ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ANNALS OF ONCOLOGY driving innovation in oncology

REVIEW

External control arms in oncology: current use and future directions

P. S. Mishra-Kalyani[1*], L. Amiri Kordestani[2], D. R. Rivera[3], H. Singh[2,3], A. Ibrahim[2], R. A. DeClaro[2,3], Y. Shen[1], S. Tang[1], R. Sridhara[3], P. G. Kluetz[2,3], J. Concato[4], R. Pazdur[2,3] & J. A. Beaver[2,3]

[1]Office of Biostatistics, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring; [2]Office of Oncologic Diseases, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring; [3]Oncology Center of Excellence, U.S, Food and Drug Administration, Silver Spring; [4]Office of Medical Policy, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, USA

Available online 10 January 2022

Although randomized control trials allow for a comparison of treatment arms with minimal concern for confounding by known and unknown factors, a randomized study is not feasible in certain disease settings. When a randomized design is not possible, incorporating external control data into the study design can be an effective way to expand the interpretability of the results of an experimental arm by introducing the ability to carry out a formal or an informal comparative analysis. This paper provides an introduction to the concepts of external controls in oncology trials, followed by a review of relevant and current research on this topic. The paper also focuses on general considerations for designing a trial that may incorporate external control data, followed by case studies of the marketing applications submitted to the Food and Drug Administration that included external control data.
Key words: external controls, real-word data, Food and Drug Administration

INTRODUCTION

The control arm of a randomized clinical trial plays a fundamental role in estimating the efficacy and safety of an investigational therapy. Concurrently randomized control arms allow for an understanding of the temporally relevant factors associated with the natural history of the disease, particularly with respect to current standards of clinical care. This contemporaneous control permits an estimation of treatment effect that is attributable to the experimental arm of interest, and randomization minimizes concern for bias by removing systematic imbalances between arms in measured and unmeasured prognostic factors. When a concurrently randomized arm is not feasible in an oncology trial due to practical or ethical concerns, single-arm trials may be appropriate. In single-arm trials, tumor response rates are an appropriate endpoint to assess treatment effect, as an individual patient's own baseline tumor measurement serves as an internal control with the assumption that most tumor types will not shrink without intervention or treatment. However, if tumor response cannot be reasonably measured due to disease characteristics (such as in certain neuro-oncologic tumors) or there is interest in estimating a comparative treatment effect within the population of interest, approaches have been explored to supplement single-arm data with data external to the clinical trial, also referred to as an external control arm.

External control arm data may be derived from prior clinical trial data (individual or pooled), or observational, real-world data (RWD), such as from registries, electronic health records (EHRs), and medical or pharmacy claims (see Figure 1). An external control arm data source should be temporally and clinically relevant to the investigational arm to minimize bias, and necessitates sufficient individual patient-level data to ensure a well-powered comparison.[1-3] Differing practical and statistical implications arise when analyzing each of these sources of data, as well as when incorporating patient-level data versus benchmark outcome rates as an external comparison. The advantages of using prior clinical trial data may include a well-defined population, with protocol-measured patient characteristics and endpoints. The same data may have disadvantages, however, such as being temporally irrelevant with respect to patient selection, prognostic biomarkers, or evolving standards of care or have variability with respect to measurement of tumor-based endpoints. Although RWD may offer a breadth of available data, its application in externally controlled trials may be problematic; a fit-for-purpose assessment is necessary as data source quality and granularity varies. In particular, tumor measurements by radiological scans may not be available or conducted with the same frequency or using the same assessment criteria as the clinical trial arm, resulting in comparison of

*Correspondence to: Dr P. S. Mishra-Kalyani, 10903 New Hampshire Avenue, Silver Spring, MD 20993, USA
E-mail: pallavi.mishra-kalyani@fda.hhs.gov (P. S. Mishra-Kalyani).

