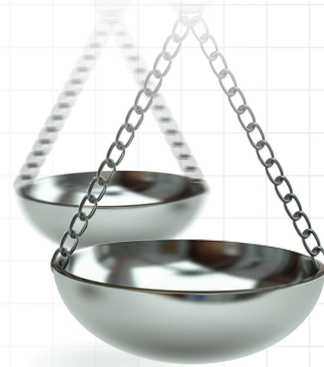




Why is target trial emulation not being used in health technology assessment real-world data submissions?

Journal of **Comparative Effectiveness Research**

Alejandra Castanon¹, Stephen Duffield², Sreeram Ramagopalan^{*.1} & Robert Reynolds³

¹ Lane Clark & Peacock LLP, London, W1U 1DQ, UK

² National Institute of Health & Care Excellence, Manchester, M1 4BT, UK

³ GSK, NY, USA

*Author for correspondence: Sreeram.ramagopalan@lcp.uk.com

“Polling of the audience attending the panel session was surprising in that it revealed that the main barrier to conducting TTE studies was lack of expertise and experience with the approach, despite its widespread use in the scientific literature.”

First draft submitted: 24 May 2024; Accepted for publication: 24 May 2024; Published online: 8 June 2024

Keywords: health technology assessment • observational data • randomized controlled trials • real-world data • target trial emulation

Target trial emulation (TTE) is a framework for the application of randomized controlled trial (RCT) design principles to the analysis of observational data, thereby explicitly tying the analysis to the trial it is emulating [1]. The TTE approach has been proven to enable inference of causal treatment effects while avoiding the critical biases that often plague observational studies [2,3]. An illustration of this comes from Dickerman *et al.* [4] who emulated a trial in real-world data (RWD) to estimate the effect of statins on the risk of small cell lung cancer among adults with low-density lipoprotein (LDL) cholesterol below 5 mmol/l. A prior observational study [5] suggested that statins may reduce the risk of small cell lung cancer by 77%, however Dickerman found that the TTE per protocol and intention-to-treat analyses reported no effect of statins on lung cancer. The TTE results were in line with those observed in meta-analyses of RCTs (Table 1). The conflicting results generated by observational studies as compared with RCTs has been one of the key reasons why regulators and health technology assessment (HTA) bodies have been reluctant to accept RWD as a measure of treatment effectiveness. By applying the TTE framework, Dickerman *et al.* demonstrated that the previous observational study had two critical flaws: using post-baseline information to assign baseline treatment, and including individuals who were using statins before baseline. When Dickerman and colleagues included these methodological flaws in their analysis, the effect estimates replicated the non-TTE observational study results (Table 1). Given the ability of the TTE framework to minimize bias in observational studies, several HTA agencies have recently advocated for TTE because application of the framework increases transparency around study design. However, despite guidance from HTA agencies being in place since 2022, to the best of our (AC, SR) knowledge no HTA RWD submission to date has explicitly employed TTE.

At an issue panel at ISPOR 2024, we debated the reasons for this. S Duffield confirmed that unlike the academic literature, where articles referencing TTE have increased exponentially since 2020 (Figure 1), a rudimentary search of National Institute of Health and Care Excellence (NICE)'s appraisals did not identify any submission explicitly using the TTE approach even though it is central to NICE's real-world evidence (RWE) framework [6]. One caveat may be that some published studies identified by systematic literature review and submitted as part of the appraisal could have used this approach.

Polling of the audience attending the panel session was surprising in that it revealed that the main barrier to conducting TTE studies was lack of expertise and experience with the approach, despite its widespread use in the scientific literature. Furthermore, many attendees expressed distrust that HTA agencies would accept TTE RWE despite explicit support of TTE by NICE and several other HTA agencies [7]. This may reflect past experience where submissions using RWD as a measure of treatment effectiveness have largely been rejected by HTA agencies [8].

Table 1. Hazard ratios and odds ratios for small cell lung cancer comparing statin therapy with no statin therapy.				
TTE per protocol HR (95% CI)	TTE intention to treat HR (95% CI)	Observational study [5] OR (95% CI)	Analyses with flaws [†] HR (95% CI)	Analyses with flaws [‡] HR (95% CI)
1.13 (0.96, 1.32)	1.08 (0.99, 1.17)	0.23 (0.20, 0.26) comparing long-term statin users (>4 years) with nonusers	0.26 (0.23, 0.3) comparing statin use for >4 vs 0 years.	0.27 (0.25, 0.2) comparing statin use for >4 vs 0 years
[†] When using post-baseline information to assign baseline treatment. [‡] When including individuals who were using statins before baseline. Data taken from [4]. HR: Hazard ratio; OR: Odds ratio; TTE: Target trial emulation.				

