Exploration of Musculoskeletal Diseases

Open Exploration

Open Access Perspective

Causal inference methods for evaluating comparative effectiveness: target trial emulation

Tobias Haugegaard[1,2], Robin Christensen[1,2]*

[1]Section for Biostatistics and Evidence-Based Research, the Parker Institute, Bispebjerg and Frederiksberg Hospital, 2000 Copenhagen, Denmark

[2]Research Unit of Rheumatology, Department of Clinical Research, University of Southern Denmark, Odense University Hospital, 5000 Odense, Denmark

***Correspondence:** Robin Christensen, Section for Biostatistics and Evidence-Based Research, the Parker Institute, Bispebjerg and Frederiksberg Hospital, 2000 Copenhagen, Denmark. Robin.Christensen@regionh.dk

Academic Editor: Tarek Tawfik Amin, Kasr Alainy School of Medicine, Cairo University, Egypt

Received: February 3, 2026 **Accepted:** June 29, 2026 **Published:** August 5, 2026

Cite this article: Haugegaard T, Christensen R. Causal inference methods for evaluating comparative effectiveness: target trial emulation. Explor Musculoskeletal Dis. 2026;4:1007132. https://doi.org/10.37349/emd.2026.1007132

Abstract

Causal inference is grounded in contrasts between potential outcomes under alternative interventions. Randomized trials are the reference standard for estimating average causal effects because randomization renders treatment assignment independent of potential outcomes, eradicating confounding by design. However, many clinically important questions in rheumatology cannot feasibly be addressed through randomized experiments due to practical, ethical, or temporal constraints. In such settings, observational data can inform decisions. This paper argues that principles derived from randomized trials should continue to anchor causal reasoning and proposes the target trial framework as a structured approach to strengthen causal inference from observational data. By explicitly specifying the protocol of the hypothetical randomized trial that would answer the question and emulating it using real-world (observational) data, investigators can clarify eligibility criteria, treatment strategies, time zero, outcomes, causal contrasts, and identifying assumptions. By describing how each part of the target trial is emulated in observational data, they increase transparency, clarify the causal estimand, and make assumptions, limitations, and sources of bias explicit. This design-based perspective helps prevent common biases, including immortal time bias, selection of prevalent users, and inappropriate conditioning on post-treatment variables, and aligns reporting with clearly defined causal contrasts and effect measures. Ultimately, credibility of causal inference from observational data depends on whether the assignment mechanism is plausibly reconstructed, design choices are made without access to outcome data, and analyses target explicitly defined causal contrasts under stated assumptions, all of which are explicitly addressed within the target trial emulation framework.

Keywords

causal inference, target trial emulation, real-world data, observational studies, rheumatology

Introduction

Causal effects are defined as contrasts between potential outcomes under alternative interventions, of which only one can be observed for each individual [1]. The basic concept of causality is defined with respect to a unit, a treatment (X), and the outcomes (Y) that would be observed under alternative treatment conditions. A unit refers to the person, place, or object upon which a treatment operates at a specific point in time. Since each observation constitutes a distinct unit, treating them otherwise conflates the temporal structure of treatment decisions with the outcome they are meant to predict, making time-varying confounding inevitable.

Here the treatment is the intervention ($X = 1$) whose effect on a particular outcome is of interest, evaluated relative to a clearly defined control condition, such as no intervention or an alternative intervention ($X = 0$). For each unit, there exist potential outcomes corresponding to each treatment condition, representing the value of the outcome that would be observed if the unit were exposed to the treatment and the value that would be observed if the unit were exposed to the control. The causal effect for an individual unit is defined as the contrast between these potential outcomes under treatment ($Y|X = 1$) and control ($Y|X = 0$). A fundamental challenge of causal inference is that, for any given unit, only one of the potential outcomes can be observed in practice; the remaining potential outcome is counterfactual. Therefore, individual-level causal effects are not directly observable, and causal inference proceeds by comparing outcomes across units to estimate average causal effects under explicit assumptions that allow observed outcomes in one group to serve as valid proxies for the unobserved counterfactual outcomes in another.

