COST–EFFECTIVENESS IN CLINICAL TRIALS:

USING MULTIPLE IMPUTATION TO DEAL

WITH INCOMPLETE COST DATA

Andrea Burton[1] MSc

Lucinda Jane Billingham[1] PhD

Stirling Bryan[2] PhD

1. Cancer Research UK Clinical Trials Unit, University of Birmingham, Birmingham, United Kingdom

2. Health Economics Facility, University of Birmingham, Birmingham, United Kingdom

Corresponding author and request for reprints from:
Andrea Burton,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick,
Gibbet Hill Road,
Coventry,
CV4 7AL, United Kingdom.
Tel: 02476150176, Fax: 02476574879,
E-mail: Andrea.Burton@warwick.ac.uk

Running title: Cost-effectiveness analysis with incomplete cost data

Abstract

Background: Cost-effectiveness has become an important outcome in many clinical trials and has resulted in the collection of resource use data and the calculation of costs for individual patients. A specific example is a Cancer Research UK phase III trial comparing chemotherapy against standard palliative care in patients with advanced non-small cell lung cancer. Resource usage from trial entry until death were collected and costs obtained on a subset of 115 trial patients. For some patients, however, the unavailability of medical notes resulted in some cost components, and hence total cost, being missing. The 82 patients with complete data were not representative of all trial patients in terms of effectiveness and thus it was necessary to address the missing data problem. **Methods:** Multiple imputation was used to impute values for the unobserved individual cost components, allowing total cost to be calculated and cost-effectiveness carried out for all patients in the cost sub-study. The results are compared with those from a complete case analysis. **Results:** After multiple imputation, the results indicated that chemotherapy had a high probability of being cost-effective for a societal willingness to pay over £20,000 per life-year gained. This was in stark contrast with the complete case analysis, which suggested that chemotherapy was not a cost-effective use of resources at any reasonable level of willingness to pay for a life-year. **Limitations**: Our findings are based on a relatively small retrospective study with all events observed. **Conclusion:** In conclusion, cost-effectiveness analysis of the complete cases only may give biased results, and therefore, in situations where there are missing costs, multiple imputation is recommended.

Keywords: **Cost-effectiveness; Missing cost data; Imputation.**

Introduction

Given the limited resources available for the delivery of health care services, decisions on funding must take into account the cost of an intervention in addition to its effectiveness and tolerability. For this reason, many clinical trials now involve the collection of resource use information on individual patients in order to compare the interventions under study in terms of both overall costs and cost-effectiveness. A specific example of this is discussed in this paper, namely the Cancer Research UK phase III randomised trial comparing the effect of mitomycin, ifosfamide and cisplatin chemotherapy (MIC) against standard palliative care in patients with advanced non-small cell lung cancer (The MIC2 trial, [1]). This trial included a sub-study of costs, which retrospectively examined the patterns and cost of care from randomisation until death for both treatment arms on a patient-by-patient basis in a subgroup of patients [2].

From a decision-maker perspective, the most informative estimate of cost is the mean cost per patient given that total health care spending must be met from a fixed budget. This is because an unbiased estimate of the expected total cost of an intervention is the mean cost of treatment per patient multiplied by the number of patients requiring the treatment. Therefore, our focus for this trial-based analysis was on the difference in mean cost between the two treatment groups and the cost-effectiveness analysis, which linked the difference in mean cost to the difference in mean survival. Patient-level cost data tend to have a highly skewed distribution, but since the mean is the statistic of interest, standard methods for analysing non-normally distributed data such as using medians or applying a

3

transformation are not appropriate [11]. Non-parametric bootstrapping provides one way of constructing confidence intervals for the differences in mean costs and cost-effectiveness [5]. The total cost of an intervention for each patient consists of a number of individual cost components that relate to different aspects of care; for example, hospital costs and General Practitioner (GP) costs. These cost components can be broken down further into the actual resources, such as the number of in-patient days.

