Volume 8 • Number 1 • 2005
VALUE IN HEALTH

An Assessment of the Impact of Informative Dropout and Nonresponse in Measuring Health-Related Quality of Life Using the EuroQol (EQ-5D) Descriptive System

Julie Ratcliffe, PhD,[1] Tracey Young, MSc,[2] Louise Longworth, MSc,[2] Martin Buxton, BA[2]

[1]School of Health and Related Research, University of Sheffield, Sheffield, UK; [2]Health Economics Research Group, Brunel University, Uxbridge, UK

ABSTRACT

Objectives: To investigate the impact of imputing EQ-5D values to allow for informative dropout and nonresponse in a longitudinal assessment of the health-related quality of life (HRQL) of liver transplant recipients.
Methods: The EQ-5D was administered at defined time intervals pre- and post-transplantation to all adults who were listed to receive liver transplants as National Health Service (NHS) treatment at each of the six Department of Health designated centers in England and Wales over a time-period of 36 months (12 month recruitment period and 24 month follow-up period). During the course of the study missing data arose for two main reasons, informative dropout and nonresponse. Informative dropout was accounted for by giving those patients who died an EQ-5D score of 0 and those patients who were too ill to respond to an EQ-5D score equivalent to the 5th percentile of respondents for each time point pretransplantation. Nonresponse was accounted for using relatively naïve approaches (last value carried forward, and upper/lower 95% confidence interval around the mean) and contrasted with a more sophisticated multiple imputation method.
Results: Adjusting for informative dropout in isolation resulted in a marked deterioration in mean scores over time pretransplant relative to the base case situation in which no such adjustments were made. Nevertheless, adjusting for informative dropout and/or nonresponders did not alter the base case conclusion of no statistically significant differences in mean EQ-5D scores over time pretransplant. In contrast, post-transplant data indicated highly statistically significant improvements in quality of life over time for the base case ($P < 0.001$) whereas no statistically significant improvements over time were found when informative dropout was allowed for in isolation ($P = 0.402$) or when informative dropout and nonresponse were allowed for simultaneously ($P = 0.105$–0.185).
Conclusions: It is important that future studies which purport to assess the HRQL over time of patients, such as these with end-stage liver disease, include an allowance for informative dropout and nonresponse within the analysis.
Keywords: EQ-5D, informative dropout, liver transplantation, quality of life.

Introduction

The measurement of health-related quality of life (HRQL) is an essential component of any cost utility analysis (CUA) in health care [1,2]. Descriptive systems are increasingly being used in this context because of their ability to provide "off the shelf" tariff values for a wide variety of generic health states which can then be used in any CUA without the requirement for collection of primary data to elicit health state values [3–5]. The EuroQol instrument (EQ-5D) represents one such descriptive system which has been applied in a variety of health-care settings throughout Europe [5,6]. Typically, patients are asked to complete the EQ-5D and their responses are converted to scores using a set of predetermined values generated from a representative sample of the general population.

The majority of CUA studies relate to an extended time period and typically health-related quality of life (HRQL) of patients needs to be assessed at several time points during the course of follow-up, often by means of a self completion questionnaire. In such studies it is common for a proportion of HRQL data to be missing.

Address correspondence to: Julie Ratcliffe, Sheffield Health Economics Group, School of Health and Related Research, University of Sheffield, Regent Court 30, Regent Street, Sheffield S1 4DA, UK. E-mail: j.ratcliffe@sheffield.ac.uk

Responses missing at a particular time point (*t*) can be categorized in three main ways [7]. First, responses may be missing completely at random, where the probability of response at time *t* is independent of both the previously observed value and the unobserved value at time *t* and other variables collected at this, or previous, time points, for example measures of disease severity. Second, responses may be missing at random where the probability of response at time *t* depends on the previously observed value or other variables collected during the study but not the unobserved values at time *t*. The third category of missing response is nonignorable nonresponse when the probability of response at time *t* depends on the unobserved values at time *t* and possibly on the previously observed values also.

Where the intervention being considered involves treatment for a severe condition, as is the case with liver transplantation, patients may be very ill before receiving the intervention and it is common for a proportion of participants to fail to complete questionnaires because of the severity of their illness or death at any time point. Such data represent nonignorable nonresponse because the missing data are directly dependent on the health status of the patient. It is also common for a proportion of patients to choose not to respond to the questionnaire at one or multiple time points. Such data may be assumed to be missing completely at random because the probability of response is independent of any previously observed values.

