Conference Report

CLINICAL INVESTIGATION

For reprint orders, please contact reprints@future-science.com

Missing Data in Clinical Trials Forum

Terry L Katz
Merck Animal Health, 2 Giralda Farms, Madison, NJ 07940, USA
Tel.: +1 973 937 5235;
Fax: +1 866 335 8370;
terry.katz@merck.com

Missing Data in Clinical Trials Forum
Loews Philadelphia Hotel, Philadelphia, PA, USA, 16–17 March 2015

Missing data, defined as values not available that would have been meaningful for analysis if they were observed, are commonplace in academic, government and industry run clinical trials. However, this data loss can result in an inadequate basis for study inferences. In 2014, multiple US FDA advisory committees challenged the quantity of missing data and the conclusions drawn from their study outcomes, resulting in approval delays or rejection of product registration. The Missing Data in Clinical Trials Forum invited experts from academia, government and industry to explore methods to avoid missing data by design, and present on tools to detect and rectify missing data during the trial. Statistical methodology to impute or model data missing at random, or not at random, was also discussed.

Keywords: avoidance • detection • effectiveness • efficacy • estimand • missing data

The conference was opened by co-chair Terry Katz (Merck Animal Health, NJ, USA) who started with a three-part definition of missing: unable to be found, not in the usual or expected place, or needed or expected but not included [1]. The optimal approach was to determine how to minimize missing data and how to detect missing data in a dataset. As an example, a numerical sequence with a single missing value was provided to the participants: 2, 5, 8, 11, 13, [blank], 30, 34, 40. The attendees first suggested leaving it blank, followed by last observation carried forward (LOCF), resulting in a second observation of '13'. Katz then used last observation carried backwards since it was equally valid (or weak) as LOCF, resulting in two observations of '30'. The audience added baseline observation carried forward, in this case '2', for consideration. Interpolation, using the midpoint between the two bracketing observations, resulted in 21.5, and the arithmetic mean resulted in 17.875. With a high degree of discordance among the choices, the audience agreed that the best method to obtain the accurate value was to query the data source for the missing value. The real value, 15, did not equal any of the presented mathematical options, as this particular dataset was not random, and corresponded to the city street-named train stops for the SEPTA Elevated-Subway line traveling underneath the conference hotel.

Current state of missing data mitigation & its impact on life science companies

Keynote speaker Bill Potter (Senior Advisor to the Director, NIH, National Institute of Mental Health, Bethesda, MD, USA) challenged the audience on 'How important is investment to minimize missing data?' Missing data can undermine causal conclusions from trials, and analytical methods cannot always compensate. But avoiding missing data can be expensive, and it is never specifically budgeted in academic trials and variably addressed for industry studies.

FUTURE SCIENCE part of fsg

Some amount of missing data can be tolerated and ignored, but defining a threshold is a challenge. High dropout rates in some therapeutic areas, such as psychology trials, used to be addressed by enrolling many extra patients to have a sufficient power to get statistical significance. With 'big data' and the new methods of pooling, this tendency has been reduced. But Clinical is still dependent on the availability of an appropriate patient population, and this is hindered by a lack of clarity of the basis of data for defining such a population. Previous studies were often flawed with a high proportion of missing data, and there is a lack of background subject level data in the public domain.

Dr Potter used an example of cardiovascular deaths in the young [2], which referenced the Case Reporting System of the National Center for the Review and Prevention of Child Death. A total of 16 states participated in working on elements to build a final analytical dataset by reclassifying, recategorizing and deriving new variables from existing data. Missing data comprised a mean of 41.7% for most key variables!

Go/No-Go studies are designed to 'tolerate' the previously observed dropout and loss to follow-up rates. But these studies have a risk of missing infrequent toxicities and make it difficult to interpret novel secondary measures or subgroups. This is especially important in NIH trials where nontraditional outcomes are explored. While new clinical trials may be sufficient for registration, NIH also wants to clearly differentiate one product from another, which is challenging when the evidence contains missing data.

Publication bias and publications with high proportions of missing data are problematic. Dr Potter examined the ticagrelor [3] and rivaroxaban [4] claims as examples of studies with key variables missing. The author referenced the Cochran Assessment [5] using an example by Taylor [6] as a key resource to defining if a manuscript has met sufficient quality measurements for publication, focusing on the subtopics of 'incomplete outcome' and 'selective reporting' as being relevant to the missing data discussion.