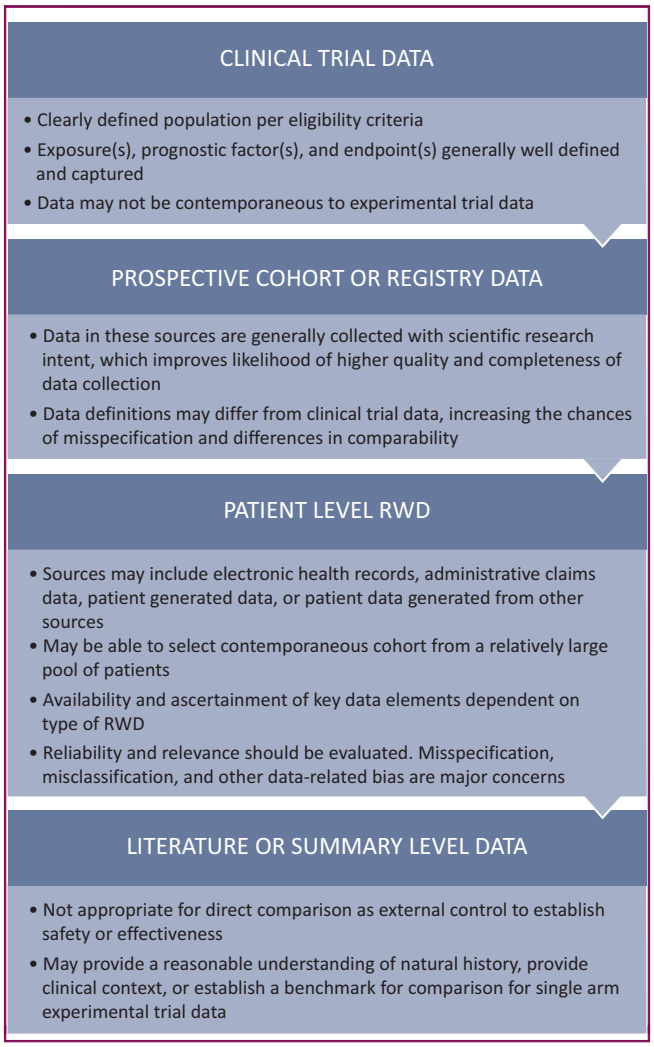

Figure 1. Considerations for selection of an external control data source.
EHR, electronic health record; RWD, real-world data.

incompatible endpoints. Additional discussion on the choice of external control data is available in the literature[4,5] as well as general clinical trial guidance.[1,6]

While the use of external control data to support the regulatory review of an experimental therapy is of interest to the Food and Drug Administration (FDA), best practices for study design and statistical analysis plans are not yet defined and challenges exist. Despite these challenges, marketing applications in oncology have included trial designs with external control comparator arm(s). The adequacy of the external control data has been variable based on the quality and it being fit for purpose in a comparative analysis. In the following sections, we outline relevant and recent research on external controls and considerations for when an external control arm may be appropriate for a study design; we also provide selected regulatory case studies of use.

CURRENT STATUS OF EXTERNAL CONTROL RESEARCH

Various contexts for the use of external control data for clinical trials in oncology have been discussed.[4,7-12] Some articles focus on general considerations for studies that include external control data focusing on the key questions and criteria that aim to ensure the quality of an external control data source and the adequacy of an analytical plan.[4] Others provide an analysis of the use of external control data, with a comprehensive description of potential applications of externally controlled trial design, accompanied by the considerations for mitigating bias by proper study design elements and analysis providing comparisons of the advantages and disadvantages with respect to selection of external control data such as temporality of the data (historical versus concurrent) and intention of data analysis (benchmark rates versus formal comparative analyses).[7]

In addition, several whitepapers from research collaborative groups provided recommendations regarding the design of oncology clinical trials that incorporate external control data. In a 2019 whitepaper,[8] a Friends of Cancer Research working group provided extensive descriptions of key design and analysis elements of trials that incorporate external control data, common limitations and pitfalls arising for unmitigated biases, as well as a case study that exchanged the concurrently randomized control arm for matched external control data from previously conducted clinical trials in a clinical trial of multiple myeloma patients. Utilizing propensity score matching, the working group was able to produce similar results using the matched external control data to those observed in the original randomized control arm.[13]