Collecting resource use and ultimately cost data on a patient-by-patient basis is often problematic, especially in a retrospective study. Missing information occurs when patient notes are not available (for example, when they are lost or they have been destroyed) or do not provide enough detail to be able to estimate the resources used. This results in some cost components being available but others unknown. Unfortunately, when at least one cost component is missing, the resulting total cost for that patient will also be missing. Analysing only the patients with an observed total cost, as in a complete case analysis, may be biased if they are not a random subset of all study patients [3]. Complete case analysis will always be inefficient, to some degree, as the sample size is reduced and the analysis will ignore any observed cost components in the excluded patients. Imputation provides a means of addressing the problem of missing data by replacing each missing value with a predicted value. With multiple imputation (MI), a set of values, rather than a single value, are drawn from the predictive distribution for each missing item. Thus, MI reflects the uncertainty in the predictions of the missing values whilst preserving the distributions and relationships in the data [4], but MI is a relatively new concept in health economic evaluations. The skewness of the cost distributions, the requirement for an estimate of the

4

mean cost and effectiveness measures and the use of bootstrapping to obtain confidence intervals complicates the application of the standard MI methods that is not experienced with non-economic data [5].

The objective of this paper is to investigate the appropriateness and practicalities of using MI to handle missing cost component data as an alternative to the standard complete case analysis, when one of the aims of the trial is to assess cost-effectiveness.

Patients and Methods

The MIC2 Costings Study

In the MIC2 trial, 351 eligible ambulatory (WHO performance status 0, 1 or 2) patients aged 75 or less with advanced non-small cell lung cancer were randomised between November 1988 and March 1996 to receive either standard palliative care (PC) or MIC chemotherapy plus standard palliative care (MIC+PC) [1]. The mean survival time was 10 months (95% CI 8 – 11) for the patients having MIC+PC compared to 7 months (95% CI 6 – 8) for those having PC only.

All 115 patients randomised into the MIC2 trial by consultants within the South Birmingham Health Authority area were included in the cost sub-study (57 on MIC+PC arm and 58 on PC arm). The baseline characteristics of the cost sub-study patients (median age 62; 73% male; 60% squamous histology; 30% World Health Organisation performance status 2) were comparable to those of the main trial patients (median age 62; 72% male;

56% squamous histology; 28% performance status 2). In addition, the effectiveness of the treatment on survival for patients in the cost sub-study (mean survival 10 months, 95% CI 6-14 on MIC+PC arm versus 7 months, 95% CI 6-9 on PC arm) was representative of that in the main trial (as specified above).

The economic evaluation for this study was from the perspective of the health sector and hence resource use information were extracted from trial entry until death from hospital, GP and hospice notes. The total cost consisted of five cost components. These were: (i) a treatment cost encompassing the cost for chemotherapy (CT) and radiotherapy (RT); (ii) a Queen Elizabeth (QE) hospital cost, where patients were predominately seen and treated (excluding CT and RT costs); (iii) a non-QE hospital cost, which included costs at any alternative hospitals that the patients attended after randomisation; (iv) a community-based GP cost; and (v) a hospice cost. Details of the specific resources associated with each of these components and their assigned unit cost relating to a common price year of 1999, are summarised in Table 1. The cost for each component was calculated as the sum of the amount of each resource used multiplied by its associated unit cost. The total cost for each patient was then calculated as the sum of the costs for the five components.

All patients have complete data for the treatment cost component and the QE cost component. Patients have six different patterns of missing data depending on which of the other three cost components were missing (Table 1). Complete data were achieved for 82 (71%) of the 115 patients in the sub-study. The remaining 33 (29%) patients had at least

6

one cost component missing. Only 60 (10%) of the total 575 components that should have been collected were missing.

The patient and tumour characteristics of the patients who had complete data were comparable to those with incomplete data, but differed in terms of the effect of treatment on survival. For the 82 complete cases in the sub-group, the MIC+PC arm had a slightly shorter survival time (mean 8.1 months, 95% CI 5.6 – 10.6) compared to those on the PC arm (mean 8.3 months, 95% CI 6.2 – 10.4), not reflecting the three month survival benefit for MIC that was seen in both the whole of the cost sub-study and the main trial. Hence the 82 complete cases were not fully represented of the cost sub-study.

