

Title Missing Data: Discussion Points from the PSI Missing Data
Expert Group

Authors PSI Missing Data Expert Group

Abstract

The Points to Consider Document on Missing Data was adopted by the Committee of Health and Medicinal Products (CHMP) in December 2001. In September 2007 the CHMP issued a recommendation to review the document, with particular emphasis on summarising and critically appraising the pattern of drop-outs, explaining the role and limitations of the “Last Observation Carried Forward” (LOCF) method and describing the CHMP’s cautionary stance on the use of mixed models.

In preparation for the release of the updated guidance document, Statisticians in the Pharmaceutical industry (PSI) held a one-day expert group meeting in September 2008. Topics that were debated included minimising the extent of missing data and understanding the pattern of missing data, defining the principles for handling missing data and understanding the assumptions underlying different analysis methods.

A clear message from the meeting was that at present, biostatisticians tend only to react to missing data. Limited pro-active planning is undertaken when designing clinical trials. Missing data mechanisms for a trial need to be considered during the planning phase and the impact on the objectives assessed. Another area for improvement is in the understanding of the pattern of missing data observed during a trial via the plotting of data; for example, use of Kaplan-Meier curves looking at time to withdrawal.

Key words: Missing Data, LOCF, MMRM, Multiple Imputation

1. Background

The Points to Consider Document on Missing Data was adopted by the Committee of Health and Medicinal Products (CHMP) in December 2001 (1). In September 2007 the CHMP issued a recommendation to review the document (2), with particular emphasis on the following.

1. Summarising and critically appraising the pattern of drop-outs.
2. Use of sensitivity analysis or the justification for their absence.
3. Explaining the role and limitations of the “Last Observation Carried Forward” (LOCF) method.
4. Describing the CHMP’s cautionary stance on the use of mixed models.

In preparation for the release of the updated guidance document, PSI (Statisticians in the Pharmaceutical Industry), a professional association of statisticians in the pharmaceutical industry, held a one-day expert group meeting in September 2008. A list of the meeting attendees and affiliations is given in Appendix 1. Topics that were debated included the following.

1. Minimising the extent of missing data and understanding the pattern of missing data.
2. Defining the principles for handling missing data.
3. Understanding the assumptions underlying different analysis methods.

The remainder of this paper summarises the questions raised, resulting discussions and consensuses reached. After a brief review of the issues associated with each topic, the major questions raised are listed, immediately followed by a summary of the discussion and agreements. The context of the discussion is largely that of longitudinal clinical trials with dropouts or withdrawals. However, many of the points raised are applicable to other situations.

2. Minimising Missing Data and Understanding the Pattern of Missing Data

The CHMP Points to Consider document on Missing Data states in Section 3: “In the design and conduct of a clinical trial all efforts should be directed towards minimising the amount of missing data likely to occur”. The expert group discussed what proactive steps could be undertaken by trialists to minimise the amount of missing data in a clinical trials, and how best to understand the observed patterns of missing data.

Q1. What practical steps can be taken to avoid the presence of missing data in a) short term and b) long term clinical trials?

Following the ICH guideline on Statistical Principles for Clinical Trials, ICH E9 (3), the expert group acknowledged that missing values represent a potential source of bias and that every effort should be undertaken to plan the study so that the amount of missing data is minimised. However, there was consensus that there will almost always be some missing data. Suggestions to minimise the amount of missing data included the following.

1. Design the study and write the protocol so that key points to be observed are clearly identified.
2. The protocol should proactively plan for missing data; for example, unambiguously state the objectives of the study, the patient population of interest and how missing data may impact any inferences to be made. To illustrate the issues a nephrology trial was considered where serum creatinine data are collected weekly for 24 weeks. In such trials it is expected that about 30% of patients will withdraw. Reasons for withdrawal include death, kidney transplantation, adverse events, loss to follow up etc. In most cases simply extending the trial or increasing sample size will not adequately address missing data. Protocols and Statistical Analysis Plans rarely discuss the expected patterns of missing data, or consider the impact of the potential patterns on the overall scientific validity of the trial. Statisticians should proactively plan for various patterns of missing data when determining the sample size, using existing knowledge of the disease and compound under investigation, and the likely impact on the overall inferences to be drawn.

