

Ophthalmic statistics note 8: Missing data – exploring the unknown

Catey Bunce^{1,2} Ana Quartilho¹, , Nick Freemantle³, Caroline J Doré⁴, on behalf of the Ophthalmic Statistics Group

¹NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

²London School of Hygiene & Tropical Medicine

³Department of Primary Care and Population Health, University College London, London, UK

⁴UCL Comprehensive Clinical Trials Unit, University College London, London, UK

Keywords:

Word count: 2324 words (excluding title page and references)

Medical research is conducted to answer uncertainties and identify effective treatments for patients. Different questions are best addressed by different types of study design – but the randomised controlled clinical trial is typically viewed as the gold standard, providing a very high level of evidence, when examining efficacy. ¹ Whilst clinical trial methodology has advanced considerably with clear guidance provided as to how to avoid sources of bias, even the most robustly designed study can succumb to missing data. ^{2 3} In this statistics note, we discuss strategies for dealing with missing data but what we hope emerges is a very clear message that there is no ideal solution to missing data and prevention is the best strategy.

Scenario one

A senior colleague asks me to critique a publication of a randomised controlled clinical trial comparing two drugs which aim to reduce intraocular pressure in patients with primary open angle glaucoma. One eye per patient has been analysed and results are provided for intraocular pressure at 6 months. The study presents data on 147 subjects treated with drug A and 145 subjects with drug B. The mean pressure in patients on drug A is lower than drug B, with an estimated treatment difference of 3.1 mmHg, 95% CI (2.5, 3.8). A P value of

< 0.001 is provided. It seems clear that drug A is more efficacious in reducing IOP at 6 months than drug B but does this mean that I am correct in deducing that A is better than B and therefore that patients should be given drug A?

Something about the numbers doesn't seem quite right. 147 vs 145 where I had expected equal numbers in the two groups. I learn (via the internet) that the researchers may have used simple randomisation in which case chance imbalances can and do occur, particularly with smaller studies.⁴ Whilst this might impact upon power, it does not itself represent an issue, however on careful scrutiny of the publication, I uncover that at the start of the study 150 patients were in each arm of the study. Three patients receiving treatment A do not provide 6 month outcome data and five patients allocated to treatment B do not appear in the final outcome analysis. The paper is not clear as to what happened to these 8 patients, however, my colleague knows the authors and agrees to drop them a note to find out what happened. He does however point out that 8 patients out of 300 is just 2.6 % which is well within the anticipated rate of loss to follow-up allowed for in the original sample size calculation.

Several months later I receive the information that I was after. Patients on treatment arm B had merely not attended their 6 month follow-up visit. A couple had moved and a couple simply were not well at the 6 month visit. They did, however, have data at 5 months and a couple had attended at 7 months. Figures at 5 months and 7 months were fairly similar to each other indicating that perhaps the 5 month data (or indeed the 7 month data) could be viewed as a reasonable estimate of 6 month data. The treatment differences seen in these patients were similar to those seen in the patients who had attended at 6 months. Patients on treatment A had, however, not attended their 6 month visit or indeed any further visits. Contact with their general practitioners revealed that they had each suffered respiratory issues. Later on, I learned that there was indeed a causal link between drug A and adverse respiratory problems. Treatment A no longer seems the best treatment – particularly for those at risk of respiratory problems.

This scenario is given to illustrate the potential for misleading conclusions to be drawn in the presence of missing data. Missing data was the focus of a statistics note by Altman and Bland.³ At the time of writing that note, the authors commented that “the topic of how to

handle missing data is not often discussed outside statistics journals". They stated also that the most common approach to deal with missing data was to simply analyse everyone with complete data only – an available case or complete case analysis – as is illustrated in scenario one. Whilst this method may be appropriate when there is little missing data – it can lead to incorrect conclusions – again as illustrated in scenario one. If an available case analysis is conducted, it is essential to examine reasons for data being missing. If the fact that an observation is missing is unrelated to both the observed and the unobserved data, the missing data are said to be missing completely at random (MCAR). By examining reasons for missingness (if possible) it may become clear that data are not MCAR but that they are missing because of reasons related to the treatments and that these reasons may differ systematically by treatment (as illustrated in scenario one). If there is not much missing data, an available case analysis with a valid assumption of data being missing completely at random, may be unbiased (i.e. it does not over estimate or under estimate a treatment difference or evidence of association), but will have lower power to detect a difference or association than if all data were present.⁵ Less data equates to less information which in turns equates to less chance of being sure that if you fail to find a significant difference it is because there truly is no difference.⁶

