



**Delfini Group**™, LLC



*Evidence- & Value-based Solutions For Health Care  
Clinical Improvement Consults, Content, Seminars, Training & Tools*

**White Paper**

# Missing Data: Considerations

February 20, 2008

**Delfini Group**™, LLC

**Michael Stuart, MD**  
President

**Sheri Strite, Principal &  
Managing Partner**

[www.delfini.org](http://www.delfini.org)

## **Our Mission –**

To assist medical leaders, clinicians and other health care professionals by ~

- Bringing science into medical practice in an **easy-to-understand** way.
- Using **simplified methods** to help navigate the complexities of such areas as evidence-based medicine and other topics.
- Building **competencies** and **confidence** in improving medical care through our well received consultations, educational programs and tools.
- Providing inspiration to others to **improve** medical care and help bring about needed change.

**INTRODUCTION**

Not having complete information on all study subjects is a common problem in research. As of the date of the writing of this paper, no one we know of has identified a preferred way to handle this problematic area. This paper is our attempt to expose the problems and provide some suggestions and guidance from a research reviewer’s point of view as well as to raise areas for continued reflection and discourse.

- Thus far, there is no true consensus on any of these issues, except for some agreement that an Intention-to-Treat analysis should be the primary analysis.
- Ultimately, missing data requires that assumptions be applied and assumptions can be wrong.

Practically speaking, most studies with a sizable amount of missing information, and which rely on assumptions that do not put the intervention through the toughest test or an otherwise reasonable test that does not favor the intervention, will warrant a Grade U: Uncertain validity and/or usefulness — or a Grade B-U at best if the study is otherwise very free from threats to validity.

The reviewer may seek to apply challenge tests to the authors’ findings of statistical significance by performing a sensitivity analysis with various assumptions and re-testing for statistical significance. (The *Delfini Intention-to-Treat Calculator* can assist with this effort.) If an otherwise valid study remains statistically significant after going through the toughest challenge to the intervention in assigning values to missing data or other reasonable sensitivity analysis, the study might be able to achieve a passing grade.

For purposes of this paper, “completed study” means that the patient completed the treatment, that complete outcomes data were collected and that the patient was followed for the prescribed period of time.

**KEY ISSUES RELATING TO “MISSINGNESS” OF DATA**

Question	Consideration	Comments
<p><b>OUTCOME IMPACT</b> Would the outcome change significantly if all persons had completed the study and we had complete information on them?</p>	<p>- Are patients who did not complete the study different from those who completed the study?</p> <ul style="list-style-type: none"> <li>▪ What would have happened to non-completers had they completed the study?</li> <li>▪ Are there differences between study arms or intragroup differences that might have affected study outcomes such as differential time to disenroll or differential rates of disenrollment due to treatment failure?</li> </ul> <p>- Are patients who are lost to follow-up (not contactable) different from completers in ways that affect outcomes?</p>	<ol style="list-style-type: none"> <li>1. Discontinuation or loss may be reflective of treatment failure or adverse events (thus completer analysis is likely to favor the intervention)</li> <li>2. A differential loss to follow-up is a clue that there may be a difference that affects outcomes — however, no differential loss is not a guarantee there are no such differences</li> <li>3. Delfini has created a conservative approach to looking at percentages of missing information — this is our best gestalt based on some reading, some number-crunching and some gut reaction. (See REVIEWER HELP below).</li> <li>4. A continuing issue is that there are gaps in standards for reporting, and missing data requires using models that rely on assumptions which may be invalid and are unverifiable</li> </ol>
<p><b>PREDICTIVENESS</b> Is any information, through what is known or what is missing, informative in a predictive way that it can be used to</p>	<p>- Is there any variable — or its lack — that is predictive of another missing variable or that lends to adjustment? (See DEFINITIONS AND CONCEPTS RELATED TO MISSINGNESS &gt; MCAR AND MAR below.)</p>	<ol style="list-style-type: none"> <li>1. Analysts attempt to determine the randomness and the interconnectedness of missing information to try and determine differences, adjust and/or fill in the blanks — but ultimately it is conjecture</li> </ol>

