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DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.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be submitted within 90 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance.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)

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

Guidance for Industry

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs

and/or

Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov
https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

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)

TABLE OF CONTENTS

Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

This guidance provides recommendations to sponsors and investigators considering the use of externally controlled clinical trials[2] to provide evidence of the safety and effectiveness of a drug product.[3] In an externally controlled trial, outcomes in participants receiving the test treatment according to a protocol are compared to outcomes in a group of people external to the trial who had not received the same treatment. The external control arm can be a group of people, treated or untreated, from an earlier time (historical control), or it can be a group of people, treated or untreated, during the same time period (concurrent control) but in another setting.[4,5]

The guidance addresses considerations for the design and analysis of externally controlled trials to study the effectiveness and safety of drugs, including discussion of threats to the validity of

[1] This guidance has been prepared by the Office of Medical Policy in the Center for Drug Evaluation and Research in cooperation with the Center for Biologics Evaluation and Research and the Oncology Center of Excellence at the Food and Drug Administration.

[2] In this guidance, the terms *clinical trials*, *clinical studies*, and *clinical investigations* are interchangeable.

[3] In this guidance, the term *drug product* includes both human drugs and biological products.

[4] FDA regulations under 21 CFR 314.126 outline the characteristics of adequate and well-controlled studies, and recognize various controls, including a *historical control*, which FDA considers to be a subset of a broader category of potential *external controls*. FDA has accepted various types of external controls, when appropriate, for a specific drug development program. See also the International Council for Harmonisation (ICH) guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials* (May 2001). We update guidances periodically. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

[5] Although multiple arms may be part of the overall trial design, this guidance discusses externally controlled trials involving analysis of a single treatment arm and a single control arm.

trial results from potential ***bias***.[6] Although various sources of data can serve as the control arm in an externally controlled trial, this guidance focuses on the use of patient-level data from other clinical trials or from ***real-world data*** (***RWD***) sources, such as registries as well as electronic health records (EHRs) and medical claims.[7] The guidance also describes considerations related to communicating with FDA and ensuring access by FDA to data from an externally controlled trial.

This guidance does not address other types of external controls, such as using summary-level estimates instead of patient-level data. This guidance does not discuss details of the design and analysis of a natural history study[8] nor the reliability and relevance of various sources of RWD[9] that could be used in an externally controlled trial. Finally, this guidance also does not discuss considerations for using external control data to supplement a control arm in a traditional randomized controlled clinical trial.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

The purpose of conducting clinical investigations of a drug product is to distinguish the effect of a drug on the target condition from other influences, such as spontaneous change in the course of the disease, placebo effect, or biased observation.[10] When properly conducted, a clinical trial—with random assignment of participants either to a treatment arm or to a placebo (or other

[6] Words and phrases in ***bold italics*** are defined in the Glossary.

[7] Given that an external control arm can involve the use of RWD, FDA is issuing this guidance to satisfy, in part, the requirements of the 21st Century Cures Act to issue guidance on the use of ***real-world evidence*** (***RWE***) in regulatory decision-making, specifically to evaluate the potential use of RWE to help support the approval of a new indication for a drug already approved under section 505(c) of the FD&C Act or to help support or satisfy postapproval study requirements.

[8] See the draft guidance for industry *Rare Diseases: Natural History Studies for Drug Development* (March 2019). When final, this guidance will represent FDA's current thinking on this topic. Natural history studies can be used for purposes such as identifying a study population, developing clinical outcome assessments or biomarkers, and serving as a comparator group in an externally controlled trial.

[9] See the following draft guidances for industry: *Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products* (September 2021); *Real-World Data: Assessing Registries to Support Regulatory Decision-Making for Drug and Biological Products* (November 2021); and *Data Standards for Drug and Biological Product Submissions Containing Real-World Data* (October 2021). When final, these guidances will represent FDA's current thinking on these topics.

[10] See 21 CFR 314.126(a).

control) arm—optimally promotes the similarity of compared groups regarding such influences, such that a conclusion can be made as to whether differences in outcomes observed between groups can be attributed to the treatment of interest. Nevertheless, for decades FDA has recognized the potential value of other types of controls, including historical controls as a type of external control.[11] Clinical trials using these other types of controls can, when appropriate, serve as the adequate and well-controlled clinical investigations generally required to provide substantial evidence of effectiveness under section 505(d) of the Federal Food, Drug, and Cosmetic Act (FD&C Act).[12]

Given that externally controlled trials do not involve randomization of the study population to the treatments being compared, the treatment and control arm populations should be as similar as possible regarding known factors that can affect the outcome being measured. These factors, discussed in more detail in section III, include important baseline characteristics (e.g., demographic factors, comorbidities), disease attributes (e.g., severity, symptoms, duration of illness), start of follow-up for the treatment of interest, concomitant therapies, and the clinical observations collected. Importantly, before choosing to conduct a clinical trial using an external control arm as a comparator, sponsors and investigators should consider the likelihood that such a trial design would be able to distinguish the effect of a drug from other factors that impact the outcome of interest and meet regulatory requirements.[13]

