Design Transparency and Data Fitness When Generating RWE for Decisions: the Importance of Target Trial Emulation

Rob Reynolds, ScD, FISPE
VP, Epidemiology
GSK R&D

gsk.com

Target Trial Emulation in HTA RWD Submissions: A Best Practice Not Embraced – Methodological Challenge or Manufacturers' Reluctance?

Disclosures

- I am employee and shareholder of GSK
- My experience with RWD / RWE is primarily for regulatory approvals (initial and/or additional indications post-market) and/or post-approval comparative safety studies at GSK and another large pharmaceutical company
- My intent today is to describe experience and lessons applying target trial emulation to RWE intended for regulatory submission

Outline

Why target trial emulation matters

Good design is more than a target trial

Examples and concluding thoughts

GSK

Target Trial Emulation: Not New, But Greater Awareness Today

Specify the trial you would have done, if you could, before designing your RW study

American Journal of Epidemiology
Vol. 183, No. 8
DOI: 10.1093/aje/kwv254
Advance Access publication:
March 18, 2016

Practice of Epidemiology

Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available

Miguel A. Hernán* and James M. Robins

* Correspondence to Dr. Miguel A. Hernán, Department of Epidemiology, 677 Huntington Avenue, Boston, MA 02115 (e-mail: miguel_hernan@post.harvard.edu).

Initially submitted December 9, 2014; accepted for publication September 8, 2015.

The target trial emulated using observational data will typically be a pragmatic trial, that is, one in which treatment strategies are compared under the usual conditions in which they will be applied (9, 10). For instance, we cannot emulate a placebo-controlled trial with tight monitoring and enforcement of adherence to the study protocol.

Seven Target Trial Protocol Components

Protocol Component	Description
Eligibility criteria	How the patient population is recruited into the trial.
Treatment strategies	Each of the clinical interventions that are to be compared.
Treatment assignment	How participants will be assigned to each treatment strategy at baseline.
Start and end of follow-up	Define when the follow-up period starts and ends for each participant.
Outcomes	Outcomes of interest and how to ascertain them.
Causal contrast	What comparative effects of the treatment strategies will be estimated.
Statistical analysis	How to estimate the intention-to-treat effect or per-protocol effect via intention-to-treat and per-protocol analyses that appropriately adjust for pre- and post-baseline prognostic factors associated with adherence and loss to follow-up.

GSK

Estimand Framework & Target Trial Emulation: A Similar Purpose

ICH
harmonisation for better health

What's Missing from the Estimand Framework?

Baseline confounders & Specifying Time Zero

Final version

Adopted on 20 November 2019

Intercurrent events may be incorporated into treatment, population and/or variable attributes

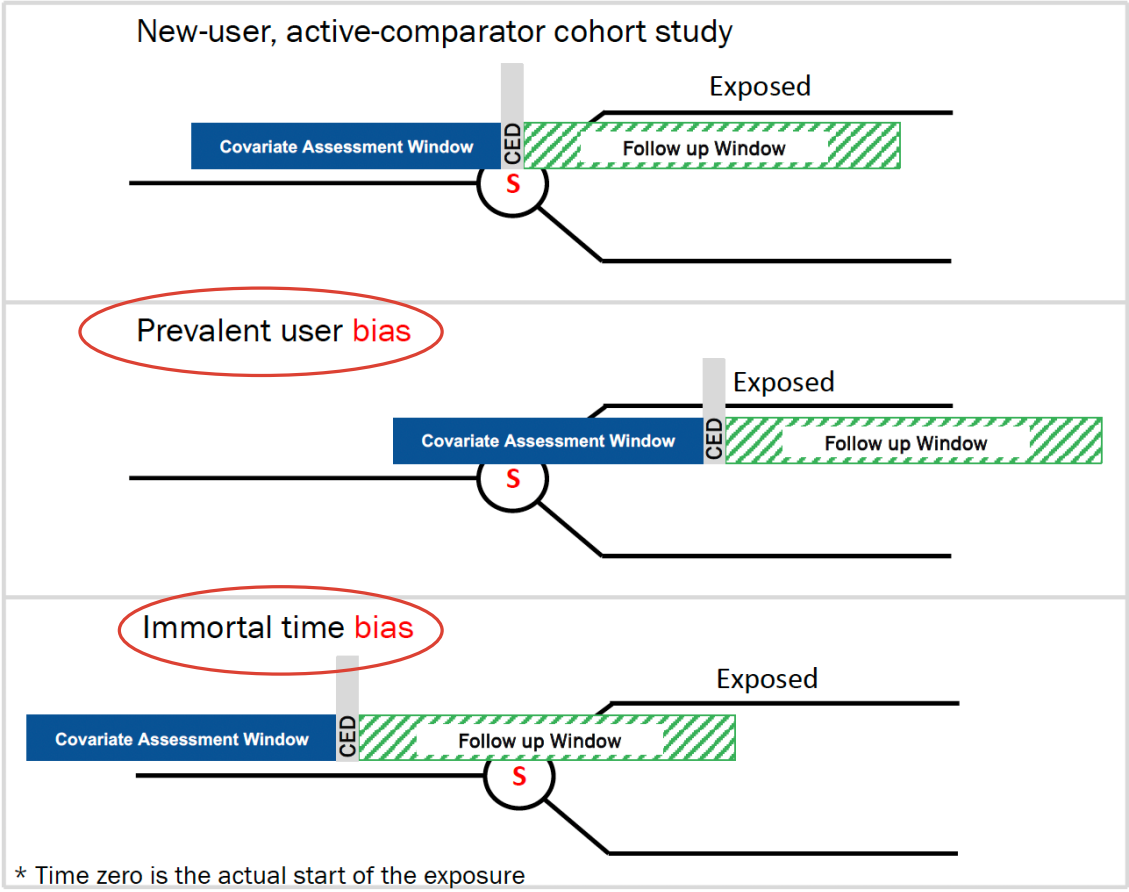
An *estimand* is a precise description of exactly what you want to find out

GSK

For comparison of target trial & estimands frameworks see J. Chen et al. "Estimands in Real-World Evidence Studies" https://arxiv.org/pdf/2307.00190.pdf

Why Use Target Trial Emulation When Designing RWE?

