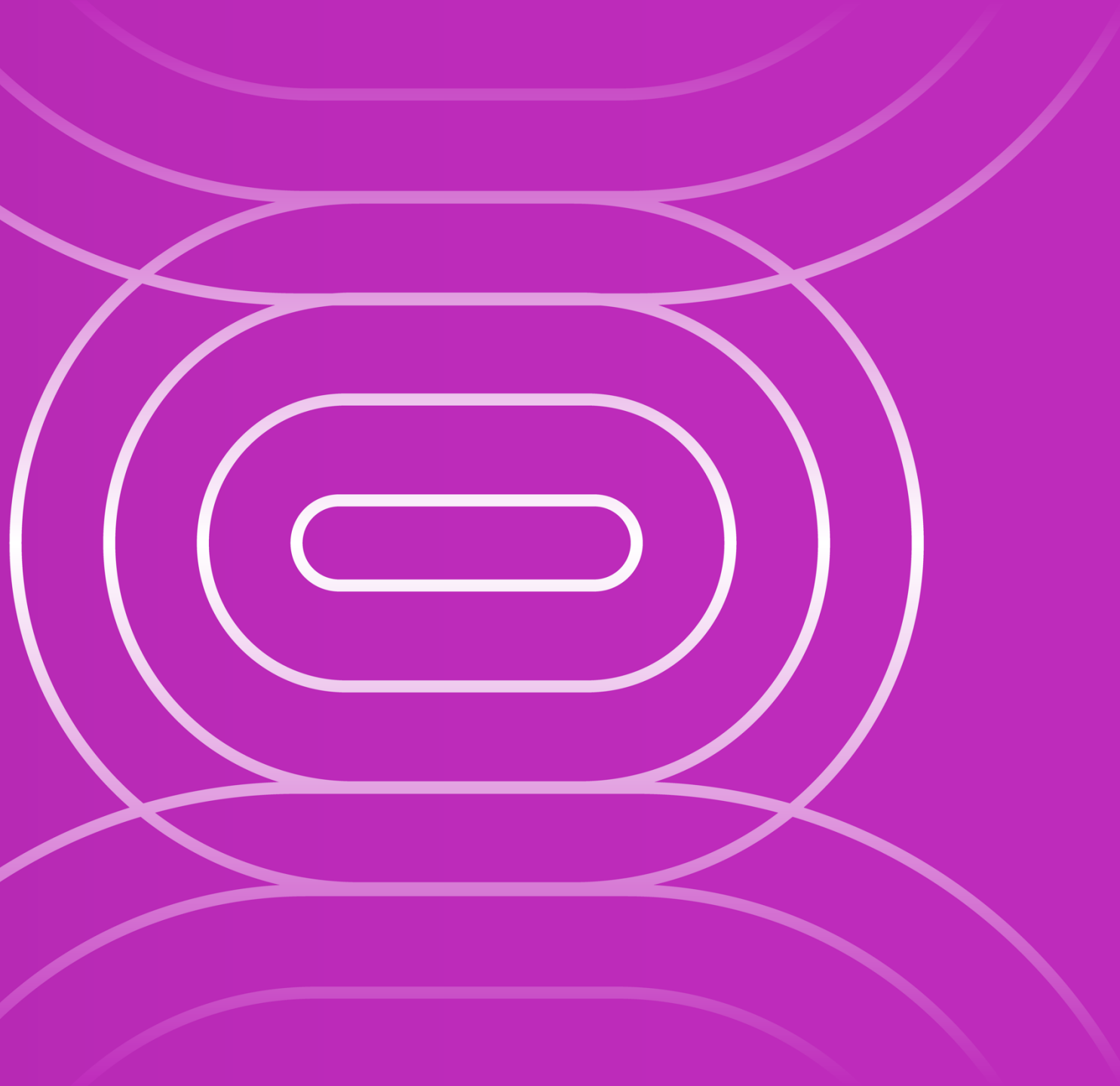


PHUSE US Connect 2026



Outcome Blinding in External Control Arm Construction Using RWD: Challenges and Best Practices

Lixiao Chen, Peishan Han, Kelvin Lin, Lorraine Fang, Lihua Yue
Bristol Myers Squibb, Lawrence Township, NJ, USA
March, 2026

Agenda

- **Background and Motivation**
- **Challenges of Outcome Blinding in RWD**
- **Risks of Outcome-Informed Decisions**
- **Best Practices and Implementation Approaches**
- **Conclusions**

Background

External control arms (ECAs) are increasingly used to contextualize single-arm clinical trials using real-world data (RWD). A critical methodological principle in constructing ECAs is ensuring that cohort construction decisions are blinded to outcome data.

