

## COMMENTARY

# Design considerations for observational studies of drugs for advanced cancer

Xabier Garcia-Albeniz<sup>1</sup> | Pär Stattin<sup>2</sup> | Marcus Westerberg<sup>2</sup> |  
Rachel Weinrib<sup>1</sup> | Per Sandstrom<sup>3</sup> | Zdravko Vassilev<sup>3</sup>

<sup>1</sup>Pharmacoepidemiology and Risk Management, RTI Health Solutions, Barcelona, Spain

<sup>2</sup>Department of Surgical Sciences, Urology, Uppsala University, Uppsala, Sweden

<sup>3</sup>Bayer US, Whippany, New Jersey, USA

## Correspondence

Xabier Garcia-Albeniz, RTI Health Solutions, Diagonal 605, Barcelona, Spain.  
Email: xgarcia@rti.org

## Funding information

Bayer

**KEYWORDS:** causal inference, real world evidence, target trial emulation

## Key Points

- Real-world evidence (RWE) studies are gaining relevance in the clinical development programmes of drugs for advanced cancer.
- RWE studies of drugs for advanced cancer can be challenging because the study drug can be part of several of the many lines of treatment (LoTs) patients can receive, and occasionally, the component drugs of a LoT are not started concomitantly.
- The target trial emulation framework can help in the design, analysis, and interpretation of RWE studies of drugs for advanced cancer.

## 1 | INTRODUCTION

Post-authorisation studies complement the clinical development programme of a drug by studying larger, more diverse populations over a longer time than studies of the programme and can rely on secondary use of routinely collected patient data (i.e., real-world data [RWD]). Often, like the randomised controlled trials (RCTs) that may precede them, their aim is to estimate causal effects<sup>1</sup> of a drug on safety and/or efficacy outcomes. Here, we present three challenges of RWD analyses of patients with advanced cancer and assess whether similar challenges occur in RCTs and how they are managed. Then we propose target trial emulation<sup>2</sup> as an analytical framework for overcoming the challenges in post-authorisation studies.

## 2 | CHALLENGE #1: THE STUDY DRUG CAN BE A COMPONENT OF SEVERAL LINES OF TREATMENT

Treatment of patients with advanced cancer is characterised by three factors: (1) Patients often receive drugs in sequence as “lines of treatment” (LoTs), which consist of one or several drugs, continued until a pre-specified duration, tumour progression, toxicity, or a decision to stop treatment is reached. Sometimes drugs in a LoT are started at different times because of logistics or the need of additional diagnostics. For example, in patients with advanced colorectal cancer treated with first-line chemotherapy plus anti-epidermal growth factor receptor (EGFR) therapy, start of anti-EGFR is sometimes delayed because its indication needs to be verified by additional testing. (2) If clinically

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 Research Triangle Institute - Health Solutions. *Pharmacoepidemiology and Drug Safety* published by John Wiley & Sons Ltd.

indicated, the patient often proceeds to a subsequent LoT when the previous LoT is no longer indicated (e.g., progression or toxicity). For example, several drugs are approved to treat metastatic castration-resistant prostate cancer – androgen receptor-targeted drugs, chemotherapy, radioactive agents, poly (ADP-ribose) polymerase inhibitors, and immunotherapy<sup>3</sup> – which can be administered almost in any sequence as several LoTs. (3) Patients receiving later LoTs have a substantially worse prognosis than those receiving earlier LoTs because the LoT reflects the natural evolution of the disease (e.g., tumour can spread to distant sites and become resistant to treatments).

The first challenge is studying a drug that can be a component of different LoTs. If the average causal effect across several LoTs is estimated and the distribution of LoTs differs between exposure groups, baseline confounding can occur (because of the prognostic nature of LoTs). Additionally, if the drug's effect is not the same in all LoTs (effect modification), the average causal effect across several LoTs may not be relevant for clinical practice – where therapeutic decisions are made on the basis of LoT – and may hinder transportability. Transportability is the process of extending inferences computed in one population to a second population with different characteristics.<sup>4</sup> If the LoT is an effect modifier and the distribution of LoT differs between the population in which the effect was estimated and the target population(s), the estimated average causal effect across several LoTs is not transportable. Differences in LoT distribution can occur due, for example, to differences in reimbursement policies, time trends in drug prescription, or prescriber preferences. The design of RCTs commonly addresses this challenge by recruiting only patients eligible to receive the drug in a specific LoT. Such restriction avoids confounding by the LoT and generates useful information to estimate effects in other populations if the effect of the drug is not homogeneous across LoTs.

