European Medicines Agency

London, 23 April 2009
Doc. Ref. CPMP/EWP/1776/99 Rev. 1 Corr*

**COMMITTEE FOR MEDICINAL PRODUCTS FOR HUMAN USE (CHMP)**

DRAFT

# GUIDELINE ON MISSING DATA IN CONFIRMATORY CLINICAL TRIALS

| | |
|---|---|
| **DISCUSSION IN THE EFFICACY WORKING PARTY** | June 1999/ November 2000 |
| **TRANSMISSION TO CPMP** | January 2001 |
| **RELEASE FOR CONSULTATION** | January 2001 |
| **DEADLINE FOR COMMENTS** | April 2001 |
| **DISCUSSION IN THE EFFICACY WORKING PARTY** | October 2001 |
| **TRANSMISSION TO CPMP** | November 2001 |
| **ADOPTION BY CPMP** | November 2001 |
| **DRAFT AGREED BY EFFICACY WORKING PARTY** | April 2009 |
| **ADOPTION BY CHMP FOR RELEASE FOR CONSULTATION** | 23 April 2009 |
| **END OF CONSULTATION (DEADLINE FOR COMMENTS)** | 31 October 2009 |

*The correction includes minor spelling amendment between lines 7 and 8.

This guideline will replace the 'Points to Consider on Missing Data in Clinical Trials' CPMP/EWP/1776/99.

Comments should be provided using this template to EWPSecretariat@emea.europa.eu

| | |
|---|---|
| **KEYWORDS** | *Baseline Observation Carried Forward (BOCF), Generalised Estimating Equations (GEE), Last observation carried forward (LOCF), Missing at random (MAR), Missing completely at random (MCAR), Missing Data, Mixed Models for Repeated Measures (MMRM), Missing not at random (MNAR), pattern mixture models.* |

7 Westferry Circus, Canary Wharf, London, E14 4HB, UK
Tel. (44-20) 74 18 84 00 Fax (44-20) 74 18 86 13
E-mail: mail@emea.europa.eu http://www.emea.europa.eu

12 **GUIDELINE ON MISSING DATA IN CONFIRMATORY CLINICAL TRIALS**

13 **TABLE OF CONTENTS**

## EXECUTIVE SUMMARY

Missing data are a potential source of bias when analysing clinical trials. Interpretation of the results of a trial is always problematic when the number of missing values is substantial. There is no universally applicable method of handling missing values, and different approaches may lead to different results. To avoid concerns over data-driven selection methods, it is essential to pre-specify the selected methods in the statistical section of the study protocol or analysis plan. It should be noted that the strategy employed to handle missing values might in itself provide a source of bias. A critical discussion of the number, timing, pattern, reason for and possible implications of missing values in efficacy and safety assessments should be included in the clinical report as a matter of routine. It will be useful to investigate the pattern of missing data in previous trials in the same or similar indications for related medicinal products. This could assist in the choice of the primary analysis method and how missing data will be handled in this analysis. It should be noted that just ignoring missing data is not an acceptable option when planning, conducting or interpreting the analysis of a confirmatory clinical trial.

Given the concerns highlighted above, how to minimise the amount of missing data and how missing data are going to be handled in the analysis are critical issues that must be considered when planning a clinical trial. A positive regulatory decision should not be based on an analysis that is biased to an important degree in favour of the experimental agent. Hence when proposing methods to handle missing data it is important that an analysis is provided for which the absence of a bias favourable to experimental treatment (i.e. conservative) can be established. The justification for selecting a particular method should not be based primarily on the properties of the method under particular assumptions (for example MAR or MCAR) but on whether it will provide an appropriately conservative estimate for the comparison of primary regulatory interest in the circumstances of the trial under consideration.

## 1. INTRODUCTION (background)

Missing data are a potential source of bias when analysing clinical trials. Interpretation of the results of a trial is always problematic when the number of missing values is substantial. This problem is only partially covered in ICH E9 (Statistical Principles of Clinical Trials).