The suitability of an externally controlled trial design warrants a case-by-case assessment, informed by issues including heterogeneity of the disease (e.g., clinical presentation, severity, prognosis), preliminary evidence regarding the drug product under investigation, the approach to ascertaining the outcome of interest, and whether the goal of the trial is to show superiority or non-inferiority.[14] Of note, if the natural history of a disease is well-defined and the disease is known not to improve in the absence of an intervention or with available therapies, historical information can potentially serve as the control group. For example, objective response rate is often used as a single-arm trial endpoint in oncology given the established understanding that tumor shrinkage rarely occurs without an intervention.[15,16]

In many situations, however, the likelihood of credibly demonstrating the effectiveness of a drug of interest with an external control is low, and sponsors should choose a more suitable design,

[11] See 21 CFR 314.126(b)(2)(v).

[12] See section 505(d) of the FD&C Act (21 U.S.C. 355(d)).

[13] See 21 CFR 314.126.

[14] A non-inferiority approach is not recommended using an externally controlled trial design. See the guidance for industry *Non-Inferiority Clinical Trials to Establish Effectiveness* (November 2016).

[15] See the ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials* (May 2001).

[16] See the guidance for industry *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics* (December 2018).

regardless of the prevalence of disease. 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,[17] prognostic differences in the study populations, knowledge of treatment assignment (lack of blinding), or other factors such as differences in concomitant therapies.

The remainder of this guidance is intended to assist sponsors in identifying and addressing commonly encountered challenges when considering the conduct of an externally controlled trial.

III. DESIGN AND ANALYSIS OF EXTERNALLY CONTROLLED TRIALS

A. Design Considerations

1. Overview

Reducing the potential for bias in externally controlled trials is best addressed in the design phase, in that well-chosen design elements increase confidence in the interpretability of study results when appropriate analytic methods are applied to estimate treatment effects. Sponsors should finalize a study protocol before initiating the externally controlled trial, including selection of the external control arm and analytic approach, rather than selecting an external control arm after the completion of a single-arm trial. Specific design elements to prespecify in the protocol (i.e., before conducting an externally controlled trial) include suitable study data sources,[18] baseline eligibility (inclusion and exclusion) criteria,[19] appropriate exposure definitions and windows, well-defined and clinically meaningful endpoints, cogent analytic plans, and approaches to minimize missing data and sources of bias.

---

[17] Scenarios that would not be suitable for externally controlled trials include when the natural history of the disease of interest is not understood sufficiently or when the disease course is considered well-understood but is variable.

[18] FDA recognizes that access to and evaluation of relevant data sources or databases are important steps in designing a control arm for externally controlled trials and in evaluating the trial's feasibility. Sponsors should document and describe in the trial protocol all data sources accessed when designing the control arm of the trial and the results of any feasibility evaluations or exploratory analyses. Sponsors should provide a justification for selecting or excluding relevant data sources and demonstrate that the choice of a final analytic dataset for the control arm aligns with the research question of interest and was not chosen to favor particular study results. FDA recommends that sponsors generate audit trails in their datasets that can track access to and analyses performed on relevant data sources. See the draft guidance for industry *Considerations for the Use of Real-World Data and Real-World Evidence to Support Regulatory Decision-Making for Drug and Biological Products* (December 2021). When final, this guidance will represent the FDA's current thinking on this topic.

[19] In this guidance, the term *eligibility criteria* refers to the requirements for entry into a clinical trial (i.e., the characteristics the participants must or must not have to be able to participate in the trial). See the guidance for industry *Enhancing the Diversity of Clinical Trial Populations — Eligibility Criteria, Enrollment Practices, and Trial Designs* (November 2020).

The estimand framework[20]—involving a precise description of the treatment effect reflecting the clinical question posed by the study objective—can be used to help design an externally controlled trial. An estimand is comprised conceptually of the study population, treatment of interest and comparator, outcome of interest, handling of ***intercurrent events***, and summary measures. Many of the elements of the estimand framework are described individually in the subsections below, and considering the elements together promotes alignment of trial objectives, conduct, analysis, and interpretation of results.

A specific design consideration for externally controlled trials involves prespecifying plans regarding how to measure and analyze data on important ***confounding*** factors and sources of ***bias***. The ability to identify confounding factors in an externally controlled trial is limited by both conceptual and practical concerns. Conceptually, when seeking to provide evidence of effectiveness using an externally controlled trial design, a thorough understanding is needed—but is often difficult to verify—regarding the natural history[21] of the disease involved and relevant prognostic factors influencing outcomes. For example, important prognostic factors for an outcome may not be known and therefore cannot be used in the process of developing the external control arm to match, as closely as possible, such factors in the treatment arm.