### 3 | CHALLENGE #2: SEVERAL LOTs CAN BE RECEIVED BEFORE THE OUTCOME CAN BE OBSERVED

The second challenge is that patients can receive additional LoTs after baseline, before the outcome of interest occurs, and such sequence of drugs may be determined by the efficacy and toxicity of previous LoTs. If we are interested in estimating the effect of a specific drug under complete adherence or if there are losses to follow-up, both baseline and post-baseline confounding can distort the identification of the causal effect. Whereas randomisation protects against baseline confounding in RCTs, it does not eliminate post-baseline confounding. Post-baseline confounding can arise from non-adherence if post-randomization prognostic factors that affect treatment decisions are unequally distributed across arms. RCTs address this challenge by specifying the post-randomization events (events occurring after treatment initiation, such as subsequent treatments that affect the interpretation or the existence of the measurements associated with the clinical question) and then by integrating such events in the treatment strategies to construct an estimand. For example, an RCT

protocol can specify the sequence of treatments allowed in each study arm (no trial would halt treatment after the randomised drug is stopped if the patient is fit to continue treatment with another LoT) and even force some control by design of subsequent treatments (e.g., by specifying whether crossing over to the experimental drug is allowed in the control group). Specifically, RCTs compare not only drugs but treatment strategies that can comprise 1 or more drugs administered concomitantly or in sequence, as well as rules to guide treatment, applied at specific decision points. A proposed approach to identify the effect under complete adherence and/or under complete follow-up consists of censoring data from individuals when they deviate from the treatment strategy (note that stopping treatment for clinical reasons is not a deviation) or from the follow-up strategy and then adjusting for pre-randomisation and post-randomisation factors that predict adherence. If baseline treatments can affect post-randomisation factors, an adjustment must be made with *g*-methods, which include inverse probability weighting, *g*-estimation of structural nested models, and the *g*-formula, rather than with conventional statistical methods.<sup>5</sup>

### 4 | CHALLENGE #3: TREATMENT STRATEGY MAY NOT BE DEFINED AT BASELINE IN A REAL-WORLD DATABASE

The third challenge is studying the effect of a drug (vs. not receiving it) that is part of a multicomponent LoT and is started later than the other components. Because mortality in patients with advanced cancer is high, immortal time bias<sup>6</sup> (a form of selection<sup>7</sup>) can occur in an observational setting because only patients who survive long enough will be able to start to use that drug. RCTs address this issue by estimating the intention-to-treat effect, effectively avoiding such immortal time bias by aligning the start of follow-up with the date of treatment assignment (date of randomization), rather than with the date of treatment initiation.

### 5 | EMULATING A TARGET TRIAL TO HANDLE CHALLENGES WITH RWD

The target trial emulation framework starts by articulating the research question in the form of a hypothetical pragmatic randomised trial. This target trial should have pre-specified eligibility criteria, treatment strategies, treatment assignment, start and end of follow-up, outcomes, causal contrast, and a data analysis plan. Subsequently, the RWD analysis will emulate the components of the target trial protocol by finding eligible individuals, assigning them to treatment strategies that are compatible with their data at baseline, following them up from treatment assignment (time zero) until outcome or end of follow-up, and conducting the same analysis as the corresponding target trial. The only difference is that the emulation requires adjustment for baseline confounders in an attempt to emulate the randomization of the target trial. By providing a well-defined estimand, this approach

**TABLE 1** Summary table containing the treatment strategies, causal contrast, and handling of analytic challenges in two real-world data studies of drugs for advanced cancer.

