

What you don't know can hurt you: a statistical commentary on missing data in trauma research

James P Byrne ,¹ Molly P Jarman ²

To cite: Byrne JP, Jarman MP. What you don't know can hurt you: a statistical commentary on missing data in trauma research. *Trauma Surg Acute Care Open* 2024;**9**:e001405. doi:10.1136/tsaco-2024-001405

Missing data are ubiquitous in clinical research.¹ This is true in trauma research, whether in a prospective trial or a retrospective database study. Unfortunately, missing data are often handled improperly or outright ignored, with detrimental consequences for the accuracy of study results. The problem arises when the likelihood of missingness to key variables is associated with other observed or unobserved variables. For example, if missing data are more common among patients with trauma with high mortality risk, simple case deletion will introduce a systematic error or bias. Conversely, missingness might obscure important associations that support or refute the proposed causal mechanism being studied.

Donohue *et al* directly confront this problem in their secondary analysis of the Study of Tranexamic Acid During Air Medical and Ground Prehospital Transport (STAAMP) trial.² Tranexamic acid (TXA) administration is associated with survival in injured patients at risk of hemorrhage.^{3,4} However, no trials have found an association between TXA and improved coagulation as measured by thromboelastography (TEG)—the proposed mechanism for this effect. Donohue *et al* examine the role of missing TEG data. Patients with missing TEG were significantly more likely to present in severe shock (SBP < 70 mm Hg) and go on to die (36% vs 7%), reflecting the logistical challenges of sampling patients in extremis. On analysis of patients in the severe shock subgroup with TEG collected, TXA was significantly associated with decreased LY30, indicating improved clot stability. These findings support the proposed mechanism behind the survival benefit of TXA among patients with severe shock in the STAAMP trial. Moreover, they provide a case study of how missing data hinders the interpretability of clinical studies.

When confronted with missing data, investigators should perform exploratory analyses to understand what pattern of censoring exists. There are typically three patterns.¹ When data are missing completely at random (MCAR), the probability that a value is censored is completely independent of both observed and unobserved characteristics. Using the example of the STAAMP trial, this would occur if TEG parameters were missing due to sporadic laboratory errors. When data are missing at random (MAR), missing values are systematically related to observed, but not unobserved, characteristics. This would occur if the patient's degree of extremis is related to the probability that TEG is not collected, rather than the TEG parameters themselves (eg, LY30). Finally, data missing not at random (MNAR)

are systematically censored in a pattern related to the unobserved characteristics. This would occur if patients with abnormal LY30 (the unobserved characteristic) were less likely to undergo TEG collection due to challenges in the setting of severe shock. Missing data are rarely MCAR. Therefore, case-wise deletion of observations with censored values is rarely appropriate. Multiple imputation (MI) is a methodology that estimates missing values based on existing correlations between known values and other observed variables.⁵ This facilitates analyses using all known data points while also accounting for the uncertainty in unknown data points, thus avoiding overly precise estimates. When data are MAR or MNAR, MI is felt to introduce less bias than more naïve methods and is the approach used by the Trauma Quality Improvement Program of the American College of Surgeons.⁶

In summary, missing data are common in trauma research. Investigators should proactively seek to understand the pattern of missingness in their data and report a statistical approach to handling it to minimize bias. Thoughtful analyses, as Donohue *et al* have done, can help to illuminate the impact of missing data on clinical observations.

Contributors I have no conflicts of interest.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; internally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

James P Byrne <http://orcid.org/0000-0002-3760-8509>
Molly P Jarman <http://orcid.org/0000-0001-7117-5869>

REFERENCES

- Haukoos JS, Newgard CD. Advanced Statistics: missing data in clinical research—part 1: an introduction and conceptual framework. *Acad Emerg Med* 2007;**14**:662–8.
- Donohue JK, Iyanna N, Lorence JM, Brown JB, Guyette FX, Eastridge BJ, Nirula R, Verccruyse GA, O'Keeffe T, Joseph B, *et al*. Missingness matters: A secondary analysis of thromboelastography measurements from A recent prehospital randomized tranexamic acid clinical trial. *Trauma Surg Acute Care Open* 2024;**9**:e001346.



► <https://doi.org/10.1136/tsaco-2023-001346>

© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Johns Hopkins University, Baltimore, Maryland, USA
²Center for Surgery and Public Health, Brigham and Women's Hospital, Boston, Massachusetts, USA

Correspondence to

Dr James P Byrne; jbyrne13@jh.edu



- 3 Guyette FX, Brown JB, Zenati MS, Early-Young BJ, Adams PW, Eastridge BJ, Nirula R, Vercruyse GA, O'Keeffe T, Joseph B, *et al*. Tranexamic acid during Prehospital transport in patients at risk for hemorrhage after injury: A double-blind, placebo-controlled, randomized clinical trial. *JAMA Surg* 2020;156:11–20.
- 4 Roberts I, Shakur H, Coats T, Hunt B, Balogun E, Barnetson L, Cook L, Kawahara T, Perel P, Prieto-Merino D, *et al*. The CRASH-2 trial: a randomised controlled trial and economic evaluation of the effects of tranexamic acid on death, vascular occlusive events and transfusion requirement in bleeding trauma patients. *Health Technol Assess* 2013;17:1–79.
- 5 Newgard CD, Haukoos JS. Advanced Statistics: missing data in clinical research--part 2: multiple imputation. *Acad Emerg Med* 2007;14:669–78.
- 6 Newgard CD, Fildes JJ, Wu L, Hemmila MR, Burd RS, Neal M, Mann NC, Shafi S, Clark DE, Goble S, *et al*. Methodology and analytic rationale for the American college of Surgeons trauma quality improvement program. *J Am Coll Surg* 2013;216:147–57.