Int J Public Health (2011) 56:63–69
DOI 10.1007/s00038-009-0113-z

ORIGINAL ARTICLE

# Lowering missing item values in quality-of-life questionnaires: an interventional study

**Uwe Müller-Bühl · Bernhard Franke · Katja Hermann · Peter Engeser**

Received: 7 June 2009 / Revised: 17 August 2009 / Accepted: 15 November 2009 / Published online: 22 December 2009


**Abstract**
*Objective* Missing item values (MIV) often occur in quality-of-life (QoL) questionnaires. This study aimed to examine whether the use of introductory exemplary questions reduces the number of MIV and what patient-related factors influence effectiveness of such a QoL form training.
*Methods* In a randomized controlled study in ten primary care practice settings, a total of 215 consecutively recruited patients with at least one chronic disease were requested to complete the Medical Outcomes Study 36 Items Short Form (SF-36) questionnaire, German version 1.0. Prior to filling out the QoL form, a sample of randomly selected patients answered three simple written questions similar in wording and appearance to the original SF-36 questionnaire.
*Results* In total, 126 (58.6%) patients completed the SF-36 questionnaire without MIV. Despite MIV the forms of 46 (21.4%) patients were still computable, i.e., scoring of scales was possible after use of the standardized SF-36 imputation algorithm. After the imputation procedure, MIV significantly hampered generating computable sum scales in 29 (26.6%) of the control group and 14 (13.2%) of the interventional group ($P < 0.05$). A univariate analysis suggested no evidence that the number of MIV was reduced by the intervention. However, intervention led to a significant decrease of MIV in males but not in females. The education status affected the number of missing data independent of intervention.
*Conclusion* This cross-sectional study showed that the prior use of three self-created questions similar in wording and appearance to the SF-36 questionnaire significantly reduces MIV in male patients. School qualification of QoL respondents inversely correlated with the number of questionnaire MIV, but independent of education status all subjects did benefit from the QoL form training.

**Keywords** Quality-of-life · Questionnaire · Missing data · SF-36

U. Müller-Bühl (✉) · B. Franke · K. Hermann · P. Engeser
Department of General Practice and Health Services Research, University Hospital, Vossstr. 2, Geb. 37, 69115 Heidelberg, Germany
e-mail: uwe.mueller-buehl@med.uni-heidelberg.de

## Introduction

Measurement of health-related quality-of-life (QoL) with generic or specific questionnaires is currently an important part of many clinical trials, in many cases even the primary study outcome. Missing data is usually a particularly severe problem because there may be both a loss of statistical power and a bias in estimation (Fayers et al. 1998). Underlying reasons for missing baseline data in a survey can be that participants inadvertently skip questions, ignore certain questionnaire items or choose not to respond. If QoL questionnaire items have been left blank, there are basically several ways in which researchers can handle the fragmentary data records. Firstly, if one is suspicious that missing item value (MIV) is caused by poor compliance of participants, the employment of helpers, such as a study nurse or research assistant, can make data collection more effective. Secondly, when the substitution of missing data is not possible by additional direct patient information, a number of statistical methods exist where a single alternative value is substituted or adjusted for a missing value by simple or multiple imputation (Engels and Diehr 2003; Donders et al. 2006; Molenberghs and Kenward 2007;

Shrive et al. 2006). Imputation involves ascribing a value to a missing data cell based on the values of other variables or of substituting a reasonable estimate for absent data elements (Little and Rubin 1989). This is the most widely practiced approach to substitute MIV. And last, one can ignore missing data completely. Then the scale score for that patient is excluded or treated as missing in the subsequent statistical analyses. This procedure presumes that data is missing completely at random, a rare condition in clinical trials (Donders et al. 2006; Wood et al. 2004).