3. Consider a two-step withdrawal process for patients: withdrawal of consent for treatment and withdrawal of consent from observation. It was acknowledged that in some disease areas (e.g., pain control, diabetes) it may be a challenge to explain the value of continuing to observe patients when not being treated or being given an alternative medication. In other disease areas (e.g., Oncology) such practices are already standard. It was also recognised that switching of treatments can be an issue with continuing to monitor patients after withdrawal of treatment. Also switching treatments can result in confounding of treatment effects, which may be difficult to interpret or of limited value for short acting treatments or subjective responses. Nonetheless if data are collected after withdrawal with the aim of improving compliance it was suggested that the amount of such data could be reduced. That is, to encourage patients to continue to be observed following withdrawal of treatment, only key information such as the primary endpoint and adverse events should be collected.
4. It was suggested that tighter control of the patient population should be maintained through stricter inclusion or exclusion criteria. That is, patients should be selected who are more likely to complete the study. The disadvantage of this approach is that it reduces the generalizability of the trial findings. However that the occurrence of missing data anyway influences the generalizability of the results obtained from the observed data.

Another suggestion that was made after the meeting was to reduce the amount of data being collected in individual trials and simplify CRFs. If only key relevant data are collected, then the chance of data being captured reliably will increase, hence reducing the amount of missing data.

Q2. What methods do you think should be routinely employed to understand the nature of missing data?

To understand the nature of missing data it is important that the relevant information is collected. In a large number of clinical trials sponsored by the Pharmaceutical

Industry, standard withdrawal or discontinuation Case Report Forms (CRFs) are employed. These have prescribed standard lists for reasons for withdrawal such as Adverse Event, Lack of Efficacy, Lost to Follow Up etc. The group felt that often statisticians do not give enough thought to the customisation of these CRFs for the disease under consideration or the study objectives; for example, how often are disease- or study-specific reasons included? To illustrate the point consider “Lost to Follow Up” in oncology trials. What does this actually mean? Should study-specific reasons be provided to better understand what happens to these patients? The understanding of patient withdrawal patterns starts with the collection of relevant information.

The expert group also felt that during the planning phase of a clinical trial it is important to identify potential predictors of missing data, both to facilitate the collection of relevant data, and for potential inclusion in the analysis. For example, consider an asthma clinical trial. In such trials FEV1 is often used as the primary endpoint. It is widely recognised that “asthma exacerbations” may also be an important endpoint. In fact, when such events occur a patient may visit their health care professional, who in turn may advise the patient to withdraw from the trial. Subsequently when designing asthma trials it may be important to collect data on “asthma exacerbations”. It is important to note that in some cases it may be difficult to distinguish whether or not the mechanism for the missing data may be treatment related (e.g., number of exacerbations may increase due to the treatment being taken). Care must therefore be taken in accommodating such observations in the final analysis, and this should properly reflect their position on the causal pathway.

Another area where the expert group felt improvements could be made was for trialists to start thinking about the mechanisms that cause missing data earlier in the process. As outlined in ICH E9, drug development spans many years and comprises an ordered program of clinical trials each with their own specific objectives. Little effort is made to understand missing data in the earlier phases of drug development. Sponsors tend to start considering the impact of missing data during late Phase II and Phase III, the pivotal clinical trials, when such issues can affect the approval of the final package by the regulatory authorities. Missing data mechanisms need to be

considered when making go/no-go decisions at the end of Phase I and early Phase II and how this may impact later phase clinical study design.

Q3. What are the relative merits of the following exploratory analyses?

- *Plotting raw data and inspection of the data?*
- *Analysis by pattern of missing data (drop-out cohort)?*
- *Logistic regression of drop-out on earlier data?*

There was consensus that graphical display is one of the most important tools available to statisticians when trying to understand the causes of missing data. Although analytical methods exist for exploring missing data, a large amount of information can be ascertained by simply plotting the data: for example, using Kaplan-Meier plots to look at time to withdrawal and plots of treatment means against time for cohorts of subjects with similar follow up times (care should be taken not to over-interpret these graphs). These should be plotted on the same time scale for ease of comparison. The key to success is thinking through the question of interest and intelligently plotting the data.

Q4. How would the approach differ if the missing data was safety data as opposed to efficacy data?