**KEY ISSUES RELATING TO "MISSINGNESS" OF DATA**

Question	Consideration	Comments
help fill in missing data?		
<p><b>ANALYSIS AND REPORTING</b> How should a researcher handle and report missing data?</p>	<ul style="list-style-type: none"> <li>- Intention-to-treat analysis should be performed in a well done manner                             <ul style="list-style-type: none"> <li>▪ All subjects are to be analyzed in the group to which they are randomized</li> <li>▪ The method for filling in missing values should be conservative and not favor the intervention</li> <li>▪ If the intervention remains statistically significant in an otherwise valid study where it has been put through the toughest challenge (missing subjects or data in the intervention group are counted as failed; missing subjects or data in the comparator group are counted as successes) the study may earn a passing grade.</li> </ul> </li> <li>- Sensitivity analyses provide greater information                             <ul style="list-style-type: none"> <li>▪ In addition to doing ITT, are data collected from contactable non-completers and reported in a meaningful way that distinguishes them from completers (examples for consideration include baseline characteristics and performance issues such as dosing, timing duration, concomitant meds, etc. including any treatment following discontinuation which may have affected efficacy or safety outcomes subsequent to discontinuation)</li> </ul> </li> <li>- Report information provides sufficient detail so that reviewers can do their own calculations</li> <li>- Report thoroughly on patient disposition dealing with completion, treatment outcomes, ADEs and contactability (in addition to supplying a CONSORT diagram, see ADDITIONAL SUGGESTIONS FOR CATEGORIZING PATIENT DISPOSITION below)</li> </ul>	<ol style="list-style-type: none"> <li>1. Prespecify statistical methods</li> <li>2. Utilize a CONSORT diagram that details the disposition of study subjects (note that CONSORT requests the number of subjects analyzed--frequently authors provide results by percentage without the actual numerators and denominators.)</li> <li>3. Report assumptions used</li> <li>4. Safety population should be only those people who received an intervention (meaning, if someone did not fill a prescription, they should be excluded from the safety population)</li> <li>5. Sensitivity analysis is helpful. However, reporting on subgroups is a de-randomized analysis just as is a completer analysis</li> <li>6. Completer/non-completer analysis will be most meaningful when the effects of a drug can be evaluated even if discontinued (e.g., curative)</li> <li>7. Last observation carried forward (LOCF) has been shown to be prone to bias:                             <ol style="list-style-type: none"> <li>a. O'Brien PC, Zhang D, Bailey KR. Semi-parametric and non-parametric methods for clinical trials with incomplete data. Stat Med. 2005 Feb 15; 24(3): 341-58. Erratum in: Stat Med. 2005 Nov 15; 24(21): 3385. PMID: 15547952</li> <li>b. Carpenter, J, Kenward, M. Guidelines for handling missing data in Social Science Research. <a href="http://www.lshtm.ac.uk/msu/missingdata/guidelines.pdf">www.lshtm.ac.uk/msu/missingdata/guidelines.pdf</a> accessed (02/20/2008)</li> <li>c. (See DEFINITIONS AND CONCEPTS RELATED TO MISSINGNESS &gt; IMPUTATION below)</li> </ol> </li> <li>8. Averaging the known data for all subjects is mathematically the same as doing a completer analysis (defining completer analysis as those who complete the study treatment and outcome assessments).</li> <li>9. Assessing outcomes through models has been reported to potentially erroneously misrepresent outcomes by a relative difference of 50% or higher. Lachin PMID 11018568</li> <li>10. If the variables are continuous, can they be put into dichotomous form for analysis? (If yes, ensure that the resulting variables are clinically meaningful.)</li> <li>11. Report complete information so the reviewer can do sensitivity analyses directly and so that all steps in the analysis are transparent</li> <li>12. Death needs to be handled equally in both groups – if deaths are not excluded in the placebo group as unrelated to treatment, likewise deaths not attributed to the intervention must not be excluded</li> <li>13. When using Kaplan Meier curves –                             <ul style="list-style-type: none"> <li>o Ensure there is no double-counting of events – and report that</li> </ul> </li> </ol>

