

# Guidance on the design, conduct, and analysis of external control arms in the generation of real-world evidence to support health technology assessment evaluations

Version 1.0 September 2026



## Record of updates

| Date           | Version | Summary of changes                                                                                                                                                                          |
|----------------|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| September 2026 | 1.0     | Publication of initial guidance on the design, conduct, and analysis of external control arms in the generation of real-world evidence to support health technology assessment evaluations. |

## Table of contents

|                                                                                                                                                                          |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Foreword.....                                                                                                                                                            | 4  |
| Abbreviations and acronyms .....                                                                                                                                         | 5  |
| Guidance on the design, conduct, and analysis of external control arms in the generation of real-world evidence to support health technology assessment evaluations..... | 6  |
| 1. Aim.....                                                                                                                                                              | 6  |
| 2. Role of ECAs in HTA.....                                                                                                                                              | 6  |
| 2.1. Values of ECAs in decision-making .....                                                                                                                             | 6  |
| 2.2. Methodological challenges with ECAs .....                                                                                                                           | 7  |
| 3. Key considerations for ECAs .....                                                                                                                                     | 7  |
| 3.1. Design considerations.....                                                                                                                                          | 7  |
| 3.2. Data considerations .....                                                                                                                                           | 9  |
| 3.3. Analysis considerations .....                                                                                                                                       | 10 |
| 4. Practical recommendations on design, conduct, and analysis of externally controlled studies in the generation of RWE to support HTA evaluations .....                 | 11 |
| 5. Limitations and future directions .....                                                                                                                               | 13 |
| 6. Conclusion.....                                                                                                                                                       | 14 |
| References.....                                                                                                                                                          | 15 |

## Foreword

Established by the Ministry of Health (MOH), the Agency for Care Effectiveness (ACE) is the national health technology assessment (HTA) and clinical guidance agency in Singapore. It produces evidence-based evaluations of health technologies (e.g. drugs, vaccines and medical technologies) to inform funding decisions by MOH committees, and publishes technology guidance documents for public hospitals and institutions in Singapore to promote the appropriate use of clinically effective and cost-effective treatments. Find out more about ACE at <https://www.ace-hta.gov.sg/about-us>.

This document provides a practical framework for the design, implementation, analysis, and reporting of external control arms constructed from real-world data (RWD) to support health technology assessment (HTA) evaluations.

Information in this document may be useful for researchers, healthcare professionals, and industry stakeholders involved in collecting RWD, generating real-world evidence, or submitting data or evidence for funding decisions. ACE will continue to review and update this document to ensure that it remains a useful resource for the Singapore healthcare system.

ACE would like to thank the following experts for their comments during the development of version 1.0 of this document (appointments listed were current when the document was first published in September 2026):

- Prof Robyn Ward, Deputy Vice-Chancellor (Research and Enterprise) and Senior Vice-President, Monash University, Australia; Chair, Commonwealth Pharmaceutical Benefits Advisory Committee (PBAC)
- Sudha Kutty, Executive Vice-President of Evidence, Products, and Services, Canada's Drug Agency (CDA-AMC), Canada
- Prof Jonathan Benger, Chief Executive, National Institute for Health and Care Excellence (NICE), United Kingdom
- Ms Natalie McMurtry, Chief Executive, Pharmaceutical Management Agency (Pharmac), New Zealand
- Prof Satvinder Singh Dhaliwal, Consultant, Agency for Care Effectiveness; Emeritus Professor, Faculty of Health Sciences, Curtin University, Australia

## Abbreviations and acronyms

| Term | Definition                       |
|------|----------------------------------|
| DDML | Doubly debiased machine learning |
| DML  | Double machine learning          |
| ECA  | External control arm             |
| EHR  | Electronic health record         |
| FDA  | Food and Drug Administration     |
| HTA  | Health technology assessment     |
| RCT  | Randomised controlled trial      |
| RWD  | Real-world data                  |
| RWE  | Real-world evidence              |

# Guidance on the design, conduct, and analysis of external control arms in the generation of real-world evidence to support health technology assessment evaluations

## 1. Aim

This guidance document provides a practical framework for the design, implementation, analysis, and reporting of external control arms (ECAs) constructed from real-world data (RWD) to support health technology assessment (HTA) evaluations. With the increasing use of real-world evidence (RWE) in HTA evaluations, ECAs offer a useful approach to generate comparative evidence, particularly when randomised controlled trials (RCTs) are not feasible or ethical.