One of the most critical but often underappreciated challenges in ECA construction is outcome blinding. In traditional randomized trials, study design decisions are made without knowledge of future outcomes. In contrast, RWD sources such as electronic health records, claims databases, and registries often involve utilizing information generated during day-to-day patient care in which **outcomes are already observed or documented at the time of cohort construction**. Without appropriate safeguards, analysts may inadvertently use outcome information when defining cohorts, assigning index dates, performing propensity score methods, leading to biased estimates and reduced credibility.

Motivation

“For example, when considering whether to use an externally controlled trial design, sponsors should decide whether it is possible to generate evidence capable of distinguishing the effect of the drug from outcomes attributable to the disease’s natural history, prognostic differences in the study populations, knowledge of treatment assignment (lack of blinding), or other factors such as differences in concomitant therapies.”

- *Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products, 2023.(fda.gov)*

Focuses on outcome blinding in the context of ECA cohort construction using RWD. We describe common challenges and risks, compare two practical implementation approaches used in our organization, and propose best practices that can be adopted to support robust and aligned with regulatory expectations of ECA studies.

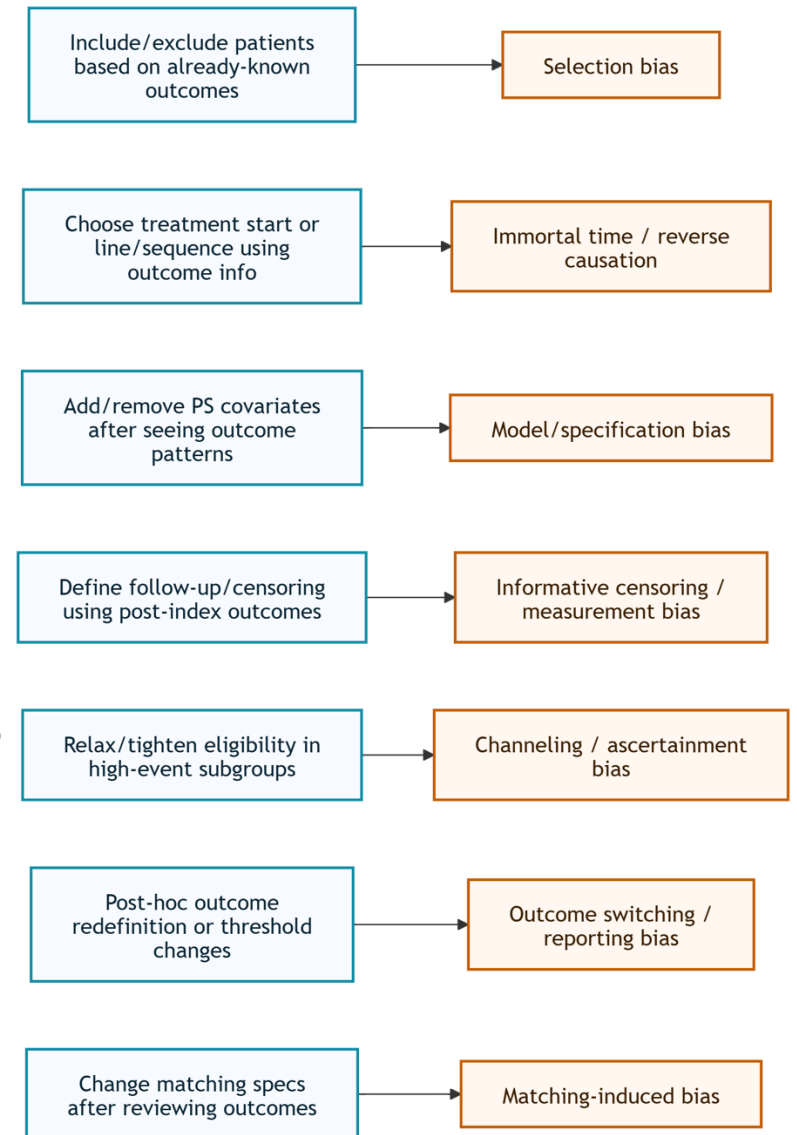
Challenges

RWD differs fundamentally from clinical trial data in its structure, completeness, and temporal availability. These differences create several challenges for maintaining outcome blinding during ECA construction. RWD typically captures complete longitudinal patient histories, including diagnoses, treatments, and outcomes, within a single data extract. As a result, analysts often have simultaneous access to baseline covariates and post-index outcomes. This coexistence makes it difficult to ensure that cohort-defining decisions are not influenced—consciously or unconsciously—by outcome information, particularly during feasibility assessments and data exploration.

