How can I deal with missing data in my study?

Abstract

Missing data in medical research is a common problem that has long been recognised by statisticians and medical researchers alike. In general, if the effect of missing data is not taken into account the results of the statistical analyses will be biased and the amount of variability in the data will not be correctly estimated. There are three main types of missing data pattern: Missing Completely At Random (MCAR), Missing At Random (MAR) and Not Missing At Random (NMAR). The type of missing data that a researcher has in their dataset determines the appropriate method to use in handling the missing data before a formal statistical analysis begins. The aim of this practice note is to describe these patterns of missing data and how they can occur, as well describing the methods of handling them. Simple and more complex methods are described, including the advantages and disadvantages of each method as well as their availability in routine software. It is good practice to perform a sensitivity analysis employing different missing data techniques in order to assess the robustness of the conclusions drawn from each approach.

(*Aust N Z J Public Health* 2001; 25: 464-9)

Derrick A. Bennett

Clinical Trials Research Unit, Department of Medicine, University of Auckland, New Zealand

Missing data in medical research is a common problem that has long been recognised by statisticians and medical researchers alike.[1,2] There are different types of 'missingness' that can occur and this may influence how the researcher should analyse the data that they have collected. When the amount of missing data are large (greater than 10%) the results of subsequent statistical analyses may be biased. The purpose of this note is to outline the types of missing data and the reasons why these may occur in medical studies. The methods available for handling the incomplete datasets will then be described.

Types of missing data

There are three types of missing data that can occur when the data are being collected. As the assumptions that can be made as to why the data are missing can affect the underlying assumptions of the statistical modelling techniques that will be employed,[3] it is important to distinguish between the types of missing data. However, it is not possible to perform a formal test to check whether the assumptions made are valid.

Missing completely at random (MCAR): The participants with complete data cannot be distinguished from participants with incomplete data. In practical terms this assumption would be equivalent to a researcher going through raw data forms and randomly discarding some of the forms. When data are MCAR, the missing values can be thought of as a random sub-sample of the actual values. In practice, it is usually difficult to meet the MCAR assumption.

Missing at random (MAR): The participants with incomplete data differ from participants with complete data, but the pattern of 'missingness' is traceable or predictable from other variables in the dataset, rather than being due to the specific variable on which the data are missing. Inferences about population parameters of interest in statistical modelling (e.g. treatment effect, odds ratio, or relative risk) do not depend on the missing data mechanism if the data are MAR. For this reason this type of missing data mechanism is sometimes called 'ignorable' because it can be ignored in the analysis process. The more relevant and related predictors (e.g. exposures or risk factors) one can include in the statistical modelling, the more likely the MAR assumption will be met.

Not missing at random (NMAR): The pattern of 'missingness' is non-random and it is not predictable from other variables in the dataset. In simpler terms. the probability that a response is missing is directly related to data that were collected or requested. This is important because inferences about treatment effects, odds ratios or relative risks in medical studies will depend not just on a model for the data, but will also depend on a model for the process that gave rise to the missing data. This is sometimes referred to as 'non-ignorable' because this cannot be ignored in the modelling process.

The differences in the three mechanisms can be illustrated in a clinical trial setting using the Australia and New Zealand Heart Failure (ANZHF) Trial.[4] This was a

Submitted: October 1999
Revision requested: November 2000
Accepted: May 2001

Correspondence to:
Derrick Bennett, Clinical Trials Research Unit, Department of Medicine, University of Auckland, Private Bag 92019, Auckland, New Zealand. Fax: +64 9 373 1710; e-mail: d.bennett@ctru.auckland.ac.nz

randomised, double-blind placebo-controlled trial of a vasodilating beta-blocker, carvedilol, in clinically stable patients with congestive heart failure and a history of coronary heart disease. Duration of follow-up was 18 months and the primary objective of the study was to determine the effects of treatment on exercise capacity, left ventricular function and left ventricular size. Data would be MCAR if the participants missed their visits totally at random. However, if the probability of missing a visit was directly related to prior observed responses the data would be MAR. If a participant taking active treatment was more likely to miss a visit than a participant taking placebo (which could be due to unforeseen side effects that the researchers were unaware of) the data would be NMAR.