Many examples of self-inflicted bias in RWE by misaligning time zero*



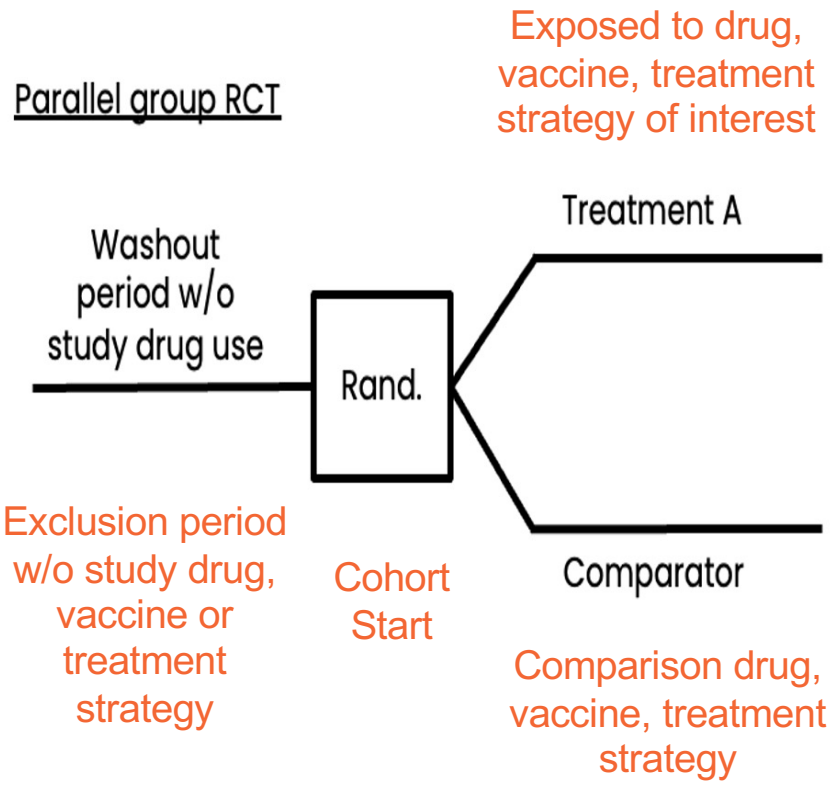
Cohort entry date (CED) is aligned with the start of follow-up at time zero,* the actual treatment start. Pre-treatment covariate assessment is in line with a causal study design.

CED is after time zero, i.e. patients are assigned treatment status by considering ongoing (*prevalent*) treatment that started in the past. Selective dropout before CED will cause bias, particularly if a treatment effect changes with time, gets weaker or stronger with longer treatment.

CED is before time zero, i.e. patients are assigned their treatment status by looking into the future.
To determine future treatment status patients will need to be alive in the future, they are *immortal* for that period. If this is differential it causes bias.

Image from Schneeweiss S & Schneeweiss M. *JID Innovations* 2023;3:100226. See also Hernan MA et al. *J Clin Epi* 79:70-5, 2016

Specifying the Target Trial (the minimum elements)



Everyone's Doing It (Not Well)…Only Fools Rush In?

57% of 200 studies (2020-2022) aiming to emulate a hypothetical target trial failed to describe and/or report their target trial specifications

JAMA Network | Open

Original Investigation | Statistics and Research Methods

Reporting of Observational Studies Explicitly Aiming to Emulate Randomized Trials
A Systematic Review

Harrison J. Hansford, BSc(Hons); Aidan G. Cashin, PhD; Matthew D. Jones, PhD; Sonja A. Swanson, ScD; Nazrul Islam, MD, PhD; Susan R. G. Douglas, BExPhys; Rodrigo R. N. Rizzo, PhD; Jack J. Devonshire, BSc(Hons); Sam A. Williams, BSc(Hons); Issa J. Dahabreh, MD, ScD; Barbra A. Dickerman, PhD; Matthias Egger, MD, MSc; Xabier Garcia-Albeniz, MD, PhD; Robert M. Golub, MD; Sara Lodi, PhD; Margarita Moreno-Betancur, PhD; Sallie-Anne Pearson, PhD; Sebastian Schneeweiss, MD, ScD; Jonathan A. C. Sterne, PhD; Melissa K. Sharp, PhD; Elizabeth A. Stuart, PhD; Miguel A. Hernán, MD, DrPh; Hopin Lee, PhD; James H. McAuley, PhD

Abstract

IMPORTANCE Observational (nonexperimental) studies that aim to emulate a randomized trial (ie, the target trial) are increasingly informing medical and policy decision-making, but it is unclear how these studies are reported in the literature. Consistent reporting is essential for quality appraisal, evidence synthesis, and translation of evidence to policy and practice.

OBJECTIVE To assess the reporting of observational studies that explicitly aimed to emulate a target trial.