Consequently, randomized experiments are the reference standard for estimating average causal effects. By construction, randomization makes treatment assignment independent of potential outcomes, so observed differences between groups can be interpreted as unbiased estimates of average treatment effects without invoking untestable assumptions about the assignment mechanism [1, 2]. Randomization resolves the fundamental "counterfactual missing-data" problem of causal inference by ensuring that, in expectation, unobserved counterfactual outcomes are comparable across groups. Formally, treatment assignment is independent of all pre-treatment variables, whether measured or unmeasured, such that simple between-group contrasts yield unbiased estimates, and any residual imbalance is attributable to chance and quantifiable. Crucially, valid inference follows from the design itself rather than from correct specification of outcome models, assumptions about functional form, or complete measurement of confounders. Randomized trials therefore protect against both known and unknown confounding and provide a coherent framework for quantifying uncertainty. A "target trial" represents the hypothetical randomized experiment conducted in the same population, comparing the same interventions and outcomes, that would directly answer the causal question of interest. It serves as the benchmark against which non-randomized evidence is evaluated [3, 4].

When evidence from randomized trial(s) is unavailable, infeasible, untimely, or unethical, well-designed non-randomized studies can provide informative evidence, particularly when they offer greater directness with respect to the population, intervention, comparator, or outcomes of interest. However, their credibility hinges on design rather than analysis. When randomization is not possible, observational studies should be explicitly designed to approximate randomized experiments by defining the causal question in advance, separating design from analysis, and achieving balance between treated and control groups using only pre-treatment information, without access to outcome data [2]. Consider an observational study comparing two biologics in a registry. A conventional approach would define exposure from dispensing records, select covariates based on the literature, and apply some regression model to estimate an adjusted effect. The target trial framework imposes a different discipline: eligibility must be defined before examining outcomes, time zero must be anchored at treatment initiation, the comparator strategy must be specified in advance, and the analysis plan must be fixed before outcome data are accessed. Causal inference in observational settings therefore depends less on post-hoc statistical adjustment and more on whether the study design successfully reconstructs a plausible assignment mechanism comparable to that of a randomized experiment [2].

In this commentary, we demonstrate how principles derived from randomized trials should continue to anchor causal reasoning, while clarifying how trial emulation approaches can be used to approximate randomized comparisons in observational data, and how their assumptions, sources of bias, and implications for decision-making should be made explicit.

Modern evidence hierarchies and the role of real-world data

Traditional evidence hierarchies have long placed randomized controlled trials (RCTs) and their meta-analyses at the apex, implicitly treating study design as a surrogate for (causal) validity [5]. However, contemporary methodological work has shown that such hierarchies are overly simplistic and risk obscuring the more fundamental determinants of evidentiary credibility, namely risk of bias, precision, consistency, and directness [6]. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group challenges the notion that certainty in evidence can be inferred from study design alone and discourages rigid interpretations of traditional evidence hierarchies [7]. Instead, GRADE treats study design as a starting point and evaluates certainty using a structured, outcome-specific framework that explicitly considers risk of bias, imprecision, inconsistency, indirectness, publication bias, and contextual factors such as magnitude of effect, dose–response relations, and residual confounding [7]. Within this framework, evidence may be rated down or up based on methodological rigor and relevance to the decision context, regardless of whether it originates from randomized or non-randomized studies. As a result, certainty in evidence reflects not the label of the design but the credibility, precision, and applicability of the estimates to the clinical question at hand [6].

In parallel, the increasing availability of real-world data has expanded the role of non-randomized evidence in addressing clinically important questions for which randomized trials are infeasible, unethical, or untimely [8]. As emphasized by Hernán and Robins [3, 9], causal inference from observational data should be understood not as a fundamentally different enterprise, but as an explicit attempt to emulate the randomized trial that would ideally answer the question of interest, the so-called target trial. When the protocol of this target trial is clearly specified, and the assumptions required for its emulation are made transparent, analyses of real-world data can provide decision-relevant estimates of effectiveness and safety [10], albeit with greater reliance on unverifiable assumptions than randomized evidence, implying a persistent risk of residual confounding [10, 11].

Taken together, these developments challenge rigid evidence hierarchies and favor a question-driven approach to evidence evaluation. Randomized trials remain the reference standard for causal inference because their design enforces exchangeability by construction. Thus, when randomization is not feasible, observational data can inform causal questions, but only if studies are deliberately designed to emulate the randomized trial that would have addressed the question of interest. Under this framework, credibility depends less on the study design and more on whether the assignment mechanism is plausibly reconstructed, design choices are made without access to outcome data, and analyses target clearly defined causal contrasts. When these conditions are met, prospective real-world data from non-randomized settings can provide valuable, and at times indispensable, evidence.