In general, if the effect of missing data is not taken into account the results of the statistical analyses will be biased and the amount of variability in the data will not be correctly estimated. It would be useful if researchers, if at all possible, record the circumstances and reasons why data are missing. There are several methods that can be used to handle missing data that are based on the assumptions of type and degree of 'missingness' in the dataset. These methods can be split into five broad categories:

1. Methods that ignore missing observations.
2. Single imputation methods.
3. Other imputation methods.
4. Likelihood-based methods.
5. Indicator methods.

These methods, as well as their advantages and disadvantages, will now be described.

1. Methods that ignore missing observations

Complete case analysis: This is the default method used in most statistical packages such as SAS and SPSS. All the patients with incomplete data are removed from the analysis. For example, a researcher has collected data on important covariates such as age, sex, cholesterol level, systolic blood pressure and diastolic blood pressure for 100 individuals. For some participants (say 30), data has not been recorded, or was unavailable for some reason, on at least one of these covariates. For an analysis involving all these covariates only 70 complete cases would be used in the analysis. If the data are MCAR the reduced dataset would represent a randomly drawn sub-sample of the original data and thus inferences made from this approach are valid. However, in practice the data are not MCAR and participants with complete data may be a biased sub-sample of all participants. As a consequence, a complete case analysis would usually produce biased results. In addition, there is a loss of power due to analysing a smaller dataset.

Available case analysis: This method uses the largest set of available cases for estimating the parameter of interest (e.g. treatment effect). In standard statistical software such as SPSS or SAS this approach is adopted for the computation of *t*-tests or Pearson's correlation. For example, suppose a researcher has collected data on systolic blood pressure (SBP) and body mass index (BMI) at baseline for 30 subjects and then at a subsequent follow visit. All baseline data is available but there is missing data at the follow-up visit, say 10 missing records for BMI and four missing records for SBP. If the researcher had wanted to compare changes in BMI and SBP from baseline to the end of using a paired *t*-test, then the computation of the test statistic would only be for those subjects who had data available at both baseline and follow-up. So for BMI the *t*-statistic would be based on 20 records and for SBP the *t*-statistic would be based on 26 records. This approach is reasonable if we have cross-sectional data. However, if the data are longitudinal, with many time points, different participants will contribute to the data at different time points depending on the pattern of 'missingness'. If the data are not MCAR then there is a lack of comparability over time points, which would lead to highly biased results.

2. Single imputation methods

The imputation methods that will be described here have had their origins in survey research.[3] In broad terms, these methods attempt to estimate the values of the missing data and 'fill-in' or impute new values. Once this has been achieved the medical researcher can then proceed as if the dataset were 'complete'. The most common methods of imputation will now be described.

Last value carried forward: This method is used when we have longitudinal data and data are found to be missing at a particular 'visit' or point in time for some participants. The researcher would then carry the last available value forward (from the last visit or time point) and impute this value for the missing values. The main problem with this method is that the researcher is assuming that there will be no change from one visit to the next. If the data are MCAR then this is reasonable. However, if this is not the case then the results can be seriously biased.

Mean substitution: If we have data available for some participants on a particular variable such as age, but not for other participants, the researcher would impute the mean value from the available participants to fill in the missing data values. If we were considering age in the previous example we would impute the average age of our available participants to the missing ages in the dataset. This approach assumes the data is MCAR but is not recommended as it can lead to under-estimates of the variance.

Regression methods: This approach involves developing a regression equation based on the complete subject data for a given variable, treating it as an 'outcome' and using all other relevant variables as predictors. For participants where the 'outcome' is missing, the predicted values from the regression equation are used as replacements. This method has similar problems to the mean substitution method but these can be overcome by adding uncertainty, usually by weighting, to the imputation of 'outcome' so that the mean value is not always imputed. This method assumes that the data are MAR. However, the weights and standard errors can become extremely complex if the data and incomplete data patterns are not simple.

Hot-deck imputation: This method involves replacing missing values with values taken from respondents with matching

covariates. The process identifies participants that are similar in terms of data observed. If we return to the ANZHF trial example, as well as the main outcomes of interest, blood pressure, heart rate, and cholesterol were measured. If we wish to impute missing blood pressure values using hot-deck imputation we could select a value to impute from patients with matching covariates (e.g. treatment, sex, age-group, heart rate and cholesterol). Hot-deck is useful as it is relatively simple and maintains the proper measurement levels of the recorded covariates. It is usually less biased than mean imputation or complete case approaches and assumes that the missing data are MAR. A disadvantage is that complex matching algorithms sometimes need to be employed in order to match respondents.

