[1] McGill University, Montreal, Canada

[2] Columbia University, New York, USA

Correspondence to: J M Brophy
james.brophy@mcgill.ca

Rethinking approaches to analysis of global randomised controlled trials

Choice of statistical models and data integrity are vital

James M Brophy, [1] Andrew Gelman[2]

A BMJ investigation by Doshi (doi:10.1136/bmj.r1201)[1] raises concerns about the integrity of two randomised trials with surrogate endpoints that supported the regulatory approval of ticagrelor.[2 3] If data integrity is compromised, inferences from these studies are unreliable. Additional concerns are the study populations had stable coronary disease and not the acute coronary syndromes ticagrelor is mainly used to treat, [2 3] and the clinical value of the surrogate marker, inhibition of platelet aggregation, remains uncertain.

Historical examples show that surrogate endpoints, even without data integrity issues, can be misleading. For example, reductions in premature ventricular contractions and glycated haemoglobin levels have both been paradoxically associated with increased mortality in clinical trials.[4 5] Regulatory agencies do not usually accept surrogate endpoints trials alone for approval, but these studies may still influence decision making by providing plausible mechanisms that make clinical endpoint results seem more credible. Ensuring data integrity in these studies is therefore crucial.

Data quality is vital in multinational randomised controlled trials of clinical endpoints used for drug approval as the decision is often based on a single pivotal phase 3 trial.[6] A controversial example is the PLATO clinical endpoint trial, which was the pivotal study for ticagrelor approval.[7] The trial recruited 18 624 patients from 862 centres in 43 countries, and data integrity issues, including selection bias in dropouts and possible biased adjudication, have been discussed in *The BMJ* previously.[8]

Influence of statistical models

The operational difficulties in performing global randomised trials are mammoth. What is less obvious is the important role that statistical models have in the integrity of study inferences. Statistical models codify scientific decision making and motivate more careful experimental design and data collection. However, although statistical methods are often presented as neutral tools to summarise data, they quietly encode strong—often unstated, unrealised, and unreasonable—assumptions that can decisively shape conclusions. Again, PLATO is a case in point. Ignoring for the moment any uncertainties regarding data quality, the estimated hazard ratio of 0.84 (P<0.001) that has been universally interpreted by regulators and clinicians as conclusive evidence of ticagrelor's superiority is highly dependent on the hidden assumptions of the chosen statistical model.

PLATO was analysed under a pooled model that assumed constant baseline risks and constant treatment effects across centres, countries, and regions, an implausible assumption given that ancillary care after myocardial infarction certainly varies across the 43 participating countries. Ignoring any systematic centre level and country level factors will lead to falsely precise confidence intervals and increased risk of spurious claims of benefit. On the other hand, a statistical model examining each regional subgroup as distinct separate entities, an unpooled model, such as the US subgroup (HR=1.27, 95% CI 0.92 to 1.75), is wasteful as it excludes most of the randomised patient data. This is the difficulty that bedevilled the FDA, which initially rejected and then approved ticagrelor without any change in the database. Unknown is the role that the previously mentioned surrogate studies may have had in subtlety influencing the regulators to change their initial opposition.

However, there is a third statistical model that offers a sortie from this quandary of being a "lumper" or "splitter." Hierarchical regression is a compromise between unpooled and completely pooled models and indeed includes those two alternatives as special cases, with the amount of pooling estimated from the data in the multicentre trial.[9]

In the bayesian hierarchical meta-analysis with partial pooling, estimated effects for individual studies are shrunk towards the global mean with improved precision or, more generally, to a prediction based on centre and country level characteristics. However, the overall summary estimate for the average hazard ratio or the hazard ratio for a predicted new hypothetical study will have wider confidence uncertainty than before, reflecting the now acknowledged additional between region or country variations.[10 11]

The PLATO hierarchical meta-analysis yields a summary estimate (HR=0.90, 0.72 to 1.10) with enough uncertainty that a reasonable interpretation is the ticagrelor signal needs confirmation in further studies, especially for US patients. A second US study has not been done. Moreover, later randomised ticagrelor trials have shown no benefit.[12 -14] This newer evidence provides more support for the cautious interpretation of the PLATO hierarchical model and less support for the confident claim of ticagrelor superiority from the pooled model.