Other recent research works focusing on external controls in oncology have described design and analytic methods for incorporation of these data in a clinical trial. The corresponding papers include original research to demonstrate the utility of external control data, as well as details of the operating characteristics of the statistical methods. For example, Amiri-Kordestani et al.[9] describe the use of data from five previously conducted clinical trials as an informal comparative cohort for a single-arm study of de-escalated therapy in human epidermal growth factor receptor 2-positive early breast cancer. The external control cohort was matched to the experimental arm of adjuvant paclitaxel and trastuzumab using propensity score methodology. After matching, consistent efficacy outcomes across treatment cohorts suggested that the de-escalated therapy could be clinically similar to treatment regimens with additional components.

Similarly, and to determine whether de-escalation or interruption of adjuvant endocrine therapy is associated with treatment outcomes, Sun et al.[10] used data from two prior studies—the Suppression of Ovarian Function Trial (SOFT) and the Tamoxifen and Exemestane Trial (TEXT)—in an informal comparison to the Pregnancy Outcome and Safety of Interrupting Therapy for women with endocrine responsIVE breast cancer (POSITIVE) study. A repeated sampling approach was applied to balance the cohorts of patients in the POSITIVE study and those from the SOFT and TEXT study with respect to baseline characteristics. The results included estimated 3-year breast cancer-free interval rates in a population with an expected natural history

https://doi.org/10.1016/j.annonc.2021.12.015

similar to those studied in the prior studies. The authors indicate that their approach could provide context with respect to design and analysis for a study population of interest, if trying to establish expected clinical outcome under standard of care.

Ventz et al.[11] considered the use of both data from previously conducted clinical trials as well as RWD for an externally controlled trial to understand treatment effect on overall survival (OS) in glioblastoma multiforme. The authors advocate for the use of several sources of external data to assess and potentially reduce potential bias, provide complete time-to-event data, and retain sufficient power when compared to traditional randomized control trial designs.

Some considerations specific to external controls that originate from RWD sources are provided by Carrigan et al.,[12] in a case study of several non-small-cell lung cancer studies in which EHR data are used to replace the concurrent control arms. For each trial, the original clinical trial eligibility criteria were applied to an EHR database to select potential external control arm patients, and then inverse probability of treatment weights were applied to the analysis comparing OS between the original trial experimental arm and the external control cohort. The results suggested that in certain diseases with measurable disease status, outcomes, and confounders that can be adequately captured in RWD, external data may provide a supplementary control information for a prospective experimental arm. This analysis was not without limitations, however, including missing data with respect to outcome assessment, potential unmeasured confounders, and concerns regarding population representativeness between the different data sources (clinical trial versus community care setting). As the authors suggest, the use of external controls in early phase trials may provide a better understanding of treatment effect than a single-arm trial to inform the assumptions for later phase trials.

DESIGN CONSIDERATIONS FOR THE USE OF EXTERNAL CONTROLS

The concept of using data collected externally to a clinical trial as part of the collective evidence to support a treatment effect is not novel. Data from previously conducted clinical studies or observational sources are often used to inform study assumptions, particularly in establishing natural history of disease or the expected standard-of-care treatment effect. However, with increasing clinical studies in specific disease subpopulation areas and availability of vast amounts of EHR data, there is a renewed interest in using data external to a clinical trial for comparative efficacy and safety analyses. Additionally, the 21st Century Cures Act has led the FDA to develop a framework for evaluating the quality and relevance, also referred to as fit for purpose, of real-world evidence, such that it may be incorporated into regulatory decision making.[14]

Prior regulation, such as 21CFR 314.126, and guidances such as ICH Harmonised Tripartite Guideline E10 titled 'Choice of Control Group and Related Issues in Clinical Trials'[1] have provided advice regarding the selection of a control arm for estimation of treatment effect, including the use of an external control arm design. In general, the use of external controls to determine comparative treatment effect is not widely applicable or appropriate for most clinical studies due to the potential for bias, such as confounding, selection bias, temporal bias, or immortal time bias amongst others. Potential applications may include a study design for a clinical question where the natural history, morbidity, or mortality of the disease is well characterized, highly predictable, the expected effect size of the investigational treatment is high, and the outcome precisely measured.[15] Even in a well-understood disease, however, lack of appropriate measurement of exposure or outcome, misaligned contemporality, or poor selection of an appropriate control can create an artificially significant demonstrated treatment effect that is not related to therapeutic intervention.

