Multiple Imputation Approaches for Missing Time-to-Event Outcomes with Informative Censoring: Practical Considerations from a Simulation Study Based on Real Data

Andrea Bellavia[(1)] [iD] , Min Guo[(2)] [iD] , Sabina A. Murphy[(1)] [iD]

(1) TIMI Study Group, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
(2) Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, United States

CORRESPONDING AUTHOR: Andrea Bellavia, PhD, TIMI Study Group, Brigham and Women's Hospital Department of Medicine, Harvard Medical School, 60 Fenwood Road, Suite 7022, Boston, MA 02115. e-mail: abellavia@bwh.harvard.edu

SUMMARY

Missing outcomes data represent a common threat to the validity and robustness of clinical trials with time-to-event outcomes. Recent extensions of multiple imputations (MI), namely controlled-MI, have been introduced as a viable approach for sensitivity analysis in the presence of informative censoring, yet they lack validation based on real data. In this study we used data from a randomized trial to generate realistic scenarios of censoring mechanisms and compare several imputation approaches for missing outcome data. Our results confirm the relevance of multiple imputations especially in studies with long follow-up and higher proportion of potentially informative censoring.

Keywords: missing data; multiple imputations; survival analysis; informative censoring.

Missing outcomes data represent a common threat to the validity and robustness of clinical trials and prospective epidemiologic studies with time-to-event outcomes. The most common reason for outcome missingness in these contexts is censoring, which generally occurs when patients are lost to follow-up or withdraw consent from an ongoing clinical study. Several studies have outlined the importance of critically evaluating missing outcome data in clinical studies, as well as the relevance of multiple imputations (MI) in this context, and researchers are expected to justify their handling of missing data [1-4]. Statistical approaches commonly used to report primary results in prospective studies, such as the non-parametric Kaplan-Meier estimator or the Cox regression model, generally provide valid estimates of the effect of interest under the assumption that censoring mechanisms are unrelated to either exposure or outcome. This assumption, commonly referred to as non-informative censoring, corresponds to an assumption of missingness at random for missing outcome data. There are several settings, however, where this assumption might not be met. In a clinical

trial, for example, occurrence of adverse or safety events could lead to drug discontinuation and potential study withdrawal. If the occurrence of a given adverse event is more likely to occur for a specific treatment regime, this would lead to a higher risk of censoring in a specific study arm (i.e. informative censoring). Recent MI extensions, namely controlled-MI, have been introduced as a viable alternative for sensitivity analysis in the presence of informative censoring [5]. In particular, γ-imputation provides a flexible tool that allows incorporating existing knowledge on the potential censoring mechanisms to evaluate how the treatment effect would change under different scenarios [6]. In brief, γ-imputation relaxes the non-informative censoring assumption of the Cox model by including a γ term, corresponding to the log-HR comparing censored and uncensored individuals, and then uses MI to impute censored observations based on this modified Cox model [7].

Existing literature and recommendations for missing outcome data are largely based on theoretical considerations and simulation studies alone. Simulations

are often conducted based on general assumptions for the censoring distributions and study population, and limited information from real data is available for clinical researchers and epidemiologists. To such end, in this study we used data from the DECLARE-TIMI 58 trial [8], a randomized, double-blind, multinational, placebo-controlled, phase 3 trial, to generate realistic scenarios of potential censoring mechanisms based on real data. This large clinical study (n=17,160) experienced very limited censoring due to withdraw consent or loss to follow-up (<2%), and we therefore evaluated it as the real target population. The detailed information on the occurrence of safety events and drug discontinuation collected was used to generate fictitious alternative settings of censoring that are likely to occur in real settings. First, we evaluated a potential scenario where all adverse events that led to drug discontinuation (affecting ~7.5% of participants) also led to trial follow-up discontinuation and therefore censoring. As several of these adverse events are associated with treatment,

such censoring would be informative. Next, under the same censoring mechanisms and trial characteristics, we generated additional fictitious trials gradually increasing the proportion of censored individuals, the length of follow-up, and the proportion of trial participants experiencing the primary event. We then applied MI and γ-based controlled-MI to impute censored observations under all settings, comparing bootstrapped HRs from imputed data to the HRs from the complete data reported in the study. When using controlled-MI, γ values are provided by the user. In our example we used reported information on adverse events to define individual-specific γ values as the inverse of the log(HR) reported in the trial safety results. In practice, when the effect of the main exposure (e.g. treatment) on censoring is unknown, it is recommended to evaluate a range of potential γ values based on clinical background and a-priori hypothesis.