Clearly there are situations where there is no information about those who are missing. In such cases, we would recommend drawing attention to the presence of missing data and the fact that it was not possible to investigate further. By doing this, readers are aware of the potential for bias. Best and worst case scenarios could be considered to show how conclusions might have differed under such circumstances. (e.g. In a study comparing drugs A and B, and where the primary outcome is treatment success, a best case scenario might be that all those lost to follow-up on treatment A were successes whilst all those on treatment B were failures. A worst case scenario would reverse these assumptions.)

An alternative to "available case analysis", where subjects with missing data are simply omitted from the analysis, is to impute data. Imputation replaces missing data with some plausible value predicted from that subject (or other subject's) data. One method of imputation which is common place in ophthalmic literature is that of "last observation carried forward" (LOCF), where any missing data is replaced with the last observed value for that patient.

Scenario two

A senior colleague draws my attention to a paper which has used LOCF. It is a paper published in a highly regarded journal. The primary outcome is visual acuity at one year after randomisation and it compares two treatments for Age Related Macular Degeneration. One hundred patients were randomised, three did not provide a measure at one year, 1 in one arm of the study and 2 in the other. The authors have looked at the reasons for the subjects being unable to provide data at one year and are satisfied that there is little to suggest that the data are not missing completely at random. Despite the proportion of data being lost to follow-up being small and within the margin expected in the sample size calculation and despite there being no overt evidence of data being anything other than missing completely at random, the authors have used LOCF. They state that this is essential in order to conduct a true Intent to treat (ITT) analysis – where all randomised patients are included. ITT is, the authors state, essential in order to preserve the benefits of randomisation and protect against bias.⁷ The paper says that the patient who withdrew from treatment arm A actually did so prior to receiving any treatment. This means that the only available observation of visual acuity was that at baseline. The two patients who withdrew from treatment arm B, moved out of the area but did attend at 11 months. Whilst it seems entirely acceptable to use the 11 month data for treatment arm B, the use of baseline data for treatment arm A seems very tenuous indeed and note that our definition of valid imputation was “a plausible” value. The authors did, however, compare their imputed analysis with the available case analysis and found little difference between results other than a slightly lower standard error for the effect estimate with the LOCF assumption, and so in this scenario, whilst the assumptions made do not appear to be sensible, the conclusion drawn from the LOCF analysis is similar to that which would have been drawn from the available case analysis.

Whilst LOCF is widely used and indeed required by the Food and Drug Administration (FDA) in the United States, it has serious and in some cases, fundamental problems.⁹

An alternative to LOCF is “simple mean imputation” replacing missing data with the average value observed in that treatment arm. This is not ideal because if there are many patients

with missing data, giving them all the same mean, will reduce variability between observations and suggest more confidence in findings than should be drawn.

Scenario three

A senior colleague draws my attention to a study comparing visual acuity in 100 patients with diabetic macular oedema after treatment with either drug A or drug B. One eye only has been included and results are presented with imputation having been conducted by replacing missing data with the average observed in each treatment arm. Five patients were missing at one year after treatment with drug A but no patients were lost to follow up from drug B. The study concludes that there is evidence that A is better than B. Something worries me however. I look at the available case analysis and it suggests that there is no evidence of a difference between A and B. I now have disagreement between the available case analysis and the analysis which imputed for missing subjects. Although the authors have not commented upon this discrepancy I now see that it highlights how assumptions made about missingness and strategies to deal with missingness have potential to mislead.

The scenarios presented thus far illustrate cases where subjects are missing at final follow-up, yet clearly missing data presents challenges to researchers in other ways

- 1) A validated questionnaire is used with a scoring algorithm provided for computing summary scores based on answers to all questions on the questionnaire. You determine that some subjects simply have not answered some of the questions. There is some data, but not all.
- 2) Many studies involve assessment of the eye by imaging equipment, such as OCT. Some of this equipment may be very expensive. No technology is immune from failure and there may be times during a trial where the equipment fails – subjects don't therefore have assessments at particular visits in the trial schedule.
- 3) Postal questionnaires – not all individuals respond despite several attempts to encourage postal return.