The expert group agreed that the principles for minimising and understanding missing data should not change for safety data, but the challenges may be very different. For example, in Phase III there are often a small number of specific adverse events of interest that are compound-specific. During the design phase careful consideration needs to be given as to how information will be collected about such events, and the impact of missing data on the inferences to be drawn. It was agreed that there is a need for more than simple summary tables of adverse event incidence rates in clinical study reports. Increased use of graphical displays and more in-depth analyses are required. Any interpretation should be linked to the Risk Management Plan (4).

3. Defining the Principles for Handling Missing Data

As discussed in the CHMP Points to Consider document on Missing Data, if missing values are handled by simply excluding any patients with missing outcomes from the analysis a large number of issues can arise which may affect the interpretation of the trial results. The following section summarises the discussions at the expert group meeting relating to the principles that should be applied when handling missing data.

Q5. Regulators have stated on numerous occasions that missing data from patients who drop out are different from other types of missing data. What are the principles for handling different types of missing data?

The expert group agreed that the key issue when handling any missing data is understanding the mechanism causing the missing data. It is essential that the proposed method of analysis, and associated handling of missing data, regardless of whether the patient discontinued or not, must be directly linked and properly reflect the original objectives of the study, including any assumptions made when designing the trial. Specifically for patients who withdraw, the group felt that the critical question is what information needs to be collected for patients who discontinue, as such patients will occur in every trial. How missing data is handled is an integral part of the description of the primary comparison. The cost of running additional trials to investigate the effect of missing data far outweighs the cost of collecting the appropriate information in the first instance.

Q6. What are the principles for sensitivity analysis in the light of missing data?

The expert group agreed that two important principles exist when considering sensitivity analyses: transparency and relevance of the assumptions. It is important to clearly describe the original assumptions when designing the study so that all stakeholders can assess their relevance. The assumptions underlying any sensitivity analyses should be divergent from the original assumptions. It was agreed that, in contrast, a series of “wrong” analyses does itself not properly constitute a sensitivity analysis.

Q7. Regulators seem to be favouring a requirement for sponsor companies to monitor patients after withdrawal. How should post-withdrawal data be handled in the statistical analysis?

The issue of collecting data after withdrawal seems to be a critical one from the regulatory perspective. The issue reinforces the need to clearly define the objectives of the study. In defining the objectives clearly and precisely it will become apparent whether collecting data from patients who withdraw is necessary to address the question of concern. It was noted that the mechanism for withdrawal may differ between on-treatment and off-treatment periods. This in turn may lead to further technical challenges when incorporating data from patients after withdrawal into the analysis.

4. Understanding the Underlying Assumptions of the Different Analysis

Methods

In recent years a large amount of literature has been published on the merits of the different approaches for handling missing data (5, 6, 7). This final session of the meeting focused on clarifying the assumptions behind the different methods and how they might relate to the objectives of the trial, specifically for a longitudinal clinical trial with dropouts or withdrawals.

Q8. What are the underlying assumptions of the a) Last Observation Carried Forward (LOCF), b) Mixed Model for Repeated Measures (MMRM) and c) Multiple Imputation (MI) methods for handling missing data in a longitudinal clinical trial with dropouts or withdrawals?

The statistical techniques developed for handling missing data usually assume that the missing data can be one of the following.

1. Missing Completely at Random (MCAR)
2. Missing at Random (MAR)
3. Missing Not at Random (MNAR)

Definitions for each of these terms are provided in Table 1.

LOCF is a single-imputation method. It makes an implicit assumption that the patients would sustain the same response seen at an early study visit for the entire duration of the trial. The assumption is untestable and potentially unrealistic. Furthermore, the uncertainty of imputation is not taken into account, and so, as discussed by Mallinckrodt et al (5), the method results in systematic underestimation of the standard errors.

MMRM and MI analyses make the assumption that data are Missing At Random. In a MMRM analysis information from the observed data is used via the within-patient correlation structure to provide information about the unobserved data, but the missing data are not explicitly imputed. A MMRM analysis uses all the available data to provide the information about the unobserved data (5). It estimates the treatment effects assuming the withdrawn patients mimic those who continued. In MI, the imputation step is separate from the modelling step, and so there is additional flexibility to explore different assumptions about the nature of the missing data. If this flexibility is not used then it may in some circumstances essentially mimic an MMRM, and so offer no advantages over that method. Further details on each of the above methods are provided in Mallinckrodt et al (5).