EVIDENCE REVIEW We searched Medline, Embase, PsycINFO, and Web of Science for observational studies published between March 2012 and October 2022 that explicitly aimed to emulate a target trial of a health or medical intervention. Two reviewers double-screened and -extracted data on study characteristics, key predefined components of the target trial protocol and its emulation (eligibility criteria, treatment strategies, treatment assignment, outcome[s], follow-up, causal contrast[s], and analysis plan), and other items related to the target trial emulation.

FINDINGS A total of 200 studies that explicitly aimed to emulate a target trial were included. These studies included 26 subfields of medicine, and 168 (84%) were published from January 2020 to October 2022. The aim to emulate a target trial was explicit in 70 study titles (35%). Forty-three studies (22%) reported use of a published reporting guideline (eg, Strengthening the Reporting of Observational Studies in Epidemiology). Eighty-five studies (43%) did not describe all key items of how the target trial was emulated and 113 (57%) did not describe the protocol of the target trial and its emulation.

Key Points

Question How are studies that explicitly aim to emulate a target trial reported?

Findings In this systematic review of 200 studies that explicitly aimed to emulate a target trial, reporting was inconsistent, and studies often did not report all necessary information related to the emulation of the target trial.

Meaning Inconsistent reporting of studies that explicitly aim to emulate a target trial may impair the appraisal, synthesis, and implementation of study findings.

+ Supplemental content

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

RWE Designed to Emulate A Trial Works

Conclusions are similar, contingent on trial design elements & fit for purpose RWD

JAMA | Original Investigation

Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses
Results of 32 Clinical Trials

Shirley V. Wang, PhD, ScM; Sebastian Schneeweiss, MD, ScD; and the RCT-DUPLICATE Initiative

RESULTS In these highly selected RCTs, the overall observed agreement between the RCT and the database emulation results was a Pearson correlation of 0.82 (95% CI, 0.64-0.91), with 72% meeting statistical significance, 66% estimate agreement, and 75% standardized difference agreement. In a post hoc analysis limited to 16 RCTs with closer emulation of trial design and measurements, concordance was higher (Pearson *r*, 0.93; 95% CI, 0.79-0.97; 94% meeting statistical significance, 88% estimate agreement, 88% standardized difference agreement). Weaker concordance occurred among 16 RCTs for which close emulation of certain design elements that define the research question (PICOT) with data from insurance claims was not possible (Pearson *r*, 0.53; 95% CI, 0.00-0.83; 50% meeting statistical significance, 50% estimate agreement, 69% standardized difference agreement).

16 RCTs were well emulated in design and measurement by RWE

16 RCTs were more difficult to emulate by RWE, due to patient selection during run-in phases or treatment counter to clinical practice

↓

Our concerns are real world, however, as are our research questions

↓

Design the hypothetical, *pragmatic* target trial

GSK

A Detour: An Applied Design and Data Feasibility Framework

The target trial is one step in designing a ”good” real-world study

2019 Clinical Pharmacology and Therapeutics

STATE OF THE ART

A Structure
Comparati
Generate V
Evidence fo

Nicolle M. Gatto[1,2,3,*], R

Real-world evidence provid
engender trust that eviden
that underlies study design
Structured Preapproval and
real-world Evidence (SPACE
validity concerns, and for d

REVIEW *2021 Clinical Pharmacology and Therapeutics*

The Structured
Purpose Data:
Framework

Nicolle M. Gatto[1,2,3,*], Ulka B. C
and Robert F. Reynolds[3,5]

To complement real-world evidence
Comparative study design framewo
elucidated a process for designing
provide a structured framework for
grade, fit-for-purpose data, which c
for a RWE program. The process wa
assessments of existing data sourc
the SPACE framework, the Structur

2023 Clinical Pharmacology and Therapeutics MINI-REVIEW

A Structured Process to Identify Fit-for-Purpose Study Design and Data to Generate Valid and Transparent Real-World Evidence for Regulatory Uses

Nicolle M. Gatto[1,2,3,*], Sarah E. Vititoe[1], Emily Rubinstein[1], Robert F Reynolds[3,4] and Ulka B. Campbell[1,2]

Generating evidence from real-world data requires fit-for-purpose study design and data. In addition to validity, decision makers require transparency in the reasoning that underlies study design and data source decisions. The 2019 Structured Preapproval and Postapproval Comparative Study Design Framework to Generate Valid and Transparent Real-World Evidence (SPACE) and the 2021 Structured Process to Identify Fit-For-Purpose Data (SPIFD)—intended to be used together—provide a step-by-step guide to identify decision grade, fit-for-purpose study design and data. In this update (referred to as “SPIFD2” to encompass both the design and data aspects) we provide an update to these frameworks that combines the templates into one, more explicitly calls for articulation of the hypothetical target trial and sources of bias that may arise in the real-world emulation, and provides explicit references to the Structured Template and Reporting Tool for Real-World Evidence (STaRT-RWE) tables that we suggest using immediately after invoking the SPIFD2 framework. Following the steps recommended in the SPIFD2

GSK AETION

The SPIFD2 Framework

The **SPIFD2 Framework** provides a step-by-step principled process for RW design & data selection with improved usability and transparency