A recent report on methods for the analysis of quality of life and survival data in health technology assessment recommended that methods for the longitudinal analysis of quality of life data allow for informative dropout [8]. This article investigates the impact on the results obtained of imputing health state values for patients who dropped out of a study for informative reasons or for those who chose not to respond to the questionnaire at one or more time points. This study draws on data from a broader study to assess the HRQL of liver transplant recipients pre- and post-transplantation, undertaken within the context of an economic evaluation of the liver transplantation program in England and Wales [9]. The majority of studies undertaken to date which have focused on the HRQL of liver transplant recipients have reported dramatic improvement after liver transplantation [9–11]. Nevertheless, no previous studies in this area have considered the impact of informative dropout or nonresponse on the results obtained.

Methods

Information on HRQL and sociodemographic data were collected using a postal questionnaire. Within the questionnaire, individuals were asked to complete the EQ-5D, a nondisease specific instrument for describing and subsequently valuing HRQL [5,6]. It is the responses to this first part of the EQ-5D that forms the basis for the analysis reported here. Respondents are asked to categorize their health status on five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression), where each dimension has three possible levels of response. A value for each health state (possible combinations of dimension levels) was derived from a representative sample of the UK general population using the time trade-off method to elicit preferences [12].

The population of interest was all individuals who were selected to receive a liver transplant as part of the National Health Service (NHS) liver transplantation program at each of the Department of Health designated centers in England and Wales during a 12-month recruitment period, from 1995 to 1996 (n = 400). The questionnaire was administered at regular time intervals during the course of their treatment. The majority of patients were surveyed at the point of their being listed for liver transplantation and at 3 months and 6 months after listing. The vast majority of patients were transplanted at between 2 and 6 months after listing. Very few patients waited more than 6 months to receive their transplant. The questionnaire was also administered to all eligible transplanted patients at 3 months, 6 months, 12 months, and 24 months post-transplantation (n = 387). At each time point, patients who, in the opinion of a clinical research nurse based at the center, were too severely ill to be approached were not sent a questionnaire. One reminder was sent to all nonrespondents at each time interval, approximately three weeks after the administration of the initial questionnaire. Individuals in the survey were asked to complete the EQ-5D from the perspective of their own health today. At every time point, each individuals' responses to the five dimensions of the EQ-5D were converted to a single value using the tariff of mean values for health states produced by the York Measurement and Valuation of Health (MVH) survey for the EQ-5D, based on the time trade-off scaling technique [13].

The EQ-5D values data were analyzed using several different approaches and comparisons were then made between the results generated using each approach. In the base case analysis (A) no adjustment was undertaken to allow for informative dropout or nonresponse. In the second analysis (B) an allowance was made for informative dropout only, by giving all patients who died pretransplant an EQ-5D tariff score of 0 and all patients who became too ill an EQ-5D tariff score which corresponded to the score attributable to the 5th percentile of respondents for each subsequent time point pretransplantation. The choice of the 5th percentile was essentially arbitrary but, of the data that were available, was considered to be most broadly representative of respondents who were most ill.

Similarly, all patients who died post-transplant were given an EQ-5D tariff score of 0 and all patients who became too ill an EQ-5D tariff score attributable to the 5th percentile of respondents for each subsequent time point(s) post-transplantation.

In the third (method C), fourth (method D), and fifth (method E) analyses an allowance was made for nonresponse only using alternative methods. A distinction was made between individuals who never responded to the survey and those who did not respond or whose responses were unusable because of an incomplete response at one or more time points. The upper (C) and lower (D) confidence intervals (CI) around the mean EQ-5D tariff score for each time-point for responders were attributed to those patients who never responded. For those patients who responded at some but not all time-points, the approach adopted was to impute using the last known value carried forward. In method E, the more sophisticated statistical technique of multiple imputation (MI) was used to estimate values for nonresponders at each time interval. MI is a Monte Carlo simulation technique where each missing data case is replaced by a set of plausible estimates, which are drawn from the predictive distribution of the missing data given the observed data. In contrast to the more naïve approaches, the technique of MI has the advantage that it includes a random component to reflect the fact that imputed values are estimated rather than treating the imputed values as if they are known with certainty. As such, MI is likely to produce more accurate estimates of the standard errors (SE) and variances of the mean utility values at each time-point than other methods of imputation [14]. A priori this was our preferred approach for handling missing data because of nonresponse.