Cold-deck imputation: This method is very similar to hot-deck imputation in its approach except that the strategy for assessing subject similarity is based on external information or prior knowledge rather than the information available in the current dataset. Continuing with our example of missing blood pressure values, a researcher may choose to impute the missing blood pressure value based on their knowledge of previous research that looked at similar variables to estimate an individual's blood pressure. The researcher may then base the imputation strategy on participants of the same gender with similar cholesterol, similar treatment, and similar heart rate. An obvious disadvantage is that this method is dependent on the quality of the available external information.

The disadvantages of all the single imputation methods described above are that they impute the same missing value every time. Hence, a statistical analysis, which treats imputed values just the same as observed values, will systematically underestimate the variance, even assuming that the precise reasons for non-response are known. Single imputation cannot represent any additional variation that arises when the reasons for non-response are not known.

3. Other imputation methods

As single imputation methods lead to an under-estimate of the variability in the dataset, other methods have been developed to combat this problem. These methods will now be described.

***Multiple imputation*:** Multiple imputation replaces each missing value by M possible values to create M complete datasets. Typically M is between 5 and 10. The investigator then uses these 'new' datasets in the analysis and combines the results into a single summary finding. This summary dataset reflects the extra variation due to the missing values. This method works well on both cross-sectional and longitudinal data and is robust to violations of non-normality of the variables used in the analysis. A summary of multiple imputation can be found in Little and Rubin[3] and the technique is discussed in detail in Rubin.[5] Schafer,[6] in a more recent review, gives an excellent summary of this method with the answers to the most frequently asked questions by medical researchers.

Markov-chain imputation: This procedure is usually applied to longitudinal (or repeated measures) data. This approach attempts to incorporate 'transient' states (such as disease progression) into the missing value mechanism. If a researcher knows a participant's response (or state of disease progression) at a previous time point, it is possible to estimate the probability of obtaining a particular response at the next time point based on the previous reading. A recent trend in clinical trials methodology is to collect data on Health Related Quality of Life (HRQoL). A common measure that is used is the Short Form-36 (SF-36)[7] questionnaire, which measures eight health domains. The questionnaire is usually completed at repeated intervals and the scores on each of the eight domains are recorded. The PF domain measures the level of physical functioning. We shall use this domain to illustrate how the Markov Chain imputation works in practice. If we call a good PF (state 1), a poor PF (state 2), progression (state 3) and death (state 4), and we know how many assessments will be made, it is possible to produce a matrix of transitions between states as shown in Table 1. We can see that of the participants who had a good PF score at the baseline assessment 76 per cent still had a good PF score (state 1) at visit 1. Eleven per cent of participants had moved from a good PF score (state 1) at baseline to a poor PF score (state 2) at visit 1. Thirteen per cent had moved from a good PF score (state 1) at baseline to disease progression (state 3) at visit 1. No participants had moved from a good PF Score (state 1) at baseline to death (state 4) at visit 1. Similar transition matrices could be generated for each assessment visit, for each of the treatment groups under consideration. If at a subsequent assessment missing values occur it is possible to take account of the state the patient was in at the previous assessment in order to determine how to impute the missing value. Using prior knowledge, we can make the assumption that a participant can move in either direction between state 1 and state 2. However, a participant cannot return to state 2 from either state 3 or state 4. A sequence of

Table 1: Transitions between health states from baseline to first follow-up visit.[a]

	State at visit 1 Good PF (state 1)[b]	State at visit 1 Bad PF (state 2)	State at visit 1 Progression (state 3)	State at visit 1 Death (state 4)
Baseline visit Good PF[c] (state 1)	76	11	13	0
Baseline visit Bad PF (state 2)	11	75	14	0
Baseline visit Progression (state 3)			91	9
Baseline visit Death (state 4)				100

Notes:
(a) Adapted from Curran et al.[22]
(b) Values are percentages.
(c) PF = physical functioning as determined by the SF-36.