Appraising the target trial framework

The target trial framework provides a disciplined approach to causal inference from observational data by requiring investigators to first specify the protocol of the randomized trial that would ideally answer the causal question of interest and then assess how closely that trial can be emulated with available data. By forcing explicit definition of eligibility criteria, treatment strategies, time zero, outcomes, and causal contrasts, the framework prioritizes design over analysis and helps prevent common self-inflicted biases, including immortal time bias, selection of prevalent users, and inappropriate conditioning on post-treatment variables. Its primary value lies not in overcoming the inherent limitations of observational data, but in clarifying the causal estimand, making assumptions explicit, and providing a transparent benchmark against which the credibility of non-randomized evidence can be judged.

Target trial emulation

In the following, we describe a practical, step-by-step tutorial for conducting and reporting a target trial emulation study, aligned with the two-step framework ([i] specify the target trial, [ii] then emulate it [4]), and with reporting expectations from the TARGET statement [12].

Step 1. State the causal question in "target trial" form

Write the question as if you were designing a pragmatic randomized trial:

- Population (eligibility at enrolment/baseline)
- Treatment strategies (initiation, dosing, duration, discontinuation rules, allowable switches)
- Comparator strategy (if not placebo, then what?)
- Outcome(s) (definition and measurement)
- Follow-up window (A trial endpoint is the operationalized use of one or more outcome measures within a specific estimand framework, defined by the timing of assessment relative to baseline [t = 0], how the measure is summarized over a prespecified follow-up period, and how it is compared between treatment groups)
- Causal contrast (Intention-to-treat [ITT] versus per-protocol, with the former most often preferred by end users of evidence-based medicine)
- Effect measures (absolute and relative)

The TARGET statement recommends that readers should be able to recognize, from the abstract and introduction, that the study is an observational emulation of a target trial and to see the causal question clearly summarized.

Step 2. Draft the target trial protocol explicitly

Create a concise "protocol" subsection (or a table) that includes, at minimum:

- Eligibility criteria
- Treatment strategies
- Assignment procedures (randomization in the hypothetical target trial)
- Outcomes
- Start/end of follow-up (time zero and censoring rules)
- Causal estimand(s) and contrasts (ITT, per-protocol)
- Identifying assumptions and analysis plan

Step 3. Map each protocol component to your observational data

Now describe, point-by-point, how you operationalize the protocol using the available real-world data:

- How eligibility is ascertained (and what you cannot capture well)
- How treatment initiation and adherence are captured (prescriptions, dispensing, administrations)
- How outcomes are defined (codes/algorithms; any validation evidence)
- How follow-up ends (administrative end, death registry, loss to follow-up)

TARGET note: Investigators should describe any imperfections in the mapping of each protocol component to the available data, including variables that could not be well operationalized and the reasons why [12]. The framework does not resolve data quality problems, but it makes them explicit before the analysis begins.

Step 4. Define time zero correctly and synchronize the "three alignments"

A core design rule is that time zero must coincide with:

1. meeting eligibility criteria,
2. assignment/classification to a treatment strategy, and
3. starting outcome counting.

When these three elements are not synchronized, patients who survive long enough to receive treatment accumulate event-free time that is attributed to the treatment group but precedes any exposure. This is immortal time bias, which produces a systematic overestimate of treatment.

TARGET note: This synchronization is repeatedly emphasized because violations generate design-related biases such as immortal time bias and selection of prevalent users.

Step 5. Choose your causal contrast

- **ITT analogue:** classify individuals by strategy at baseline and follow regardless of later adherence, adjusting for baseline confounders.
- **Per-protocol effect:** requires accounting for deviations from the assigned strategy by censoring at deviation and using appropriate methods to adjust for prognostic factors related to adherence (often g-methods/weighting).

It is important to recognize that the ITT and per-protocol effects address fundamentally different causal questions rather than representing competing estimators of the same quantity. The ITT effect reflects the consequence of assigning a treatment strategy regardless of subsequent adherence, which is relevant to prescribing and policy decisions. The per-protocol effect reflects the consequence of following the strategy as specified throughout follow-up, which is more relevant to understanding the biological effect of treatment. The choice between them should be guided by the clinical question rather than analytical convenience.

TARGET note: Report which estimand is primary (subject for the main reporting in the manuscript and why).