SPIFD2 Steps

Step 1: State the research aim, question, and objectives
1a: Overarching research aim
1b: Research question
1c: Primary objective(s)
Step 2: Describe the hypothetical target trial
2a: Conceptual definition
2b: Target trial definition
Step 3: Describe RW emulation of the hypothetical target trial
Step 3a: Best-case operationalization under routine clinical care
(See SPIFD2 Table S2 for potential confounder identification)
Step 4: Identify a fit-for-purpose dataset
4a: Minimal criteria for valid operationalization in RWD source
4b: RWD minimal criteria ranking with regard to uniqueness and importance
4c: Detailed data source feasibility assessment findings and summary heat map
4d: Practical considerations
Step 5: Document final RW operationalization, rationale, validity concerns, & approaches to address these concerns
5a: Final RW operationalization in STaRT-RWE
5b:. Rationale for final operationalization of RW study design element or variable
5c: Validity concerns related to RW operationalization, selected RW data source and/or study design
5d: How these validity concerns are/will be addressed

Source: Gatto et al, CPT 2023

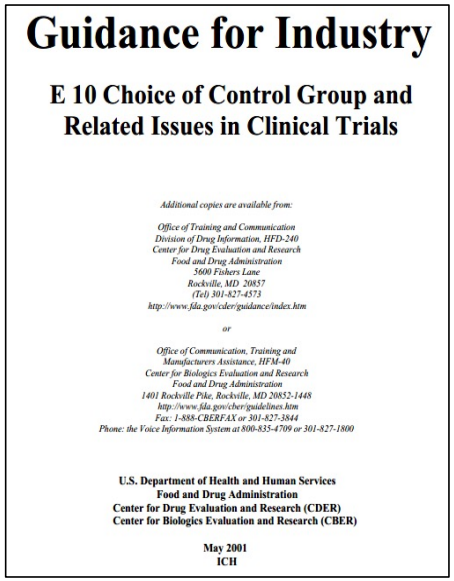
Scientific Feasibility: Key Principles

- Facilitating causal inference requires a principled approach to study design following scientific best practices
- Conceptualizing the target trial facilitates robust operationalization of key study elements
- Published tools and downloadable templates aid in critical decision-making and enable reproducibility and evaluation
- Detailed documentation and registration prior to implementation enables trust

Two Examples Relevant to Regulatory and HTA Decisions

Applying Target Trial Emulation to the Use of RWD

External Control (Comparator) Arms



Guidance for Industry

E 10 Choice of Control Group and Related Issues in Clinical Trials

Additional copies are available from:

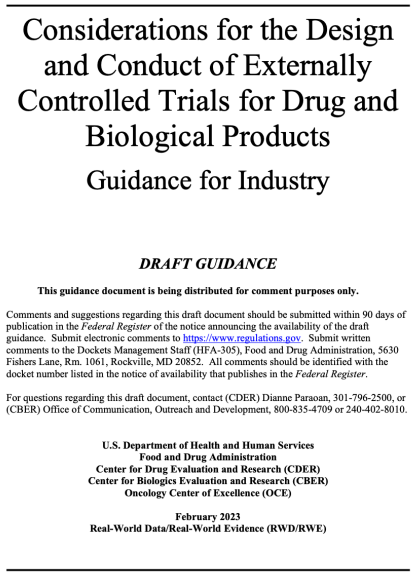
Office of Training and Communication
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
5600 Fishers Lane
Rockville, MD 20857
(Tel) 301-827-4573
http://www.fda.gov/cder/guidance/index.htm

or

Office of Communication, Training and Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
http://www.fda.gov/cber/guidelines.htm
Fax: 1-888-CBERFAX or 301-827-3844
Phone: the Voice Information System at 800-855-4709 or 301-827-1800

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

May 2001
ICH



Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

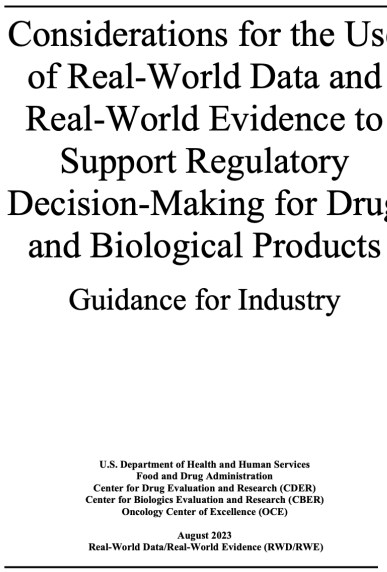
For questions regarding this draft document, contact (CDER) Dianne Paraoan, 301-796-2500, or (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

February 2023
Real-World Data/Real-World Evidence (RWD/RWE)

Compares a group of subjects receiving the test treatment with a group of patients external to the study, rather than to an internal control group consisting of patients from the same population assigned to a different treatment. Compares through benchmarking or patient-level w statistical adjustment.

Post-Approval Effectiveness Study

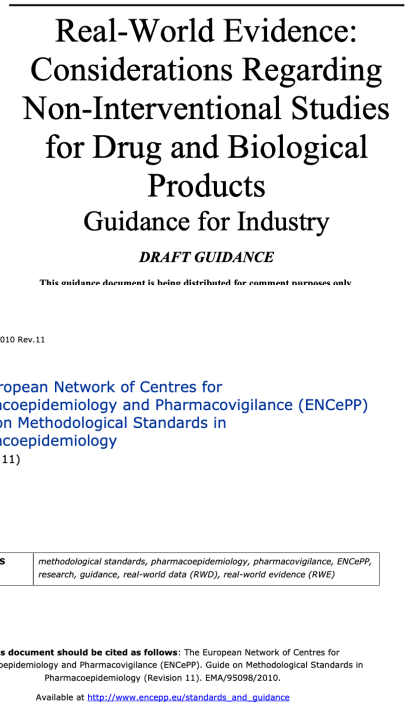


Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products

Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

August 2023
Real-World Data/Real-World Evidence (RWD/RWE)



Real-World Evidence: Considerations Regarding Non-Interventional Studies for Drug and Biological Products

Guidance for Industry

DRAFT GUIDANCE

EMA/95098/2010 Rev.11

The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology

(Revision 11)

KEYWORDS	*methodological standards, pharmacoepidemiology, pharmacovigilance, ENCePP, research, guidance, real-world data (RWD), real-world evidence (RWE)*

This document should be cited as follows: The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (Revision 11). EMA/95098/2010.