responses for one participant's visits is observed as [1, 1, 2, 1, -, -, 2, 2], where each number represents a particular disease state and a dash represents a missing value. It is possible to impute the missing values using Markov Chain techniques. Taking into account that the participant was in state 2 at the next observed value we know that we must impute a 1 or a 2 for the two missing values as a participant cannot return to state 2 from progression or death. The missing values would be imputed as follows:

a. Possible transitions for state 1 are 1→1 (76%), 1→2 (11%) gives a total of 87%.
b. Adjust percentages so that they add up to a 100% so 1→1(76/87=87%), 1→2 (11/87=13%).
c. Using a random number generator in a statistical package (such as SAS) a number can be selected randomly from a uniform distribution from 0 to 100.
d. If the number selected is less than 87 we impute a 1, otherwise we impute a 2.

It is possible to improve the imputation by using the next value after the missing observation as well as the previous value outlined above. More complex versions of this approach are utilised in the BUGS software.[8]

4. Likelihood-based methods

These methods are more robust than the imputation methods described as they have good statistical properties. The most common methods employed are:

Expectation-Maximisation (E-M) approach: This method uses the fact that the missing data contain relevant information to be used in the estimation of the parameter of interest (e.g. treatment effect, odds ratio). In addition, the estimate of the parameter also helps in finding likely values of the missing data. The E-M algorithm is an iterative procedure, which aims to estimate the missing values and consists of two steps in each iteration, the Expectation step (E-step) and the Maximisation step (M-step). In the E-step the distribution of the missing values based on the known values for the observed data and the current estimate of the parameters is found. This is, in effect, the algorithm's 'best guess' on what to impute for the missing data based upon the model specified and the existing data points. In M-step it substitutes the expected values (typically means and covariances) for the missing data obtained from the E-step and then maximises the likelihood function as if no data were missing to obtain new parameter estimates. The new parameter estimates are substituted back into the E-step and a new M-step is performed. The procedure iterates through these two steps until convergence is obtained. Convergence occurs when the change of the parameter estimates from iteration to iteration becomes negligible. This method assumes that data are MAR rather than MCAR. The parameter estimates based on the E-M algorithm are reliable, as are the standard errors after some adjustment.[3] This method is routinely available in standard statistical packages such as SAS and SPSS but it can be quite slow in achieving convergence of the parameter estimates if there is a large amount of missing data. A more detailed account of this method is given by Schafer[9] and Little and Rubin.[3]

Raw Maximum Likelihood methods: This method uses all of the available information about the observed data, including means and variances for each available covariate to generate estimates of the missing values using maximum-likelihood. Raw maximum likelihood methods are also model-based in that they are implemented as part of a fitted statistical model. Raw maximum likelihood method only produces variances and means for the covariates that have been measured and the statistical package then uses these as imputes for further analyses. This approach is similar to the E-M approach, except that raw maximum likelihood has no E-step and typically converges faster. This method is available in SAS and SPSS and is also found in structural equation modelling programs such as AMOS[10] and LISREL.[11] If the data are MAR this method can generate means and standard errors that are less biased than the available case, complete case, and single imputation approaches.[12] This method has also been shown to outperform the simpler methods even if the data has non-ignorable 'missingness'.[12]

5. Indicator methods

This final approach to handling missing data attempts to incorporate the missing data patterns into the data analysis. The two most common methods are:

Indicator method: When only a single covariate has a missing value the indicator method involves creating an additional 'missing' category for this covariate. This method is popular among epidemiologists,[13] however Greenland[13] states that this method may lead to appreciable biases if the covariate is an important confounder of the effect of interest. If there are several covariates with missing values the indictor method is modified so that for each variable with missing values, the researcher creates a missing value indicator to accompany the variable in all analyses. This missing value indicator takes the value 1 wherever the original variable is missing, and 0 otherwise. For example, if we have data for age, then our indicator variable M_1=0 if we have age available and M_1=1 if age is missing. In a regression analysis the researcher would include the missing indicator M_1 and the product of age and one minus the missing indicator (i.e. age*(1-M_1)). This method is also biased under most conditions (e.g. linear regression), and the amount of bias is based on whether the variable is a primary study covariate or a confounder.[14] An additional problem with this method is that if there are a large number of covariates with missing values then a large number of corresponding indicator variables need to be created. If the sample size is small there will be an increased risk of selecting unimportant covariates and failing to include important ones. Another important consideration is the number of events or subjects per variable (EPV) included in the statistical model. The EPV is the ratio of the number of subjects in the study to the number of covariates that are to be included in the model. Harrell et al.[15] concluded that for regression modelling the EPV should be at least 10 times the number of potential covariates that could be included in the model. If the

sample size is small, and there are many covariates with missing data, the construction of extra indicator variables to include in the model may violate this criterion. Peduzzi et al[16] have shown that the model estimates from studies with an EPV of less than 10 are unreliable.