Risks

Bias

Using outcome information during cohort construction can introduce multiple forms of bias, undermine internal validity and trigger regulatory and HTA concerns. FDA's draft guidance warns that viewing external control outcomes can lead to data-dependent design choices, increasing the risk of biased estimates. Therefore, FDA recommends that key study design and analysis decisions—such as cohort definitions and propensity score specifications (e.g., IPTW or matching)—be finalized under outcome blinding. This helps ensure that ECA construction is not influenced by observed outcome patterns and maintains a fair comparison to the trial arm.



❖ What is IPTW and Matching Method?

Inverse probability of treatment weighting (IPTW) and matching are commonly used propensity score-based methods to reduce confounding when comparing treatment groups in observational data. The propensity score summarizes each patient's probability of receiving the treatment given observed baseline covariates. In matching, treated patients are paired with control patients with similar propensity scores, creating a matched cohort in which baseline characteristics are more comparable across groups. In IPTW, each patient is weighted by the inverse of the probability of receiving the treatment they received, generating a pseudo-population in which the distribution of measured baseline covariates is balanced between groups. By improving covariate balance, both approaches help mitigate selection bias and confounding due to measured variables, thereby supporting a less biased estimation of treatment effects.

Best Practices and Implementation Approaches

1. Staged Two functional Team Workflows

In this approach, two **independent** groups from different functional areas are involved.

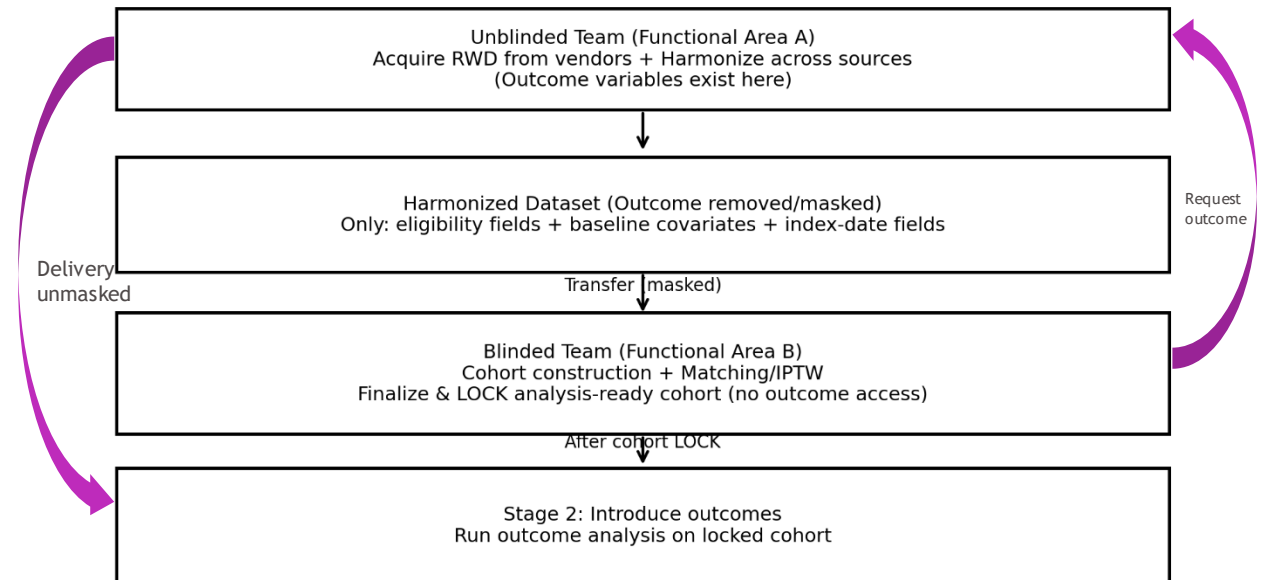
A Data Analytics team(unblinded) is responsible for acquiring RWD from data vendors and harmonizing data across multiple sources in one functional area.