There are many possible sources of missing data, affecting either complete subjects or specific items. There are many possible reasons for missing data (e.g. patient refusal to continue in the study, treatment failures or successes, adverse events, patients moving) only some of which are related to study treatment. Different degrees of data incompleteness can occur, i.e. measurements may be available only at baseline, or may be missed for one or several follow-up assessments. Even if a patient completes the study, some data may remain simply uncollected. In general this document concentrates on how to handle the situation where data is missing due to patients withdrawing from a trial.

Missing data violate the strict ITT principle which requires measurement of all patient outcomes regardless of protocol adherence. This principle is of critical importance as confirmatory clinical trials should estimate the effect of the experimental intervention in the population of patients with greatest external validity and not the effect in the unrealistic scenario where all patients receive treatment with full compliance to the treatment schedule and with a complete follow-up as per protocol. Full set analysis generally requires the imputation of values or modelling for the unrecorded data. Actually, even the per protocol analyses might also require the use of some method of handling missing data. This process can have, depending upon the amount and type of missing data, a crucial influence on the final results of a clinical trial and on the certainty with which conclusions can be drawn.

The extent to which missing values lead to biased conclusions about the size and existence of any treatment effect is influenced by many factors. Among these are the relationship between missingness, treatment assignment, and outcome; and the type of measure employed to quantify the treatment effect and the expected direction of changes over time for patients in the trial. All relevant factors should be considered to determine appropriate strategies for missing data handling (see Section 6).

It should be noted that the strategy employed to handle missing values might in itself provide a source of bias and that there is no universal best approach for all situations. The acceptability of an approach will depend on the assumptions made and whether it is reasonable to make these assumptions in the particular case of interest. It is very important when designing a study that the likely pattern of missing data is taken into account when specifying the primary analysis and the predefined sensitivity analyses. The amount of missing data and the strategies selected to handle missing data can influence the required sample size, the estimate of treatment effect and the confidence with which data can ultimately be interpreted. As such, how to minimise the amount of missing data and how missing data are going to be handled in the analysis are critical issues that must be considered when planning a clinical trial.

This document is not an extensive review of all the available methods. Instead general recommendations on acceptable frameworks for handling missing data in a regulatory setting are outlined. A positive regulatory decision should not be based on an analysis that is biased in favour of the experimental agent. Hence when proposing methods to handle missing data it is important that an analysis is provided for which the absence of a bias favourable to experimental treatment (*i.e*. the method is considered to be conservative) can be established.

## 2. SCOPE

This guideline provides advice on how the presence of missing data in a confirmatory clinical trial should be addressed in a regulatory submission. The pattern of missing data (including reasons for and timing of the missing data) observed in previous related clinical trials should be taken into account when planning a confirmatory clinical trial. This information should be used to minimise the amount of missing data present in a confirmatory clinical trial and assist in the choice of the method used to analyse the primary endpoint. The method chosen should aim to provide a conservative estimate of the treatment effect. In other words the treatment effect should not be biased to an important degree in favour of the experimental treatment. It is also recommended that appropriate sensitivity analyses are pre specified that explore the impact the methods used to handle missing data have on the interpretation of the results of the trial.

## 3. LEGAL BASIS

The Guideline should be read in conjunction with Annex I to Directive 2001/83/EC, as amended, and all other relevant EU and ICH guidelines. These include, but are not limited to:

CPMP/ICH/363/96, ICH Topic E9 Step 4 Note for Guidance on Statistical Principles for Clinical Trials.

## 4. THE EFFECT OF MISSING VALUES ON DATA ANALYSIS AND INTERPRETATION

The following problems may affect the interpretation of the trial results when some missing data are present.

### 4.1 Power and Variability

The sample size and the variability of the outcomes affect the power of a clinical trial. Power is greater the larger the sample size and the smaller the variability.

If missing values are handled by simply excluding any patients with missing values from the analysis this will result in a reduction in the number of valid cases available for analysis and therefore normally result in a reduction of the statistical power. Clearly the greater the number of missing values the greater the likely reduction in power.

In addition, non-completers might be more likely to have extreme values (treatment failure leading to dropout, extremely good response leading to loss of follow-up). Therefore, the loss of these non-completers could lead to an underestimate of variability and hence artificially narrow the confidence interval for the treatment effect, potentially increasing the power of the study. However, this increase would be artificial and if the methods used to handle missing data do not adequately take this into account the resulting confidence interval would not be considered a valid summary of the uncertainty of the treatment effect.