The most effective way to prevent the problems of missing data is to make an effort to collect accurate and particularly complete questionnaire item data. This study aimed to determine how many questionnaire items had missing values when answered by primary care patients and whether a simple intervention before patients fill out a QoL form can lead to an effective reduction in missing data. Moreover, a question of interest was what patient-related factors influence the number of MIV and which patients benefit from the interventional questionnaire training. The SF-36 health survey questionnaire was selected as the study object because it has widespread use across diverse populations and possesses an imputation algorithm to reduce MIV before scoring. The SF-36 form comprises 36 items in eight subscales assessing different dimensions of self-perceived general health status. It is easy to use, and fulfils stringent criteria of reliability and validity (Bullinger and Kirchberger 1998).

## Patients and methods

### Sample recruitment

In ten German general practices, all patients aged 18–85 years were asked to participate in the study. The recruitment of the general practices ensued from quality circles for general practitioners. Out of a group of 46 GPs who were asked to participate in the study, 10 GPs (22%), 3 females and 7 males, finally took part. The sample represents a diversity of practice size and location (rural, urban). Patient criteria for inclusion in the study were that the patient is known in the practice, diagnosis of at least one chronic disease, adequate knowledge of the German language, close cooperation, patient is orientated to time and place, personal doctor encounter.

### Study design

Each patient personally contacting the general practice on a scheduled day was asked for study participation. After providing his consent, the patient was randomly assigned to either the control group or the interventional group. Following the GPs encounter, each participant had to fill out the SF-36 form personally without the help of any other person. Prior to filling out the original SF-36 questionnaire, individuals of the interventional group under the guidance of a research assistant answered three simple written questions which were similar in wording and appearance to the original SF-36 questionnaire (see "Appendix"). These introductory questions were arbitrary formulated and should simulate the original questions of the SF-36 form as far as possible. The research assistant explained the different structure of the exemplary questions and the possibilities of reply. Particularly, he emphasized the requirement of a complete item marking. Along with the patient he reviewed whether each answer corresponds to the patient's real intention. Subsequently, subjects had to fill out the original SF-36 questionnaire without any further help.

### Analysis of missing item values

First, the number of completed items and the amount of MIV in all questionnaires were determined and the effect of intervention was statistically calculated. Second, after applying the standardized imputation procedure to all questionnaire data according to the manual direction, the number of computable scales was scored. SF-36 scoring is possible if a valid answer is given for at least half of the items in a scale, except for the two-item scales bodily pain and social functioning where both of the respective items need a valid answer (Bullinger and Kirchberger 1998). In all other scales, if more than 50% of the items were left blank, the imputation algorithm identifies the scale as a 'not computable scale'. If the sum scales could not be scored, the SF-36 form was classified as not evaluable. The assessment of health status was made according to the number of chronic diseases declared from the patient chart. The International Classification of Primary Care (ICPC) was used for coding the diseases (Lamberts and Wood 1987). The ICPC classification consists of a biaxial structure. Each chapter of the first axis corresponds to a body system or a problem area. The second axis describes the doctor–patient encounter elements in detail. Only the alphabetic codes of the chapter axis were registered.

In the statistical analysis, only MIV of original form marks were regarded, because some respondents additionally put marks in locations which are not scheduled for marking, for example in heading lines or in spaces between answer categories.

### Statistical analysis

Bivariate comparisons for variables were conducted using Pearson's Chi-square test and Student's *t*-test for

categorical and continuous variables, respectively, in order to demonstrate differences between patients with and without missing data. A univariate ANOVA was employed to identify factors which might moderate the number of missing data. For all analyses, $P < 0.05$ was considered to indicate statistical significance. All statistical procedures were performed with SPSS 15.0 software (SPSS Inc., Chicago, IL, USA).

## Results

A total of 242 subjects were asked to participate in the study. Fifteen (6.2%) patients withholded approval without further explanation. Twelve (5%) patients did not meet the inclusion criteria. All of the 215 patients who agreed to participate filled out the SF-36 questionnaire. Table 1 shows similar patient characteristics of both the control group and the interventional group according to age, gender, education status and distribution of diagnoses.