Step 6. Pre-specify confounders using a causal rationale

Because observational assignment is not random, the analysis typically relies on conditional exchangeability given measured baseline covariates (no unmeasured confounding conditional on those covariates). TARGET recommends justifying covariate choice, often via a causality diagram, and describing how those variables are measured.

Practical implementation:

- Define baseline covariates collected at or before time zero
- Avoid adjusting for post-baseline variables that may be consequences of treatment (risk of collider bias or adjusting away part of the effect)

Step 7. Handle ambiguity in treatment assignment with appropriate design tools

If treatment strategies are not distinguishable at time zero (for example, "initiate within 2 days"), you may need:

- **Cloning:** at time zero, each eligible individual is simultaneously represented in all treatment strategies they are compatible with, creating one copy per strategy. Each copy follows its assigned strategy and is censored at the point where the individual's observed data become inconsistent with that strategy. Because all copies share identical baseline characteristics, the cloning step preserves comparability at time zero; inverse probability of censoring weights subsequently correct for the fact that censoring is related to prognosis [3, 13].

- **Censoring:** a participant is removed from follow-up at the moment its observed treatment history becomes incompatible with the assigned strategy. This artificial censoring is by design informative and must therefore be corrected for in the weighting step [3, 13].
- **Weighting:** inverse probability of censoring weights are estimated from a model for the probability of remaining uncensored at each time point given prior covariate and treatment history. Upweighting copies that remain on-strategy creates a pseudo-population in which adherence is independent of prognosis, recovering the distribution of outcomes that would have been observed under full adherence to each strategy [3, 13].

Step 8. Choose an estimator consistent with your estimand

Common approaches (depending on estimand and data structure):

- **Outcome regression or standardization:** a regression model is fitted for the outcome as a function of treatment strategy and baseline covariates. Predicted outcomes are then computed for each individual under each strategy and averaged across the study population. This yields the marginal absolute risk under each strategy, standardized to the covariate distribution of the whole population [3, 4, 13, 14].
- **Propensity score methods, including inverse probability weighting:** the propensity score is the probability of receiving a given treatment strategy conditional on baseline covariates. Weighting each individual by the inverse of their propensity score creates a pseudo-population in which treatment assignment is independent of measured covariates, mimicking randomization. The weighted analysis then estimates marginal treatment effects without directly modelling the outcome-covariate relationship [3, 4, 13, 15, 16].
- **G-methods for per-protocol and time-varying settings:** G-methods—including the parametric g-formula, marginal structural models estimated via inverse probability weighting, and g-estimation—are a family of estimators designed for settings where confounders are time-varying and themselves affected by prior treatment [15]. In this setting, standard regression adjustment is insufficient because conditioning on a time-varying confounder that lies on the causal pathway can both remove part of the treatment effect and introduce collider bias. G-methods instead model the treatment and censoring processes to produce estimates that reflect the outcomes that would have been observed had all individuals followed the assigned strategy throughout follow-up [3, 13–15, 17].

TARGET note: Report the software and model details sufficient for reproducibility, including functional forms and handling of missing data. For analysts working in R, relevant packages include *TrialEmulation* for implementing the cloning, censoring, and weighting approach; *WeightIt* and *ipw* for propensity score and inverse probability weighting methods; *gfoRmula* for the parametric g-formula.

Step 9. Plan and conduct sensitivity analyses and robustness checks

At minimum:

- Alternative operationalizations of exposure, eligibility, and outcomes
- Alternative confounding adjustment strategies
- Quantitative bias analysis or negative/positive controls where relevant

TARGET note: These are integral to interpretation and expects them to be reported, not relegated to vague statements.

Step 10. Report results like a pragmatic trial

TARGET is explicit that the results should look trial-like. Key reporting elements:

1. **Participant flow:** numbers assessed, eligible, assigned/classified; flow diagram strongly recommended (consolidated standards of reporting trials [CONSORT]-like).

2. **Baseline table by strategy:** ideally show balance before/after adjustment.
3. **Follow-up summary:** duration, censoring reasons, competing events as relevant.
4. **Outcome reporting:** frequencies and, for time-to-event, absolute risks over time with precision.
5. **Effect estimates:** report both absolute and relative measures when applicable, each with measures of precision.

It should be highlighted as a practical benefit of this approach: it encourages reporting absolute risk under each strategy (often more decision-relevant than hazard ratios alone). Since GRADE evidence profiles require absolute effect estimates over a defined time horizon for each outcome of interest, the implementation of the target trial emulation framework into common practice in rheumatology would facilitate integration into systematic reviews and guideline panels [18].