If values for missing data are imputed or modelled then all subjects can be included in the analysis in line with the ITT principle.

### 4.2 Bias

Bias is the most important concern resulting from missing data

If patients are excluded from the analysis this may affect:

- The comparability of the treatment groups.
- The representativeness of the study sample in relation to the target population (external validity).

A consequence of this may be a bias in the estimation of the treatment effect.

While the reduction of the statistical power is mainly related to the number of missing values, the risk of bias in the estimation of the treatment effect from the observed data depends upon the relationship between missingness; treatment and outcome (see also Section 6.1):

- In principle missing values will not be expected to lead to bias if they are not related to the real value of the unobserved measurement (e.g. poor outcomes are no more likely to be missing than good outcomes).
- Conversely, if the unmeasured observation is related to the real value of the outcome (e.g. the unobserved measurements have a higher proportion of poor outcomes), this will lead to bias even if the missing values are not related to treatment (i.e. missing values are equally likely in all treatment arms).
- Missing observations will lead to bias if they are related to both the treatment and the unobserved outcome variable (e.g. missing values are more likely in one treatment arm because it is not as effective).

In most cases it is difficult or impossible to elucidate whether the relationship between missing values and the unobserved outcome variable is completely absent. Thus it is sensible to adopt a conservative approach, considering missing values as a potential source of bias.

The causes of bias that critically affect interpretation will depend upon whether the objective of the study was to show a difference or demonstrate equivalence/non-inferiority.

## 5. GENERAL RECOMMENDATIONS

Unfortunately, there is no methodological approach for handling missing values that is universally accepted in all situations. Nevertheless there are some rules which should be considered when handling missing data.

## 5.1 Avoidance of missing data

There is no rule regarding the maximum number of missing values that could be acceptable. The quantity of missing data may be affected by a number of factors:

a) The nature of the outcome variable: the occurrence of missing values is expected to be lower when the outcome variable is mortality (e.g. cardiovascular trials), than when the outcome is more difficult to assess and requires the active participation of patients and/or sophisticated methods of diagnosis.

b) The length of the clinical trials: the longer the follow up the greater the likelihood of missing values.

c) The therapeutic indication: missing values are more frequent in those diseases where the adherence of patients to the study protocol is usually low (e.g. Psychiatric disorders).

d) The treatment modalities: e.g. surgical versus medical treatment.

Several major difficulties arise as a result of the presence of missing values and these are aggravated as the number of missing values increases. Thus, it is extremely important to avoid the presence of unobserved measurements as much as possible, by favouring designs that minimise this problem, as well as strengthening data collection regardless of the patient's adherence to the protocol and encouraging the retrieval of data after the patient's drop-out. In some circumstances, in particular where retrieved dropout information represents the progression of the patient without (or before) impact of further therapeutic intervention, these data will give the best approximation to the ITT population.

Where possible, outcome data from after withdrawal should be collected. Also data should be collected on other therapies received post drop-out. Specifically full details of the type of therapy given, including when and for how long it was used and at what dose, should be collected. This information will allow the value of any outcome data collected after withdrawal to be put into context.

## 5.2 Design of the study. Relevance of predefinition

It is very important when designing the study and specifying the statistical methods to be used, to anticipate the number of missing values likely to be observed in the trial. Experience from exploratory trials and from trials of other products in similar indications should inform expectations for missing data when planning the trial. Careful planning will help specify a plausible approach to handling missing data and also help specify a range of sensitivity analyses that could explore the impact of departures from the expected missing data pattern. Indeed, an estimate of the foreseen and acceptable amount of missing data is highly recommended: firstly because this may have repercussions for the variability and the expectations of the effect size and hence the sample size calculation, secondly because proper planning should minimise the risk that the strategy for missing data handling itself introduces bias, and thirdly because the uncertainty in interpreting the results introduced increases (and hence the number of sensitivity analyses required may need to increase – See Section 7) as the number of missing values increases.