The result of SF-36 scoring was consistent with those found in German chronic sick patient samples (Bullinger and Kirchberger 1998). According to the demands of Bullinger and Kirchberger (1998) the physical functioning ($r = 0.82$), role limitation-physical ($r = 0.77$) and bodily pain ($r = 0.79$) scales correlated most highly with the physical component sum score, and likewise the mental component sum score correlated most highly with the mental health ($r = 0.87$), role limitation-emotional ($r = 0.75$) and social functioning ($r = 0.67$) scales.

### Effect of intervention on the number of missing item values

The average number of MIV per item ranged between 1.8% (interventional group, physical functioning) and 19.0% (control group, role-emotional; Table 2). On average, a patient in the interventional group had 1.9% less MIV than a patient in the control group.

A univariate analysis was undertaken to identify patient-related factors of data absence. In general, there was only slight evidence that the number of MIV was reduced by the intervention variable ($P < 0.10$; Table 3). The more detailed analysis of patient-related factors affecting MIV suggested that a reduction in MIV depended on patient gender. An interventional effect was found for men, but not for women. Table 3 shows that of all the variables, only the education status affected the number of missing data independent of intervention. The higher the school qualification of respondents, the lower was the number of questionnaire MIV.

### Effect of intervention on number of missing computable scales

Table 4 shows that 126 (58.6%) patients completed the SF-36 questionnaire without MIV. For 46 patients

**Table 1** Characteristics of study patients

| | Control group ($n = 109$) | Interventional group ($n = 106$) | Total ($n = 215$) |
|---|---|---|---|
| Age | | | |
| Mean (range) | 58.5 (18–82) | 56.1 (18–85) | 57.3 (18–85) |
| Gender | | | |
| Males | 55 | 52 | 107 |
| Females | 54 | 54 | 108 |
| School qualification[a] | | | |
| <8 years | 7 | 4 | 11 |
| 8–9 years | 51 | 51 | 103 |
| 10 years | 34 | 33 | 67 |
| 12–13 years | 17 | 17 | 34 |
| Diagnosis chapters[b] | | | |
| Circulatory (K) | 92 | 81 | 173 |
| Digestive (D), endocrine, metabolic (T) | 69 | 78 | 147 |
| Musculo-skeletal (L) | 80 | 59 | 139 |
| General, blood, eye, ear, respiratory, skin (A, B, F, H, R, S) | 57 | 73 | 130 |
| Neurological, psychological (N, P) | 23 | 28 | 51 |
| Urological, genital, pregnancy (U, X, Y, W) | 17 | 15 | 32 |

[a] Since the German school system cannot be directly compared with those in other countries, an approximate classification by the years of schooling is given

[b] According to ICPC classification (see section 'Analysis of missing item values')

**Table 2** Effect of intervention on the number and percent of missing item values in 215 patients filling out SF-36 questionnaire