The primary aim of this guidance is to offer actionable recommendations to researchers, policymakers, and other stakeholders involved in the collection, generation, and submission of RWE for HTA purposes. By consolidating regulatory guidance, HTA methodologies, and current literature (1-7), this guidance seeks to promote consistent, transparent, and robust use of ECAs in decision-making. Key objectives include:

- Establishing best practices for the design and execution of ECAs;
- Providing analytical approaches to address bias and confounding; and
- Recommending standards for transparent reporting to enhance credibility.

While ECAs may provide important supplementary evidence, they do not replace RCTs or single-arm trials as external controls alone are generally insufficient to demonstrate effectiveness reliably. The guidance focuses on ECAs constructed from RWD sources such as disease registries, claims data or electronic health records. It does not cover ECAs from clinical trial data, methodological considerations of natural history studies or assess the reliability and applicability of different RWD sources for use in externally controlled trial (8). In addition, it does not address the use of external control data in RCTs, to supplement the control arm (3).

## 2. Role of ECAs in HTA

ECAs are increasingly used as a pragmatic approach for generating RWE, particularly when traditional RCTs are not feasible or appropriate due to ethical or practical constraints. ECAs involve the use of a comparison group derived from external data sources, such as historical records or past observational studies, to evaluate the outcomes of an intervention.

### 2.1. Values of ECAs in decision-making

2.1.1. ECAs may be considered in clinical contexts where conducting RCTs is impractical, with appropriate justification and rigorous methods to mitigate bias and confounding. This includes rare diseases, rapidly evolving therapeutic areas such as oncology (9), and situations where ethical considerations preclude withholding treatment (3). By leveraging RWD from sources such as disease

registries, electronic health records (EHRs), and administrative claims, ECAs enable the generation of comparative evidence to support evaluations (3, 10).

2.1.2. Regulatory bodies and HTA agencies have increasingly acknowledged the potential value of ECAs but remain cautious about their limitations. For example, the United States Food and Drug Administration (FDA) has published guidance highlighting the importance of careful design, data quality, and pre-specification (3) in externally controlled trials. Similarly, the HTA agencies in Europe and Canada have emphasised the need for rigorous methodological approach to address the inherent limitations of ECAs, such as missing data, confounding factors and sources of bias (1, 11).

## **2.2. Methodological challenges with ECAs**

2.2.1. Despite their potential, ECAs have limitations and are not recommended for non-inferiority trials (12). The absence of randomisation increases the risk of bias due to differences in baseline characteristics between the intervention group and ECA. Furthermore, challenges in data consistency, evolving standards of care, and unmeasured confounding can impact the validity of findings.

2.2.2. To address these challenges, it is essential to ensure that ECAs are rigorously designed and analysed (6, 10). This includes defining comparable populations, using carefully selected data, well-defined and clinically meaningful endpoints, and employing appropriate statistical methods. Recent methodological advancements offer options to reduce bias from observed confounding (for example, propensity score methods, outcome regression, g-methods) and emerging techniques that combine machine learning with causal inference principles (for example, doubly debiased machine learning) (4, 13-15). Appropriate summary measures should also be selected and reported to support valid interpretation of comparative effectiveness estimates. Foundational literature on causal inference and propensity scores remains central to implementing ECAs (16, 17). Effective design and execution ensure that ECAs provide credible evidence to complement RCTs and support HTA evaluations.

## **3. Key considerations for ECAs**

The robustness of ECAs depends on adherence to key principles across three primary domains: study design, data selection, and analytical considerations. Each of these elements plays a crucial role in ensuring the validity, reliability, and applicability of findings from ECAs in HTA evaluations. This section outlines the critical considerations that must be addressed to generate credible evidence for decision-making.