Methods for imputing missing cost component data

The imputation procedure uses all the known covariates thought to be associated with the missingness mechanism and cost, together with the interrelationships between the cost components, to help predict the values for the missing data. The incomplete response variables were the three cost components of non-QE hospital cost, GP cost and hospice cost. The observed covariates considered were sex (dichotomous), age (continuous), performance status (ordinal), survival time (continuous; log transformed), place of death (categorical), randomised treatment arm (dichotomous), number of CT cycles (continuous), number of fractions of RT (continuous), time of entry into the trial (continuous). The fully observed QE hospital cost (continuous; log transformed) was also included to preserve its relationships with the incomplete cost components.

There are several possible imputation strategies [3,4,6-8]. Briggs et al [8] provide a comprehensive review of possible imputation approaches for handling incomplete cost data, in the context of estimating mean costs but they do not consider the estimation of cost-effectiveness. The MI method adopted here is based on data augmentation [6]. It is an iterative process that alternates between estimating the parameters for this distribution and using these values to predict the missing values. Once the process stabilises, imputed values for the missing data are randomly obtained from its predictive distribution. Auto correlation function plots, where the iteration number, k, is plotted against the lag-k Pearson's correlation coefficient between the simulated parameter value at any cycle and its value k cycles later, can be used to assess the convergence of this process [6]. The multiple imputations were performed for the incomplete cost components at the patient level using the MI procedure in SAS statistical software version 8 (SAS Institute, SAS Circle, Cary, NC, USA). In this study, five imputations were obtained as this should give an efficiency of 99% compared to using an infinite number of imputations [4]. A total run length of 12500 iterations was used with imputations made after every 2500th iteration to ensure that the imputations were independent [6].

The MI data augmentation procedure used here assumes that the data have a multivariate normal distribution. Suitable transformations were necessary for this assumption to hold. The distributions for the three incomplete cost components were highly skewed and the hospice and non-QE cost components had a semi-continuous distribution with a large proportion of patients having a zero cost (63% and 45% respectively), since they did not use these facilities. A two-stage process was adopted [9]. The hospice cost and non-QE cost

8

were each recoded as two variables: a binary variable, identifying whether the cost was zero or not, and a continuous variable for the actual cost when non-zero, but set to missing for cases with observed zeros. Values for the binary and continuous variable were imputed and then multiplied to give an imputed cost of zero when the binary variable was zero and equal to the value from the imputed continuous variable when the binary variable was one. The new binary variables were treated as normally distributed in the imputation procedures and the imputed values rounded to zero or one using a cut-off value of 0.5 [9]. The continuous variables for the non-zero values for the hospice cost, non-QE cost and the GP cost components were highly skewed. A scaled logit transformation, as suggested in Schafer's Norm program [10], was chosen to give normally distributed and plausible values. The scaling values were chosen such that only positive values were imputed and no values would exceed four standard deviations from the mean cost. After imputation the complete data were transformed back to their original scale prior to any analyses being performed. Each completed dataset was then analysed using standard methods for cost-effectiveness as detailed below.

Methods for cost-effectiveness analysis

A total cost per patient was obtained for the 82 complete cases and for the five sets of 115 cases after MI. For the complete cases, the difference in mean cost between treatments is estimated directly from the data. Non-parametric bootstrapping [5] was used to calculate an accelerated bias-corrected confidence interval (BCa) [5, 12] for the difference in mean cost that reflected its skewness. A total of 2000 bootstrap samples of size 82, equal to the number of complete cases, were obtained, from which 2000 estimates of the difference in

9

mean cost were calculated and an associated BCa confidence interval determined. The data were bootstrapped by treatment arm to ensure that the treatment balance remained constant. The bootstrap analysis was carried out using R statistical software [13].

After MI, the overall difference in mean cost between treatments was calculated using Rubin's rules [4] as the average of estimates from each of the five MI datasets. The standard confidence interval for data from MI, as defined by Rubin [4], is symmetric and therefore was not considered appropriate for skewed cost data, as the analysis required a BCa confidence interval. The ABC algorithm [14] was applied to each imputed dataset to obtain approximations to the BCa confidence interval using the '*abcnon*' function within the 'bootstrap' library in the R statistical software and determine the confidence densities for the full range of cost differences [14]. The five confidence densities were then averaged over all imputations as per Rubin's rules [4] to give the required MI confidence density function from which the required percentiles of the distribution were determined by calculating the area under the combined density curve.