| Subscales | Control group | | Interventional group | | *P* value |
|---|---|---|---|---|---|
| | *n* | % | *n* | % | |
| All (*n*) | 109 | | 106 | | |
| Physical functioning | 4.3 | 3.9 | 1.8 | 1.7 | n.s. |
| Role limitation due to physical health problem | 15.3 | 13.9 | 7.8 | 7.3 | n.s. |
| Bodily pain | 6 | 5.5 | 4.5 | 4.2 | n.s. |
| General health | 9.8 | 9 | 2.4 | 2.3 | <0.01 |
| Vitality | 12.8 | 11.7 | 3.5 | 3.3 | <0.01 |
| Social functioning | 10.5 | 9.6 | 7 | 6.6 | n.s. |
| Role limitation due to emotional problems | 19 | 17.4 | 10.7 | 10.1 | n.s. |
| Mental health | 13 | 11.9 | 2.8 | 2.6 | <0.01 |
| Health transition | 3 | 2.8 | 1 | 0.9 | n.s. |
| Males (*n*) | 54 | | 54 | | |
| Physical functioning | 2.3 | 4.2 | 0.3 | 0.6 | n.s. |
| Role limitation due to physical health problem | 10.5 | 19.4 | 0.5 | 0.9 | <0.01 |
| Bodily pain | 4 | 7.4 | 0 | 0 | <0.05 |
| General health | 5 | 9.3 | 0.2 | 0.5 | <0.01 |
| Vitality | 7.5 | 13.9 | 0.3 | 0.5 | <0.01 |
| Social functioning | 8 | 14.8 | 2 | 3.8 | <0.05 |
| Role limitation due to emotional problems | 11.7 | 21.6 | 1 | 1.9 | <0.01 |
| Mental health | 5.8 | 10.7 | 0 | 0 | <0.01 |
| Health transition | 1 | 1.9 | 0 | 0 | n.s. |
| Females (*n*) | 55 | | 52 | | |
| Physical functioning | 2 | 3.6 | 1.5 | 2.7 | n.s. |
| Role limitation due to physical health problem | 4.8 | 8.1 | 7.3 | 13.4 | n.s. |
| Bodily pain | 2 | 3.6 | 4.5 | 8.6 | n.s. |
| General health | 4.8 | 7.3 | 2.2 | 4.2 | n.s. |
| Vitality | 5.3 | 9.5 | 3.3 | 6.3 | n.s. |
| Social functioning | 2.5 | 4.5 | 5 | 9.6 | n.s. |
| Role limitation due to emotional problems | 7.3 | 13.3 | 9.7 | 18.5 | n.s. |
| Mental health | 7.2 | 13.0 | 2.8 | 5.4 | n.s. |
| Health transition | 2 | 3.6 | 1 | 1.9 | n.s. |

*Number n* missing/*n* items per subscale

(21.4%), MIV were found, but scales were nonetheless computable, i.e., scoring of scales was possible following the imputation procedure. Overall, questionnaire completion was achieved in the control group in approximately 75% of subjects. In 29 (26.6%) of the control group questionnaires, too many items were incompletely filled out or not at all, not allowing generation of computable data by use of the imputation procedure. With 14 (13.2%) not computable questionnaires, the interventional group showed significantly fewer not computable scales compared to the control group ($n = 29$, 26.6%; $\chi^2 = 8.55$, $P < 0.05$). However, as a result of the intervention, a significant increase in computable scales was only found in males.

## Discussion

Ideally researchers should aim to receive complete QoL data sets at the outset. The approach of our study was to eliminate the individual problems of participants associated with the filling out procedure of the QoL form. The findings of the study suggest that despite the imputation procedure, MIV hindered a complete SF-36 questionnaire analysis in 20% of cases. This percentage of missing data seems to be relatively high compared to other patient trials or population surveys using SF-36. In a UK normative general population sample as well as in adult survivors of childhood cancer, 88% of subjects completed all SF-36 items. The percentage of MIV per item ranged between 0.5