KEY ISSUES RELATING TO "MISSINGNESS" OF DATA

Question	Consideration	Comments
		<p>you have checked for this</p> <ul style="list-style-type: none"> <li>o Do not use if the rate of the outcome is not otherwise constant between groups (e.g., one would expect more deaths in a surgical group early as compared to deaths in a medically treated group)</li> <li>o Realize that censoring is not ITT analysis as the numerator and the denominator are changed – conduct an ITT analysis as well</li> <li>o Investigators should provide complete details regarding censoring</li> <li>o Other suggestions provided by John M. Lachin, Sc.D., Professor of Biostatistics and Epidemiology, and of Statistics, The George Washington University (personal communication):             <ul style="list-style-type: none"> <li>▪ Evaluate censoring by examining both administrative censoring and censoring due to loss-to-follow-up. Administrative censoring (censoring of subjects who enter a study late) may not result in significant bias. Censoring because of loss-to-follow-up is more likely to pose a threat to validity</li> <li>▪ Compare characteristics of losses (e.g., withdrawing consent, adverse events, loss to follow-up) versus completers (including administratively censored) within groups.</li> <li>▪ Compare characteristics of losses (not administratively censored) between groups.</li> <li>▪ Adjust group effect for factors in which groups differ.</li> </ul> </li> <li>o At a practical level, most users will not get this level of detail in a reported study. Therefore, we advocate a conservative approach. We think the bottom line is that differential loss is a threat; however, loss that is not differential overall may mask resulting differences in prognostic variables between groups.</li> </ul>
<p><b>REVIEWER HELP</b> How should a reviewer react to significant missing data?</p>	<p>- Is there sufficient patient drop-out or missing data points that the study's validity is threatened?</p> <p>- Look beyond the terms, "loss to follow up," (e.g., discontinuations, excluded for protocol violation, etc.)</p> <p>- Consider the following (percentages are approximations):</p> <ul style="list-style-type: none"> <li>▪ Likely minimal threat: &lt; 5% and no differential loss*</li> <li>▪ Possible threat: &gt;= 5% but &lt;10% and no differential loss*</li> <li>▪ Acceptable for efficacy: &gt;= 5%, but with reasonable sensitivity analysis, and analysis continued to agree with authors' findings about statistical significance</li> <li>▪ Threat: &gt;=5% with differential loss*, or &gt;= 10% without differential loss, and without worst-case sensitivity analysis, or otherwise</li> </ul>	<ul style="list-style-type: none"> <li>▪ See above. If the author has not followed these suggestions, and if the study is otherwise a valid and clinically useful study, you might be able to reanalyze the results yourself based on these suggestions</li> <li>▪ ITT analysis requires dichotomous variables. If the outcomes are presented as continuous variables, can you assign the outcomes in such a way that they are made dichotomous? If not, you cannot reanalyze using a 2X2 table.</li> <li>▪ Generally, our usual advice is to start with the hardest challenge for the intervention – (missing values in intervention group are counted as failed; missing values in the comparator group are counted as successes) to see if the intervention still shows a statistically significant benefit.             <ul style="list-style-type: none"> <li>o There may be other reasonable choices which do not favor the intervention depending upon the context of the study and which get closer to truth. Example: In a study of prehypertension,</li> </ul> </li> </ul>

**KEY ISSUES RELATING TO “MISSINGNESS” OF DATA**

Question	Consideration	Comments
	<p>reasonable sensitivity analysis, conducted by authors or reviewers</p> <ul style="list-style-type: none"> <li>▪ *Differential loss:                             <ul style="list-style-type: none"> <li>○ For small to medium study (e.g., less than 300 total randomized), differential loss must be low to non-existent (e.g., 2% or less difference in missing data points between groups)</li> <li>○ For large study (e.g., more than 300 total randomized), differential loss must be minimal (e.g., 5% or less difference in missing data points between groups)</li> </ul> </li> </ul> <p>Note: We do not have evidence that supports the above loss level distinctions, but it is a conservative approach drawn from various recommendations plus Delfini experience generally. Consider percentages along with effect size. A large effect size is unlikely to change from reanalyzing with a small amount of loss. However, there are instances where a small loss (even less than 5%) can affect the p-values. So the above percents are offered as general alerts only.</p>	<p>missing in both the candesartan and the placebo group were counted as failed. Since blood pressure is unlikely to spontaneously improve from anything active in the placebo (in that both the placebo and the study drug could account for placebo effect) this appears to be a clinically reasonable choice (Julius S et al. TROPHY Study. N Engl J Med. 2006 Apr 20; 354(16):1685-97. Epub 2006 Mar 14. PMID: 16537662. .</p> <ul style="list-style-type: none"> <li>○ You are not looking for “truth” since you cannot know the true answer. The question is “What more could you possibly know differently after putting the intervention through the toughest test aside from safety and how much efficacy?” If the answer to that is nothing else, then you may feel comfortable concluding that the intervention is efficacious.</li> </ul> <ul style="list-style-type: none"> <li>▪ If the P-values are no longer statistically significant, consider other sensitivity analyses; however, Delfini would generally not grade such a study above a Grade B-U</li> <li>▪ Safety issues may not be revealed in a study or may not show up for many years – consider the study duration</li> <li>▪ Safety issues reported in long term follow-up past discontinuation of treatment may be a) confounded, and/or b) diluted</li> </ul>