Whilst examples discussed here relate to tightly controlled clinical trials, it is evident that missing data is likely to be more of an issue outside the rigour of a randomised trial e.g. electronic patient records, observational studies.

The strategies presented here for missing data are simple, yet many better methods are now well described in the statistical literature – multiple imputation and model based approaches such as mixed models and weighted generalized estimating equations exist.⁸ A word of caution is provided by Streiner, however

“the easy methods are not good and the good ones ... are not easy”.⁹

Whatever approach is adopted, missing data are “what it says on the tin”, “missing”, and as eloquently summarised by Bland and Altman “there is no satisfactory solution to this”.³ Greater efforts should be made at the design stage to limit the likelihood of data being missing and one simple yet very rewarding approach can be to talk to patients in advance of conducting a study. At a thyroid eye disease patient day organised by the NIHR Biomedical Research Centre for Ophthalmology patients were quite vocal about the need for researchers to carefully consider treatment schedules when designing studies. Patients with thyroid eye disease may find their condition leads to fatigue and disfigurement, resulting in their not wanting to venture out of the house – a trial requiring monthly visits, when standard practise is 6 monthly is likely to suffer recruitment and retention issues.¹⁰

Ensuring that everyone involved in research understands the potential for missing data to undermine the scientific integrity of the study and ultimately do a disservice to patients and public might also be a simple yet rewarding approach.

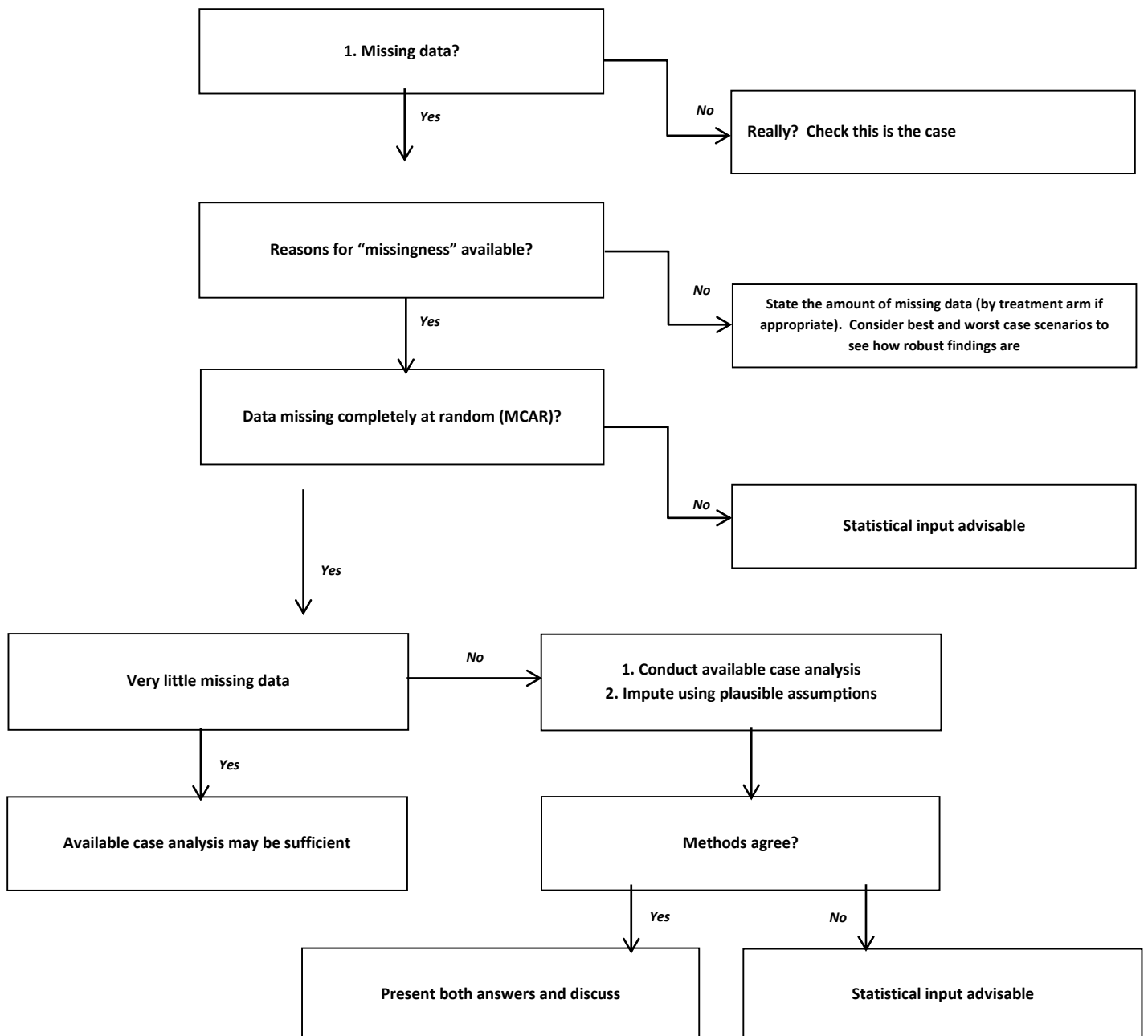
Some missing data is however inevitable: patients are human and humans get ill; go on holiday; look after sick children; drop out of studies; machines fail; post can go missing. Missing data undermines internal validity and causes loss of power and simple methods of accounting for missing data can produce biased estimates of the treatment effect. Data will be missing for a reason and researchers are strongly encouraged to record why a value is missing. This paper hopefully highlights the need to be explicit in relation to the potential impact of missing outcome data and outlines some helpful strategies to consider when this does occur. Available case analysis used to be the standard and whilst there are indeed occasions where this isn't helpful, simple imputation can also lead to erroneous conclusions. The final comment we leave with Streiner