From a practical perspective, fit-for-use data on suspected confounding factors (e.g., history of cigarette smoking, performance status) may be missing for some patients or participants or may be measured differently in the external control arm compared to the treatment arm. Accordingly, before deciding whether an externally controlled trial is a suitable design to answer the research question of interest, sponsors should confirm that recognized, important prognostic characteristics can be assessed in the data sources that will be used in an externally controlled trial. Specifically, the source population for the external control arm should be as comparable as possible to the treatment arm population, given that controlling for differences between the two study arms (see section III.C) becomes more challenging with increasingly dissimilar populations.

Although unmeasured confounding, lack of blinding, and other sources of bias cannot be eliminated in externally controlled trials, an assessment of the extent of confounding and bias, along with analytic methods to reduce the impact of such bias, are critically important in the conduct of such trials. Given the challenges outlined, externally controlled trials are more likely

[20] For further information, see the ICH guidance for industry *E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials* (May 2021).

[21] Changes over time in the understanding of the natural history of a disease can also introduce bias in an externally controlled trial. For example, diagnosis of patients with a genetic disorder may have been based historically on the development of signs and symptoms, whereas the development and increased use of genetic testing in clinical trials can diagnose patients at earlier stages of disease (see, for example, EA Nannenberg, IAW van Rijsingen, PA van der Zwaag, MP van den Berg, JP van Tintelen, MWT Tanck, MJ Ackerman, AAM Wilde, and I Christiaans, 2018, Effect of Ascertainment Bias on Estimates of Patient Mortality in Inherited Cardiac Diseases, Circ Genom Precis Med, 11(10):e001797). In such situations, a historical control arm would have shorter diagnosis-to-death intervals than a treatment arm, even if the drug of interest has no impact on survival.

to provide convincing results when the effect size on a well-characterized outcome of interest is anticipated to be large.[22]

2. Characteristics of Study Populations

In the absence of randomization, a major concern for externally controlled trials is that attributes of patients[23] likely to influence outcomes in an external control arm will differ from corresponding attributes of participants in a treatment arm of the trial. Examples of baseline attributes of participants or patients in treatment and control groups that can be dissimilar include demographic and related factors (e.g., age, sex, race, socioeconomic status, geographic region). Additional attributes that could be dissimilar but often are more challenging to address include disease characteristics (e.g., severity, duration, specific signs and symptoms, performance status), prognostic or predictive biomarkers,[24] comorbidities, and prior and current treatments received. When accounting for baseline characteristics, specific challenges can include (1) whether relevant confounding factors are known and well-characterized; (2) whether such confounding factors are captured; (3) whether these factors have been assessed with appropriate methods and measured similarly across compared groups; and (4) whether the study's analytic methods sufficiently address the differences in clinical characteristics between the compared groups.

A specific consideration involves how well the eligibility criteria can be applied to the external control arm in order to obtain a population comparable to the treatment arm. In addition, unless a concurrent control group is being used, sponsors should consider whether diagnostic criteria for the condition of interest and other relevant baseline factors, or the approaches used to ascertain data on such factors, have changed during the time of data collection. Accordingly, the protocol for an externally controlled trial should include specific plans for evaluating eligibility criteria to determine if the criteria can be applied in a manner that allows for selection of similar patients in the treatment and external control groups, recognizing the limitations of information available in many RWD sources.

3. Attributes of Treatment

In properly designed and conducted randomized trials, observed differences in efficacy and safety outcomes can generally be attributed to the investigational drug, but confidence in such attribution is diminished in externally controlled trials because of concerns over potentially

---

[22] See the ICH guidance for industry *E10 Choice of Control Group and Related Issues in Clinical Trials* (May 2001) and 21 CFR 314.126(b)(v).

[23] In this guidance, the term *patient* refers to a person whose health care information (e.g., regarding a disease) is included in a study, whereas the term *participant* refers to a healthy person or a person with a disease who participates in a study.

[24] Prognostic and predictive biomarkers are used to assess the rate of disease progression or response to therapy, respectively. For additional discussion, see BEST (Biomarkers, EndpointS, and other Tools) Resource, available at https://www.ncbi.nlm.nih.gov/books/NBK338448, as well as the guidance for industry *Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products* (March 2019).

important imbalances with respect to treatment between the treatment arm and the external control arm that are either not documented or cannot be accounted for. Such imbalances can involve factors related to the treatment of interest (e.g., adherence, dose, timing of initiation, and duration of treatment) and receipt of additional treatments. These concerns are expected when the data in the external control arm are from an RWD source, and although the remainder of this section focuses on such data sources, potential imbalances can also exist when the data come from other clinical trials.