**DEFINITIONS AND CONCEPTS RELATING TO “MISSINGNESS”**

Missing Completely at Random (MCAR)	Information is missing for reasons completely due to chance. It is hard to know if this is actually true.
Missing at Random (MAR)	The missingness of information is informed by other variables (e.g., reading ability and completing a survey). Researchers try to adjust for this, but it is unverifiable.
Imputation	<p>Assigning values for missing information: various methods include —</p> <ul style="list-style-type: none"> <li>▪ Single imputation (a single value is assigned to a variable)                             <ul style="list-style-type: none"> <li>○ Last observation carried forward (LOCF) assigns the last observed outcome — this is a method prone to bias                                     <ul style="list-style-type: none"> <li>▪ It assumes stability, which may not be so</li> <li>▪ It doesn't account for intra-subject variation</li> <li>▪ In a progressive illness or the case of deterioration, it will falsely present outcomes better than they, in fact, are</li> </ul> </li> <li>○ Worst case scenario against intervention (also known as an extreme informative model)</li> </ul> </li> <li>▪ Multiple imputation draws a random sample from its distribution – this is unverifiable</li> <li>▪ Partial imputation is a mix of single and multiple imputation and only imputes until the groups are made even, which could be a false assumption – this is unverifiable</li> <li>▪ Regression models – use unverifiable assumptions</li> </ul>
Intention-to-treat Analysis	Analyzing all subjects randomized in the groups to which they were assigned

**ADDITIONAL SUGGESTIONS FOR CATEGORIZING PATIENT DISPOSITION (EXAMPLE)**

A **CONSORT Diagram Generator** for reporting patient disposition is available at <https://swolpin.cirg.washington.edu/CSD/> (accessed 02/20/2008).

Some considerations we are interested in are as follows, by treatment arm, number of patients —

1. Randomized
2. Received intervention
3. Did not receive intervention
4. Experienced ADE and not discontinued
5. Discontinued due to treatment failure
6. Discontinuation due to ADEs
7. Discontinued due to other
8. Migrated to different arm
9. Completed study without full data
10. Completed study with full data
11. Did not complete study
12. Contactability of patient and whether follow-up continued (and even if so, the data will be confounded - is this separated out for sensitivity analysis?)
13. Lost to follow-up
14. Outcomes
15. Excluded from analysis
16. Number we accept in the author's analysis
17. Number we wish reanalyzed and how

Where significant data points are missing, reviewers need to see sensitivity analyses that address these various considerations.

We also want to know, by treatment arm, information about —

- Adherence
- Protocol violations
- Incorrect treatment assignment
- Anything else about missing values that is pertinent

## REFERENCES AND RECOMMENDED READING

*Delfini* Clicks at [www.delfini.org](http://www.delfini.org):

**Attrition Bias: Intention-to-Treat Basics**

[http://www.delfini.org/delfiniClick\\_PrimaryStudies.htm#models](http://www.delfini.org/delfiniClick_PrimaryStudies.htm#models)

**Intention-to-Treat Analysis: Censoring**

[http://www.delfini.org/delfiniClick\\_PrimaryStudies.htm#ITTVioxxCurve](http://www.delfini.org/delfiniClick_PrimaryStudies.htm#ITTVioxxCurve)

**Intention-to-Treat Analysis: Misreporting and Migraine**

[http://www.delfini.org/delfiniClick\\_PrimaryStudies.htm#itt](http://www.delfini.org/delfiniClick_PrimaryStudies.htm#itt)

**Missing Data Points: Difference or No Difference**

[http://www.delfini.org/delfiniClick\\_PrimaryStudies.htm#missingdifference](http://www.delfini.org/delfiniClick_PrimaryStudies.htm#missingdifference)

### Other References and Recommended Reading

Carpenter, J, Kenward, M. Guidelines for handling missing data in Social Science Research. [www.lshtm.ac.uk/msu/missingdata/guidelines.pdf](http://www.lshtm.ac.uk/msu/missingdata/guidelines.pdf) accessed (02/20/2008)

Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ*. Vol 319. Sept 1999: 670-674.

**NOTE: Delfini disagrees that differential loss to follow-up is not an issue and believes that a loss to follow-up of greater than 5 percent without an appropriate ITT analysis may be considered threatened.**

Lachin JL. Statistical considerations in the intent-to-treat principle. *Control Clin Trials*. 2000 Oct;21(5):526. PMID: 11018568

O'Brien PC, Zhang D, Bailey KR. Semi-parametric and non-parametric methods for clinical trials with incomplete data. *Stat Med*. 2005 Feb 15;24(3):341-58. Erratum in: *Stat Med*. 2005 Nov 15;24(21):3385. PMID: 15547952

Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *The Lancet*. Vol 359. March 2, 2002: 781-785. PMID: 11888606

**NOTE: Delfini stresses that the approach taken for missing values should not give an advantage to the intervention.**

Shih W. Problems in dealing with missing data and informative censoring in clinical trials. *Curr Control Trials Cardiovasc Med*. 2002 Jan 8;3(1):4.

PMID: 11985778