Available at http://www.encepp.eu/standards_and_guidance

External Control (Comparator) Arms Using RWD

GSK

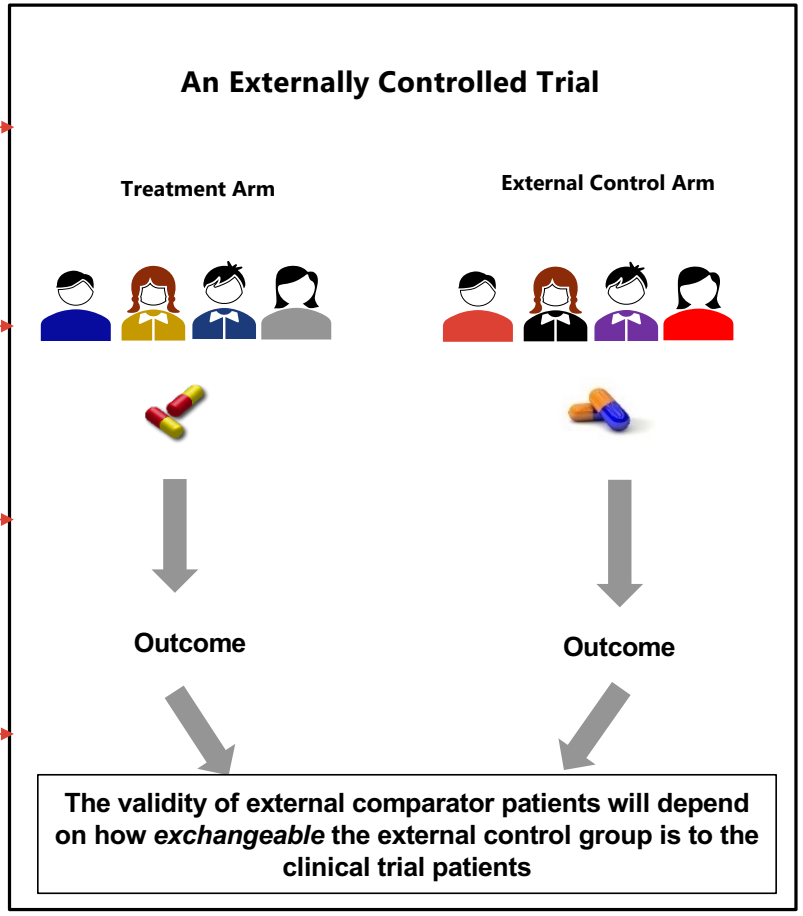
When using RWD for an external control, the target is an *actual* trial

Patients selected from a different source population to the clinical trial treatment arm

Treatment administration & compliance measures may be different

Follow-up frequency and duration, and clinical care may be different

Outcome type & assessment methods may be different



When creating an external control arm, potential bias is addressed by emulating the treatment arm of a single arm (or randomized arm) trial wherever possible

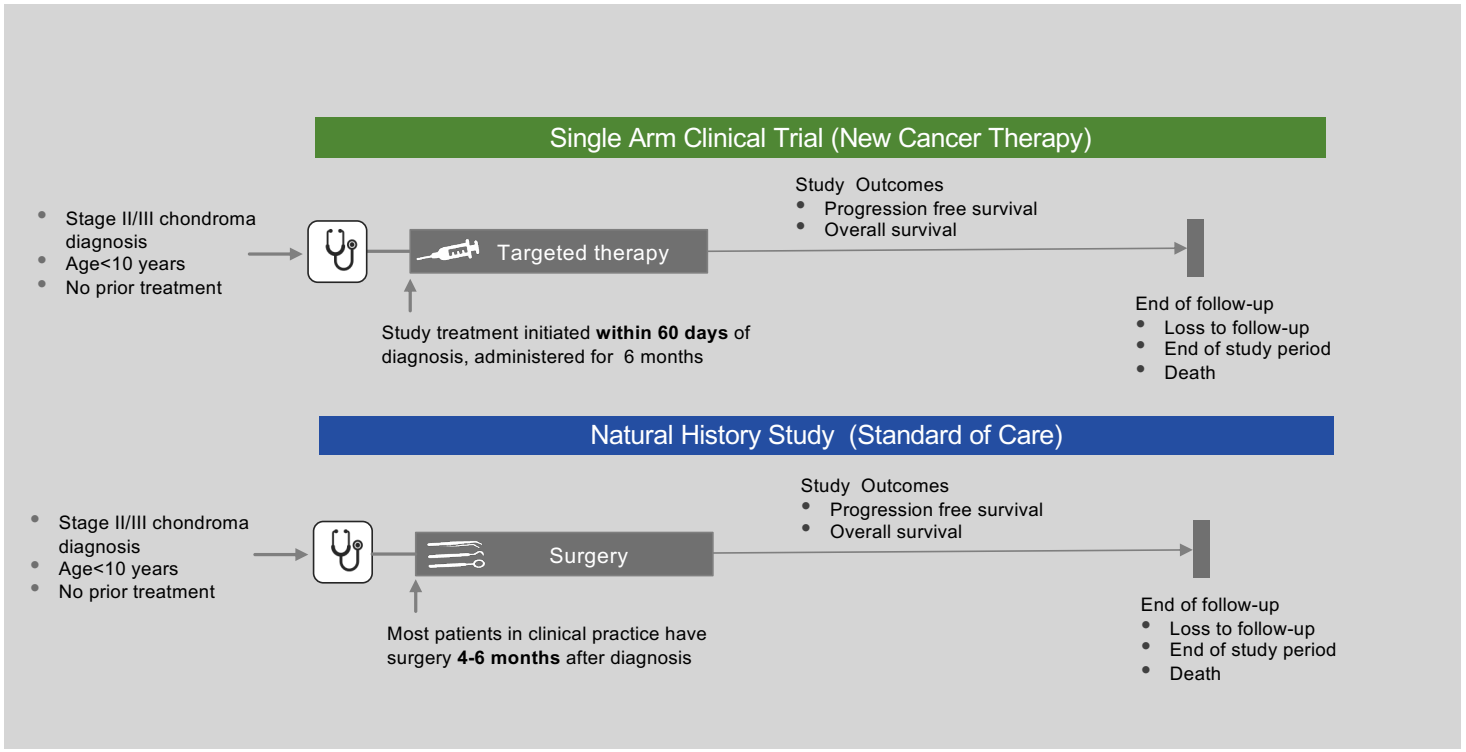
Comparability to trial population: similar patient selection, key variables measured, measurement is emulated or translated to real-world