	Study on the safety of Ra-223 in patients with mCRPC <sup>9</sup>	Study on the effectiveness of erlotinib in patients with advanced pancreatic cancer <sup>12</sup>
Treatment strategies	<p><b>Group 1: Initiate Ra-223.</b> Patients can stop Ra-223 after 6 cycles or earlier in the event of toxicity, cancer progression, or worsening of the overall health status. Patients can start other systemic drugs for mCRPC after the initiation of Ra-223, when clinically indicated, but they can never be used while taking Ra-223. ADT with first-generation antiandrogens can be used at any time</p> <p><b>Group 2: Initiate other standard of care (docetaxel, cabazitaxel, enzalutamide, abiraterone, others).</b> Patients are allowed to stop the standard of care and continue with other LoTs, with the exception of Ra-223, when clinically indicated. ADT with first-generation antiandrogens can be used at any time.</p>	<p><b>Group 1.</b> Initiate gemcitabine as first-line treatment. <b>Initiate erlotinib</b> (any dose) within the grace period: up to 12 weeks after gemcitabine initiation.</p> <p><b>Group 2.</b> Initiate gemcitabine as first-line treatment within the grace period. <b>Do not initiate erlotinib.</b> Under both strategies, the decision to discontinue gemcitabine or erlotinib, as well as to initiate any additional therapies, is left to the patient's and physician's discretion.</p>
Causal contrasts	Observational analogue of the per-protocol effect	
Handling of challenge #1: a study drug can be a component of several LoTs	LoT-specific cohorts were created	This challenge was absent because erlotinib was only approved for first-line treatment. A single first-LoT cohort was created
Handling of challenge #2: several LoTs can be received before observing the outcome	Specification of treatment strategies that consider post-baseline events. Use of artificial censoring to select the person-time compliant with the treatment strategies and inverse probability of treatment weighting to estimate the effect under complete adherence	
Handling of challenge #3: treatment strategy not defined at baseline in a real-world database	This challenge was absent	Clones were used to classify patients into treatment strategies according to their baseline data

Abbreviations: ADT, androgen deprivation therapy; LoT, line of treatment; mCRPC, metastatic, castration-resistant prostate cancer.

helps to identify and address possible sources of bias,<sup>8</sup> and by comparing well-defined courses of action, it helps decision-makers. Appropriate handling of start of follow-up (time zero) and appropriate handling of time-varying confounding are among the most useful aspects of this approach. We present solutions to the challenges above through two applications.

The first two challenges were present in PRECISE, a post-authorisation study that, as the primary objective, estimated the effect of Radium-223 on the incidence of fractures compared with other standard of care in Swedish men with metastatic castration-resistant prostate cancer.<sup>9</sup> As part of our target trial, we specified the treatment strategies considering baseline drugs and decision rules based on post-baseline events, which included observed drug toxicity, tumor progression and overall health status (Table 1). Since safety was the focus of the study, the causal contrast (or estimand) of interest was the effect under complete adherence to such strategies.<sup>10</sup> To diminish confounding by LoT and improve transportability of our estimates, we created a cohort of patients eligible for a first LoT (containing Ra-223 or other standard of care). Because Ra-223 could be used as any LoT during the study period, we also created cohorts of patients eligible for a second, third, and fourth LoT. In each LoT, we aligned time zero, eligibility assessment, and exposure assignment.<sup>11</sup> Because patients in our database could be eligible for more than 1 LoT, we could let them contribute to 1 specific LoT or to all LoTs.<sup>2</sup> For efficiency, we let patients contribute to all LoT-specific cohorts

that they were eligible for. The observational analogue of the per-protocol effect was implemented by selecting the person-time compliant with each of the exposure strategies and using inverse probability weighting to adjust for potential baseline and post-baseline confounding. Because the effect of Ra-223 on the risk of fractures across LoTs was considered homogeneous, cohorts were pooled.<sup>9</sup>