Optimizing trial design & monitoring to minimize occurrence of missing data

Conference co-chair Carol Robertson-Plouch (Eli Lilly, IN, USA) opened this topic with a need to design the study to minimize the occurrence of missing data. The author emphasized that avoidance of missing removes the need for analysis-based data modeling, which the author presented as "building a better barndoor instead of letting the Statistical Horses out!"

Plouch was concerned with early discontinuation from loss to follow-up, safety issues or lack of efficacy. Missing primary efficacy/safety data are most prominent, but also incomplete diaries, missing vital signs or missing visit information. Later the author presented examples of sponsors being called to 2013–2014 US FDA advisory boards in part connected to missing data. Afrezza (Mannkind/Sanofi) had challenges on adherence to protocol and high dropout/missing data rates. Bronchitol (Pharmaxis) was challenged as having 'informatively missing data' and Anoro Ellipta (GSK) for patients discontinuing early compared with the placebo group. Missing data were also challenged for Farydak (Novartis), Northera (Chelsea), Oral Testosterone (Clarus) and Xarelto (Janssen).

Celestine Hicks (AstraZeneca, DE, USA), as part of her presentation on the Regulatory Landscape, discussed the interaction with FDA as an opportunity to 'fess-up' to a missing component. To lessen the likelihood of a major missing component for a late-stage breakthrough therapy, Hicks performs data sweeps for the data monitoring committee (DMC) and routine updates to FDA.

Patrick Zbyszewski (Onconova, NJ, USA) discussed on-going data collection and monitoring in preparation for interim analyses for safety, efficacy or futility. For events that occurred after the data cut off, but before the DMC, the author provides both the 'cleaned' data and the new events and safety findings for full transparency to the DMC. The author creates a projected visits report based on the randomization date and last visit to estimate the next visit, and uses color-coding to facilitate tracking.

Conference co-chair Terry Katz (Merck Animal Health, NJ, USA) looked at classical missing data as a 'gap' that has a potential for being filled by referencing the other puzzle pieces. Drop-outs, however, are a truncated bridge-to-nowhere where the potential path for that subject is unpredictable. Detection of missing, and the timing when the missing occurs, can be crucial. A missing chess piece before the game starts is obvious, but a card missing from a deck is hidden and often not detected until the game has been played. Missing should not be limited to the response variables, and there is a need to look at explanatory variables including strata, covariates and exposure.

The mechanism underlying why a value was not collected was less critical for data management than designing tools to minimize missing data. A visit missed for cause (patient ill with progressive disease or adverse effect) as opposed to a visit missed for an overabundant snowfall has different interpretation for the statistical/clinical analysis, but for data management both are noncollected data. As a contrast, 'lost data' were collected data that could have been innocently lost (such as paper case report forms [CRF] misplaced or data entry to wrong patient), or excluded for protocol

deviations or not usable since the collected blood was hemolyzed. 'Loss-to-follow-up' could be related to the patient feeling too ill to return for an examination visit, or feeling too well to return for an examination visit or a simple move to a new location. Missing explanatory variables, measurement of the wrong population, use of a weak surrogate and having weak *a priori* information such that the optimum T_{max} was not a planned timepoint, were other types of missing discussed.

Candace Shelton (Celerion, PA, USA) concentrated on early stage data, and the change from paper CRF to electronic systems, eSource and patient reported outcomes. The author follows a 'fit-for-purpose' approach by SCDM [7] to capture data where and when first generated, and leverage automated quality checks. Phase 1 units typically have repetitive events with standardized data collection regardless of sponsor. Data transfer is another potential pitfall and the author recommends having a data transfer plan, a reconciliation plan, test transfers and treating every subsequent transfer as if it was the first.

Collectively, these speakers offered many strategies to reduce missing data. Protocol simplification can reduce issues by limiting the number of patient visits to the Investigator, with reasonably wide visit windows. Data collection is aided by short trials, use of Investigator sites with historically low missing rates, incentives for patients and Investigators and noncluttered standardized (CRF) with a minimal number of data values. Reminders to the patient by the Investigator (phone call, email, appointment cards, web calendar) increase the likelihood of a patient remembering to come on time, and a quick call to a patient's cell phone if they are late may be able to salvage the visit within the applicable window.