Pattern-mixture models: Pattern mixture models do not need to specify the precise form of the missing data process (i.e. they do not make assumptions that the data are MCAR, MAR, or NMAR prior to statistical analysis). These types of models stratify the population by the missing data patterns, implying a model for the whole population that is a mixture of the different patterns. If we consider a simple example of a longitudinal study in which we have data at several time points it is often found that the number of missing observations increase with time. It is possible to have 1,2,..,k patterns of missing data, where k is usually equal to the number of time points under consideration. In order to deal with this the researcher can create a set of indicator variables I_{ik} which take the value one if the missing data are of pattern 1,2,..,k and zero otherwise. An example of such a missing data pattern can be illustrated for a longitudinal study with three time points. The pattern (0,0,0) represents that there is no data missing at each of the three time points, and the pattern (1, 0, 0) represents that the data are missing for a particular participant at time point 2 but data are available at time points 1 and 3 for that participant. The missing data patterns, as characterised by the indicator variables $I_{ik,}$ can be incorporated into the statistical model of interest. The investigator can then determine if the missing data pattern has predictive power in the model by fitting the pattern mixture indicators along with the covariates of interest and all interactions between the pattern mixture indicators and the covariates of interest.

Conclusions

The missing data methods described in this article deal with incomplete data from the perspective of reduction of bias in terms of estimation of parameters of interest (e.g. treatment effects, odds ratios etc). The methods are not primarily concerned with the prediction of values for specific individuals in the dataset. A medical researcher with statistical knowledge can easily implement some of the imputation methods outlined in this article. Generally, any method of analysis that can handle unequally spaced and different numbers of observations on each participant can handle MCAR data. Likelihood-based longitudinal methods can

Table 2: Summary of methods available.

Method	Bias	Handling of variability	Availability
Complete case (CC)	Unbiased if MCAR otherwise highly biased	Under-estimates variability in the dataset	SAS, SPSS, BMDP
Available case (AC)	Unbiased if MCAR otherwise highly biased	Under-estimates variability in the dataset	SAS,SPSS, BMDP
Mean imputation (MI)	Unbiased if MCAR otherwise highly biased	Under-estimation of the variance but less than CC and AC	SAS,SPSS, BMDP
Last value carried forward (LVCF)	Unbiased if MCAR otherwise highly biased	Under-estimation of the variance but less than MI	User-defined
Regression methods (RM)	Unbiased if MCAR or MAR otherwise highly biased	Under-estimation of variance but can be overcome by using weights	SOLAS or user defined
Hot-deck imputation	Unbiased if MCAR or MAR otherwise highly biased	Under-estimation of variance but less than MI, LVCF and RM	SOLAS or user defined
Cold-deck imputation	Unbiased if MCAR or MAR otherwise highly biased	Under-estimation of variance but depends on quality of external information	User defined
Multiple imputation	Unbiased if MCAR or MAR otherwise highly biased	Produces good estimates of variability in the dataset	SOLAS, SAS macros available
Markov-Chain imputation	Unbiased under MCAR, MAR, and NMAR conditions	Produces good estimates of the variability in the dataset	BUGS
E-M Algorithm	Unbiased if MCAR or MAR otherwise highly biased	Produces good estimates of the variability in the dataset	SAS, SPSS, BMDP
Raw Maximum Likelihood	Unbiased if MCAR or MAR otherwise highly biased. Depends on model fitted being specified correctly	Produces accurate estimates of variability in the dataset	Some SAS procedures, AMOS, LISREL
Indicator method imputation	Unbiased if MCAR or MAR otherwise highly biased. Bias may vary depending on whether the missing data occurs for a covariate or a confounder	Does under-estimate the variability in the dataset. Under-estimation of similar magnitude to CC approach	User defined
Pattern mixture models	Unbiased under MCAR, MAR and NMAR	Produces good estimates of the variability in the dataset	User defined