The third challenge was present in Petito et al.,<sup>12</sup> who estimated the effect of adding erlotinib to gemcitabine on survival in patients with advanced pancreatic cancer. In routine care, erlotinib may be initiated a few weeks later than gemcitabine, and advanced pancreatic cancer is highly lethal. To avoid the potential for immortal time bias, Petito et al.<sup>12</sup> defined time zero as the time when patients initiated gemcitabine and met the eligibility criteria. The exposure strategies allowed erlotinib to be started up to 12 weeks after time zero (Table 1). Time zero, treatment assignment, and eligibility were aligned by classifying patients into treatment strategies that their baseline data were compatible with, effectively creating two copies (clones) of each eligible individual and by assigning one copy to each exposure strategy.<sup>13</sup> Then, the person-time compliant with the assigned strategy was selected, and inverse probability weighting was used to adjust for potential baseline and post-baseline confounding, which is the same approach used in PRECISE.<sup>12</sup>

In conclusion, we recommend the following target trial emulation framework in RWD studies of drugs for advanced cancer: create LoT-specific cohorts; align the beginning of follow-up (time zero),

eligibility criteria, and exposure assignment; and use g-methods to estimate effects under complete adherence and under complete follow-up, especially in safety-focused studies, like PASs. In the absence of unmeasured confounding, measurement error, and model misspecification (including models used to handle missing data), this approach will help identify the causal effect of well-specified treatment strategies.<sup>14</sup>

#### ORCID

Xabier Garcia-Albeniz <https://orcid.org/0000-0002-9814-2343>

Pär Stattin  <https://orcid.org/0000-0002-8306-0687>

Marcus Westerberg <https://orcid.org/0000-0002-8906-6967>

Rachel Weinrib  <https://orcid.org/0000-0002-7753-0996>

#### REFERENCES

- Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. *Chance*. 2019;32(1):42-49. doi:10.1080/09332480.2019.1579578
- Hernan 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
- NCCN. NCCN Clinical Practice Guidelines In Oncology (NCCN Guidelines<sup>®</sup>) for Prostate Cancer. National Comprehensive Cancer Network, V.4.2022.
- Dahabreh IJ, Hernán MA. Extending inferences from a randomized trial to a target population. *Eur J Epidemiol*. 2019;34(8):719-722. doi:10.1007/s10654-019-00533-2
- Hernan MA, Robins JM. Per-protocol analyses of pragmatic trials. *N Engl J Med*. 2017;377(14):1391-1398. doi:10.1056/NEJMs1605385
- Time WA. *Observation and Inference: An Introduction to the Methods of Epidemiology*. Epidemiology Resources Inc; 1991.
- Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615-625.
- Robins JM, Hernán MA. Target trial emulation. *Causal Inference: What If*. Boca Raton; 2020:277-287.
- Stattin P, Westerberg M, Lissbrant I, et al. Real-world outcomes in patients with metastatic, castration-resistant prostate cancer treated with Radium-223 in routine clinical practice in Sweden. *Clin Genitourin Cancer*. 2022;21:107.e1-107.e9. doi:10.1016/j.clgc.2022.09.002
- Hernan MA, Hernandez-Diaz S. Beyond the intention-to-treat in comparative effectiveness research. *Clin Trials*. 2012;9(1):48-55. doi:10.1177/1740774511420743
- García-Albeniz X, Hsu J, Hernan MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol*. 2017;32(6):495-500. doi:10.1007/s10654-017-0287-2
- Petito LC, García-Albéniz X, Logan RW, et al. Estimates of overall survival in patients with cancer receiving different treatment regimens: emulating hypothetical target trials in the surveillance, epidemiology, and end results (SEER)-Medicare linked database. *JAMA Netw Open*. 2020;3(3):e200452. doi:10.1001/jamanetworkopen.2020.0452
- Cain LE, Robins JM, Lanoy E, Logan R, Costagliola D, Hernan MA. When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int J Biostat*. 2010;6(2): Article 18.
- Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes (Lond)*. 2008;32(Suppl 3):S8-S14. doi:10.1038/ijo.2008.82

**How to cite this article:** Garcia-Albeniz X, Stattin P, Westerberg M, Weinrib R, Sandstrom P, Vassilev Z. Design considerations for observational studies of drugs for advanced cancer. *Pharmacoepidemiol Drug Saf*. 2023;1-4. doi:10.1002/pds.5612