If the appropriateness and feasibility of the clinical study question of interest is suitable for utilizing an external control arm, careful attention should be paid to determining if the proposed data are fit for purpose, developing a detailed study protocol, and the pre-specification of a comprehensive statistical analysis plan. Figure 2 depicts some of the key elements that can be considered when determining whether a proposed external control data source is fit for purpose. Precisely defining the patient population, through eligibility criteria that can be applied to both the investigational arm of interest as well as the external control arm, can reduce subjectivity of the comparison.

Availability of important demographic and prognostic characteristics, as well as measurement of exposure data for treatments of interest (investigational, control, supportive, or concomitant) and endpoints of interest, is of paramount concern for a comparative analysis. In particular, it must be well established that the definition and ascertainment of these data elements are identical to ensure feasibility of a comparative analysis that is not riddled with concerns of bias due to measurement error or misclassification.[16] Further, the strength of evidence provided by the external control data is directly related to the quality and comprehensiveness of the data. A determination of quality may depend on the category of data (e.g. RWD versus prior trial) as well as the completeness and validation of individual data elements. Moreover, it is important that selection of data eligible for an external control arm is transparent and traceable with an audit trail available for FDA inspection.

Careful planning in the design phase before study initiation along with the implementation of a detailed protocol can ensure comparability of patient populations and data, such that the planned statistical analyses that adjust for residual biases can be reasonably relied on for better estimation of treatment effect in isolation.

A pre-specified statistical analysis plan for any study that incorporates external control data increases the integrity of the data analysis and results. The analysis plan should

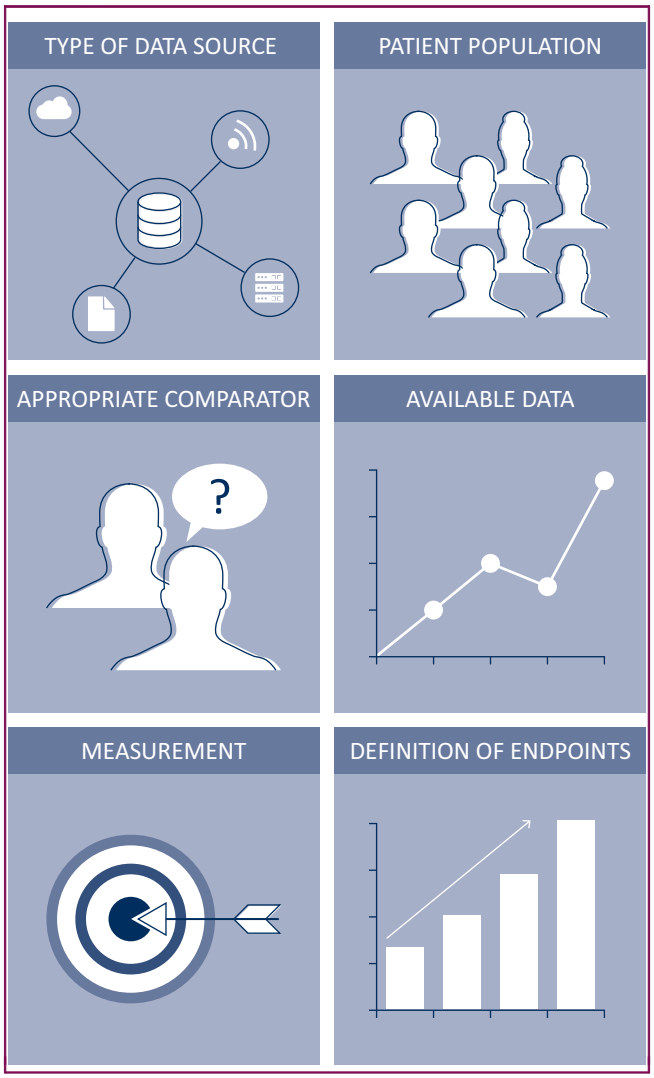

Figure 2. **Key factors for determining if data are fit for purpose.**

include statistical methods to account for various types of potential bias, including major threats from lack of randomization (e.g. selection bias) and confounding amongst others. An assessment of the similarity of the patient populations in each arm using pre-specified criteria to measure balance,[17,18] before and after any statistical procedures or adjustments, could further minimize the concern of bias. These comparisons of the populations of interest before and after analytical adjustments to account for bias and confounding are ideally conducted before the analysis of any outcome data.