Various methods to check the dataset for missing were demonstrated, including electronic CRFs and programmatic checks. Looking for a blank in a CRF page or common listing was shown to be less effective than tabulations with missing cells highlighted by a large red mark. An effective electronic data capture 'pop-up' window when a variable was skipped or entered outside the acceptable range enabled the data enterer a chance to correct immediately after typing. For greatest efficiency, the investigator should collect data bed-side so that missing or odd values can be rechecked while the investigator is still with the patient. Dashboards and graphics could show missing values and a novel filter was displayed which showed the patient IDs with missing or partial data.

Emerging technology for sample tracking, patient recruitment & retention

Three interrelated presentations were made by Maria Minasian (Bristol-Myers Squibb, NJ, USA), Taisa Skubiak (Bristol-Myers Squibb, NJ, USA) and Lori Post (Yale University School of Medicine, CT, USA).

Minasian looked at current methods to track patient samples, and expressed that they were insufficient to prevent delays or loss of samples. These samples can include blood, urine or tissue. Often, loss of a sample results in a missing value for a key efficacy parameter, resulting in a lower quality dataset. UPC bar-coding to identify each sample was the first step, followed by harvest logs and chain-of-custody tracking in a spreadsheet to link to the unique patient ID, sampling site and investigator. Innovative advanced tracking tools linked to the electronic data capture system enabled real-time automatic updates and query resolution.

Skubiak mentioned that 80% of sites fail to meet enrollment timelines, either by not reaching their enrollment commitments or reaching enrollment goals well after the scheduled date. The author quoted the Tuffs Center for Study of Drug Development [8] that two-thirds of sites fail to meet enrollment goals, and in 2010, less than one in four screened patients were retained for the duration of the clinical trial. Skubiak referenced many articles to show a pattern of increasingly complex protocols burdening patients with a large number of procedures, resulting in a more limited pool of interested patients for enrolling or fulfilling the trial obligations. Reducing a protocol to the minimum number of procedures and time points would aid recruitment/retention and avoid the cost to the sponsor for delays and patient loss.

Decentralized recruitment had twice the industry money investment (in 2012) than centralized recruitment. Decentralized includes physician referrals, broadcasted newspaper ads, research center posters and internet/television ads. Centralized recruitment concentrates more on targeted review of electronic medical records and hospital grand rounds than media advertisements. Newer approaches, such as social media and internet searches, are reaching 60% of the potential patients. Using an internet questionnaire provides an initial layer of screening without the sponsor, investigator or patient expending much time. This allows a directed follow-up for patients passing the first screening as an effective, low cost approach.

Skubiak showed a retention checklist that effectively presents the sponsor and investigator with approaches to keep the patient engaged. Seeing the patient as soon as they arrive, using appointment cards, reminder calls and emails and coordinating reimbursed transportation were all presented strategies. Confirming phones numbers and email at each visit reduces the likelihood of the patient moving to a new locale without warning. A checklist section for red flags of a patient withdrawing was also presented.

Dr Post went deeper into patient recruiting and retention strategies, including setting a minimum threshold of >90% recruitment. The author covered one of the more controversial topics by using highly trained Hispanic recruiters in a region that had many Hispanic patients. The ability to relate to a similar heritage can be a necessity, though at times the author uses a complementary recruiter when the patient pool is known to verbally open-up to an opposite, such as a gay man opening up more to a woman. Hire-for-content was the author's mantra by using clean cut workers with the elderly and young tattooed workers for drug intervention studies. Regardless of study, sponsors need to hire 'good employees' who are sensitive, passionate, resilient and dedicated. Extensive training on the subject matter and protocols, using role play and piloting with real patients with observers, is critical to teach recruiters how to act when something goes wrong, and how to interact with the healthcare team without interfering.