GSK

Example 1: Single arm clinical trial (the “target”) with an ECA

Why single arm clinical trial study design?

- Disease with high mortality
- Rare disease
- High risk for complications with standard of care



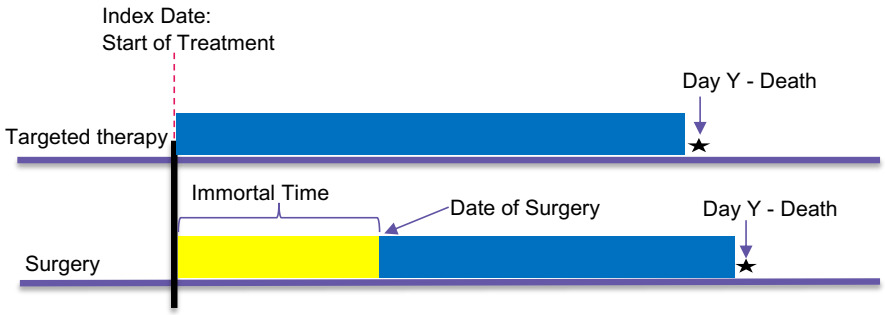
Example 1: Emulate Elements of the Actual Trial Using RWD

Protocol component	Single Arm Clinical Trial (the Target Trial)	External Control (Comparator) Arm using RWD
Design	Multicenter single arm trial assessing the efficacy of targeted therapy in pediatric patients with stage II/III chondroma	A real-world external control evaluating the effectiveness of surgery in pediatric patients with stage II/III chondroma (for benchmarking or statistically adjusted comparison to single arm)
Eligibility criteria	Pediatric patients aged <10 years newly diagnosed with stage II/III chondroma without prior treatment	Pediatric patients aged <10 years newly diagnosed with stage II/III chondroma without prior treatment
Treatment strategies	Targeted therapy	Surgery
Treatment assignment	Patients received targeted therapy treatment strategy	Patients received surgery treatment strategy as part of standard of care (SOC). Adjustment of baseline confounders and informative censoring through different analytic methods.
Treatment implementation	Treatment initiated after study enrollment	Treatment initiated after diagnosis based on SOC
Follow-up	Follow-up starts at time zero, when an individual is assigned to the treatment strategy	Follow-up starts at time of diagnosis of stage II/III chondroma which does not correspond to treatment assignment (the cloning, censoring and weighting approach is used to adjust for immortal time bias)
Censoring	Loss to follow-up, study withdrawal, end of study period	Loss to follow-up, end of study period
Outcomes	Disease progression and death from all causes	Disease progression and death from all causes
Causal contrast	Intent-to-treat-effect (effect of being randomized to treatment strategies at baseline regardless of whether the individuals adhere to them during follow-up)	Intent-to-treat-effect (effect of receiving one of treatment strategies regardless of whether the individuals adhere to them during follow-up)
Estimand	Difference in progression free survival and overall survival between treatment arm and the ECA	Difference in progression free survival and overall survival between treatment arm and the ECA