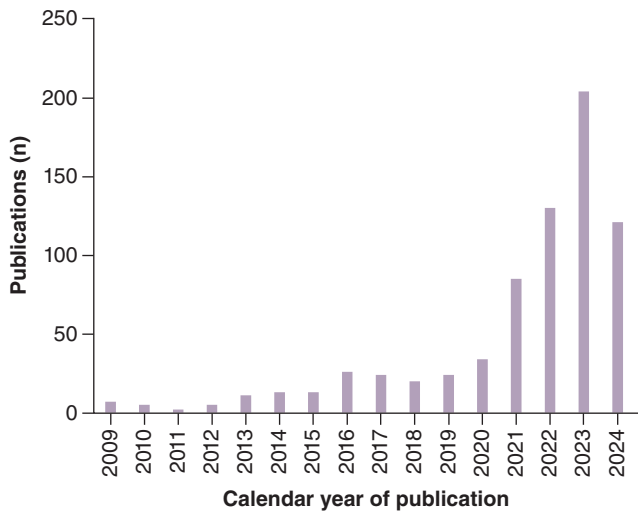


Figure 1. Peer reviewed publications mentioning target trial emulation. Based on a search conducted on 17 May 2024 using the following terms: ‘target trial’ OR ‘target trial emulation’ OR ‘trial emulation’.

R Reynolds provided a manufacturer’s perspective. While an advocate for the TTE approach, he cautioned that TTE is not a failsafe, and there may be barriers such as the fact that TTE cannot address questions that cannot be framed in terms of a pragmatic trial (i.e., self-controlled or test negative designs in vaccines). R Reynolds also highlighted the fact that TTE is not the only framework to provide best practice around RWD study design and execution. Notably, the estimand [9] and the SPIFD2 frameworks [10] could also be used to minimize bias. SPIFD2 calls for the selection of fit for purpose RWD, articulation of the hypothetical target trial and transparent documentation of the rationale for decisions made which can then be used to populate the STaRT-RWE or HARPER templates [11,12]. The estimand framework on the other hand provides a systematic approach to defining the treatment effect under investigation in a clinical trial. This covers characterizing PICOs (population, intervention, comparator, outcome) but also includes describing how handling intercurrent events will be handled and how outcomes will be summarized and compared between treatments. R Reynolds pointed out, however, that the estimand framework was developed with a focus on RCTs and therefore is lacking guidance on the description of baseline confounders and the need to specify time zero. S Duffield noted that both the estimand and SPIFD2 frameworks have also not been explicitly used in HTA RWD submissions.

Additionally, R Reynolds noted that there are currently no standard reporting guidelines for TTE (although these are in development [13]) and therefore while TTE may be performed by manufacturers, it may not be documented as such in submissions to HTA bodies. In support of the fact that manufacturers are performing TTE, an example of RWE for post-approval effectiveness that explicitly used the TTE framework was shared [14]. In this example, challenges such as defining the time of treatment implementation and required deep clinical knowledge to enable accurate and appropriate definitions. A recent demonstration project also highlighted this as a challenge when conducting HTA and offered some pragmatic suggestions when defining the hypothetical trial to be emulated [15].

To illustrate the importance of transparency in RWD analyses, S Duffield discussed a case study regarding a recent submission to NICE. The example was an external control arm for a single arm trial (SAT) conducted in a rare subgroup of non-small cell lung cancer [16]. The ECA included US cohort data and routinely collected cancer registry data from the UK (albeit for a very small number of individuals). The resulting adjusted indirect comparison was based on a blended comparator and the committee was concerned that the mix of comparator

treatments and subsequent therapies might not be relevant to the UK. Further, results were based on complete case analysis and confounders were taken from a prespecified list which led the committee to question whether all relevant confounders were captured and the magnitude of missing data. The committee also commented that it was not possible to assess whether best-practice study design principles had been followed and whether confounding and residual bias had been robustly addressed. S Duffield noted that the committee's concerns might have been addressed with the completion of the NICE DataSAT tool and by describing how the RWD emulated the SAT, by for example describing it using a TTE protocol.