Clinical trial protocols typically include a plan for collecting data on use of concomitant or supportive therapies (including non-prescription products) that could affect the outcomes of interest, along with detailed data on the characteristics and administration of such therapies. Examples include drug formulation, dose, strength, route, timing, frequency, and duration—and for certain medications, specific rules for dose modifications, interruptions, or discontinuations are specified in the protocol. In contrast, documentation of such data in routine clinical care may not be complete or accurate, and RWD may therefore lack comprehensive details describing the administration of a treatment or information on the use of concomitant or supportive therapies. For example, suitable data on additional treatment modalities (e.g., radiotherapy and surgical interventions when treating patients with cancer) may not be available in certain data sources. In addition, management of treatment- or disease-related adverse events may not be predefined or described consistently compared to a trial protocol.

Additional factors can influence the treatment and delivery of care that patients receive as well as the assessment of outcomes related to those treatments when data from clinical care are analyzed. Examples include differences in health-seeking behaviors, insurance coverage (including prescription drug plans), adoption of clinical practice guidelines, availability of novel treatments, and use of companion diagnostic testing (e.g., a genetic test used in conjunction with a corresponding therapeutic product). Access to emergency department or intensive care, availability and coordination of subspecialty care, and academic versus community health care settings can also be markedly different within or across health care systems or geographic areas. These and other health care delivery factors—at the level of the patient, provider, or health system—can influence treatment selection. Such factors should be identified and accounted for adequately in externally controlled trials; otherwise, a different design approach (e.g., randomized controlled trial) should be considered.

4. Designation of Index Date (Time Zero)

A specific and difficult challenge when designing externally controlled trials is specifying the index date (also called *time zero* or *zero time*), which is the start of the observation period for assessing endpoints. Given the lack of randomization in externally controlled trials, differences in the way the index date is determined across trial arms may lead to biased effect estimates. The index date for the treatment and control arms in a randomized trial is usually designated as the time when eligibility criteria are determined to have been met and a decision was made regarding the intended treatment strategy for each participant. For an externally controlled trial that relies on RWD, however, the index date for the control arm can be assigned in various ways.

If there are temporal differences in this date relative to treatment initiation or other important landmark times by treatment arm, any observed treatment effects may be biased.

Determination of the index date in the treatment arm and the external control arm should avoid analyses that include a period of time (immortal time) during which the outcome of interest could not have occurred in one of the two arms. If the index date is not established appropriately across compared arms in an externally controlled trial, bias due to immortal time can occur. For example, consider an externally controlled trial that involves a time-to-event mortality endpoint and an index date established as the time of having failed prior therapy. If analyses of participants in the treatment arm include only those who actually receive the drug of interest, then any period of time between eligibility determination (i.e., failed prior therapy) and treatment initiation is immortal time; that is, the person must survive the period to receive the drug and be accounted for in the analysis. In contrast, if patients in the external control arm do not receive subsequent therapy after determination of eligibility (i.e., failed prior therapy), these patients would be included in the analysis regardless of survival. Accordingly, patients with very short survival times would be included in the control arm but not in the treatment arm, introducing a bias that makes the drug seem more effective than it actually is.[25]

When assessing bias that may be introduced related to immortal time in an externally controlled trial, the clinical circumstances related to assigning the index date should be considered. Specifically, if a treatment strategy is assigned immediately after a discrete and identifiable clinical event, the index date for the compared groups may be reasonably determined by the time of occurrence of that event. For example, if treatment is started after an acute myocardial infarction, stroke, or heart failure hospitalization, these events may be more suitable to identify the index date for both the treatment arm and external control arm. In contrast, when the event that prompts the treatment of interest is not discrete and readily identifiable, such as worsening of heart failure symptoms or poor control of hypertension, determining a suitable index date can be difficult or may not be possible. Identifying an index date can also be especially challenging in situations in which no treatment is the treatment strategy for the external control arm.

5. Assessment of Outcomes

The lack of blinding to treatments in externally controlled trials can pose challenges when considering certain outcomes, in that knowledge of the particular treatment by patients, caregivers, clinicians, or investigators can potentially lead to a biased estimate of the effect of treatment. Accordingly, whenever possible and for suitable endpoints, the outcome should be assessed blinded to treatment status. In some cases, this activity may require re-adjudication of the externally controlled data, such as by blinded independent central review. Bias can also be introduced if outcome assessments in the treatment arm and the external control arm differ based on the sources of data involved or the criteria used to establish outcomes. Sponsors should seek to assess outcomes consistently across the treatment arm and the external control arm for the results of an externally controlled trial to be credible.

[25] In a randomized trial, potential periods of immortal time are expected to be balanced across treatment groups.