There is no universally applicable method of handling missing values, and different approaches may lead to different results. To avoid concerns over data-driven selection methods, it is essential to pre-specify the selected methods in the statistical section of the study protocol or analysis plan. This section must include a detailed description of the selected methods and a justification of why the methods to be applied are expected to be an appropriate way of summarising the efficacy results of the study and to result in an absence of bias in favour of experimental treatment. The sensitivity analyses to be performed should also be pre-specified.

It is considered of particular importance to ensure that the selected method is a conservative approach that is not expected to favour the study's working (alternative) hypothesis (e.g. demonstration of superiority to placebo or demonstration of non-inferiority to active control). Some methods of handling missing data underestimate the variability associated with the treatment effect and therefore produce artificially narrow confidence intervals.

The process of imputation or modelling might be relevant to not only the main variables, but also the secondary efficacy, safety and baseline variables and covariates.

### 5.3 Final report

A detailed description of the pre-planned methods used for handling missing data, any amendments of that plan and a justification for the amendment should be included in the clinical study report.

A critical discussion of the number, timing, pattern, reason for and possible implications of missing values in efficacy and safety assessments should be included in the clinical report as a matter of routine. Graphical summaries (e.g. Kaplan-Meier plots) of the dropout patterns should be provided so that it can be clearly seen if there is a differential dropout pattern between treatment groups. These graphical summaries should identify the reason for dropout.

Data explorations and accompanying explanations that investigate missing data imbalance in all relevant factors and whether patients with and without missing values have different characteristics at baseline can also be informative. Data presentations should be such that it is possible to determine the contribution of each patient to the statistical analysis. For example, if single imputation methods are used the imputed values must be listed and identified.

If the pattern of missing data is different to that envisaged at the design stage, and the planned sensitivity analyses are inadequate, further sensitivity analyses should be provided that are tailored to the missing data pattern observed.

When a patient drops out of a trial full reporting of all reasons for their discontinuation should be given. This should allow identification of the most important reason that caused them to discontinue and this may influence how this subject is treated in the missing data analysis. Any follow-up information collected post drop out could be helpful in justifying how this patient is handled in all analyses.

As stated before, sensitivity analyses should investigate the robustness of the conclusions of a study and it is essential that, under clearly stated assumptions, at least one analysis which gives a demonstrably conservative estimate of the treatment effect. Also the confidence interval for this analysis should appropriately reflect the uncertainty associated with the estimated treatment effect.

Because of the unpredictability of some problems, it may be acceptable to allow in the study protocol the possibility of updating the strategy for dealing with missing values in the statistical analysis plan, or during the blind review of the data at the end of the trial. Relevant deviations from and amendments of the pre-specified plan should be clearly documented and justified. In addition, the time-point at which these deviations and amendments were decided and implemented in relation to the blinding of the data must be clearly identified. Methods for the documentation of these changes can be found in ICH E9. If unexpected missing data patterns are found in the data it will be necessary to conduct some post hoc sensitivity analyses in addition to those predefined in the statistical analysis plan (see section 7). In this case the reasons why these analyses have been conducted should be carefully explained and thoroughly justified. Proper planning will minimise the number of such analyses required, avoiding concerns over data-driven

selection of methods.

The final report must include documentation of any deviation from the expected number of missing values, a discussion of whether the pre-defined analysis is still sensible, plus appropriate sensitivity analyses.

## 6. HANDLING OF MISSING DATA

### 6.1 Theoretical Framework

The framework in the literature for the applicability of the different methods to handle missingness is based on a classification according to the following missingness mechanisms:

- If the probability of an observation being missing does not depend on observed or unobserved measurements then the observation is Missing Completely At Random (MCAR). A typical example is a patient moving to another city for non-health reasons. Patients who drop-out of a study for this reason could be considered a random sample from the total study population and their characteristics are similar.
- If the probability of an observation being missing depends only on observed measurements then the observation is Missing At Random (MAR). This assumption implies that the behaviour of the post drop-out observations can be predicted from the observed variables, and therefore that response can be estimated without bias using exclusively the observed data. For example, when a patient drops out due to lack of efficacy reflected by a series of poor efficacy outcomes that have been observed, the appropriate value to assign to the subsequent efficacy endpoint for this patient can be calculated using the observed data.
- When observations are neither MCAR nor MAR, they are classified as Missing Not At Random (MNAR) or non-ignorable i.e. the probability of an observation being missing depends on unobserved measurements. In this scenario, the value of the unobserved responses depends on information not available for the analysis (i.e. not the values observed previously on the analysis variable or the covariates being used), and thus, future observations cannot be predicted without bias by the model. For example, it may happen that after a series of visits with good outcome, a patient drops-out due to lack of efficacy. In this situation the analysis model based on the observed data, including relevant covariates, is likely to continue to predict a good outcome, but it is usually unreasonable to expect the patient to continue to derive benefit from treatment.

As already stated in section 4.2., it is impossible to be certain whether there is a relationship between missing values and the unobserved outcome variable or to judge whether that missing data can be adequately predicted from the observed data. It is not possible to know whether the MAR, never mind MCAR, assumptions are appropriate in any practical situation. A proposition that no data in a confirmatory clinical trial are MNAR seems implausible.

Because it is considered that some data are MNAR, the properties (e.g. bias) of any methods based on MCAR or MAR assumptions cannot be reliably determined for any given dataset. Therefore the method chosen should not depend primarily on the properties of the method under the MAR or MCAR assumptions but on whether it is considered to provide an appropriately conservative estimate in the circumstances of the trial under consideration.

### 6.2 Complete case analysis

One approach used to handle incomplete data is to ignore them and to perform the statistical

analysis with complete data only (complete case analysis). Some problems associated with this approach are discussed in section 3. Furthermore, complete case analysis violates the intention to treat principle and is subject to bias. Therefore this approach cannot be recommended as the primary analysis in a confirmatory trial. The approach may be considered in other circumstances, *e.g.*

- In exploratory studies, especially in the initial phases of drug development.
- In confirmatory trials as a secondary supportive analysis (sensitivity analysis) to illustrate the robustness of conclusions.

### 6.3 Methods of handling of missing data

Factors that affect the acceptability of individual methods include differences between the treatment groups in the proportion and timing of withdrawals, the direction of any spontaneous changes over time, and the reason for the withdrawals. All of these factors must be comprehensively displayed and their influence discussed when the method used to impute missing data is justified. This highlights the importance of proper planning (see section 5.2).

**Single imputation methods**

To cope with situations where response collection is interrupted after one point, one widely used method is last observation carried forward (LOCF). This analysis uses the last measured response as an endpoint by itself, not necessarily attached to a particular study time point. LOCF is one example of a single imputation method.

LOCF only produces unbiased treatment estimates under the MCAR assumption, but this approach can still provide a conservative estimate of the treatment effect in some circumstances.

To give some particular examples, if the patient's condition is expected to deteriorate over time (for example Alzheimer's disease) an LOCF analysis is very likely to give overly optimistic results for both treatment groups and if the withdrawals on the active group are earlier (e.g. because of adverse events) the treatment comparison may be biased in favour of the test product. Hence in this situation an LOCF analysis is not considered appropriate. Indeed in Alzheimer's disease, finding a method that gives an appropriate estimate of the treatment effect will usually be difficult and multiple sensitivity analyses may be required.

However, in other clinical situations (e.g. depression), where the condition is expected to improve spontaneously over time, LOCF (even though it has some sub optimal statistical properties) might be conservative in the situations where patients in the experimental group tend to withdraw earlier and more frequently. Establishing a treatment effect based on a primary analysis which is clearly conservative represents compelling evidence of efficacy from a statistical perspective.

Baseline observation carried forward (BOCF) is another single imputation approach that is sometimes used. The use of BOCF may be appropriate in, for example, a chronic pain trial where when a patient withdraws from treatment it may be reasonable to assume that their pain return to their baseline level and that the patient does not, in the long-term, derive benefit from treatment.

Another simple approach for imputing missing data is to replace the unobserved measurements by values derived from other sources. Possible sources include information from the same subject collected before withdrawal, from other subjects with similar baseline characteristics, a predicted value from an empirically developed model, historical data. Examples of empirically developed models are unconditional and conditional mean imputation, best or worst case imputation (assigning the worst possible value of the outcome to dropouts for a negative reason (treatment failure) and the best possible value to positive dropouts (cures)), regression methods and hot-deck imputation.