### **3.1. Design considerations**

3.1.1. A robust design begins with a clear research question that specifies the population, intervention, comparator, outcomes, and time period. This foundational step ensures that the study is appropriately scoped and aligned with the decision-making requirements of HTA (3, 10).

3.1.2. Determining the appropriateness of an ECA requires examination of the clinical context. ECAs are most appropriate when the disease course is stable and well understood (for example, the natural progression, symptom patterns, and prognosis remain consistent over time and across populations), outcomes are well-defined, objectively measured (for example, overall survival), clinically meaningful, the standard of care is stable over the period from which external control data are extracted, and relevant to the HTA context. A detailed understanding of the natural history of the disease, prognostic factors and disease heterogeneity is essential to account for confounders, both known and unknown, and to inform appropriate study design decisions. Conversely, ECAs are less suitable in contexts with rapid shifts in practice, subjective endpoints, or poorly recorded prognostic factors (5, 6).

3.1.3. Population alignment is essential to ensure patient comparability between the intervention group and ECAs. The treatment and control arm should be as similar as possible with respect to known factors that affect the outcomes being measured. This could mean matching patient populations based on clinically relevant variables such as demographics (for example, age, sex, geographic), disease characteristics (for example, stage, severity, biomarkers), co-morbidities, treatment history, and other baseline characteristics (for example, lifestyle factors). Eligibility criteria from the intervention trial should be applied to the external control data to the extent possible. This includes defining objectively and to the extent possible, the index date (excluding any period of immortal time), baseline covariate assessment windows, and follow-up approaches consistently between the datasets.

3.1.4. Temporal alignment – ensuring contemporary comparators – is critical. Historical controls are vulnerable to bias if diagnostic practices or therapies have changed. Where possible, external controls should be sampled from the same or a similar time period, treatment practice, context and health system as the intervention trial (6).

3.1.5. Outcomes must be assessed consistently in both the intervention and ECAs for the results to be credible, prioritising objective outcomes that can be tracked reliably with RWD (2). The timing and frequency of outcome assessments should be specified for the intervention and control arms, and consider how the differences may influence interpretation of the findings. Where possible, the same data sources and criteria for establishing outcomes should be used. Treatment definitions must be standardised to ensure that terminologies such as no treatment, treatment discontinuation, and standard of care are consistently defined across studies. Differences in the capture of intercurrent and unplanned events should also be addressed as additional therapies and other post-baseline events may be better captured in a clinical trial with scheduled study visits than in routine clinical care, potentially affecting the interpretation of treatment effects. When alignment is infeasible, it is essential to document any differences and apply analytic adjustments to address measurable disparities where possible, and report the potential limitations (3). To reduce bias, outcomes should be evaluated independently of treatment status, and the validity of the selected outcomes should be confirmed with independent blinded reviewers when appropriate (2). Threshold-

crossing approaches could be considered, whereby pre-specified effect size thresholds help identify treatment effects of sufficient magnitude to reduce the likelihood that findings are attributed to bias or confounding alone (18).

3.1.6. Study protocols must be pre-specified and registered prior to conducting any analysis to reduce selection bias. They should include predefined eligibility criteria, exposure definitions and windows (especially index date), endpoint definitions and assessments, covariate lists, and statistical analysis plans, and methods for measuring and analysing confounding factors and sources of bias. The protocol should also specify the selection of the ECA, data feasibility and the analytic methods. Any subsequent changes must be appropriately justified and documented. Communication with regulators and HTA agencies can clarify expectations and improve acceptability.

3.1.7. When feasible, hybrid designs that incorporate elements of randomisation, internal controls, or prospective observational cohorts may improve causal inference while leveraging RWD. Synthetic control arms – constructed from pooled and harmonised RWD – can be powerful but require careful validation (10).

## **3.2. Data considerations**

3.2.1. Data sources should be selected from reliable sources, guided by the policy-relevant research question and the need for clinical and temporal alignment with the intervention group. Each source presents a trade-off between breadth, depth, and clinical granularity:

- Registries often capture detailed clinical variables, but may be limited in coverage;
- Claims data capture large populations with robust billing information but lack clinical details; and
- EHRs provide clinical depth but vary in structure and completeness (5).