To illustrate the process of going through the 10 steps of a target trial emulation mentioned above, Figure 1 goes through the 2009 Hetland study comparing Infliximab, Etanercept and Adalimumab head-to-head, based on observational data from the Danish nationwide rheumatology biologics registry (DANBIO) register [19].

Conclusion and perspectives for evidence-based rheumatology

Causal inference in rheumatology increasingly relies on integrating evidence from randomized trials and real-world data, rather than treating them as competing sources of knowledge. While randomized trials remain the reference standard for establishing causal effects, they cannot feasibly address all clinically relevant questions in heterogeneous, chronic, and evolving rheumatic diseases [19]. In this context, the target trial framework provides a coherent design-based approach for extracting decision-relevant causal estimates from observational data while making assumptions, limitations, and sources of bias explicit [20, 21].

For evidence-based rheumatology, the value of the target trial framework lies not in elevating non-randomized studies to the status of randomized trials, but in clarifying what causal question is being answered, under which conditions, and with what degree of credibility. By enforcing explicit specification of eligibility criteria, treatment strategies, timing, outcomes, and causal estimands, target trial emulation reduces common design-related biases and improves transparency, interpretability, and reproducibility. Importantly, it aligns observational analyses more closely with the logic of pragmatic trials and encourages reporting formats that are directly relevant to clinical and guideline decision-making, including absolute risks under alternative treatment strategies.

The target trial emulation framework carries implications for the evaluation of new therapeutic targets, where the exposure of interest is often a biomarker measured at a single time point. Mendelian randomization and target trial emulation are complementary approaches in this setting: Mendelian randomization uses genetic variants as natural experiments to probe causal relevance, and target trial emulation operationalizes the clinical question of whether a biomarker-defined population benefits from a specific intervention. In both cases, the biomarker measurement must genuinely precede treatment assignment, and when the biomarker itself drives prescribing, confounding by indication is structural rather than incidental. Specifying time zero in relation to the biomarker measurement and justifying its role using a causal diagram allows investigators to distinguish whether the question concerns the predictive value of the biomarker, its causal role, or its capacity to define a population in whom a particular treatment is more effective.

As rheumatology increasingly turns to registries, electronic health records, and other real-world data sources to evaluate long-term effectiveness, safety, and treatment sequencing, adopting a target trial mindset offers a principled way forward [9]. It provides a shared conceptual language for clinicians, methodologists, and guideline panels to assess when non-randomized evidence can complement randomized trials, when it can provisionally inform decisions, and when uncertainty remains too great [8].