An attractive approach for imputing missing data may be to employ a different pre-specified imputation technique for each different reason for withdrawal, rather than the same technique for all patients. While this would represent a relatively novel approach, there is no objection to this in principle. The strategy has more flexibility in handling different reasons for and timings of withdrawal and consequently the possible relationship between missing data and the outcome of interest. If used appropriately, it may better address the question of primary regulatory interest. The method also offers an intuitive framework for conducting a range of sensitivity analyses.

A potential disadvantage of single imputation methods is that these methods risk biasing the standard error downwards by estimating a central value and ignoring its uncertainty. Therefore, the confidence intervals for the treatment effect calculated using single imputation methods may be too narrow and give an artificial impression of precision that doesn't really exist. This possibility should be addressed when results from these analyses are presented.

In conclusion, single imputation methods, including LOCF and BOCF, can be accepted as a primary analysis in confirmatory trials provided that the applicant has justified that the estimated treatment effect is not expected to be biased in favour of experimental treatment (see section 5.3) and the associated confidence interval does not underestimate the variability of this estimate to an important extent.

**Mixed models, Multiple imputation methods and Generalised estimating equations**

The risk of underestimating the variance when imputing can be avoided by some techniques such as multiple imputation. Multiple imputation methods generate multiple copies of the original data set replacing missing values by randomly generated values, and analyse them as complete sets.

Some statistical approaches to handling missing data do not employ formal imputation. For continuous responses, linear mixed models are sometimes used to handle missingness when a series of outcomes are measured repeatedly over time (mixed-effect models for repeated measures (MMRM)). For categorical responses and count data, the so-called marginal (e.g. generalized estimating equations (GEE)) and random-effects (e.g., generalized linear mixed models (GLMM)) approaches are in use. Likelihood-based methods (MMRM and GLMM) and some extended GEE (*i.e*. weighted GEE) models are applicable under MCAR and MAR assumptions.

In many cases, there is a variety of different settings for each method which could lead to different conclusions (e.g. type of variance-covariance matrix for MMRM, method for imputation on MI approaches, assumptions to model the un-observed measurements used in the MNAR methods, etc.). Therefore, the precise option settings must be fully justified and predefined in advance in detail, so that the results could be replicated by an external analyst.

The methods above are unbiased under the MAR assumption and can be thought of as aiming to estimate the treatment effect that would have been observed if all patients had continued on treatment for the full study duration. Therefore, for effective treatments these methods have the potential to overestimate the size of the treatment effect likely to be seen in practice and hence to introduce bias in favour of experimental treatment in some circumstances. In light of this the point estimates obtained can be similar to those from a complete cases analysis. This is problematic in the context of a regulatory submission as confirmatory clinical trials should estimate the effect of the experimental intervention in the population of patients with greatest external validity and not the effect in the unrealistic scenario where all patients receive treatment with full compliance to the treatment schedule and with a complete follow-up as per protocol. The appropriateness of these methods will be judged by the same standards as for any other approach to missing data (i.e. absence of important bias in favour of the experimental treatment) but in light of the concern above, the use of only these methods to investigate the efficacy of a medicinal product in a regulatory submission will only be sufficient if missing data are negligible. The use of these

methods as a primary analysis can only be endorsed if the absence of important bias in favour of the experimental treatment can be substantiated.

Generally, MNAR data is difficult to rule out, and it is not clear whether even a small amount of MNAR data could have an impact on the study results in a particular experiment. Therefore approaches that investigate different MNAR scenarios such as pattern mixture, selection and shared parameter models should be explored. A combined strategy incorporating several methods for handling missingness may also be considered. As described above, methods that do not assume MCAR or MAR such as pattern mixture models may offer a flexible framework to explore the impact of treating different types of missing data as MNAR and evaluating the impact different modelling strategies have on the estimated treatment effect.