The data were assumed to follow a multivariate normal distribution and multiple imputation was carried out. The predictive distribution reflects the uncertainty about the missing data. The number of estimates required is determined by the rate of missing information for the quantity being estimated [10]. In this study the rate of missing information determined that five estimates were required. The missing EQ-5D scores were estimated taking into account the time-point of questionnaire administration and the characteristics of each of the nonresponders, including their age, sex, and type of liver disease, and matching them to equivalent individuals in the group of responders. Each completed data set was then separately analyzed using complete data statistics and the results (mean estimates and standard deviations) were combined using simple rules developed by Rubin to produce overall mean estimates and standard deviations reflecting missing data uncertainty because of nonresponse [15].

The data were analyzed using the statistical package SPSS version 10 [16] and the computer software package NORM [17], was used for multiple imputation. All of the statistical tests used in the analysis were parametric. To determine differences in EQ-5D scores between each time point, a repeated measures analysis of variance was performed.

Results

During the course of the study, two patient groups accounted for missing data, the first led to informative dropout and the second led to nonresponse. Within the informative dropout group, 31 individuals became too ill to complete the questionnaire at some time point, of whom 12 individuals subsequently died, and 81 individuals dropped out of the study because death. Within the nonresponse group, 114 individuals chose not to respond to the questionnaire at one or more time intervals and 16 individual's responses were unusable as they were incomplete. The survey population, respondents, and nonrespondents by time-point are presented in Table 1.

The results of the analyses of the pretransplant data for dependent samples are presented in Table 2 (and graphically in Fig. 1). A comparison of mean EQ-5D pretransplant scores at listing and 3 months after listing in the base case analysis (A) revealed a deterioration in HRQL between listing and 3 months after listing. Adjusting for informative dropout in isolation (Method B) resulted in a marked deterioration in mean scores over time pretransplant relative to the base case situation in which no such adjustments were made. Nevertheless, adjusting for informative dropout and/or nonresponse, regardless of the mechanism used to adjust for nonresponse, did not alter the base-case

Table 1 Survey population, respondents, and nonresponders by time-point

Time point	Total eligible population	No. of patients from total eligible population omitted as too ill	No. of patients from total eligible population omitted because of death	No. of quests sent out	No. of quests not returned or unusable (missing data)	Useable responses N (%) relative to the total eligible population
Listing	400	17	0	383	104	74%
3 months after listing	119*	7	2	110	37	69%
3 months after transplant	387†	2	22	363	78	80%
6 months after transplant	387	2	39	346	86	78%
12 months after transplant	387	0	55	332	60	84%
24 months after transplant	387	0	69	298	80	79%

*Of the patients listed, 281 had been transplanted by 3 months.
†Of the 400 patients listed, 13 were delisted before transplantation.

conclusion of no statistically significant differences in mean EQ-5D scores over time pretransplant.

Comparisons of dependent samples post-transplant (Table 3 and Fig. 2) showed highly statistically significant improvements over time when no allowance for informative dropout was made (method A, $P < 0.001$). In marked contrast, no statistically significant improvements over time were found when informative dropout was allowed for (method B, $P = 0.402$). Allowing for nonresponders in isolation (methods B, C, and E) reinforced the findings of the main analysis regardless of the methods used to impute EQ-5D scores for nonresponders. Simultaneous adjustment for informative dropout and nonresponse (methods B + C, B + D, and B + E) resulted in no statistically significant improvements in mean EQ-5D scores post-transplant regardless of the methods used to impute scores for nonresponders.

Table 2 Comparison of mean EQ-5D pretransplant scores over time (Repeated measures analysis of variance) at listing and 3 months after listing

Method	Listing mean 95% CI (N)	3 months after listing mean 95% CI (N)	*P*-value
A: (unadjusted analysis)	0.462 0.422–0.502 (279)	0.365 0.274–0.456 (73)	0.077
B: (adjusting for informative dropout)	0.428 0.387–0.469 (296)	0.329 0.245–0.413 (82)	0.081
C: (nonresponders upper 95% CI)	0.473 0.444–0.502 (383)	0.396 0.345–0.446 (110)	0.184
D: (nonresponders lower 95% CI)	0.451 0.420–0.480 (383)	0.334 0.280–0.388 (110)	0.190
E: (nonresponders missing imputation)	0.470 0.436–0.510 (383)	0.382 0.335–0.405 (110)	0.064
B + C: (adjusting for informative dropout + nonresponders upper 95% CI)	0.459 0.429–0.488 (400)	0.372 0.320–0.424 (119)	0.069
B + D: (adjusting for informative dropout + nonresponders lower 95% CI)	0.438 0.408–0.468 (400)	0.361 0.309–0.413 (119)	0.072
B + E: (adjusting for informative dropout + nonresponders multiple imputation)	0.456 0.415–0.494 (400)	0.368 0.284–0.449 (119)	0.093