To assess the cost-effectiveness of the treatments the bootstrapped differences in mean cost between treatments were plotted against the associated bootstrapped differences in mean survival on a cost-effectiveness plane. This resulted in 2000 points for the results for complete case analysis and five sets of 2000 estimates after MI. In addition, a net monetary benefit (NMB) was calculated for each of the complete cases as

$$NMB_i (\lambda) = \lambda E_i - C_i,$$

where E_i and C_i are the observed survival and cost for patient i and λ is the societal willingness to pay for a year of life [15]. Using a λ value of £24,000, which is based on a willingness to pay of £2000 per month, estimates for the difference in the mean NMB between treatments were calculated, along with a BCa confidence interval using non-parametric bootstrapping. A cost-effectiveness acceptability curve (CEAC) [16], showing the probability of a treatment being cost-effective for a wide range of λ values (from £1000 to £100,000), was constructed by plotting the proportion of the 2000 bootstrap samples where the estimated mean difference in NMB(λ) was greater than zero, against the corresponding value of λ. Similar procedures were adopted to analyse the MI datasets, such that the difference in mean NMB between treatments was calculated as the average of the estimates from each of the five MI datasets [4] and associated confidence interval as previously described for the difference in mean cost. The proportion of 2000 bootstrapped samples with estimated mean difference in NMB(λ) greater than zero was obtained for each imputed dataset and the average proportion over all MI datasets obtained as per Rubin's rules [4]. A CEAC was then produced by plotting this average proportion against the value of λ

Results

The complete case analysis resulted in a higher mean cost for those patients randomised to MIC+PC of £2804 (95% confidence interval (CI) £1236 to £4290) compared to PC. When MI was used, a smaller difference between treatments in terms of the mean cost of £2384 was seen (95% CI £833 to £3954).

The scatter plots of the mean estimates of incremental cost and effectiveness plotted on the cost-effectiveness plane for both complete case and MI analysis (Figure 1) provided further evidence that the difference in mean costs was positive, that is the patients allocated to the MIC+PC arm of the trial had a higher cost than PC alone. However, visual inspection suggested that the complete cases were not representative of the whole study patients in terms of effectiveness. In the complete case dataset, compared to the MI datasets, a higher proportion of the bootstrap samples had a negative incremental survival estimate, implying that the palliative patients survive longer. Thus, a more positive cost-effectiveness result was implied by the analyses where MI had been applied.

Taking a threshold of willingness to pay of £24,000 per life-year gained and using the complete cases, the mean NMB for chemotherapy was -£3346 (95% CI -£9248 to £3593), suggesting that chemotherapy is not cost-effective for this patient group. Conversely, after MI, the NMB for chemotherapy was £1186 (95% CI -£4463, £9683), which implies that chemotherapy is cost-effective, although the confidence interval includes zero. The CEAC plotted using data derived by MI show that chemotherapy has higher probability of being the more cost-effective treatment option for values of λ above approximately £20,000 per life-year gained (Figure 2). The CEAC for the complete cases was considerably lower, as expected given the earlier results. This suggests a very different policy conclusion: that chemotherapy is not a cost-effective use of resources, at least up to a willingness to pay for life-year of £100,000.

Discussion

Collecting resource data at a patient level is often problematic, especially in retrospective studies due to missing data when patient notes are lost or destroyed. Additional missing information occurs when it is known that resources have been used but there is insufficient detail to estimate the magnitude. Any unrecorded resources will not necessarily be known to be missing if there is no indication in the notes. The analyses presented here are based on the assumption that the data obtained from the patient notes are a complete record of the resources used by the patient.

This paper has investigated the use of MI for handling incomplete cost data in a cost-effectiveness analysis. Methods for multiple imputation were originally developed within the survey context and have rarely been used in health economic evaluations. With the introduction of MI techniques within standard software, the more sophisticated techniques of MI are now becoming accessible. MI is more efficient than complete case analysis, incorporating the 29% patients with partial cost data in the analysis. Inferences from MI properly reflect the uncertainty in the predictions of the missing values, whilst preserving the distributions and relationships in the data. Data augmentation does not require specialist programming as it is available in standard statistical software. It is also advocated as the natural and proper approach to MI [4, 7] and has been shown to perform well against many alternatives [17]. The applied imputation procedure assumes that the incomplete cost data are missing at random [4] such that the missingness of the cost components are associated

only with the observed data, either the observed covariates or effectiveness, but is not related to the unobserved cost data. Unfortunately it is impossible to reliably prove or disapprove this assumption without obtaining the complete data for comparison. The complete cases were similar to the incomplete cases in terms of any observed covariates, but differences were apparent in the effectiveness results. Including the effectiveness and all available covariates in the imputation model makes the missing at random assumption and the conclusions from these results more plausible [6].