**Table 3** Effect of intervention on the number of missing item values in 215 patients filling out SF-36 questionnaire

| | All | | Control group | | Interventional group | | ANOVA[b] | |
|---|---|---|---|---|---|---|---|---|
| | *n* | M ± SD | *n* | M ± SD | *n* | M ± SD | *F* | *P* |
| Total MIV | 215 | 2.4 ± 5.0 | 109 | 3.3 ± 5.9 | 106 | 1.4 ± 3.5 | 3.37 | 0.07 |
| Age | | | | | | | | |
| ≤70 | 150 | 1.7 ± 4.0 | 76 | 2.7 ± 5.2 | 74 | 0.7 ± 1.7 | 1.91 | 0.17 |
| >70 | 65 | 3.9 ± 6.5 | 33 | 4.8 ± 7.2 | 32 | 3.0 ± 5.6 | | |
| Gender | | | | | | | | |
| Males | 108 | 2.1 ± 4.9 | 54 | 3.9 ± 6.4 | 54 | 0.3 ± 0.7 | 5.53 | 0.02 |
| Females | 107 | 2 7 ± 5.1 | 55 | 2.8 ± 5.4 | 52 | 2.5 ± 4.7 | | |
| School qualification[a] | | | | | | | | |
| <8 years | 11 | 4.9 ± 6.8 | 7 | 7.0 ± 7.8 | 4 | 1.3 ± 1.9 | 0.73 | 0.53 |
| 8–9 years | 103 | 3.5 ± 6.2 | 51 | 4.7 ± 7.2 | 51 | 3.4 ± 5.2 | | |
| 10 years | 67 | 0.7 ± 2.0 | 34 | 1.1 ± 2.8 | 33 | 0.3 ± 0.6 | | |
| 12–13 years | 34 | 1.3 ± 2.8 | 17 | 2.1 ± 3.7 | 17 | 0.4 ± 0.8 | | |
| General health | | | | | | | | |
| ≤3 diagnoses | 171 | 2.1 ± 4.9 | 89 | 3.0 ± 5.9 | 82 | 1.1 ± 3.4 | 0.31 | 0.58 |
| >3 diagnoses | 44 | 3.4 ± 5.1 | 20 | 4.9 ± 6.0 | 24 | 2.2 ± 3.9 | | |

[a] Independent of group, school qualification influences the number of MIV ($F = 5.19$, $P < 0.01$)

[b] Effects reported are interaction effects with intervention

**Table 4** Effect of intervention on number of missing computable scales after standardized imputation procedure of SF-36 questionnaires data ($n = 215$)

| Missing item values (MIV) | All | | | | | | Control group | | | | | | Interventional group | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *n* | % | Females | % | Males | % | *n* | % | Females | % | Males | % | *n* | % | Females | % | Males | % |
| | 215 | 100 | 107 | 49.8 | 108 | 50.2 | 109 | 50.7 | 55 | 50.5 | 54 | 49.5 | 106 | 49.3 | 52 | 49.1 | 54 | 50.9 |
| No MIV | 126 | 58.6 | 57 | 53.3 | 69 | 63.9 | 54 | 49.5 | 29 | 52.7 | 25 | 46.3 | 72 | 67.9 | 28 | 53.8 | 44 | 81.5 |
| MIV, sum scales computable | 46 | 21.4 | 25 | 23.4 | 21 | 19.4 | 26 | 23.9 | 14 | 25.5 | 12 | 22.2 | 20 | 18.9 | 11 | 21.2 | 9 | 16.7 |
| MIV, sum scales not computable | 43 | 20.0 | 25 | 23.4 | 18 | 16.7 | 29 | 26.6 | 12 | 21.8 | 17 | 31.5 | 14 | 13.2 | 13 | 25.0 | 1 | 1.9 |

The proportion of males with not computable sum scales is lower in the interventional group [$\chi^2(2) = 14.01$, $P < 0.01$]; there are no gender differences in the control group [$\chi^2(2) = 1.30$, $P = 0.52$]

and 2.9% (Jenkinson et al. 1993; Reulen et al. 2006). However, there are some reasons which could explain the difference from the result of our study. The studies of Jenkinson et al. (1993) and Reulen et al. (2006) involved postal surveys with a response rate for the questionnaire booklets of 72%. One can suppose that the 28% non-responders through a selection bias differ systematically from the responders, probably providing more missing data (Barclay et al. 2002). Further, the English version of the original SF-36 health survey questionnaire starts with an exemplary question 'How strongly do you agree or disagree with each of the following statements? (a) I enjoy listening to music. (b) I enjoy reading magazines'. The answer is not regarded in the final appraisal process. The German version of the SF-36, which we used, does not comprise a comparable introductory question. Moreover, a considerably higher age of our study participants compared to the English survey population may be a further factor affecting the higher number of MIV (Bullinger et al. 2003). More consistent with the results of our study, in a randomized trial in community psychiatry, the number of MIV in baseline data of a self-reported satisfaction score was between 18.3 and 20.4% (White and Thompson 2005). It is conceivable that in baseline QoL measurements MIV more often occur than is usually reported. Equipment of QoL questionnaire software with the imputation algorithm may improve but not completely solve this problem. Also, reviews of longitudinal studies published in major medical