Once data harmonization is complete and qualified subjects are identified, only data elements required for eligibility criteria and baseline covariates are made available to a separate blinded team responsible for cohort construction. Outcome variables are explicitly excluded or masked at this stage.

Then blinded team will get analysis-ready cohort finalized and locked, will also create inverse probability of treatment weighting (IPTW) or matching datasets.

Finally, Blinded team will require outcome data for analysis from Unblinded team to complete the study. This separation ensures that cohort definitions are finalized without knowledge of outcome distributions and closely mirrors the prospective nature of clinical trial conduct.

Approach 1: Staged Programming Workflows (Two Independent Groups)



Best Practices and Implementation Approaches

2. Controlled Access Within One Team

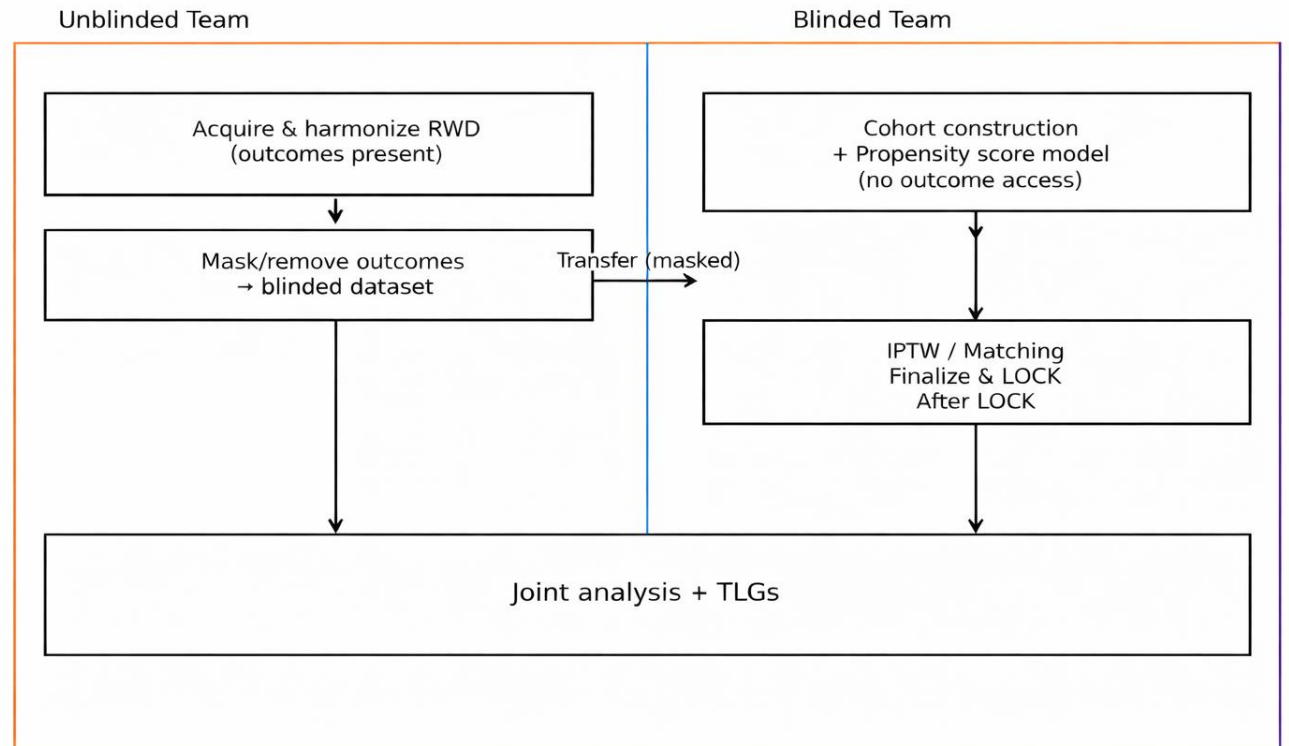
An alternative approach consolidates responsibilities within a single functional group while maintaining internal outcome blinding. In this model, team members, including statisticians and programmers, are divided into unblinded and blinded sub-teams.