Difficulties arose in obtaining a CI for the difference in mean cost and mean NMB that reflected the skewness in these estimates, properly accounted for the missing data uncertainty and had good coverage properties. Rubin's approach [4] would have produced a symmetric confidence interval and simply pooling the bootstrap samples from each imputed dataset and calculating an overall BCa confidence interval would have failed to sufficiently account for the between imputation variability. Efron's approach [14] to combining confidence densities produced unsymmetrical confidence intervals, incorporated sufficient variability and therefore should have good coverage properties.

In this study, with only 10% of the cost components being imputed and a fraction of missing information of 6%, the analysis of the difference in mean costs was dominated by the observed data and, thus, similar results were obtained using both missing data methods. The difference in mean costs between treatments was so large that the loss of power as a result of the smaller sample size for complete cases had no effect on the overall conclusion of the cost analysis. In another study with a smaller treatment effect or a larger proportion

of missing data, imputation may have a greater effect on the cost analysis results. MI, however, had an impact on the results from the cost-effectiveness analysis, although differences were not statistically significantly. The complete case analysis suggested a different policy decision to the one after MI. The differences in the conclusions seen were due to the complete cases being unrepresentative of the whole study patients in terms of effectiveness and hence not satisfying the missing completely at random assumption required for a complete case analysis to provide valid and unbiased results [4]. In contrast, MI enabled the inclusion of all study patients with the assumption that the patients with missing cost data could be explained by the observed data [6].

The MI method of data augmentation considered in this paper is a parametric approach that assumes the data are normally distributed. Inferences based on this MI approach can be robust to some departures from normality [6,7]. However the individual incomplete hospice and non-QE hospital cost components were semi-continuous with a large number of patients having a zero cost, and thus assuming a normal distribution would distort the distributional aspects and relationships within the data [9]. The two-stage approach used here to handle these semi-continuous data provides sensible imputed values that reflect the original distribution, when the overall proportion of zeros is not too extreme [9], as is the case here. However, the two stage process leads to more instability in the imputation model which required a much larger number of iterations within the data augmentation procedure than is generally necessary [9]. If only a few cases had non-zero values, then alternative approaches may be considered using imputation procedures based for example on Tobit selection models [9] or blocked general location models [18]. A scaled logit transformation

[10] was employed in preference to alternative transformations as this enabled values to be imputed within the plausible range for each of the incomplete cost components. Back transforming the data prior to analysis enabled the cost data to be analysed on its original skewed scale. Alternative MI procedures such as MI with chained equations and using predictive mean matching [19] can alleviate the problem of skewness within continuous variables by avoiding relying on the normality assumption. Bayesian methods may also be considered, which can incorporate costs data from either a lognormal or gamma distributions [20].

Individual cost components were imputed at a patient level in preference to the overall total cost per patient. This utilised more of the collected data and hence avoided the wasteful deletion of information when one or more components were incomplete. Imputing the cost components produced imputed values that were "unobtainable", in the sense that some imputed values were not multiples of the unit cost associated with that cost component, but the effect should be balanced on both treatment arms. To avoid this potential problem, imputation undertaken at the actual resource use level instead of the cost estimate for the component would be desirable [8]. This would ensure that only "obtainable" total costs were produced and also enable changes in unit costs to be incorporated without having to repeat the whole imputation process and hence be applicable to the centres that incur different costs. However, implementation may be problematic due to the potentially large number of different resources with missing data, especially when the sample size is relatively small. Also the individual resource use data tends to be less normally distributed than the cost component data with a large proportion being semi-continuous and hence

doubling the number of variables for which modelling is required. Thus further investigation of the imputation procedures at the resource level is required. The imputation of cost components is, in our opinion a good compromise between the imputation of total cost and resource usage.