It was acknowledged that if the underlying mechanisms that cause missing data are non-informative the resulting impact on the statistical analysis is far easier to handle, compared to informative missingness. The data being analysed, however, cannot provide evidence to distinguish between these two situations.

Q9. When might the assumptions for each of the methods be considered valid ?

Table 2 outlines when it might be appropriate or inappropriate to use LOCF, MMRM or MI techniques in a longitudinal clinical trial with dropouts or withdrawals. In such

studies it is important to recognise that MMRM and MI in their most basic form, both assume the multivariate normal distribution when providing information about the missing data. Invalid inferences can be drawn when the assumption is not met. There are however generalizations and modifications of these approaches, which while based on the same basic principles, are valid under other distributional assumptions. One nice feature about the MI technique however, is that the method can be applied to other types of response variables.

One of the main issues when determining how to handle missing data is that the true missing data mechanism will always be unknown and not testable from the data. No amount of clever modelling can overcome this issue. If the mechanism for missingness is informative then it will not be possible to fully evaluate the impact of the treatment of missing data in the analysis and this must be carefully considered in the interpretation of the data. Subsequently, the key issues are what questions are being answered from the analysis for the trial, and under what assumptions does the proposed analysis answer the questions. Doubts about aspects of the assumptions can be addressed through appropriate sensitivity analyses.

5. **Conclusion**

The Points to Consider Document on Missing Data was adopted by the CHMP in December 2001. Since the issuance of the guidance document there has been increased debate within the statistical community about the merits of the different approaches used to handle missing data such as LOCF, MMRM and MI. Subsequently in September 2007 the CHMP issued a recommendation to review the document, with particular emphasis on summarising and critically appraising the pattern of drop-outs, explaining the role and limitations of LOCF and describing the CHMP's cautionary stance on the use of mixed models. It was clear from the one-day PSI sponsored expert group meeting that the 2001 guideline places a great deal of emphasis on the merits of the different statistical methods available for handling missing data, and not enough on the principles that should be considered when designing trials.

The expert group also concluded that currently biostatisticians tend to react to missing data. Comprehensive, proactive planning is rarely undertaken when designing trials. It is imperative that the precise objectives of the trial are documented and the potential impact of missing data thoroughly considered during the planning phase. Missing data mechanisms for a trial need to be considered. Sensitivity analyses investigating the robustness of the inferences to the different assumptions made should also be considered. Another identified area for improvement is in the understanding of the pattern of missing data observed during a trial via the plotting of data; for example, use of Kaplan-Meier curves of time to withdrawal. Finally it was concluded that the handling of missing data is a difficult area. If the mechanism for the missing data is non-informative then the issue can be addressed by using relatively straightforward statistical techniques. However, if the mechanism for the missing data is informative then the issues are complex, and appropriate sensitivity analysis is called for.

References

1. CPMP Points to Consider Document. Missing Data (2001) Available at <http://www.emea.europa.eu/pdfs/human/ewp/177699EN.pdf>
2. Recommendation for the Revision of the Points to Consider on Missing Data (2007). Available at <http://www.emea.europa.eu/pdfs/human/ewp/43998007en.pdf>
3. ICH E9 Expert Working Group. Statistical Principles for Clinical Trials: ICH Harmonized Tripartite Guideline (1999). *Statistics in Medicine*, 18 : 1905 – 1942.
4. Guideline on Risk management Systems for Medicinal products for Human Use. Available at <http://www.emea.europa.eu/pdfs/human/euleg/9626805en.pdf>
5. Mallinckrodt CH, Lane PW, Schnell D, Peng Y, Mancuso J (2008) Recommendations for the Primary Analysis of Continuous Endpoints in Longitudinal Clinical Trials. *DIJ*, 42: 303-319.
6. Mallinckrodt CH, Watkin JG, Molenberghs G, Carroll RJ (2004) Choice of the primary analysis in longitudinal clinical trials. *Pharmaceutical Statistics*, 3: 161-169.
7. Lane P (2008) Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharmaceutical Statistics*, 7, 93-106.