One study design that incorporates external control data is an augmented or hybrid design, in which the concurrent control arm of a randomized clinical trial is supplemented by external data. For example, by increasing the randomization ratio to 2 : 1 or 3 : 1 and supplementing the concurrently randomized control arm with an external data source, a study design can achieve a 1 : 1 overall ratio for the comparison of experimental therapy to combined control. This approach would both reduce the patient burden for prospective enrollment, provide some concurrently randomized control data and reduce the risk of relying exclusively on external control data, and allow comparability of the external control arm and prospective experimental data by comparing patient characteristics or early clinical outcomes between external and concurrently randomized controls.[19] Although there has been previous discussion of statistical methods for borrowing data from external sources for augmented control arms in oncology,[20,21] this topic is a specific area of interest in the literature for pediatric clinical trials in various disease areas. Hybrid or augmented clinical trial designs in oncology could build on such established work in pediatric trials, which provide some guidance for study design and analysis plans.[22,23]

The required specifications of study design and analysis plan will vary by intended purpose of external control data and selected data source. As discussed earlier, external control data have previously been used in oncology marketing applications reviewed by the FDA to establish natural history of disease, with or without an established standard of care, provide descriptive information regarding the treatment effects of individual components of a combination treatment effect, or to provide supportive analyses of direct comparison to an experimental arm.[24-30] To date, no primary efficacy analysis of a study used to support approval of a marketing application in oncology has included a formal comparison to an external control arm. In the following section, we describe recent use cases, summarized in Table 1, including external control data in marketing applications for oncology, organized by intended use of the data.

CASE STUDIES

External controls to establish natural history of disease

Selumetinib for pediatric patients with neurofibromatosis type 1 with inoperable plexiform neurofibromas. In April 2020, the FDA approved selumetinib for pediatric patients, 2 years of age and older, with neurofibromatosis type 1 (NF1) who have symptomatic, inoperable plexiform neurofibromas (PNs). NF1 is a rare disease that occurs in ∼1 in 3000 births.[24] PN is a benign tumor that occurs in 20%-50% of patients with NF1 and can cause disfigurement, pain, and motor and neurologic dysfunction.[25,26] Approval of selumetinib was based on the SPRINT Phase II Stratum 1 study (NCT01362803), a single-arm, multicenter trial (FDA 2020) that demonstrated a durable overall response rate (ORR) as per the Response Evaluation in Neurofibromatosis and Schwannomatosis criteria and supported by observed clinical improvements in PN-related symptoms and functional impairments among 50 pediatric patients with inoperable PN.[27,28]

Given the rarity of the disease and uncommon occurrence of spontaneous regression of NF1 PN, two previously conducted trials were also submitted as supportive external control data for the SPRINT Phase II Stratum 1 data. The first study was a natural history study of NF1 in patients <35 years of age with a clinical diagnosis of NF1 or a confirmed

https://doi.org/10.1016/j.annonc.2021.12.015

Table 1. Regulatory case studies

Drug	Disease setting	Source of external control data	Regulatory use of external control data
Selumetinib	Neurofibromatosis type 1 with inoperable plexiform neurofibromas (pediatric)	Previously conducted clinical trials	Establish natural history of disease
Erdafitinib	Unresectable urothelial cancer harboring select FGFR genetic alterations	Patient-level EHR data from US community-based cancer clinics	Establish natural history of disease
Pembrolizumab and lenvatinib	Advanced endometrial carcinoma that is not MSI-H or dMMR	Previously conducted clinical trials	Isolation of treatment effect
Several immunooncology combination therapies	Untreated, locally advanced or metastatic renal cell carcinoma	Previously conducted clinical trials	Isolation of treatment effect
Blinatumomab	Precursor B-cell ALL in complete remission with detectable MRD	Retrospective observational cohort study	Comparative efficacy analysis

ALL, acute lymphoblastic leukemia; dMMR, mismatch repair deficient; EHR, electronic health record; FGFR, fibroblast growth factor receptor; MRD, minimal residual disease; MSI-H, microsatellite instability-high.