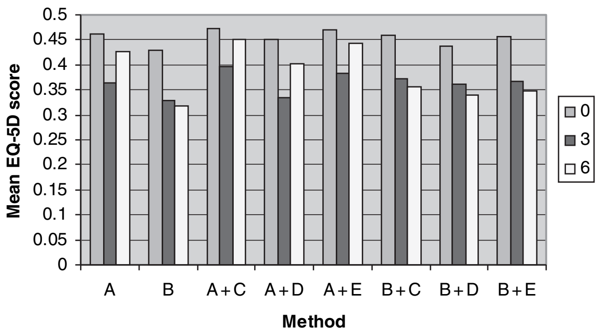

Figure 1 Comparison of mean EQ-5D scores pretransplant.

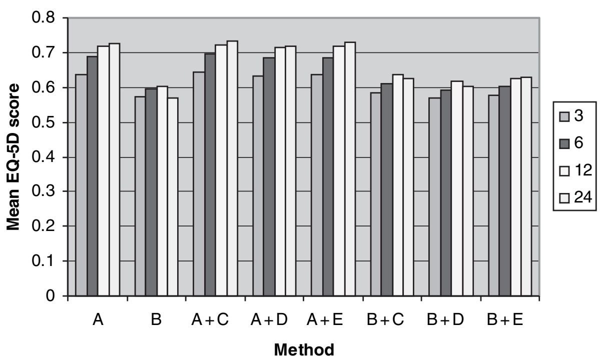

Figure 2 Comparison of mean EQ-5D scores post-transplant.

Table 3 Comparison of mean EQ-5D post-transplant scores over time (Repeated measures analysis of variance) at 3, 6, 12, and 24 months post-transplant

Method	3 months post-transplant mean 95% CI (N)	6 months post-transplant mean 95% CI (N)	12 months post-transplant mean 95% CI (N)	24 months post-transplant mean 95% CI (N)	*P*-value
A: (unadjusted analysis)	0.636 0.605–0.667 (285)	0.689 0.657–0.721 (260)	0.717 0.687–0.747 (272)	0.725 0.688–0.762 (218)	<0.001
B: (adjusting for informative dropout)	0.574 0.534–0.617 (309)	0.595 0.549–0.638 (301)	0.601 0.560–0.642 (327)	0.569 0.506–0.629 (287)	0.402
C: (nonresponders upper 95% CI)	0.644 0.624–0.669 (363)	0.697 0.669–0.721 (346)	0.722 0.697–0.747 (332)	0.733 0.698–0.762 (298)	<0.001
D: (nonresponders lower 95% CI)	0.632 0.604–0.656 (363)	0.683 0.658–0.708 (340)	0.714 0.682–0.739 (332)	0.718 0.688–0.749 (298)	<0.001
E: (nonresponders missing imputation)	0.635 0.601–0.672 (363)	0.686 0.640–0.711 (340)	0.719 0.685–0.741 (332)	0.730 0.663–0.744 (298)	<0.001
B + C: (adjusting for informative dropout + nonresponders upper 95% CI)	0.586 0.559–0.613 (387)	0.610 0.580–0.640 (387)	0.635 0.592–0.667 (387)	0.625 0.580–0.661 (387)	0.169
B + D: (adjusting for informative dropout + nonresponders lower 95% CI)	0.569 0.542–0.596 (387)	0.593 0.563–0.623 (387)	0.618 0.605–0.676 (387)	0.603 0.551–0.648 (387)	0.105
B + E: (adjusting for informative dropout + nonresponders multiple imputation)	0.576 0.538–0.602 (387)	0.601 0.572–0.641 (387)	0.626 0.589–0.661 (387)	0.629 0.540–0.776 (387)	0.185

Discussion

This study has highlighted the marked differences which can occur in mean utility values within longitudinal studies when proper allowance is made for informative dropout and/or nonresponse relative to a situation in which such adjustments are not made. The majority of studies undertaken to date which have focused on the HRQL of liver transplant recipients have reported dramatic improvement after liver transplantation, despite the possibility of problems post-transplant including rejection, infection, and the side effects of lifelong immunosuppression [8,14,15]. The results from our base case analysis mirror the findings of previous studies in that, in general, HRQL after liver transplantation improves dramatically relative to the pretransplant situation. Whilst such analyses typically provide a detailed descriptive profile of the quality of life of the respondents, they do not provide a full account of the quality of life experience of all patients who have received the intervention.