Well-defined, reliable, and clinically meaningful outcomes that are typically used in randomized trials may be particularly difficult to ascertain and evaluate in an RWD source that is being considered for an externally controlled trial. For example, radiologic endpoints in controlled oncology trials (e.g., objective response rate and progression-free survival) are based on prespecified imaging assessment frequency and standardized measurement criteria for response evaluation criteria in solid tumors (RECIST). In routine clinical care, however, radiologic assessment frequency is variable, and formal tumor measurement may not routinely be performed or documented, making a valid assessment of progression-free survival or objective response rate using external control data, such as data from EHRs, challenging.[26] A similar consideration applies to the assessment of motor milestones, such as the ability to sit or walk, which are usually not recorded with the same rigor during routine clinical care compared to approaches used in clinical trials. As another example, a randomized trial may include specific testing to detect or confirm a particular clinical entity (e.g., severe inflammatory bowel disease activity confirmed by endoscopy), whereas various strategies may be used in clinical care to identify and confirm the same event. In some cases, and depending on the outcome, the occurrence of an event (e.g., worsening heart failure status according to a specific classification system) may not have been evaluated in clinical care or, if evaluated, may not have been recorded. As a general consideration, outcomes of interest are more likely to be recorded in clinical records when events are objective and/or require immediate medical attention (e.g., stroke or myocardial infarction).

When considering outcomes in externally controlled trials, sponsors should also evaluate the consistency of timing of outcome assessments in the treatment arm compared to the external control arm. In general, the timing and frequency of outcome assessments in RWD will have been determined during clinical care[27] and may have been influenced by the patient's clinical status, whereas outcome assessments in the treatment arm are protocol-specified. In addition, even when external control arm data are from another clinical trial rather than from an RWD source, the approach to outcome ascertainment may differ from the treatment arm. Accordingly, sponsors should first establish for what total duration of time and at what intervals the outcome of interest should be assessed in the analysis of data from an externally controlled trial. Based on such determinations, sponsors can then evaluate whether the availability and timing of outcome assessments are sufficient and comparable across both arms of the externally controlled trial for the research hypothesis being tested.

Additional challenges when considering the selection of outcomes to be assessed in an externally controlled trial include changing diagnostic criteria over time for what constitutes abnormal clinical, radiographic, serologic, or other outcomes. Whereas both trial arms would be similarly affected in a traditional randomized trial, extensive heterogeneity or substantial changes in

[26] See EA Eisenhauer, P Therasse, J Bogaerts, LH Schwartz, D Sargent, R Ford, J Dancey, S Arbuck, S Gwyther, M Mooney, L Rubinstein, L Shankar, L Dodd, R Kaplan, D Lacombe, and J Verweij, 2009, New Response Evaluation Criteria in Solid Tumours: Revised RECIST Guideline (version 1.1), Eur J Cancer, 45(2):228–247.

[27] Registries (one type of RWD) may collect data at predetermined and regular intervals, whereas EHRs and medical claims data would usually not.

diagnostic criteria can introduce bias when analyzing outcomes using a non-contemporaneous external control arm (or when using a reasonably contemporaneous external control arm that reflects a different diagnostic standard of care). As another challenge that can introduce bias, biomarkers used as surrogate outcomes in clinical trials may be used for different purposes in clinical care, or biomarkers used in clinical care may not be well-characterized in terms of comparability to assays used in clinical trials.

Further challenges may arise from differential capture of intercurrent events that may preclude the measurement of or impair the interpretability of the treatment effect on the outcome of interest. For example, initiation of ancillary therapy after treatment with the drug of interest is started may be protocol-determined and recorded during study visits in a clinical trial, whereas data from routine clinical care may not accurately capture additional therapies, which may confound interpretation of the effect of treatment on the study outcome.

Other considerations apply when an outcome in an externally controlled trial is based on certain clinical outcome assessments.[28] For example, the potential lack of standardization and training in the definitions and use of such assessments in routine clinical care settings—if the assessments are used at all—compared to what occurs in clinical trial settings, can lead to higher variability or bias in the measurements from an external control arm. Accordingly, clinical outcome assessments that are acceptable in randomized trials may not be fit for use in externally controlled trials.

B. Data Considerations for the External Control Arm

1. Data from Clinical Trials

Using data from another clinical trial for an external control arm can have advantages compared to using data collected during routine clinical care, based in part on the rigor of protocol-based (and therefore more consistent) data collection. Such use would only be appropriate, however, when comparability exists between the two trial arms regarding participant eligibility criteria, treatment administration, patterns of care (e.g., location of treatment sites), recording of concomitant medications, and assessments of adverse events and outcomes. A particular concern for bias would be the selection of an external control arm from a completed trial whose outcomes are already known. This would be especially problematic if the results of the external control arm are inconsistent with prior experience. Furthermore, when using data from other clinical trials as an external control arm, sponsors should consider the extent of and reason for any missing data and how the interpretability of study results may be affected.