This paper has considered the use of imputation only in relation to one study and further investigation is needed to explore the generalisability of the findings reported here. In this example, all patients had died so the survival times were known for all patients and all costs had been incurred, i.e. uncensored, however, specific methods for dealing with censored survival and cost data are available [21].This paper demonstrates that extending cost-effectiveness analysis to MI data is reasonably straightforward and allows the known effectiveness data in the larger sample to be included. Further development work looking into the inclusion of all trial patients (that is, to include patients who were not part of the cost sub-study) in a cost-effectiveness analysis could also be considered. In this case, however, using data from 23% of the trial patients to impute costs for the remaining 77% of patients may introduce too much uncertainty into the results than is gained in efficiency from using these extra cases.

Imputation involves making potentially untestable assumptions and therefore it is always preferable to have complete data. Every endeavour should be made to minimize the amount of missing data. With economic evaluations increasingly being undertaken alongside clinical trials, the prospective collection of resource use data should become more routine, reducing the problem of partially missing cost data. When missing cost data occur, MI

provides a realistic and practical solution to enable all patients to be included in the overall cost-effectiveness analysis to produce valid and unbiased results. MI should in general be preferred to complete case analysis.

References

1. Cullen MH, Billingham LJ, Woodroffe CM, et al. Mitomycin, ifosfamide and cisplatin in unresectable non-small-cell lung cancer: effects on survival and quality of life. J Clin Oncol 1999;17:3188-3194.

2. Billingham LJ, Bathers S, Burton A, et al. Patterns and costs of care in advanced Non-small cell lung cancer in a trial of chemotherapy versus supportive care. Lung Cancer 2002;37:219-225.

3. Little RJA, Rubin DB. Statistical analysis with missing data. John Wiley and Sons: New York, 1987.

4. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley and Sons: New York, 1987.

5. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. Stat Med 2000;19: 3219-3236.

6. Schafer JL. Analysis of incomplete multivariate data. Chapman and Hall: London, 1997

7. Schafer JL, Graham JW. Missing data: Our view of the state of the art. Psychol Methods 2002;7(2):147-177.

8. Briggs A, Clark T, Wolstenholme, Clarke P. Missing… Presumed at random: Cost analysis of incomplete data. Health Econ 2003;12(5):377-392.

9. Schafer JL, Olsen MK. Modelling and imputation of semicontinuous survey variables. The Methodology Center, Penn State University, USA, Technical report 1999. Available at **http://methodology.psu.edu/publications/fcsmfinal.pdf.** . Accessed Jan 6 2005.

10. Schafer JL. NORM Version 2.02 for windows: Multiple imputation of incomplete multivariate data under a normal model. Available at **http://www.stat.psu.edu/~jls/misoftwa.html.** Accessed Jan 6 2005.

11. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluations of healthcare interventions. Health Technol Assess 1999;3(2).

12. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Stat Med 2000;19: 1141-1164.

13. R Development Core Team (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. Accessed Jan 6 2006.

14. Efron B. Missing data, imputation and the bootstrap. JASA 1994;89:463-475.

15. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ 2002;11: 415-430.

16. Lothgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Econ 2000;9: 623-630.

17. Scheffer J. Dealing with missing data. Res Lett Inf Math Sci 2002;3:153-160. Available at **http://www.massey.ac.nz/~wwiims/research/letters/volume3number1**. Accessed Jan 6 2006.

18. Javaras KN, Van Dyk DA. Multiple imputation for incomplete data with semicontinuous variables. JASA 2003;98:703-715.

19. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stats Med 1999;18;681-694.

20. Thompson SG, Nixon RM. How sensitive are cost-effectiveness analyses to choice of parametric distributions? Med Decis Making 2005;25(4):416-423.

21. Heitjan DF, Kim CY, Li H. Bayesian estimation of cost effectiveness from censored data. Stats Med 2004;23:1297-1309.

Table 1: Distribution of the completeness of each cost component (with associated resources and unit costs).
Key: ✓ = Complete cost component, ✗ = Incomplete cost component