**Survival analysis**

When the outcome measure is time to event, survival models which take into account censored observations may be used. However, standard survival methods assume that there is no relationship between the response and the missing outcome. This generally cannot be assumed and violations from this assumption could lead to biased results. Therefore whether it is reasonable to assume non-informative censoring should be discussed in a study report and in situations where the amount of missing data/patient withdrawals could influence whether the treatment effect is established or could influence the size of the treatment effect, a range of sensitivity analyses should be provided. Further considerations for handling this type of data are outlined in Appendix 1 to the CHMP guideline on the evaluation of anticancer medicinal products in man (EMEA/CHMP/EWP/267575/2006Corr.).

**Responder analysis**

Commonly, the primary analysis of a continuous variable is supported by a responder analysis. How missing data are going to be categorised in this analysis should be pre-specified and justified. There will be some experimental situations when conducting a responder analysis could be a viable option for the primary analysis.

## 7. SENSITIVITY ANALYSES

In this context, sensitivity analyses can be defined as a set of analyses where the missing data is handled in a different way in each analysis. This will show the influence of different methods of handling missing data on the study results. These analyses can help to justify the choice of the particular method applied as the primary approach.

In all submissions with non-negligible amounts of missing data sensitivity analyses should be presented as support to the main analysis. Because the performance of any analysis presented (in terms of bias and precision) cannot be fully elucidated, presentation of trial results without adequate investigation of the assumptions made for handling missing data cannot be considered comprehensive.

If the results of the sensitivity analyses are consistent and lead to reasonably similar estimates of the treatment effect this provides some assurance that the lost information had little or no effect on the overall study conclusions. In this situation the robustness of the results is clear and the missing values will not generally be considered to be a serious source of concern. A broader range of sensitivity analyses will give greater reassurance on the robustness of the trial results. Conversely, whilst not all sensitivity analyses must necessarily give statistically significant results, if they produce inconsistent results (e.g. a markedly smaller estimate of treatment effect), their repercussions on the conclusions of the trial must be discussed. In certain circumstances, the

influence of missing data is so great that it might not be possible to reliably interpret the results from the trial.

The sensitivity analyses required will need to be defined on a case-by-case basis, though will usually comprise the analyses already described in Section 6 above.

Some ways of performing a sensitivity analysis are:

- Compare the results of the full set analysis to those of the complete case analysis.
- As discussed in section 6.2 it is not possible to guarantee that at least some of the missing data are not MNAR. Therefore, further sensitivity analyses that treat certain types of missing data as MNAR should be provided. It may be appropriate to treat data missing for different reasons in different ways. A range of analyses should be provided that explore these possibilities. For each of these analyses a clear explanation of what values have been imputed should be given. This may be done using multiple imputation methods incorporating pattern mixture approaches.
- Compare the impact different model settings have on the results. If different results are obtained from models using the same missing mechanism assumption (e.g. MI versus MMRM both assuming MAR) full details of the differences between these models that explain the different results obtained should be provided. In any case the impact different settings of a model have on the results obtained should be explained in detail.
- Utilise retrieved drop-out data if not already done for the primary analysis. If a patient has received other therapies after withdrawing from the study a positive value for the primary endpoint at the end of the trial could be due, in part at least, to the switching of therapies for this patient. Analyses that down play the positive outcome to give a more realistic view of the product being evaluated should be conducted.
- In a responder analysis, an analysis that treats all missing values as failures or treats missing values due to a certain reason (e.g. due to adverse events) as failures and for other reasons (e.g. excellent response to treatment for majority of trial with missing final value) as successes.
- A worst case analysis: compare the results of two analyses, one assigning the best possible outcome to missing values in the control group and the worst possible to those of the experimental group. If this extreme analysis is still favourable then it can be confidently concluded that the results are robust to the handling of missing data.

Each sensitivity analysis should be designed to assess the effect on the results of the particular assumptions made in the handling of the missingness. The sensitivity analysis should be planned and described in the protocol and/or in the statistical analysis plan and any changes must be documented and justified in the study report.

## REFERENCES

Appendix 1 to the CHMP guideline on the evaluation of anticancer medicinal products in man (EMEA/CHMP/EWP/267575/2006Corr.).

ICH E9 Statistical Principles of Clinical Trials (CPMP/ICH/363/96).