GSK

Cloning, censoring, or weighting to minimize bias

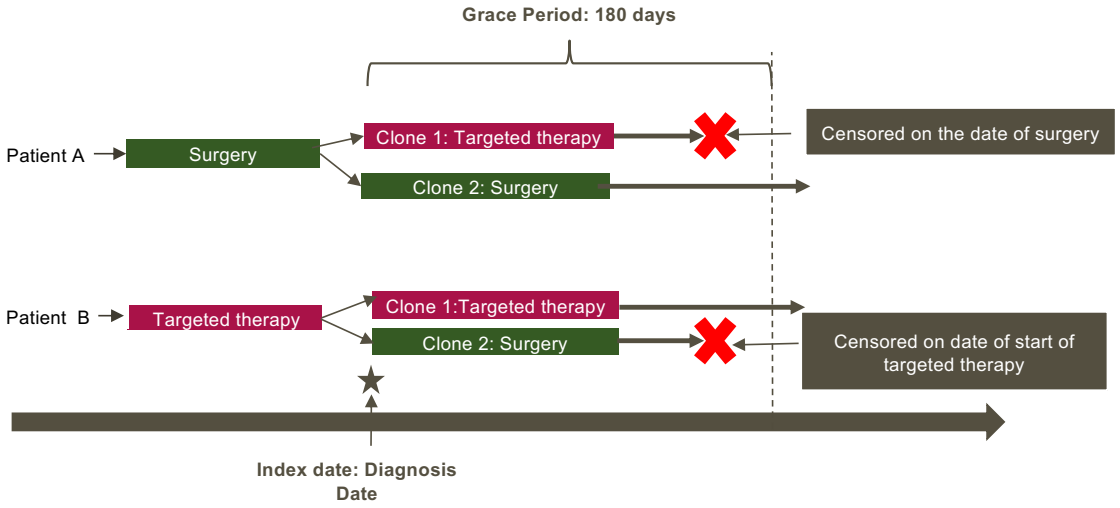
More than the target trial: different analytic strategies to adjust for immortal time bias



Conventional analyses looking at treatment vs surgery are prone to immortal time bias

Example of One Analytic Approach

- A cloning approach can minimize the potential for immortal time bias
- IPCW adjusts for the informative censoring



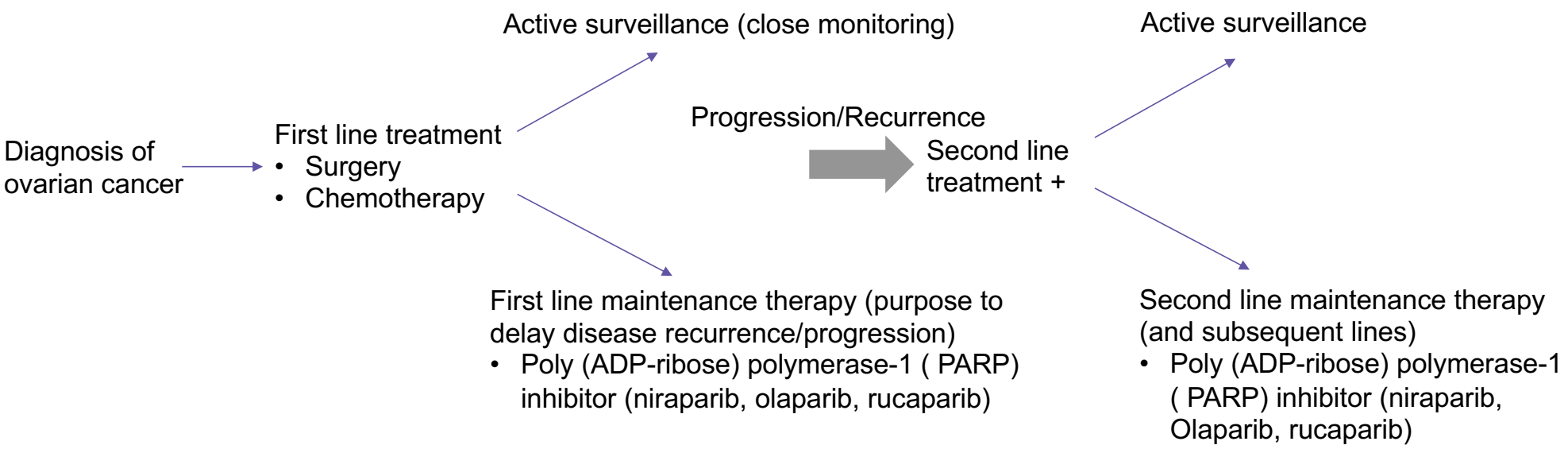
Maringe C et al. Reflection on Modern Methods. *Int J of Epidemiology* 1719-1729, 2020

RWE for Post-Approval Effectiveness

GSK

Example 2: Maintenance vs. Active Surveillance for Recurrent Ovarian Cancer Management

Describe outcomes in patients with recurrent ovarian cancer treated with 2L maintenance therapy (MTx) or under active surveillance (AS) post completion of 2L therapy



RL Coleman et al. Real-world overall survival in second-line maintenance niraparib monotherapy vs active surveillance in *BRCA* wild-type patients with recurrent ovarian cancer.. *JCO* **41**, 5592-5592(2023). DOI:10.1200/JCO.2023.41.16_suppl.5592