In many situations, data for the treatment and control arms in an externally controlled trial will have been collected during different time periods. Lack of concurrent data collection may be of

[28] A *clinical outcome assessment* is a measure that describes or reflects how a patient feels, functions, or survives. Types of such assessments include measures of *patient-reported outcomes*, *observer-reported outcomes*, *clinician-reported outcomes*, and *performance outcomes*. See BEST (Biomarkers, EndpointS, and other Tools) Resource, available at https://www.ncbi.nlm.nih.gov/books/NBK338448.

particular concern when the assessment and management of a disease (including supportive care) changes over time, such as use of predictive or prognostic biomarkers in the patient population. For example, prior trials involving certain cancers may not have information regarding newer biomarkers or specific gene alterations of interest or tumor mutational burden. Accordingly, sponsors should assess whether use of data from a specific clinical trial is justified as an external control arm when planning an externally controlled trial.

2. Data from RWD Sources

The concerns described in the preceding section regarding comparability of participant characteristics, timing and frequency of data collection, and patterns of care should be addressed when using RWD collected on patients for non-research purposes as external control arms. In addition, specific concerns regarding missing data from RWD sources obtained as part of routine clinical practice can threaten the validity of the results of an externally controlled trial. For example, patients who initially met eligibility criteria may be lost to follow-up (e.g., due to changing their health care provider) from the external control arm. Furthermore, availability of a dataset containing patients with the disease of interest does not guarantee that there is sufficient information on relevant clinical characteristics (e.g., prognostic factors for the outcome of interest) to permit an appropriate comparison.

3. Considerations for Assessing Comparability of Data Across Trial Arms

The table below summarizes important considerations, discussed above, regarding the comparability of data between the treatment arm and the external control arm. The relevance of each consideration can vary on a case-by-case basis, depending on attributes of the treatment arm, the selected data source for the external control arm, and the stage of the trial (design, conduct, or analysis).

374 **Table. Summary of Considerations for Assessing Comparability of Data**[29]
375

Focus of Comparison	Considerations for Data Comparability
Time periods	Various aspects of clinical care may change over time, such as the standard of care for the condition of interest, types of treatments, supportive care regimens, and criteria for determining disease response or progression. Such temporal differences are difficult to address using statistical analyses alone. It is important to consider whether and how different time frames in the treatment arm and the external control arm impact the interpretability of study findings.
Geographic region	Standards of care and other factors (e.g., access to care) that affect health-related outcomes can vary across geographic regions and health care systems. A balance of participants or patients across geographic regions and health care systems in an externally controlled trial, when possible, can help reduce the impact of confounding based on such differences.
Diagnosis	The criteria used to establish a diagnosis may differ based on practice variation or may have changed in the interval between when the treatment arm of the trial was conducted and when the data for the external control arm were collected. Sponsors should consider the diagnostic standards used and whether relevant clinical tests to establish a diagnosis were conducted and reported equally across the compared arms.
Prognosis	Based on demographic and clinical characteristics—and if sufficient knowledge of relevant prognostic factors is available—prognostic indicators for the participants or patients in each arm of the trial should be evaluated and shown to be of sufficient similarity to permit an unbiased assessment of the treatment-outcome association.
Treatments	Attributes of the treatment of interest—including drug formulation, dose, route of administration, timing, frequency, and duration as well as specific rules for dose modifications, interruptions, discontinuations, and adherence—will have been prespecified or measured in the treatment arm. In contrast, specific aspects of a comparator treatment (as applicable) in the external control arm may not have been protocol-driven depending on the data source. Accordingly, sponsors should assess whether the external control arm data can be meaningfully compared to the treatment arm data.
Other treatment-related factors	Various treatment-related considerations, when relevant, include (1) previous treatments received (e.g., lines of therapy in patients with cancer), (2) medications received concomitantly that can affect the outcome of interest, or (3) predictive biomarkers (e.g., genomic testing) related to the treatment of interest. When differentially distributed across groups being compared, such factors can threaten an assessment of the drug-outcome association.
Follow-up periods	Designation of the index date should be consistent between the treatment arm and the external control arm, and the duration of follow-up periods should be comparable across compared arms.
Intercurrent events	The relevance of intercurrent events across treatment arms should be assessed, including differential use of additional therapies after initiation of the treatment of interest.
Outcome	Whether endpoints used in an externally controlled trial can be reliably and consistently measured across the external control arm and the treatment arm will be influenced by several factors, including the definitions of the endpoints, the data

[29] Some of the considerations will be relevant to multiple rows.

Focus of Comparison	Considerations for Data Comparability
	source for the external control arm, and the potential for the outcome to be influenced by knowledge of treatment received. In addition, sponsors should be able to apply the same criteria for the evaluation and timing of outcome assessments across both arms of the externally controlled trial.
Missing data	The extent of missing data in the external control arm should be assessed before conducting an externally controlled trial to evaluate feasibility (when such data are available). When analyzing results from such a trial, the extent of missing data in both the treatment and external control arms should be assessed to examine the potential impact of missing data.

The considerations listed in the table above are directed at understanding and managing potential threats to the validity of externally controlled trials. Additional considerations regarding the comparability of trial arms may be relevant for a specific externally controlled trial.