Table 1: MCAR, MAR and MNAR definitions

Missing Completely at Random (MCAR)	The missing value mechanism is unrelated to the observed or unobserved responses, or to other measurements such as baseline values and treatment group. In particular, the probability that an observation is missed does not depend on how big or small it would have been if observed or on the size of the previous or subsequent observations on the same or any subject. Under MCAR any method of analysis that would have been valid for the complete data, such as ANCOVA, remains valid for the observed data.
Missing at Random (MAR)	The missing value mechanism may be dependent on <i>observed</i> measurements, including responses, but given these measurements, there is no remaining dependence on <i>unobserved</i> responses. The concept of Missing at Random (MAR) is most simply explained in the context of patient dropout in a longitudinal study. Suppose that two patients share the same treatment and covariates, and exactly the same response measurements up to the point at which one drops out and the other remains. Then the missing data from the subject who drops out are MAR if they have the same statistical behaviour as the observations from the subject who remains. Under MAR a valid analysis can be constructed that does not require knowledge of the specific form of the missing value mechanism.
Missing Not at Random (MNAR)	Even after accounting for observed measurements, there remains dependence between the missing value mechanism and the unobserved responses. Under MNAR a valid analysis does require knowledge of the specific form of the missing value mechanism, but in practice we will almost never know this mechanism.

Table 2: Summary of when LOCF, MMRM and MI should be considered for a longitudinal clinical trial with a continuous response variable

Statistical Method	Situations where technique can be considered	Situations where technique should not be considered
LOCF	<p>Stable disease following first post-treatment observation</p> <p>Short-term trials</p> <p>Last observation seen is part of the outcome</p>	<p>Diseases with marked improvement or deterioration over time (e.g., Alzheimers).</p> <p>Relapsing or remitting diseases (e.g., Generalized Anxiety Disorder)</p> <p>Disease involving transient treatment effects.</p> <p>When MMRM is used since it nullifies the repeated-measures aspects of the technique.</p>
MMRM	<p>Trials where the objective is to make inferences about treatment effects if patients stayed on treatment, but where no post-withdrawal data has been collected</p> <p>An unstructured covariance matrix should always be employed. Time should always be fitted as a class variable.</p> <p>Nearly always cross the baseline response with time (a).</p>	<p>When withdrawal patients do not mimic patients who continue in the study given same background history..... etc</p> <p>Trials where the objective is to make inferences about treatment effects if patients stayed on treatment but where off-treatment data is included in the analysis.</p> <p>If multivariate normal assumption does not hold for providing information about the missing data. That is, if the underlying distribution “needs” shape or tails.</p>
MI	<p>Method provides increased flexibility because the imputation part is separated from modelling part. Extra variables and complexity can be incorporated such as treatment withdrawals, outcomes etc. In particular</p>	<p>When Monte Carlo simulation not appropriate.</p> <p>If multivariate normal assumption does not hold for providing information about the missing data. That is, if the underlying distribution “needs” shape or tails.</p>

post-randomization
variables predictive of
dropouts can be
incorporated.

Different imputation
schemes are needed for
different treatment groups.

a) If the baseline response is not crossed with time then an increase (rather than a decrease) in the variability of the treatment comparison data can be observed. The reason for this is that the correlation between the baseline score and the outcome variable nearly always decreases with time; that is, the serial correlation decays. If baseline is fitted as a main effect then the estimated regression coefficient is averaged across all visits, and is larger than the correct baseline regression coefficient for the final time. This means that the analysis over-corrects for the endpoint of interest which is typically a comparison at the final visit. So even with no missing data one can get an over-corrected estimate of treatment difference.

Appendix 1: PSI Discussion Group Membership and Affiliation

Tomasz Burzykowski	MSOURCE Medical Development
James Carpenter	London School of Hygiene and Tropical Medicine
Corneel Coens	EORTC
Daniel Evans	Pfizer
Lesley France	AstraZeneca
Mike Kenward	London School of Hygiene and Tropical Medicine
Peter Lane	GlaxoSmithKline
James Matcham	Amgen
David Morgan	Ipsen
Alan Phillips	ICON Clinical Research
James Roger	GlaxoSmithKline
Brian Sullivan	Statistical Solutions
Ian White	MRC Biostatistics Unit
Ly-Mee Yu	Centre for Statistics in Medicine, University of Oxford