journals reported that 90% of analyzed QoL studies had missing data which are often inadequately handled in the statistical analysis (Fielding et al. 2008). To make the impact of missing data on patients outcome estimation more transparent, missing data is divided into 'missing data completely at random', 'missing data not at random' and 'missing data at random' (Rubin 1976). This classification is important for interpretation of outcome data and helps to avoid biased results, but it provides no information about reasons of absence.

In our study, the introduction of exemplary questions prior to the filling out procedure resulted in a significant increase in computable questionnaire scales. However, this effect depended on sociodemographic patient factors. As consistent with a German QoL study using the SF-12 questionnaire, low patient education status was a predictor of missing data (Morfeld et al. 2003). It is of interest that all subjects did benefit from the introductory questionnaire training independent of their level of school qualification. Obviously subjects with low education status profit more than higher qualified respondents by employment of introductory exemplary QoL questions. In contrast, in a study of Bullinger et al. (2003) a higher education status was correlated with a higher proportion of MIV, but this result was not evaluated by an adequate univariate analysis.

Gender was a strong differentiating feature predicting absent data cells. In females, contrary to males, no positive response of interventional training could be registered. There is no reasonable explanation for that gender difference. A more deficient learning effect of females is implausible. Bullinger et al. (2003) described a similar gender-dependent difference with significantly more MIV in females than in males. On the other hand, the findings of a longitudinal postal respiratory questionnaire survey from Barclay et al. (2002) showed that QoL questionnaire non-responders were significantly younger and more likely to be male. However, patient characteristics of this survey seem to be too different from our study sample and therefore difficult to compare, for example patients were 10 years younger than in our study.

We do not believe that the gender difference suggested in our study is a random artifact. But it might be a limitation of this study that too small a sample of participants was chosen. Although a power calculation based on 10% MIV in a pilot study was performed, there is no reasonable explanation at hand to explain why the intervention worked better in men. Probably a larger study sample would allow a more detailed view into the many influences on filling out a questionnaire completely. A further limitation might be that in the practices patients were recruited consecutively to the study before they were assigned at random to either the intervention or control group. Approximate 6% eligible patients were not interested in participating in the study and refused their approval. But in our opinion this relatively low percentage of non-respondents is part of most clinical trials and does not bias the study results.

According to Rubin (1976) we believe that the missing data of our study is 'missing at random', because the reasons for absence are associated with known patient characteristics. Taking this into account, we think that the results of our cross-sectional study are plausible and essential for patients dealing with a QoL questionnaire. In longitudinal studies the reasons for missing QoL data may be more different and not only due to structure or comprehension of QoL instruments. Further, the reiterated filling out of QoL forms may involve certain training effects and adaptation of patients to questionnaires as well.

The SF-36 scoring algorithm contains imputation strategies for missing data and thus enables estimating scores for respondents with incomplete data (Ware and Sherbourne 1992). Our study did not aim to improve effectiveness and validity of this proven useful QoL instrument. However, it is conceivable that the procedure we used possesses similar effects in less convenient QoL instruments than SF-36 with a more complex structure and less practical comfort. Moreover, since the occurrence of MIV in short QoL instruments may potentially provide more biased results, one or more introductory exemplary questions could possibly be especially effective in these instruments.