C. Analysis Considerations

1. General Considerations

Before conducting an externally controlled trial, sponsors should develop a statistical analysis plan that prespecifies analyses of interest, such as analyses of primary and secondary endpoints, calculations of statistical power and sample size, and plans to control the chance of erroneous conclusions (e.g., to control the overall type I error probability). The statistical analysis plan should be submitted along with the protocol to the relevant review division before initiation of enrollment in the clinical trial for the experimental treatment. In addition, decisions regarding the study design and statistical analysis plan for an externally controlled trial should be blinded to any observed external control data (e.g., from an existing RWD source), with the exception of planned feasibility analyses, such as evaluating the availability of key variables or missing data. During the conduct of an externally controlled trial, and specifically when analyzing data already collected, changes to the statistical analysis plan are discouraged. If such changes are nonetheless implemented, any revisions should be date-stamped and the corresponding rationale provided and discussed with the relevant FDA review division.

FDA does not recommend a particular approach to analyzing data from externally controlled trials. No single statistical or analytical method will be suitable for all trials involving external control arms, and potential approaches should be discussed with the appropriate FDA review division. Sponsors should provide a justification for the analytic methods selected as well as a description of the strengths and limitations of the methods used to assess the effect of treatment. In general, the analytic method used should identify and manage sources of confounding and bias, including a strategy to account for differences in baseline factors and confounding variables between trial arms.

Various statistical methodologies may be appropriate for these types of comparisons, each with a corresponding level of complexity regarding approaches to account for bias. The assumptions involved should be made explicit, and sensitivity analyses as well as model diagnostics should be

conducted to examine such assumptions. Importantly, however, adding complexity to an analytical framework usually requires making additional assumptions, which often cannot be substantiated and may impair the interpretability of results.

Even when employing analytic methods to balance the trial arm populations, sponsors should propose additional analyses to evaluate the actual comparability between the external control and treatment arms for important covariates. Determining similarity across trial arms will require selection of specific population characteristics to compare, a method for the comparison, and criteria to demonstrate similarity. For example, an a priori threshold[30] could be set to determine whether the external control population has a statistical distribution of covariates that is similar to the treatment arm population after a balancing method, such as weighting, has been applied.

Consideration should also be given, based on available scientific data, to the anticipated effect size for analyses of the primary endpoint. Especially when the anticipated effect size is modest, an externally controlled trial may not be an appropriate study design because of concerns for bias affecting the results. In addition, sponsors should develop a priori plans for assessing the impact of confounding factors and sources of bias, with quantitative or qualitative bias analyses used to evaluate these concerns. Such prespecified analyses can assist in the interpretation of study results.

2. Missing Data

The proposed analytical methods should include a strategy for dealing with missing data, including data that may not be available in a chosen data source based on the type and frequency of assessments conducted during the patient encounter, patients no longer being followed, or other reasons. Analytical methods (such as strategies for imputing missing data) may be used in such situations, but these methods require assumptions regarding the pattern of missing information.[31] Assumptions about missing data can be unverifiable and may be difficult to justify, in addition to other assumptions required for estimation of treatment effect in a non-randomized setting.

[30] FDA does not endorse a single approach for determining thresholds. As one example, a threshold value could be selected for standardized mean differences as a metric that summarizes the statistical distribution of important prognostic covariates.

[31] The terms *missing completely at random, missing at random,* and *missing not at random* describe assumptions about why data are missing. When observations of a variable are *missing completely at random*, the missing observations are a random subset of all observations, such that the missing and observed values have the same underlying distributions, and bias from missing data is not a threat to the study. *Missing at random* indicates systematic differences may exist between the observed and unobserved values of a variable, but other observed variables could be used to address such differences and mitigate bias. *Missing not at random* indicates that the missing data are directly related to the treatment or outcome under investigation, and bias can be introduced. See AR Donders, GJ van der Heijden, T Stijnen, and KG Moons, 2006, Review: A Gentle Introduction to Imputation of Missing Values, J Clin Epidemiol, 59(10):1087–1091.

To understand the potential impact of missing data, externally controlled trials should be designed to capture and analyze information relevant to the missing data (e.g., available characteristics of patients with and without missing data). Analytical methods may be used, as mentioned above, to address potential bias caused by the missing data in the primary analysis. In addition, sensitivity analyses should be used to evaluate the potential impact of plausible violations in missing data assumptions on the results of the primary and other key analyses.

In some cases, data may be missing because of an intercurrent event, which may interfere with the measurement of outcomes and estimation of the treatment effect. The study analysis plan and an appropriate estimand should account for any intercurrent event that can be considered potentially related to both the treatment and outcome of interest, recognizing that certain intercurrent events may be difficult to detect in external control datasets. For example, in contrast to data collected according to a research protocol, RWD sources may not capture the time of occurrence of an intercurrent event, precluding accurate assessment of time-to-event endpoints such as progression-free survival.