The statistical model is not a passive, trivial entity but rather an important element that mediates how data are transformed into regulatory approval, guideline endorsement, and clinical decisions. We believe hierarchical regressions offer more robust estimates and provide a more complete and open

accounting of uncertainty. Hence, we recommend they become standard in multicentre randomised trials, especially when data are aggregated across heterogeneous healthcare systems. Although researchers may choose to reject a particular hierarchical model for a multicountry trial, there should be a consensus that when conclusions differ drastically according to reasonable choices of statistical model, the data can hardly be considered robust and the results demand replication.

In an era where global trials shape global practice, assuring both data integrity and their accompanying statistical models are trustworthy is essential. Statistical models must be critically examined with the same thoroughness as the data they seek to summarise.

1   Doshi P. Ticagrelor doubts: inaccuracies uncovered in key studies for AstraZeneca's billion dollar drug. *BMJ* 2025;389:. doi: 10.1136/bmj.r1201.

2   Gurbel PA, Bliden KP, Butler K, etal. Response to ticagrelor in clopidogrel nonresponders and responders and effect of switching therapies: the RESPOND study. *Circulation* 2010;121:-99. doi: 10.1161/CIRCULATIONAHA.109.919456. pmid: 20194878

3   Gurbel PA, Bliden KP, Butler K, etal. Randomized double-blind assessment of the ONSET and OFFSET of the antiplatelet effects of ticagrelor versus clopidogrel in patients with stable coronary artery disease: the ONSET/OFFSET study. *Circulation* 2009;120:-85. doi: 10.1161/CIRCULATIONAHA.109.912550. pmid: 19923168

4   Cardiac Arrhythmia Suppression Trial (CAST) Investigators. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *N Engl J Med* 1989;321:-12. doi: 10.1056/NEJM198908103210629. pmid: 2473403

5   Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* 2007;356:-71. doi: 10.1056/NEJMoa072761. pmid: 17517853

6   Food and Drug Administration. Guidance for industry providing clinical evidence of effectiveness for human drug and biological products. 1998. https://www.gmp-compliance.org/files/guidemgr/clineff.pdf

7   Wallentin L, Becker RC, Budaj A, etalPLATO Investigators. Ticagrelor versus clopidogrel in patients with acute coronary syndromes. *N Engl J Med* 2009;361:-57. doi: 10.1056/NEJMoa0904327. pmid: 19717846

8   Doshi P. Doubts over landmark heart drug trial: ticagrelor PLATO study. *BMJ* 2024;387:. doi: 10.1136/bmj.q2550. pmid: 39662999

9   Gelman A, Hill J. *Analytical methods for social research: Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press, 2006.

10   IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open* 2016;6:e010247. doi: 10.1136/bmjopen-2015-010247. pmid: 27406637

11   Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;342:. doi: 10.1136/bmj.d549. pmid: 21310794

12   Goto S, Huang CH, Park SJ, Emanuelsson H, Kimura T. Ticagrelor vs. clopidogrel in Japanese, Korean and Taiwanese patients with acute coronary syndrome -- randomized, double-blind, phase III PHILO study. *Circ J* 2015;79:-60. doi: 10.1253/circj.CJ-15-0112. pmid: 26376600

13   Gimbel M, Qaderdan K, Willemsen L, etal. Clopidogrel versus ticagrelor or prasugrel in patients aged 70 years or older with non-ST-elevation acute coronary syndrome (POPular AGE): the randomised, open-label, non-inferiority trial. *Lancet* 2020;395:-81. doi: 10.1016/S0140-6736(20)30325-1. pmid: 32334703

14   Kutcher SA, Dendukuri N, Dandona S, Nadeau L, Brophy JM. Ticagrelor compared to clopidogrel in acute coronary syndromes trial (TC4): a bayesian pragmatic cluster randomized controlled trial. *CMAJ* 2025;197:-18. doi: 10.1503/cmaj.241862. pmid: 40164463