3.2.2. It is essential to verify that the data has sufficient detail at the individual patient level to support the comparison and ensure the data is temporally and clinically relevant to the intervention arm. Tools alongside clinical expert input enable evaluation of whether existing datasets are appropriate for developing an ECA (19-22).

3.2.3. Data quality and provenance must be comprehensively documented. This includes detailing the origin of the data, inclusion and exclusion criteria for the source cohort, data curation procedures, linkage steps, and any transformations applied. Transparency in data governance should address ethical approvals, data access restrictions, and de-identification methods (1, 3).

3.2.4. Variable definitions and harmonisation are central to ensuring comparability between datasets. Standardised definitions for exposures, covariates, and outcomes reduce measurement bias. Where possible, use common data models or mapping frameworks to align disparate data sources (23). Explicit

mapping tables and code lists should be included in appendices or supplementary materials to ensure reproducibility.

3.2.5. Missing data is an inherent challenge in RWD and must be addressed using robust statistical methods. Common approaches including multiple imputation, inverse-probability weighting, and sensitivity analyses for missing-not-at-random assumptions should be considered (24). Avoid ad-hoc deletion of records without justification, as it can introduce bias.

3.2.6. It is important to assess whether external control data are representative of the target population. This involves comparing baseline characteristics between the intervention group and ECA and reporting absolute differences. Visualisations and tables that show overlap in propensity score distributions or covariate balance are useful for assessing the feasibility of adjustment to address confounding (16, 25).

### 3.3. Analysis considerations

3.3.1. An analytical plan must be prespecified, detailing how confounding factors and sources of bias will be measured and analysed. A systematic literature search of prognostic variables, confounders, and effect modifiers must be conducted along with input from key opinion leaders to identify the appropriate covariates for statistical conditioning (1, 2). Analytic methods must be chosen to address measured confounding and reduce bias while acknowledging that unmeasured and residual confounding and bias may remain. However, the analytical plan cannot fully overcome inherent limitations in poor-quality, incomplete, or systematically biased data. Common statistical methods for addressing confounding and bias include:

- Propensity score methods – matching, stratification, and inverse-probability of treatment weighting are widely used to balance observed covariates between trial and external cohorts and to emulate target trial comparisons (5, 7, 16, 26);
- Outcome regression and g-methods (for example, G-computation and marginal structural models) can handle time-varying confounders and treatment strategies, particularly in longitudinal datasets;
- Combining propensity-based approaches with outcome modelling – 'doubly robust' estimators provide protection against model misspecification (4); and
- Recent methodological advances incorporate machine learning within causal frameworks. Doubly debiased machine learning (DDML) and double machine learning (DML) can flexibly model high-dimensional relationships while retaining valid inference under certain conditions; however, they require careful tuning, cross-fitting, and transparent reporting (4, 13).

3.3.2. Causal inference statistical methods use observational data to estimate the counterfactual, allowing for broader application when RCTs are infeasible or unethical, but they require strong statistical assumptions to overcome confounding and bias. These assumptions need to be checked and detailed thoroughly, noting the limitations of observational data in making inferences and causal conclusions.

3.3.3. Intercurrent events in the ECA may impact the interpretability of treatment effect and must be carefully accounted for in the analysis. Sensitivity analyses are essential to investigate the influence of unmeasured confounding and assess the robustness of findings. Common techniques include quantitative bias analysis, E-value calculations, negative control outcomes/exposures, and bounding approaches (27-30). Additionally, reporting a range of plausible effect estimates under varying assumptions (for example, adjusting the inclusion and exclusion criteria, applying different outcome or exposure definitions or varying the timing and occurrence of intercurrent events in the ECA) helps to evaluate the robustness of conclusions.

3.3.4. Subgroup and heterogeneity analyses can provide insight into treatment-effect variation and external validity. However, these analyses should be prespecified with clinically meaningful subgroups and interpreted cautiously, recognising the limitations of observational data in supporting inference including assessing multiplicity, interactions, and causal relationships.