3. Misclassification of Available Data

Misclassification[32] (mischaracterization) of data in externally controlled trials, especially in an external control arm using RWD sources, can occur when the value of a measurement is assigned to an incorrect category for subsequent analysis, potentially affecting estimates of the observed drug-outcome association. For example, EHR data collected during routine clinical care may include information on lifestyle characteristics, such as alcohol use. Beyond concerns about potentially inaccurate reporting by patients about their alcohol intake because of stigma or other factors, differences in the approach used to classify alcohol use within or across various sources of data can lead to misclassification. In routine clinical practice, for example, different health care providers may use different quantitative or qualitative descriptions of alcohol use, such that two patients with the same actual intake may be assigned to two different categories in the RWD source.

If misclassification is extensive—especially when information on treatments, outcomes, or confounding factors are involved—a biased assessment of the drug-outcome association may occur. For example, the scenario described above regarding misclassification of alcohol intake would be relevant when alcohol use is a potentially important confounding factor (covariate) in an analysis of an externally controlled trial. Although analytical modeling methods could be used to assess the potential impact of misclassification, the best strategy to avoid bias is to use objective and reliable measurements for the data of interest. For example, RWD sources that include information on alcohol intake collected using structured questionnaires are generally more reliable than patient-reported and clinician-documented values obtained during routine patient care.

[32] Misclassification errors can be non-differential when the probability of misclassification is equal across study arms or differential when the probability of misclassification differs across study arms. Misclassification can introduce bias regarding the drug-outcome association when involving the drug of interest, covariates, or outcomes of interest.

4. Additional Analyses

Sponsors can also use specific sensitivity analyses to test the vulnerability of trial results to assumptions in the analysis plan. For example, if the primary analysis of a time-to-event endpoint assumes proportional hazards, an appropriate sensitivity analysis could be estimation by a statistical method that does not assume proportional hazards. Finally, prespecified supplementary analyses can provide further understanding of the treatment effect. An example would be supplementary analyses in prespecified subgroups based on prognostic factors for the outcome.

IV. CONSIDERATIONS TO SUPPORT REGULATORY REVIEW

A. Communication with FDA

Sponsors should consult with the relevant FDA review division early in a drug development program about whether it is reasonable to conduct an externally controlled trial instead of a randomized controlled trial. As part of these discussions, sponsors should provide a detailed description of the (1) reasons why the proposed study design is appropriate, (2) proposed data sources for the external control arm and an explanation of why they are fit for use, (3) planned statistical analyses, and (4) plans to address FDA's expectations for the submission of data.

B. Access to Data and Documents

Sponsors must include in their marketing applications relevant patient-level data (i.e., data on each participant and patient in the externally controlled trial), as required under FDA regulations,[33] for both the treatment and external control arms. If sponsors do not own the data used for the external control arm, they should structure their agreements with the data owners to ensure that patient-level data can be provided to FDA in support of the marketing application. Sponsors should also ensure that FDA has access to ***source documents*** and ***source data*** for the external control arm as part of an FDA inspection or upon request.[34]

[33] See 21 CFR 314.50(f) and 601.2.

[34] See the guidances for industry *Use of Electronic Health Record Data in Clinical Investigations* (July 2018) and *Electronic Source Data in Clinical Investigations* (September 2013).

GLOSSARY

Bias: Any systematic error in the design, conduct, analysis, or interpretation of a study that results in an erroneous estimate of a treatment's effect on the outcome of interest.

Confounding: Distortion of the measure of the effect of a treatment on an outcome due to another factor that is associated with both the treatment and the outcome.

Intercurrent Event: An event occurring after treatment initiation that affects either the interpretation or the existence of the measurements associated with the clinical question of interest. Examples include switching or discontinuing treatment, using rescue medications, or experiencing terminal events such as death.

Real-World Data (**RWD**): Data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources.

Real-World Evidence (**RWE**): Clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD.

Source Data: All information in original records and certified copies of original records of clinical findings, observations, or other activities (in a clinical investigation) used for the reconstruction and evaluation of the study. Source data are contained in source documents (i.e., original records or certified copies).[35]

Source Documents: Original documents, data, and records (e.g., hospital records; clinical and office charts; laboratory notes; memoranda; subjects' diaries or evaluation checklists; pharmacy dispensing records; recorded data from automated instruments; copies or transcriptions certified after verification as being accurate copies; microfiches; photographic negatives; microfilm or magnetic media; x-rays; subject files; and records kept at the pharmacy, at the laboratories, and at the medico-technical departments involved in the clinical trial).[36]

[35] See the guidance for industry *Electronic Source Data in Clinical Investigations* (September 2013).

[36] See the guidance for industry *E6(R2) Good Clinical Practice: Integrated Addendum to ICH E6(R1)* (March 2018).