



## Invited Commentary | Oncology

# Trial Emulation and Real-World Evidence

Rolf H. H. Groenwold, MD, PhD

Increasingly, articles are being published that advocate for the use of observational data to estimate the effects of medical treatments in daily practice. In contrast to evidence from randomized clinical trials (RCTs), observational studies provide evidence that applies to the real world—or so it is claimed.

There are indeed various reasons why the results of RCTs may not apply directly to daily clinical practice. Traditionally, ideal conditions are created in RCTs to demonstrate treatment efficacy: strict inclusion and exclusion criteria, masking of participants and researchers, and close monitoring of the safety of participants and their adherence to the treatment protocol. Moreover, participants and their treating physicians in RCTs are explicitly asked to participate (ie, they provide informed consent), which can lead to a further selection.

Observational studies, in contrast, are based on what is called real-world data, such as those from electronic health records, and often better represent daily practice. However, due to the absence of randomization and masking (of patients and of physicians), it is always questionable whether the observed results are unbiased. Incomparability of treatment groups (confounding) and (selective) dropout are serious limitations of observational studies.<sup>1</sup> Also, the quality of electronic health records data tends to be inferior compared with those collected in RCTs, although there are examples in which they appear to be on par.<sup>2</sup> Another difference between RCTs and observational studies is that the former usually estimate intention-to-treat effects, whereas the latter focus on per-protocol effects.

Observational studies and RCTs of different adjuvant chemotherapy strategies in patients with stage III colon cancer have found different results. Boyne and colleagues<sup>3</sup> investigated whether those differences could be because of methodological issues. They used observational data to mimic an RCT as much as possible, a technique called target trial emulation.<sup>4</sup>

Target trial emulation is not as simple as it might sound. Less than 20% of participants in the initial cohort in the study by Boyne et al<sup>3</sup> were included for analysis. The studied treatment strategies did not exactly match those of the RCT, and the sample size of the observational study (485) was much smaller than that of the RCT (12 834), limiting the power to detect a difference between the effect estimates from the different studies, should it exist. Boyne and colleagues<sup>3</sup> tried to separate methodological explanations for differences between observed results, notably those due to immortal time bias. In a naive observational analysis, exposure levels were based on the actual duration of treatment. Because this information becomes available during the study follow-up and therefore is not yet available at study entry, this introduces the risk of immortal time bias; participants have to survive until a certain time to classify as having received a certain exposure level. As a result, those who received the shorter adjuvant chemotherapy treatment regime had a worse prognosis. Immortal time bias is one of many possible sources of bias that become apparent by articulating what a target trial looks like (ie, target trial emulation).

Results of the target trial-emulated observational study by Boyne and colleagues<sup>3</sup> were indeed in line with those of the RCT, more so than the results of a naive analysis of the observational data. Does this show that target trial emulation yields observational studies that are as credible as RCTs? If results from observational studies concur with those of RCTs, this may suggest that the design and analysis of the former are valid. Unfortunately, this is not true.

First, there may be alternative explanations for comparability of treatment effect sizes across different studies (with different designs). These include chance, cancellation of biases, and choices made when analyzing the data (eg, which effect is estimated), to name a few.

**+ Related article**

Author affiliations and article information are listed at the end of this article.

**Open Access.** This is an open access article distributed under the terms of the CC-BY License.

Second, a premise of the claim about the validity of an observational study is that results should in fact be the same, but that need not be the case. There are various clinical reasons why RCTs and observational studies can yield different results. Selection is important. Treatment effects may differ between those included in an RCT and patients seen in daily practice; this phenomenon is often referred to as effect modification. While standardization of trial results can overcome differences in distributions of effect-modifying variables,<sup>5</sup> this only applies to variables that are measured. Furthermore, physicians who participate in a trial may not be a representative sample of daily practice,<sup>6</sup> eg, because they are more experienced or more research oriented. What is more, the Hawthorne effect does not affect observational studies as it does RCTs, and related to this, treatment adherence likely differs between the 2 approaches as well.