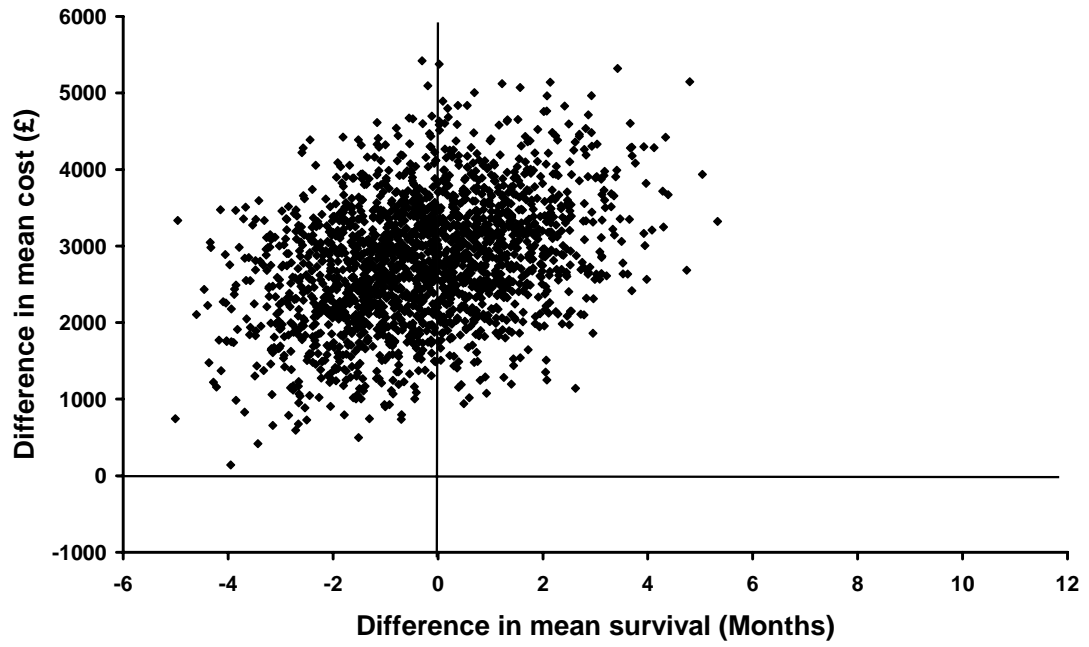| Cost component		Completeness of each Cost component for each pattern of missing data					
Resource	Unit Cost	1	2	3	4	5	6
Treatment		✓	✓	✓	✓	✓	✓
MIC chemotherapy	£310.52 per course						
Palliative radiotherapy	£115 per fraction						
QE hospital		✓	✓	✓	✓	✓	✓
In-patient days	£222 per day (£359 for 1st day of emergency stay)						
Out-patient visit	£63 per visit						
Consultant home visit	£54 per visit						
District nurse home visit	£64 per visit						
Non-QE hospital		✓	✓	✓	✗	✗	✗
In-patient days	£222 per day (£359 for 1st day of emergency stay)						
Out-patient visit	£63 per visit						
Consultant home visit	£54 per visit						
District nurse home visit	£64 per visit						
GP		✓	✓	✗	✓	✗	✗
Surgery visit	£18 per visit						
GP Home visit	£54 per visit						
District nurse home visit	£64 per visit						
Hospice		✓	✗	✓	✓	✓	✗
In-patient days	£250 per day						
Out-patient visit	£94.80 per visit						
Hospice home visit	£64 per visit						
Total Cost		✓	✗	✗	✗	✗	✗
Number of patients (%)		82(71%)	5(4%)	6(5%)	2(2%)	13(11%)	7(6%)

Figure Legends

Figure 1 Scatter plots of the difference in mean cost and effectiveness on a cost-effectiveness plane. a) Complete case analysis. b) Multiple imputation.

Note: Each dot represents the difference in mean cost estimate against the difference in mean survival for each bootstrap sample.