3.3.5. Transparent reporting, reproducibility, and collaboration enhance credibility. Detailed statistical analysis plans and all deviations from pre-specified protocols should be documented, and analytic code should be shared where possible. Where privacy or data use agreements prevent sharing raw data, sharing synthetic datasets or detailed summary-level diagnostics can facilitate independent assessment.

## **4. Practical recommendations on design, conduct, and analysis of externally controlled studies in the generation of RWE to support HTA evaluations**

4.1. The following steps provide systematic guidance for designing, conducting, analysing, and reporting studies that utilise ECAs, addressing critical aspects from protocol development through to results interpretation and reporting. Each step incorporates specific methodological considerations and analytical approaches designed to minimise bias, enhance comparability between intervention and comparator groups, and ensure robust evidence generation for decision-making (28).

### **Step 0. Assess the suitability of an ECA**

- Determine whether the disease and treatment setting are suitable for an ECA, considering whether the natural history, prognosis, and standard of care are sufficiently well characterised and can support valid comparisons using RWD.
- Assess whether suitable RWD are available to provide a clinically relevant external control with sufficient information on patient characteristics, treatments, outcomes and follow-up.

### **Step 1. Design of the protocol**

- Justify the use of an ECA in the study protocol, including why an RCT is not feasible.
- The selection criteria of external control, data sources, and analytical plans should be prespecified in study protocols to reduce the risk of post-hoc bias.

### **Step 2. Select external control sources**

- Evaluate whether the data source is accurate, complete, and reliable for addressing the research question, including the consistency of data capture, coding practices, and the extent of missing or incomplete information that could bias results.
- Assess that the dataset includes a population, outcomes, and follow-up period that align with the research question, including sufficient sample size for planned analyses.
- Consider whether the care setting, geographic region, and healthcare delivery context are appropriate and comparable to the intervention setting.
- Where possible, ensure the external control data is from the same or a similar time period as the intervention arm to minimise temporal confounding, if this is not feasible, clearly document any changes in standard of care across time and its potential impact on the results.

### **Step 3. Assess comparability between external control and intervention arms**

- Assess similarity in eligibility criteria, disease definitions, prognosis, and baseline characteristics, and ensure that key covariates, treatments, and outcomes are pre-specified to support adjustment and comparison (3).
- Evaluate how treatments are defined and administered in each cohort, including duration of exposure, concomitant therapies, treatment pathways, and how “time zero” is defined to ensure alignment of index dates.
- Confirm that follow-up duration is comparable and that endpoints are defined and measured consistently across data sources, including assessment methods and timing.
- Consider differences in handling of intercurrent events (for example, treatment switching, discontinuation), and assess and document whether differences in data completeness or missingness could introduce bias in outcome estimation.

### **Step 4. Generate external control**

- Screen external control samples from the source data using consistent inclusion and exclusion criteria as the intervention group.
- Use statistical methods, such as propensity score-based approaches (for example, matching or weighting) when generating external controls, to adjust for baseline covariates and potential confounding factors. Standardised mean differences should also be used to assess the balance of baseline covariates between the treatment and ECAs. The choice of method should be clearly justified.

### **Step 5. Treatment effect estimation**

- Adjust for important and clinically relevant baseline covariates in the primary analysis to account for potential imbalances in characteristics between the intervention and ECAs.

### **Step 6. Analysis**

- Perform sensitivity analyses based on a priori assumptions, such as various methods for adjusting for confounding (for example, matching specifications and covariate sets) and for handling missing data, to assess the robustness of the study results.
- Estimate the magnitude and uncertainty associated with the bias in treatment effects by using quantitative bias analysis to model plausible distributions of unmeasured confounders and incorporate them into the external control comparison (31).

### **Step 7. Interpret results and draw conclusions**

- Draw well-supported conclusions by conducting a comprehensive interpretation of the results to ensure the transparency of the study from the:
  - a. primary analyses;
  - b. sensitivity analyses; and
  - c. quantitative bias analyses.

### **Step 8. Reporting and transparency**

- Report the study methods, assumptions, analyses, results, and limitations transparently to enable assessment, interpretation, and reproducibility of the study findings.