In conclusion, while the TTE framework has been shown to minimize bias in observational studies and is advocated for by several HTA agencies, its use in HTA submissions has been limited to date. This may be because RWE guidance in the HTA setting has only been around for just under two years, perhaps meaning that manufacturers have not had enough time to employ the approach for submissions. The lack of reporting guidance on TTE may also be a reason why TTE is not explicitly mentioned in HTA submissions. However, the main barrier from an ISPOR audience appears to be a lack of expertise and experience with the approach since its use has only become prominent in the scientific literature recently. Ultimately, the key to successful RWE submissions lies in transparently documenting the rationale for decisions made in data source selection and analysis, which TTE supports. Broadly, there appears to be a need in the health economics and outcomes research community to develop expertise in this area, something which stakeholders should collaboratively seek to address.

Financial disclosure

R Reynolds is an employee of GSK. S Ramagopalan and A Castanon are employees of Lane Clark and Peacock LLP. The authors have no other 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 apart from those disclosed.

Competing interests disclosure

The authors have no competing interests or relevant affiliations with any organization or entity 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.

Writing disclosure

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

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. *Eur. J. Epidemiol.* 32, 473–475 (2017).
2. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am. J. Epidemiol.* 183, 758–764 (2016).
3. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Evaluating metformin strategies for cancer prevention: a target trial emulation using electronic health records. *Epidemiology* 34, 690 (2023).
4. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat. Med.* 25, 1601–1606 (2019).
5. Khurana V, Bejjanki HR, Caldito G, Owens MW. Statins reduce the risk of lung cancer in humans: a large case-control study of US veterans. *Chest* 131, 1282–1288 (2007).
6. National Institute for Health and Care Excellence. NICE real-world evidence framework. Methods for real-world studies of comparative effects (2022). <https://www.nice.org.uk/corporate/ecd9/chapter/methods-for-real-world-studies-of-comparative-effects>
7. Vanier A, Fernandez J, Kelley S *et al.* Rapid access to innovative medicinal products while ensuring relevant health technology assessment. Position of the French National Authority for Health. *BMJ Evid. Based Med.* 29, 1–5 (2024).
8. Cox O, Sammon C, Simpson A, Wasiak R, Ramagopalan S, Thorlund K. The (harsh) reality of real-world data external comparators for health technology assessment. *Value Health* 25, 1253–1256 (2022).
9. Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer on the ICH E9(R1) addendum. *BMJ* 384, e076316 (2024).

10. Gatto NM, Vittoe SE, Rubinstein E, Reynolds RF, Campbell UB. A structured process to identify fit-for-purpose study design and data to generate valid and transparent real-world evidence for regulatory uses. *Clin. Pharmacol. Therapeut.* 113, 1235–1239 (2023).
11. Wang SV, Pottegård A, Crown W *et al.* HARmonized protocol template to enhance reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: a good practices report of a joint ISPE/ISPOR Task Force. *Value Health* 25, 1663–1672 (2022).
12. Wang SV, Pinheiro S, Hua W *et al.* STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ* 372, m4856 (2021).
13. Hansford HJ, Cashin AG, Jones MD *et al.* Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. *BMJ Open* 13, e074626 (2023).
14. Coleman RL, Perhanidis J, Kalilani L, Zimmerman NM, Golembesky A, Moore KN. Real-world overall survival in second-line maintenance niraparib monotherapy vs active surveillance in BRCA wild-type patients with recurrent ovarian cancer. *J. Clin. Oncol.* 41(Suppl. 16), 5592 (2023).
15. Moler-Zapata S, Hutchings A, O'Neill S, Silverwood RJ, Grieve R. Emulating target trials with real-world data to inform health technology assessment: findings and lessons from an application to emergency surgery. *Value Health* 26, 1164–1174 (2023).
16. National Institute for Health and Care Excellence. Amivantamab for treating EGFR exon 20 insertion mutation-positive advanced non-small-cell lung cancer after platinum-based chemotherapy. Technology appraisal guidance. TA850 (2022). <https://www.nice.org.uk/guidance/ta850/chapter/3-Committee-discussion>