Statistical modeling for missing

Craig Mallinckrodt (Eli Lilly, IN, USA) defined two general categories of estimands, defined as what is trying to be estimated. Efficacy is the demonstration of clinical benefit when the drug product is used as directed; also called *de jure* benefit, and most aligned with confirmatory clinical trials. Effectiveness is the demonstration of clinical benefit with the drug product as it was actually taken; also called *de facto* benefit, which is best tested in more naturalistic settings [9]. These can have different purposes, such as drug labeing (*de jure*/efficacy) versus public health (*de facto*/effectiveness). For example, missing data caused by the use of rescue medications after a drug failure are aligned to an effectiveness trial as it mimics real-life scenarios. Choice of estimand can result in the analysis focusing on a different time point for the endpoint, which can control the effect of patient nonadherence to the protocol.

Missing data mechanisms were defined as missing completely at random (MCAR) where neither observed nor unobserved outcomes of the dependent variable explain dropouts, missing at random (MAR) where observed outcomes explain dropouts but unobserved do not, and missing not at random where both observed and unobserved outcomes of the dependent variable explain dropouts. MAR is often assumed in clinical trials since it provides an unbiased estimate of the missing data.

Roderick Little (University of Michigan, MI, USA) discussed the statistical handling of missing data as recommended by the National Research Council [10]. Complete-case analysis, which deletes cases with missing data, is generally inappropriate for a regulatory setting since it essentially requires MCAR assumptions. Single imputation methods like LOCF are valid under missing not at random, but not with MCAR or MAR, and may be anticonservative, so they are not recommended by the NRC panel. Preferred methods include inverse probability-weighted methods which assigns a missingness weight to the complete cases to make them more representative of all cases. Augmented inverse probability-weighted adds model-based residuals for robustness to model misspecification. Multiple imputation, including weighted general estimating equations creates multiple filled-in datasets using draws from predictive distribution and applies combining rules. Analysis models may use fewer variables than the full imputation models. Sensitivity analyses using pattern-mixture models, consisting of repeating the inference at different plausible values, are easy to interpret and explain to clinicians.

Conclusion

Missing Data in Clinical Trials Forum brought together experts supporting pharmaceutical, academic and government clinical trials in the USA. General consensus was to design a trial to avoid missing data. Simpler protocols, well-trained recruiters and actions to avoid missed visits or dropouts led to higher patient retention and protocol compliance. When a missing data threshold was reached where imputation was necessary, multiple imputation with pattern-mixture models was far better than using simplistic methodology like LOCF.

Financial & competing interests disclosure

The author has no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1 An Encyclopædia Britannica Company . www.merriam-webster.com/dictionary/missing

2 Vetter V, Dugan N, Haley D *et al.* Development of a data set of national cardiovascular deaths in the young. *Am. Heart J.* 168(4), 568–576 (2014).

3 DiNicolantonio J, Tomek A. Misrepresentation of vital status follow-up: challenging the integrity of the PLATO trial and the claimed mortality benefit of ticagrelor versus clopidogrel. *Int. J. Cardiol.* 169(2), 145–146 (2013).

4 Krantz M, Kaul S. The ATLAS ACS 2-TIMI 51 trial and the burden of missing data: (Anti-Xa Therapy to

Lower Cardiovascular Events in Addition to Standard Therapy in Subjects With Acute Coronary Syndrome ACS 2-Thrombolysis In Myocardial Infarction 51) *J. Am. Coll. Cardiol.* 62(9), 777–781 (2013).

5 Cochrane Database of Systemic Reviews. http://community.cochrane.org

6 Taylor F, Huffman M, Macedo A *et al.* Statins for the primary prevention of cardiovascular diseases. *Cochrane Database Syst. Rev.* 1, CD004816 (2013).

7 Society for Clinical Data Management. eSource implementation in clinical research: a data management perspective. www.scdm.org/sitecore

8 Tufts Center for the Study of Drug Development. Drug developers actively improving efficiency of clinical trials. csdd.tufts.edu/news/complete_story/rd_pr_apr_2011

9 Mallinckrodt C, Kenward M. Conceptual considerations regarding endpoints, hypotheses, and analyses for incomplete longitudinal clinical trial data. *Ther. Innov. Regul. Sci.* 43(4), 449–458 (2009).

10 Little R, D'Agostino R, Cohen M *et al.* The prevention and treatment of missing data in clinical trials. *N. Engl. J. Med.* 367(14), 1355–1360 (2012).