4.2. It is important to recognise that ECAs complement, rather than replace randomised evidence. Wherever feasible, ECAs can be used to augment trial data, support subgroup analyses, or inform long-term outcomes and safety profiles.

## **5. Limitations and future directions**

5.1. The guidance primarily describes the use of reliable, high-quality RWD but does not extensively address the challenges arising from inconsistencies in data quality, completeness, or granularity across different datasets. These limitations may restrict the applicability of the guidance in contexts where RWD infrastructure is less developed or inconsistent, highlighting the need to assess and harmonise disparate data sources effectively.

5.2. Additionally, the guidance focuses on frequentist statistical approaches for the generation of ECAs. While Bayesian methods offer the advantage of probabilistic insights and the ability to incorporate prior distributions into the analysis, they also introduce unique challenges that warrant additional considerations.

5.3. Separately, future methodological research should review approaches to assess and mitigate unmeasured confounding, and develop standards for reporting and validation that are acceptable to both regulators and HTA agencies (32). Establishing mechanisms for stakeholder feedback and collaboration will be critical to refining the guidance and addressing emerging challenges. International collaboration, particularly on data harmonisation and the development of reporting standards, will further strengthen the credibility and utility of ECAs in evaluations (1, 23).

5.4. ECAs will not be appropriate for all research questions. Limitations include the potential for residual confounding, measurement error, and incomplete capture of important prognostic variables in RWD (6, 10). It is important to report these limitations, as statistical methods cannot fully address them (33, 34), particularly when ECAs involve fundamentally different populations. Assessing how such limitations may influence the conclusions is therefore essential.

## 6. Conclusion

ECAs constructed from RWD may provide a valuable tool for generating comparative evidence when randomised controls are not feasible. When designed thoughtfully, analysed rigorously, and reported transparently, ECAs may enhance the evidentiary base for HTA decision-making and support timely patient access to innovative therapies. This guidance offers practical steps to enhance the credibility of ECAs, emphasising pre-specification, data quality, robust causal inference methods, and transparency. Continued methodological development in regulatory and HTA evidence expectations will be key to realising the full potential of ECAs in healthcare decision-making.

## References

1. Vanier A, Fernandez J, Kelley S, Alter L, Semenzato P, Alberti C, et al. Rapid access to innovative medicinal products while ensuring relevant health technology assessment. Position of the French National Authority for Health. *BMJ Evidence-Based Medicine*. 2024;29(1):1-5.
2. Vidalis A, Dumoulin O, Godbole M, Proenca CC. The role and value of real-world evidence in health technology decision-making in France, Germany, Italy, Spain, and the UK: Insights on external control arms. *International Journal of Technology Assessment in Health Care*. 2025;41(1):e25.
3. Food and Drug Administration (FDA). Considerations for the design and conduct of externally controlled trials for drug and biological products guidance for industry. 2023
4. Loiseau N, Trichelair P, He M, Andreux M, Zaslavskiy M, Wainrib G, et al. External control arm analysis: An evaluation of propensity score approaches, G-computation, and doubly debiased machine learning. *BMC Medical Research Methodology*. 2022;22(1):335.
5. Yap TA, Jacobs I, Andre EB, Lee LJ, Beaupre D, Azoulay L. Application of real-world data to external control groups in oncology clinical trial drug development. *Frontiers in Oncology*. 2022;11(695936):1-9.
6. Jaksa A, Louder A, Maksymiuk C, Vondeling GT, Martin L, Gatto N, et al. A comparison of 7 oncology external control arm case studies: Critiques from regulatory and health technology assessment agencies. *Value in Health*. 2022;25(12):1967-76.
7. Food and Drug Administration (FDA). Framework for FDA's real-world evidence program. FDA; 2018.
8. REAL World Data In ASia for HEalth Technology Assessment in Reimbursement (REALISE) working group. Use of real-world data and real-world evidence to support drug reimbursement decision-making in Asia. Abridged version. 2021.
9. Mishra-Kalyani PS, Amiri Kordestani L, Rivera DR, Singh H, Ibrahim A, DeClaro RA, et al. External control arms in oncology: Current use and future directions. *Annals of Oncology*. 2022;33(4):376-83.
10. Burcu M, Dreyer NA, Franklin JM, Blum MD, Critchlow CW, Perfetto EM, et al. Real-world evidence to support regulatory decision-making for medicines: Considerations for external control arms. *Pharmacoepidemiology and drug safety*. 2020;29(10):1228-35.
11. Canada's Drug Agency. Methods guide for health technology assessment. 2025.
12. Snapinn SM. Noninferiority trials. *Trials*. 2000;1(1):19.
13. Chernozhukov V, Chetverikov D, Demirer M, Duflo E, Hansen C, Newey W, et al. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*. 2018;21(1):C1-C68.
14. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*. 2011;46:399-424.
15. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005;61(4):962-73.
16. Stuart EA. Matching methods for causal inference: A review and a look forward. *Statistical Science*. 2010;25(1):1-21.
17. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *American Journal of Epidemiology*. 2006;163(12):1149-56.