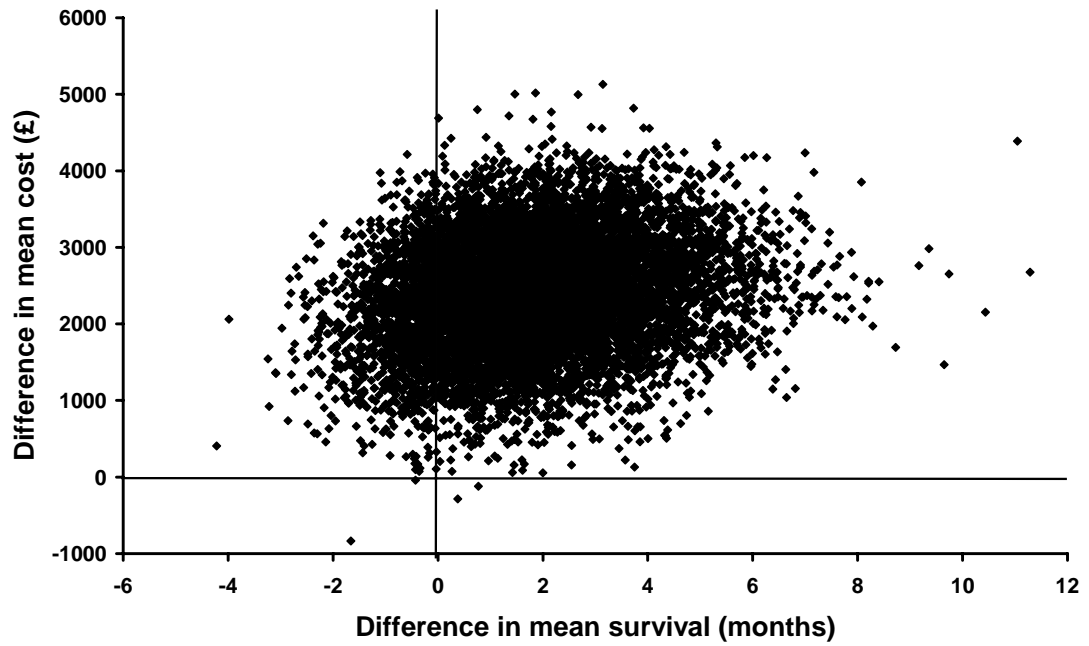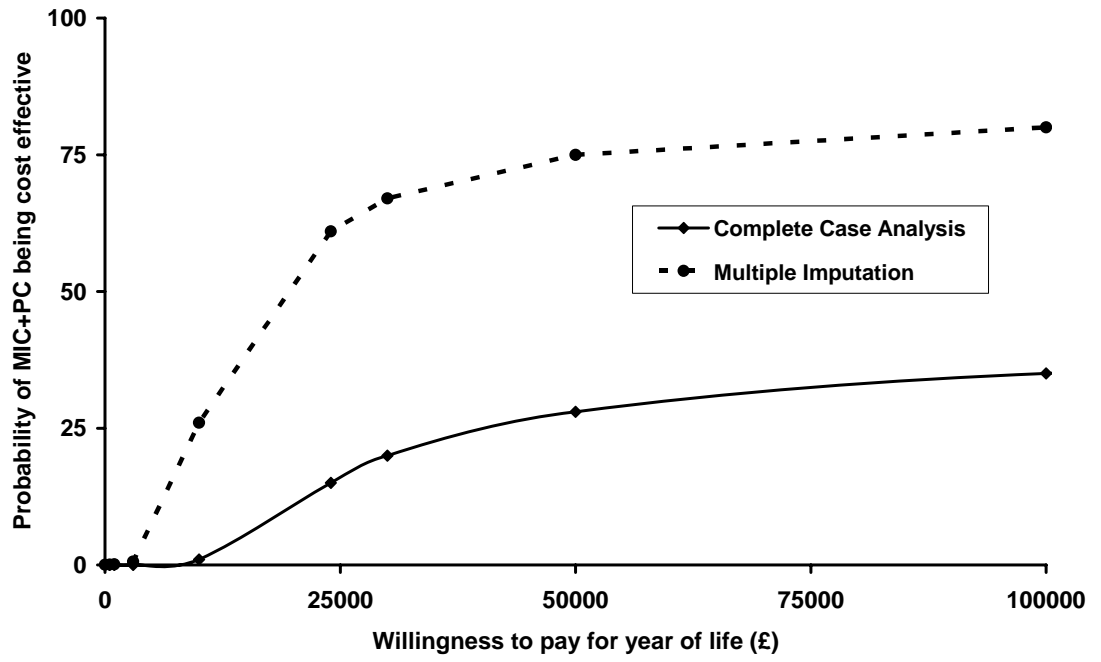**Figure 2 Cost-effectiveness acceptability curves**

Figure 1

(a)

(b)

Figure 2

Acknowledgements