The unblinded sub-team performs data acquisition and harmonization.

Following harmonization, outcome variables and outcome-related fields are removed or masked from the datasets. These masked datasets are then transferred to a restricted-access environment for use by the blinded sub-team.

The blinded team retains no access to outcome data until the analysis-ready cohort is constructed and finalized. Role-based access controls, data masking, and separate analytic environments are used to enforce this separation.

Approach 2: Unblinded → Blinded (Lock) → Joint Analysis (TLGs)



Comparison of Two Approaches

Dimension	Single-Team Approach (Better)	Staged Programming Model
Overall Efficiency	More efficient overall. Draft baseline datasets can be shared early. Enables early cohort logic development, IPTW/matching testing, and parallel processing of outcome data.	Less efficient. Blinded team must wait for unblinded team to finish full processing and validation. Sequential workflow causes delays.
Early Development & Feedback	Allows early cohort construction testing and rapid feedback to unblinded team while outcome processing continues in parallel.	Blinded team blocked from accessing baseline/I/E variables until full validation is complete, delaying development and feedback.
Data Quality & Completeness	Unblinded team can inspect completeness early using draft transfers and request customized vendor data to reduce missingness.	Data issues discovered late. Limited opportunity to request additional/customized fields mid-process.
Parallel Workstreams	Outcome processing continues while blinded team develops cohort logic and QC in parallel.	Workstreams are sequential. Blinded team waits for fully masked outcome dataset before starting.
Resource Utilization	Teams can rotate roles. Reduced idle time. Fewer total resources required.	Less flexible. One team may need to wait, increasing idle time and total resource demand.
Suitability for RWE	Well suited to rapidly evolving real-world evidence workflows with iterative data delivery.	More rigid, like clinical trial workflows; not ideal for dynamic RWE.

Documentation & Regulatory Expectations

D

Documentation & Transparency

- Pre-specify cohort definitions, I/E criteria, prognostic factors, index-date rules, and analysis methods in Study protocols and analysis plans.
- Document assumptions, decisions, and deviations with justification.
- Maintain programming specs, decision logs, and data flow diagrams for transparency.
- Use reproducible code, version control, and traceable outputs.

R

Regulatory Considerations

- Regulators emphasize transparency, pre-specification, and bias mitigation in RWD.
- Outcome blinding reinforces trial emulation and reduces data-driven decisions.
- Robust blinding improves acceptability of ECAs supporting efficacy/safety claims in regulatory submissions.
- Clearly describe blinding in protocols, SAPs, and final reports.

Conclusion

- Outcome blinding is a foundational methodological principle in the construction of external control arms using real-world data. Given the inherent availability of outcomes in RWD sources, proactive safeguards are required to prevent bias and preserve credibility. Staged programming workflows, controlled access to outcome data, and thorough documentation collectively provide a practical framework for maintaining outcome blinding.
- By adopting these best practices, study teams can enhance the methodological rigor of ECAs, align more closely with regulatory expectations, and strengthen confidence in RWD-based evidence to support clinical development and decision-making.

Q&A and Contact

- Lixiao Chen: Lixiao.chen@bms.com
- Jenny Han: Jennypeishan.han@bms.com