“the solution is to consult with a statistician; most of them are (relatively) friendly”.⁹

Lessons learnt

- **Prevention is best, even in relation to missing data.**
- **Report missing data where it occurs and reasons for missingness wherever possible.**
- **Statistical methods do exist for handling missing data, but assumptions made by such methods must be rigorously evaluated.**

Figure 1. Flow diagram of missing data



References

1. [No authors listed] Definition of levels of evidence (LoE) and overall strength of evidence (SoE). *Global Spine J.* 2015 Jun;5(3):262. doi: 10.1055/s-0035-1556954.
2. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. for the CONSORT Group. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine.* 2001;134:663-94
3. Altman DG, Bland JM. Missing data. *BMJ.* 2007 Feb 24;334(7590):424.
4. McPherson GC, Campbell MK, Elbourne DR. Use of randomisation in clinical trials: a survey of UK practice. *Trials.* 2012 Oct 26;13:198
5. Bell ML, Fairclough DL Practical and statistical issues in missing data for longitudinal patient-reported outcomes. *Stat Methods Med Res.* 2014 Oct;23(5):440-59.
6. Bunce C, Patel KV, Xing W, Freemantle N, Doré CJ; Ophthalmic Statistics Group.
Ophthalmic statistics note 2: absence of evidence is not evidence of absence. *Br J Ophthalmol.* 2014 May;98(5):703-5
7. Streiner DL. Statistics Commentary Series: Commentary #10-Dealing With Drop-Outs.
J Clin Psychopharmacol. 2015 Mar 20
8. White IR, Horton NJ, Carpenter J, Pocock SJ Strategy for intention to treat analysis in randomised trials with missing outcome data.*BMJ.* 2011 Feb 7;342:d40
9. Streiner DJ Statistics commentary series: commentary #3--last observation carried forward. *J Clin Psychopharmacol.* 2014 Aug;34(4):423-5.
10. Smith HB, Porteous C, Bunce C, Bonstein K, Hickey J, Dayan CM, Adams G, Rose GE, Ezra DG. Description and evaluation of the first national patient and public involvement day for thyroid eye disease in the United Kingdom. *Thyroid.* 2014 Sep;24(9):1400-6.