In relation to the informative dropout group, such data represent nonignorable nonresponse because the missing data are directly dependent on the health status of the patient. Where informative dropout is not accounted for, the responses achieved are likely to be biased upwards because of the absence of any extremely low values. The extent of the bias achieved is dependent on the proportion of patients within the overall sample who drop out of the study because of death or severe ill health. In this study allowing for informative dropout resulted in a marked reduction in mean EQ-5D scores over time post-transplant and effectively canceled out the steady improvement in HRQL over time reported in the base case.

In relation to the nonresponse group, the data are assumed to be missing completely at random. Allowing for nonresponse in isolation resulted in similar mean EQ-5D scores to those achieved in the base case, regardless of the method of imputation used. Nevertheless, naive methods of correcting for nonresponse (e.g., upper and lower 95% CI around the mean, methods C and D) assume that all nonresponders would have responded in a similar fashion to respondents, thereby producing broadly equivalent estimates of utility, with little estimated variance. The uncertainties involved in the estimation of these utility values are ignored. MI has the advantage that it includes a random component to reflect the fact that imputed values are estimated rather than treating the imputed values as if they are known with certainty. Thus, the technique of MI accounts for the amount of uncertainty involved in the imputation of missing values. This thereby produces a more accurate estimate of the SE and variances of the mean utility values at each time-point.

To provide a full account of the quality of life experience of all patients, it is important that future studies which purport to assess the HRQL over time of patients with end-stage liver disease and other chronic conditions include an allowance for informative dropout and nonresponse within the analysis.

This work was undertaken whilst Julie Ratcliffe was employed at the Health Economics Research Group, Brunel University. We would like to acknowledge the assistance of the cost-effectiveness of liver transplantation (CELT) study team in the collection of the data from which this article emanates. We thank three anonymous referees for their helpful comments.

Source of financial support: The research was financially supported by the Department of Health's Policy Research Programme. The views expressed and any errors or omissions are the responsibility of the authors alone.

References

1 Drummond M, O'Brien B, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes (2nd ed.). Oxford: Oxford University Press, 1997.
2 Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-Effectiveness in Health and Medicine. Oxford: Oxford University Press, 1996.
3 Kaplan RM, Anderson JP. The general health policy model: an integrated approach. In: Spilker B, ed., Quality of Life Assessments in Clinical Trials. New York: Raven Press, 1990.
4 Torrance GW, Feeny DH, Furlong WJ, et al. Multiattribute utility function for a comprehensive health status classification system. Health Utilities Index Mark 2. Med Care, 1994;34:702–22.
5 Brooks R. EuroQol-: the current state of play. Health Policy 1996;37:53–72.
6 EuroQol group. EuroQol a new facility for the measurement of health related quality of life. Health Policy 1990;16:199–208.
7 Little DB, Rubin RJA. Statistical Analysis with Missing Data (2nd ed.). Wiley series on probability and statistics. New York: John Wiley & Sons Inc., 2002.
8 Billingham LJ, Abrams KR, Jones DR. Methods for the analysis of quality of life and survival data in health technology assessment. Health Technol Assess 1999;3:10.
9 Ratcliffe J, Longworth L, Young T, et al. Assessing health related quality of life pre and post liver transplantation: a prospective multi-centre study. Liver Transpl 2002;8:263–70.
10 Hellgren A, Berglund B, Gunnarsson U, et al. Health-related quality of life after liver transplantation. Liver Transpl Surg 1998;4:215–21.
11 Bryan S, Ratcliffe J, Neuberger J, et al. An assessment of patient's quality of life following liver transplantation. Qual Life Res 1998;7:115–20.
12 Dolan P. Modelling values for EuroQol health states. Med Care 1997;35:1095–108.
13 Dolan P, Gudex C, Kind P, Williams A. The TTO method, results from a general method survey. Health Econ 1996;5:141–54.
14 Briggs A, Clark T, Wolstenholme J, Clarke P. Missing . . . presumed at random: cost-analysis of incomplete data. Health Econ 2003;12:377–92.
15 Rubin DB. Multiple Imputation for Non-Response in Surveys. New York: Wiley, 1987.
16 Statistical Package for the Social Scientist Version 10.0.5. SPSS Inc. 1989–1999.
17 Schafer JL. Software for Multiple Imputation, 2000. Available from: http://www.stats.psu.edu/~jls/misoftw.html