if the baseline HRQoL score is negative and the majority of the area above/under the curve is on the negative side. If a patient is at baseline in the worst EQ-5D score of -0.59 , AS3 can produce a maximum of 1.59 QALYs in 1 year, if the HRQoL score at 6 months is one and remains at that level until 12 months.

In trials dealing with elective care AS4 is the most frequently used approach for the calculation of QALYs in the literature [1]. When applying AS3 and AS4 in the critical care setting; however, the assumption that without ICU or HDU treatment all patients would live the whole year at the baseline level of HRQoL does not hold in practice. With treatments available outside ICUs or HDUs many patients would have died immediately and some would have lived for unknown durations at unknown levels of HRQoL. Thus those assumptions may result in an overestimate of QALYs to be experienced without ICU or HDU treatment.

In critical care, the rule of rescue; that is, the duty to safe endangered life where possible benefit can be seen, applies and, therefore, we do not know what would happen to the patients in terms of length and quality of life by just “conventional” treatment. This is also difficult to establish, as it would, in most cases, be unethical to organise a trial, where patients would be randomised to obtain ICU/HDU or conventional treatment. Consequently, studies evaluating the effectiveness of critical care must always be based on assumptions. To be able to compare the cost-effectiveness of treatments across a variety of medical specialties, similar assumptions should be used irrespective of if care is provided on an emergency or elective basis. Because it is difficult to say which assumption set is the most realistic one in the critical care setting, there is a clear need for sensitivity analyses using different assumption sets when reporting results of studies.

Because both the choice of the HRQoL instrument and the assumption set regarding progression of disease have a significant

effect on the total number of QALYs gained, it is evident that also the cost/QALY varies within a wide range. Our results should be seen as a theoretical example of how different instruments and assumptions play a major role in the calculation of QALYs and stress the fact that when reporting results of cost-utility studies in the critical care setting, special attention should be paid to describing how the QALYs have been calculated. Otherwise the comparison of the results of different studies is futile.

Source of financial support: Funding for this project was provided by the Helsinki University Central Hospital and the Yrjö Jahnsson Foundation. Harri Sintonen is the developer of the 15D instrument and one of the developers of the EQ-5D instrument.

REFERENCES

- [1] Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;14:487–96.
- [2] Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health Econ* 2004;13:1203–10.
- [3] Sznajder M, Aegerter P, Launois R, et al. A cost-effectiveness analysis of stays in intensive care units. *Intens Care Med* 2001;27:146–53.
- [4] Åhlström A, Tallgren M, Peltonen S, et al. Survival and quality of life of patients requiring acute renal replacement therapy. *Intens Care Med* 2005;31:1222–8.
- [5] Kaarlola A, Tallgren M, Pettilä V. Long-term survival, quality of life, and quality-adjusted life-years among critically ill elderly patients. *Crit Care Med* 2006;34:2120–6.
- [6] Kantola T, Mäklin S, Koivusalo AM, et al. Cost-utility of molecular adsorbent recirculating system treatment in acute liver failure. *World J Gastroentero* 2010;16:2227–34.
- [7] Stavem K. Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease. *Qual Life Res* 1999;8:45–54.
- [8] Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Ann Med* 2001;33:358–70.
- [9] Vainiola T, Pettilä V, Roine RP, et al. Comparison of two utility instruments, the EQ-5D and the 15D, in the critical care setting. *Intens Care Med* 2010;36:2090–3.
- [10] Billingham LJ, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. *Health Technol Asses* 1999;3:55–69.
- [11] Karlsson S, Ruokonen E, Varpula T, et al. Long term-outcome and quality-adjusted life years after severe sepsis. *Crit Care Med* 2009;34:1268–74.
- [12] Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;35:1095–108.
- [13] Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;33:328–36.
- [14] Hofhuis J, Hautvast JL, Schrijvers AJ. Quality of life on admission to the intensive care: can we query the relatives? *Intens Care Med* 2003;29:974–9.
- [15] Elliott D, Lazarus R, Leeder SR. Proxy respondents reliably assessed the quality of life of elective cardiac surgery patients. *J Clin Epidemiol* 2006;59:153–9.
- [16] McPhail S, Beller E, Haines T. Two perspectives of proxy reporting of health-related quality of life using the Euroqol-5D, an investigation of agreement. *Med Care* 2008;46:1140–8.