18. Eichler H-G, Bloechl-Daum B, Bauer P, Bretz F, Brown J, Hampson LV, et al. "Threshold-crossing": A useful way to establish the counterfactual in clinical trials? *Clinical Pharmacology & Therapeutics*. 2016;100(6):699-712.
19. National Institute for Health and Care Excellence. Data Suitability Assessment Tool (DataSAT) 2022 [Available from: <https://www.nice.org.uk/corporate/ecd9/chapter/appendix-1-data-suitability-assessment-tool-datasat>.
20. Gatto NM, Campbell UB, Rubinstein E, Jaksa A, Mattox P, Mo J, et al. The structured process to identify fit-for-purpose data: A data feasibility assessment framework. *Clinical Pharmacology & Therapeutics*. 2022;111(1):122-34.
21. Langan SM, Schmidt SAJ, Wing K, Ehrenstein V, Nicholls SG, Fillion KB, et al. The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology (RECORD-PE). *British Medical Journal*. 2018;363:k3532.
22. Wang SV, Pottegård A, Crown W, Arlett P, Ashcroft DM, Benchimol EI, 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 in Health*. 2022;25(10):1663-72.
23. Observational Health Data Sciences and Informatics. Common Data Model (CDM) resources 2020 [Available from: [https://www.ohdsi.org/web/wiki/doku.php?id=resources:2020\\_data\\_network](https://www.ohdsi.org/web/wiki/doku.php?id=resources:2020_data_network).
24. Kang H. The prevention and handling of the missing data. *Korean Journal of Anesthesiology*. 2013;64(5):402-6.
25. Hernán MA, Robins JM. Causal inference: What if. Boca Raton: Chapman & Hall/CRC; 2024.
26. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70(1):41-55.
27. Lash TL, Fox MP, Cooney D, Lu Y, Forshee RA. Quantitative bias analysis in regulatory settings. *American Journal of Public Health*. 2016;106(7):1227-30.
28. Liu J, Yao M, Wang M, Jie W, Liu Y, Luo X, et al. Design, conduct, and analysis of externally controlled trials. *JAMA Network Open*. 2025;8(9):e2530277-e.
29. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: A tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21(3):383-8.
30. Gao C, Zhang X, Yang S. Doubly robust omnibus sensitivity analysis of externally controlled trials with intercurrent events. *Biometrics*. 2025;81(2).
31. Gupta A, Hsu G, Kent S, Duffield SJ, Merinopoulou E, Lockhart A, et al. Quantitative bias analysis for single-arm trials with external control arms. *JAMA Network Open*. 2025;8(3):e252152-e.
32. Velummailum RR, McKibbin C, Brenner DR, Stringer EA, Ekstrom L, Dron L. Data challenges for externally controlled trials: Viewpoint. *Journal of Medical Internet Research*. 2023;25:e43484.
33. Senn S. Causal estimates for external controls. How reasonable are the assumptions? *Journal of the Royal Statistical Society Series A: Statistics in Society*. 2025.
34. Zhou X, Zhu J, Drake C, Pang H. Causal estimators for incorporating external controls in randomized trials with longitudinal outcomes. *Journal of the Royal Statistical Society Series A: Statistics in Society*. 2024;188(3):791-818.