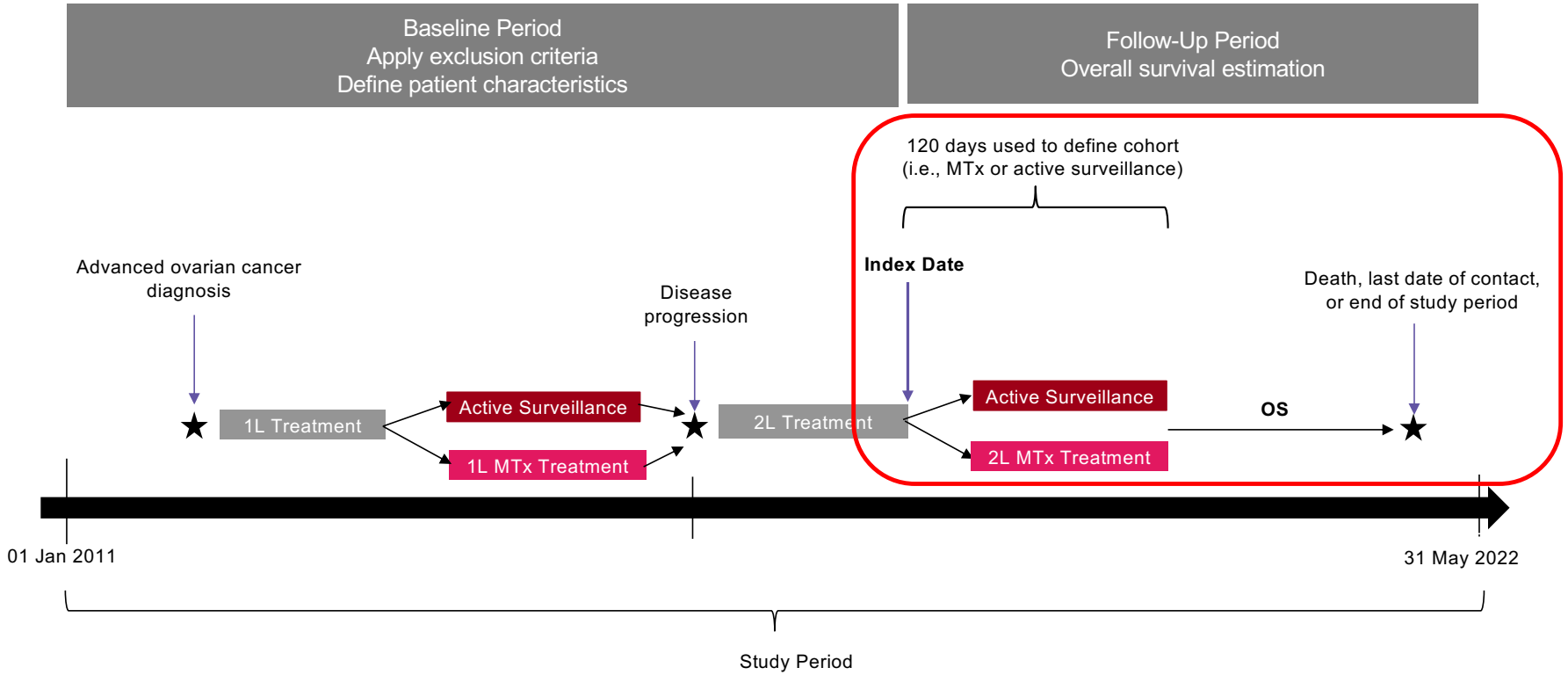
Example 2: Specifying the target trial and inclusion criteria

Protocol component	Target Trial	Emulated trial using RWD
Design	Randomized trial	Cohort study with real world follow-up
Eligibility criteria	Women with recurrent ovarian cancer who had completed 2 lines of platinum-based treatment	Women with recurrent ovarian cancer diagnosed who had received 2 lines of treatment
Treatment strategies	2L maintenance therapy or active surveillance	2L maintenance therapy vs. active surveillance
Treatment assignment	Patients were randomly assigned to *either treatment strategy* (maintenance therapy or active surveillance)	Patients not randomly assigned to *either treatment strategy*. Randomization emulated via cloning of patients in both arms
Treatment implementation	*Treatment initiated soon after randomization following completion of 2L treatment*	*120 days grace period after end of 2L treatment*
Outcome	Death from all causes	Death from all causes
Start of follow-up	Started at randomization which corresponded to treatment assignment	Started at end of 2L treatment which does not correspond to treatment assignment
Censoring	Loss to follow-up, study withdrawal	Loss to follow-up, end of study period
Causal contrast	Intent-to-treat-effect	Intention-to-treat effect
Estimand	Difference in overall survival between the two treatment arms	Difference in the overall survival among patients assigned to each treatment strategy

RL Coleman et al. Real-world overall survival in second-line maintenance niraparib monotherapy vs active surveillance in *BRCA* wild-type patients with recurrent ovarian cancer.. *JCO* **41**, 5592-5592(2023). DOI:10.1200/JCO.2023.41.16_suppl.5592

Example 2: Patients are followed from completion of 2L treatment

Index date is defined as the last date of the 2L non-maintenance treatment



MTx-maintenance therapy,1L-first line treatment, 2L-second line treatment, TTNT- time to next treatment, OS- overall survival

RL Coleman et al. Real-world overall survival in second-line maintenance niraparib monotherapy vs active surveillance in *BRCA* wild-type patients with recurrent ovarian cancer.. *JCO* **41**, 5592-5592(2023). DOI:10.1200/JCO.2023.41.16_suppl.5592

Concluding Thoughts: Yes, Use Target Trial Emulation

Factors relevant to promoting target trial emulation

Enablers

- FDA, EMA, NICE are explicit about value of TT emulation
- Good study design, data & study reporting frameworks are available
- TARGET (TrAnsparent ReportinG of observational studies Emulating a Target trial) guideline in development
- Training is available on causal diagrams, e.g., directed acyclic graphs (DAGs)

Barriers

- A tool specifically for reporting target trials is not available (ROBINS-I, a tool to evaluate potential for bias, is)
- Most researchers designing RWE are not trained as trialists
- Effort to create a target trial protocol / outline may seem inefficient
- TT emulation isn't a fail-safe, nor does it address self-controlled or test-negative designs in vaccines

A credible RWE study requires good design *plus* the right data and execution

GSK