Third, the RCT need not correspond with the target trial. For example, a target trial is not subject to the Hawthorne effect or to selection due to informed consent procedures. In a target trial, there are no ethical or legal reasons not to include, eg, patients who cannot consent, children, or pregnant women. Both RCTs and target trials must follow laws and regulations regarding medical practice (eg, only trained surgeons are allowed to perform surgery), but in a (conceptual) target trial, we need not consider laws and regulations regarding medical research.

RCTs and their observational counterparts should be compared, but given the multitude of methodological and clinical reasons why results may be different, comparison should be made first and foremost in terms of the design and analytical choices that are made. Target trial emulation is a means of improving the quality of studies of medical treatments and interpretability of results and can serve as a conceptual benchmark for these choices. Only when observational studies are methodologically sound (which is an assessment that should not be based on comparing results with those of an RCT of the same treatment), can we look further and consider what clinical value the study adds.

Direct comparisons between results of RCTs and their observational counterparts have limited value because there are multiple explanations for differences between their results.<sup>7</sup> Evidence that originates from daily practice does not necessarily provide valid evidence for daily practice. Using (real-world) data from daily practice for studies of comparative effectiveness can introduce many sources of bias, such as confounding, missing data, and misclassification. Observational studies based on real-world data are not a test of the applicability of the results of RCTs and, vice versa, RCTs are not a litmus test of the validity of observational studies. However, a thorough breakdown of possible explanations (methodological and clinical) for observed differences in results could provide insight into the applicability of the results of RCTs and the possible sources of bias in observational studies.

---

## ARTICLE INFORMATION

**Published:** March 30, 2021. doi:[10.1001/jamanetworkopen.2021.3845](https://doi.org/10.1001/jamanetworkopen.2021.3845)

**Open Access:** This is an open access article distributed under the terms of the [CC-BY License](https://creativecommons.org/licenses/by/4.0/). © 2021 Groenwold RHH. *JAMA Network Open*.

**Corresponding Author:** Rolf H. H. Groenwold, MD, PhD, Department of Clinical Epidemiology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA, Leiden, the Netherlands ([r.h.h.groenwold@lumc.nl](mailto:r.h.h.groenwold@lumc.nl)).

**Author Affiliations:** Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands.

**Conflict of Interest Disclosures:** Dr Groenwold reported receiving grants from the Netherlands Organization for Scientific Research and Leiden University Medical Center.

## REFERENCES

1. Collins R, Bowman L, Landray M, Peto R. The magic of randomization versus the myth of real-world evidence. *N Engl J Med*. 2020;382(7):674-678. doi:[10.1056/NEJMs1901642](https://doi.org/10.1056/NEJMs1901642)

2. van Dijk WB, Fiolet ATL, Schuit E, et al. Text-mining in electronic healthcare records can be used as efficient tool for screening and data collection in cardiovascular trials: a multicenter validation study. *J Clin Epidemiol*. 2020;132:97-105. doi:[10.1016/j.jclinepi.2020.11.014](https://doi.org/10.1016/j.jclinepi.2020.11.014)
3. Boyne DJ, Cheung WY, Hilsden RJ, et al. Association of a shortened duration of adjuvant chemotherapy with overall survival among individuals with stage III colon cancer. *JAMA Netw Open*. 2021;4(3):e213587. doi:[10.1001/jamanetworkopen.2021.3587](https://doi.org/10.1001/jamanetworkopen.2021.3587)
4. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764. doi:[10.1093/aje/kwv254](https://doi.org/10.1093/aje/kwv254)
5. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 trial. *Am J Epidemiol*. 2010;172(1):107-115. doi:[10.1093/aje/kwq084](https://doi.org/10.1093/aje/kwq084)
6. van Staa TP, Dyson L, McCann G, et al. The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. *Health Technol Assess*. 2014;18(43):1-146. doi:[10.3310/hta18430](https://doi.org/10.3310/hta18430)
7. Franklin JM, Glynn RJ, Suissa S, Schneeweiss S. Emulation differences vs biases when calibrating real-world evidence findings against randomized controlled trials. *Clin Pharmacol Ther*. 2020;107(4):735-737. doi:[10.1002/cpt.1793](https://doi.org/10.1002/cpt.1793)