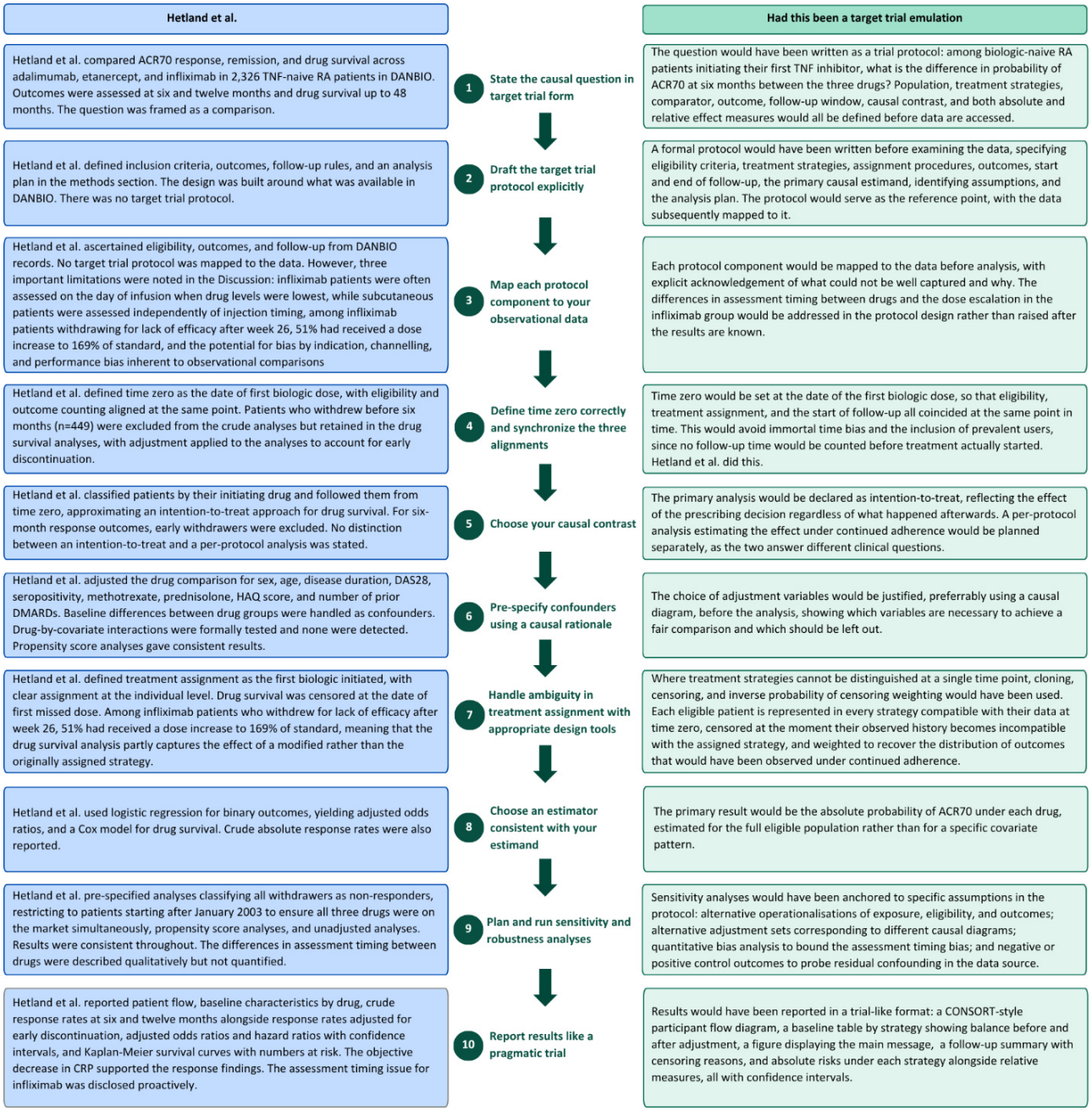

Figure 1. Applying the ten steps to an already well-conducted observational study [19].

In this way, the target trial framework strengthens the methodological foundations of evidence-based rheumatology by shifting emphasis from analytic sophistication to causal clarity, from design labels to estimands, and from implicit assumptions to transparent reasoning [22].

Abbreviations

GRADE: Grading of Recommendations Assessment, Development and Evaluation

ITT: intention-to-treat

RCTs: randomized controlled trials

Declarations

Author contributions

TH: Project administration, Software, Visualization, Writing—original draft, Writing—review & editing, Funding acquisition. RC: Conceptualization, Project administration, Supervision, Validation, Writing—original draft, Writing—review & editing, Funding acquisition. Both authors read and approved the submitted version.

Conflicts of interest

Robin Christensen reports providing consulting on biostatistical matters, including trial design and statistical inference, for ZPD A/S (ScanDroiti™), Image Analysis Ltd (trading as IAG, Image Analysis Group, UK), Compass Communications Ltd., and Ascendis Pharma A/S, with payments made directly to him as an individual. He serves as Statistical Editor for Osteoarthritis and Cartilage (OAC) and Acta Orthopaedica (ActaO), with payments made directly to him. He is a founding member of the OMERACT Technical Advisory Group, a member of the GRADE Working Group, serves on the Statistical Advisory Board for BMJ Open, and is on the editorial boards of Arthritis Care & Research (AC&R) and Arthritis Research & Therapy (AR&T). Robin Christensen, who is an Associate Editor of Exploration of Musculoskeletal Diseases, had no involvement in the decision-making or the review process of this manuscript. The other author declares no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

Not applicable.

Funding

Section for Biostatistics and Evidence-Based Research, the Parker Institute, Bispebjerg and Frederiksberg Hospital is supported by a core grant from the Oak Foundation (OFIL-24-074). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Publisher's note

Open Exploration maintains a neutral stance on jurisdictional claims in published institutional affiliations and maps. All opinions expressed in this article are the personal views of the author(s) and do not represent the stance of the editorial team or the publisher.