In conclusion, the results presented here suggest that—in patient baseline data records—completeness of QoL forms might be improved by the implementation of simple introductory questions. However, regarding a gender difference on filling out a questionnaire completely, the power calculation of future studies should comprise an adequately large sample size. The future German version and SF-36 questionnaires translated into other languages should not abandon at least one exemplary question at the beginning of the SF-36. The efficacy of three opening questions was tested in our study in only one QoL instrument. Whether this procedure, however, is similarly effective and also associated with reduction of MIV in other QoL questionnaires and different patient samples, remains to be clarified.

## Appendix

Table 5

**Table 5** Introductory questions which 106 randomly selected subjects had to answer prior to filling out original SF-36 form

| | Excellent | Very good | Good | Fair | Poor |
|---|---|---|---|---|---|
| In general, would you say your appetite is | 1 | 2 | 3 | 4 | 5 |
| | All of the time | Most of the time | Some of the time | A little of the time | None of the time |
| How much skin pruritus have you had during the past 4 weeks? | 1 | 2 | 3 | 4 | 5 |
| | Definitely true | Mostly true | Don't know | Mostly false | Definitely false |
| How did you sleep during last week? | | | | | |
| I was dead to the world | 1 | 2 | 3 | 4 | 5 |
| I was often wide awake | 1 | 2 | 3 | 4 | 5 |
| I often awoke from bad dreams | 1 | 2 | 3 | 4 | 5 |

## References


Barclay S, Todd C, Finlay I, Grande G, Wyatt P (2002) Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaires studies of GPs. Fam Pract 19:105–111

Bullinger M, Kirchberger I (1998) German manual SF-36 questionnaire. Hofgrefe, Göttingen

Bullinger M, Morfeld M, Kohlmann T et al (2003) SF-36 health survey in rehabilitation research. Findings from the North German Network for Rehabilitation Research within the Rehabilitation Research Funding Program. Rehabilitation 42:218–225

Donders AR, van der Heijden GJ, Stijnen T, Moons KG (2006) Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 59:1087–1091

Engels JM, Diehr P (2003) Imputation of missing longitudinal data: a comparison of methods. J Clin Epidemiol 56:968–976

Fayers PM, Curran D, Machin D (1998) Incomplete quality of life data in randomized trials: missing items. Stat Med 17:679–696

Fielding S, Maclennan G, Cook JA, Ramsay CR (2008) A review of RCTs in four medical journals to assess the use of imputation to overcome missing data in quality of life outcomes. Trials 9:51. doi:10.1186/1745-6215-9-51

Jenkinson C, Coulter A, Wright L (1993) Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. Br Med J 306:1437–1440

Lamberts H, Wood M (1987) ICPC, International classification of primary care. Oxford University Press, Oxford

Little RJ, Rubin DB (1989) The analysis of social science data with missing values. Soc Methods Res 18:292–326

Molenberghs G, Kenward MG (2007) Missing data in clinical studies. Wiley, New York

Morfeld M, Dietsche S, Bürger W, Koch U (2003) Der SF-12—Das Problem der Missing Data. Diagnostica 49:129–135

Reulen RC, Zeegers MP, Jenkinson C, Lancashire ER, Winter DL, Jenney ME, Hawkins MM (2006) The use of SF-36 questionnaire in adult survivors of childhood cancer: evaluation of data quality, score reliability, and scaling assumptions. Health Qual Life Outcomes 4:77. doi:10.1186/1477-7525-4-77

Rubin DB (1976) Interference and missing data. Biometrika 72:359–364

Shrive FM, Stuart H, Quan H, Ghali WA (2006) Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. BMC Med Res Methodol 6:57

Ware J, Sherbourne C (1992) The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. Med Care 30:473–483

White IR, Thompson SG (2005) Adjusting for partially missing baseline measurements in randomized trials. Stat Med 24:993–1007

Wood AM, White IR, Thompson SG (2004) Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials 1:368–376