easily accommodate data that is MAR, but are biased for data that is NMAR. A useful review is given by Cnaan et al.[17] of longitudinal data analysis methods for clinical trials and epidemiological studies. Statistical software packages are now able to deal with some of these more complex methods of imputation described in this article such as the SAS procedure MIXED,[18] BMDP procedure 5V[19] and the SPSS Missing Values Analysis procedure.[20] Specialist software such as SOLAS can be used to implement both single imputation and multiple imputation.[21] This software can deal with data in several formats including SAS, S-Plus and SPSS datasets. In conclusion, the researcher with missing values in their study has several options when deciding how to deal with this common problem. It may also be useful to perform a sensitivity analysis using different approaches to handling the missing data in order to assess the robustness of the results.

References

1. Woodward M, Smith WC, Tunstall-Pedoe H. Bias from missing values: Sex differences in implication of failed venepuncture for the Scottish Health Study. *Int J Epidemiol* 1991;20:379-83.
2. Bernhard J, Celia DF, Coates AS, et al. Missing quality of life data in cancer clinical trials: Serious problems and challenges. *Stat Med* 1998;17:517-32.
3. Little RJ, Rubin DB. *Statistical Analysis with Missing Data*. New York: Wiley, 1987.
4. Australia and New Zealand Heart Failure Collaborative Group. Randomised, placebo-controlled trial of carvedilol in patients with congestive heart failure due to ischaemic heart disease. *Lancet* 1997;349:375-80.
5. Rubin DB. Multiple imputation after 18+years. *JAm Stat Assoc* 1996;91:473-89.
6. Schafer J. Multiple imputation: A primer. *Stat Methods Med Res* 1999;8: 3-15.
7. Ware JE, Snow KK, Kosinki M, Gandek B. *SF-36 Health Survey: Manual and Interpretation Guide*. Boston(Mass): New England Medical Center, Health Institute, 1993.
8. Spiegelhalter D, Thomas A, Best NG, Gilks WR. BUGS: Bayesian Inference Using Gibbs Sampling. Cambridge(UK), 1994.
9. Schafer JL. *Analysis of Incomplete Data*. London: Chapman and Hall, 1997.
10. Arbuckle JL. *Amos*: analysis of moment structures [computer program]. Version 3.5 for Windows. Chicago: SmallWaters, 1995.
11. Joreskog KG, Sorbom D. *Linear Structural Equations LISREL 7*: A guide to the program and applications. Chicago: SPSS, 1989.
12. Wothke W. Longitudinal and group modelling with missing data. In: Little TD, Schabel KU, Baumert J, editors. *Modelling Longitudinal and Multilevel Data: Practical Issues, Applied Approaches and Specific Examples*. Mahwah(New Jersey): Laurence Erlbaum, 1998.
13. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiological regression analyses. *Am J Epidemiol* 1995;142:1255-64.
14. Jones MP. *Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression*. Technical Report. Iowa City(Ia): Department of Statistics, University of Iowa, 1994.
15. Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati DB. Regression modelling strategies for improved prognostic modelling. *Stat Med* 1984;3:143-52.
16. Peduzzi P, Concato J, Feinstein AR, Holford TR. The importance of events per independent variable (EPV) in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol* 1995;48:1503-10.
17. Cnaan A, Laird NM, Slasor P. Using the general linear mixed model to analyse unbalanced repeated measures longitudinal data. *Stat Med* 1997;16:2349-80.
18. Littell R. *SAS System for Mixed Models*. Cary: SAS Institute, 1996.
19. Dixon WJ. *BMDP Statistical Software Manual*. Berkeley: University of California Press, 1990.
20. Norussis MJ. *SPSS for Windows 7.5: Guide to Data Analysis*. Chicago: Prentice Hall, 1997.
21. *SOLAS*: For Missing Data Analysis 2.0 [User Reference]. Cork(Ireland): Statistical Solutions, 1999.
22. Curran D, Fayers PM, Molenberghs G, Machin D. Analysis of incomplete quality of life data in clinical trials. In: Staquet MJ, Hays RD, Fayers PM, editors. *Quality of Life Assessment in Clinical Trials: Methods and Practice*. Guildford(UK): Oxford University Press, 1998. p. 249-80.