References

1. Little RJ, Rubin DB. Causal Effects in Clinical and Epidemiological Studies Via Potential Outcomes: Concepts and Analytical Approaches. Annu Rev Public Health. 2000;21:121–45. [DOI] [PubMed]
2. Rubin DB. The design *versus* the analysis of observational studies for causal effects: parallels with the design of randomized trials. Stat Med. 2006;26:20–36. [DOI] [PubMed]

3. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available: Table 1. Am J Epidemiol. 2016;183:758–64. [DOI] [PubMed] [PMC]
4. Hernán MA, Wang W, Leaf DE. Target Trial Emulation. JAMA. 2022;328:2446–7. [DOI] [PubMed]
5. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017; 390:415–23. [DOI] [PubMed]
6. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21:125–7. [DOI] [PubMed] [PMC]
7. Zhang Y, Akl EA, Schünemann HJ. Using systematic reviews in guideline development: The GRADE approach. Res Synth Methods. 2018;10:312–29. [DOI] [PubMed]
8. Schünemann HJ, Tugwell P, Reeves BC, Akl EA, Santesso N, Spencer FA, et al. Non-randomized studies as a source of complementary, sequential or replacement evidence for randomized controlled trials in systematic reviews on the effects of interventions. Res Synth Methods. 2013;4:49–62. [DOI] [PubMed]
9. Hernán MA. Methods of Public Health Research — Strengthening Causal Inference from Observational Data. N Engl J Med. 2021;385:1345–8. [DOI] [PubMed]
10. Hernán MA, Dahabreh IJ, Dickerman BA, Swanson SA. The Target Trial Framework for Causal Inference From Observational Data: Why and When Is It Helpful? Ann Intern Med. 2025;178:402–7. [DOI] [PubMed] [PMC]
11. Agoritsas T, Merglen A, Shah ND, O'Donnell M, Guyatt GH. Adjusted Analyses in Studies Addressing Therapy and Harm. JAMA. 2017;317:748–59. [DOI] [PubMed]
12. Cashin AG, Hansford HJ, Hernán MA, Swanson SA, Lee H, Jones MD, et al. Transparent reporting of observational studies emulating a target trial: the TARGET Statement. BMJ. 2025;390:e087179. [DOI] [PubMed]
13. Hernán MA, Robins JM. Causal Inference: What If? Boca Raton: Chapman & Hall/CRC; 2020.
14. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. Int J Epidemiol. 2009;38:1599–611. [DOI] [PubMed] [PMC]
15. Robins JM, Hernán MA, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. Epidemiology. 2000;11:550–60. [DOI] [PubMed]
16. ROSENBAUM PR, RUBIN DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41–55. [DOI]
17. Smith LH, García-Albéniz X, Chan JM, Zhao S, Cowan JE, Broering JM, et al. Emulation of a target trial with sustained treatment strategies: an application to prostate cancer using both inverse probability weighting and the g-formula. Eur J Epidemiol. 2022;37:1205–13. [DOI] [PubMed]
18. Skoetz N, Goldkuhle M, van Dalen EC, Akl EA, Trivella M, Mustafa RA, et al. GRADE guidelines 27: how to calculate absolute effects for time-to-event outcomes in summary of findings tables and Evidence Profiles. J Clin Epidemiol. 2020;118:124–31. [DOI] [PubMed]
19. Hetland ML, Christensen IJ, Tarp U, Dreyer L, Hansen A, Hansen IT, et al. Direct comparison of treatment responses, remission rates, and drug adherence in patients with rheumatoid arthritis treated with adalimumab, etanercept, or infliximab: Results from eight years of surveillance of clinical practice in the nationwide Danish DANBIO registry. Arthritis Rheum. 2009;62:22–32. [DOI] [PubMed]
20. Amris K, Christensen R, Wæhrens EE, Duhn PH, Henriksen M. Impact of enhanced pain knowledge on core outcomes in fibromyalgia patients with high self-reported pain education needs: a target trial emulation using the DANFIB registry. RMD Open. 2025;11:e005576. [DOI] [PubMed] [PMC]
21. Rasmussen LE, Hoffmann AG, Blanche P, Espersen F, Justesen TF, Rasmussen LE, et al. Surgeon Training and Revision Rates After Patellofemoral Arthroplasty. JAMA Netw Open. 2025;8:e2517825. [DOI] [PubMed] [PMC]
22. Christensen R, Ranstam J, Overgaard S, Wagner P. Guidelines for a structured manuscript: Statistical methods and reporting in biomedical research journals. Acta Orthop. 2023;94:243–9. [DOI] [PubMed] [PMC]