Table 1 summarizes results from the different evaluated settings. Results indicate that both standard

Table 1. Estimates of treatment effects (dapagliflozin vs placebo) for primary endpoints of DECLARE-TIMI 58 using multiple imputations and controlled multiple imputations, under different replications of the trial data summarizing settings with informative censoring

		MACE			HHF/CVD	
Original trial DATA[a]		0.93 (0.85, 1.03)			0.83 (0.73, 0.95)	
		MI	**Controlled-MI[c]**		**MI**	**Controlled-MI**
Modified trial DATA[b]		0.924 (0.834, 1.023)	0.937 (0.853, 1.030)		0.833 (0.719, 0.964)	0.824 (0.710, 0.955)
Simulated trial DATA[d]	**Real HR**			**Real HR**		
A: Trial data with higher censoring proportion	~15%	0.931	0.932 (0.842, 1.033)	0.928 (0.839, 1.027)	0.826	0.821 (0.739, 0.912)	0.817 (0.736, 0.908)
	~30%	0.962	0.970 (0.872, 1.08)	0.956 (0.859, 1.065)	0.829	0.832 (0.745, 0.93)	0.823 (0.738, 0.917)
B: Trial data with longer follow-up	5y	0.936	0.932 (0.851, 1.022)	0.931 (0.849, 1.02)	0.826	0.821 (0.739, 0.912)	0.817 (0.736, 0.908)
	10y	0.936	0.947 (0.882, 1.016)	0.939 (0.876, 1.006)	0.847	0.866 (0.772, 0.972)	0.852 (0.755, 0.96)
C: Trial data with higher events proportion	~20%	0.925	0.926 (0.837, 1.024)	0.924 (0.836, 1.022)	0.841	0.845 (0.761, 0.937)	0.844 (0.761, 0.936)
	~50%	0.917	0.913 (0.826, 1.01)	0.911 (0.823, 1.007)	0.824	0.827 (0.745, 0.917)	0.825 (0.744, 0.915)

MACE=Major Adverse Cardiovascular Events; HHF/CVD= hospitalization for heart failure or cardiovascular death

[a] Original data without imputations

[b] All participants experiencing drug discontinuation due to adverse event (~7.5 %) are censored on the day of the adverse event

[c] For individuals experience adverse events associated with treatment (potential informative censoring), gamma values are defined as the inverse of the log(HR) for that adverse event. For all other censored individuals, gamma=0 (i.e. non-informative censoring)

[d] Results based on 10 replications of the trial data under the different detailed setting. The table shows the average real HRs from the 10 replications and the average HRs estimated from MI and controlled-MI

MI and controlled-MI provide robust inference under all evaluated scenarios, with slightly improved performances of controlled-MI only with heavy censoring (~30%) and longer follow-up (10 years). These results confirm the relevance of multiple imputations approaches for missing outcome data in prospective studies. We join previous researchers in recommending their inclusion as sensitivity analyses in study protocols and statistical analysis plans, especially in studies with long follow-up and higher proportion of potentially informative censoring. Controlled-MI, and in particular γ-imputation, can be used to assess the extent to which potential informative censoring would affect primary trial results, thus providing a powerful tool for sensitivity analysis for prospective studies in the presence of missing outcome data and censoring for which reasons are unknown.

To our knowledge, this is the first study comparing MI and controlled-MI approaches for missing outcome data using extensive simulated scenarios that are based on real data and censoring mechanisms occurring in applied clinical and epidemiologic research. By generating fictitious replications of a complete clinical study and using observed information on adverse events mechanisms to generate censoring, this study can help practitioners appreciate the advantages of different imputation approaches under realistic settings in prospective studies in clinical epidemiology.

DISCLAIMER

FUNDING

TRIAL REGISTRATION

ClinicalTrials.gov NCT01730534.

REFERENCES

1. Bell ML, Fiero M, Horton NJ, Hsu CH. Handling missing data in RCTs; a review of the top medical journals. BMC Med Res Methodol. 2014;14:118.
2. The Prevention and Treatment of Missing Data in Clinical Trials. Washington (DC)2010.
3. Groenwold RH, Donders AR, Roes KC, Harrell FE, Jr., Moons KG. Dealing with missing outcome data in randomized trials and observational studies. Am J Epidemiol. 2012;175(3):210–7.
4. Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med. 2009;3(2):e51–3.
5. Tan PT, Cro S, Van Vogt E, Szigeti M, Cornelius VR. A review of the use of controlled multiple imputation in randomised controlled trials with missing outcome data. BMC Med Res Methodol. 2021;21(1):72.
6. Jackson D, White IR, Seaman S, Evans H, Baisley K, Carpenter J. Relaxing the independent censoring assumption in the Cox proportional hazards model using multiple imputation. Stat Med. 2014;33(27):4681–94.
7. Cro S, Morris TP, Kenward MG, Carpenter JR. Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: A practical guide. Stat Med. 2020;39(21):2815–42.
8. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, et al. Dapagliflozin and Cardiovascular Outcomes in Type 2 Diabetes. N Engl J Med. 2019;380(4):347–57.