NF1 mutation (without a requirement to have PN-related morbidity at enrollment).[29] The second set of external control data was a placebo arm of Study 01-C-0222, a multicenter, double-blinded, randomized study comparing a different investigational agent to placebo in children and young adults ($\geq$3 and $\leq$25 years) with a clinical diagnosis of NF1 and unresectable, progressive PN with the potential to cause significant morbidity.[30]

The patient populations of these two external control data sources were small and heterogeneous, with dissimilarities to each other as well as the experimental arm of SPRINT Phase II Stratum 1.[27,28] Both studies helped to confirm that the occurrence of spontaneous regression was uncommon, such that observed responses in SPRINT were reasonably deemed the effect of treatment with selumetinib. Additionally, though the outcome data of these studies were provided, there was limited information regarding patient demographic and clinical characteristics, and there was no pre-specified protocol or statistical analysis plan. Thus, no formal statistical comparisons could be made. The external control data were considered supportive in understanding the natural history of NF1 with inoperable PN as it provided important clinical context for the observed response rate.[27,28] The ORR of 66% (with 82% of responses durable for $\geq$12 months), coupled with the associated effects on PN-related morbidities seen in SPRINT Phase II Stratum 1, was of sufficient magnitude and durability to be deemed direct evidence of clinical benefit.

Erdafitinib for patients with unresectable urothelial cancer harboring select FGFR genetic alterations. In April 2019, the FDA granted accelerated approval to erdafitinib for patients with locally advanced or metastatic urothelial carcinoma with susceptible fibroblast growth factor receptor (FGFR) genetic alterations, specifically FGFR2 or FGFR3, and have progression of disease during or following platinum-containing chemotherapy based on trial BLC2001. This trial was a single-arm, multicenter study in patients with metastatic or surgically unresectable urothelial cancer harboring select FGFR genetic alterations (NCT02365597). The primary endpoint was ORR as per blinded independent central review according to RECIST 1.1.[31]

Among patients with relapsed or refractory urothelial cancer, it is estimated that FGFR genetic alterations occur in $\sim$20% of patients. Though the FGFR genetic alterations are a molecularly defined subset of urothelial cancer, data establishing the natural history of disease or tumor response to other available therapy in this selected group of patients are limited. To obtain accelerated approval, erdafitinib needed to demonstrate an improvement over available therapies and it was important to understand how patients with FGFR-altered urothelial cancer responded to other available treatments such as immunotherapy. An RWD-based external control arm was included in the application to establish that patients with FGFR-mutated urothelial cancer do not have the expected tumor response to immunotherapy (an available therapy with survival benefit).[31]

The external control data source was real-world longitudinal patient-level EHR data from US community-based cancer clinics with linage to next-generation sequencing results.[31] The patient population included patients with confirmed diagnosis of advanced bladder cancer with primary sites in the bladder, renal pelvis, ureter, or urethra. The application included an analysis of association between FGFR mutations or fusions and anti-programmed death-ligand 1 (PD-L1)/programmed cell death protein 1 (PD-1) therapy outcomes, as well as comparison of OS among patients with FGFR genetic alterations who received erdafitinib in BLC2001 compared to those treated with available treatment options in the real-world external control arm using propensity score weighting using a model that included observed key demographic and prognostic characteristics.

There were several limitations to these analyses including a small sample size of only 27 patients.[31] In addition, important prognostic factors, such as tumor stage, Eastern Cooperative Oncology Group status, and PD-L1 expression status, had substantial missing data. Also, the patients in the BLC2001 study were subject to more stringent eligibility criteria than the external control arm and thus were likely to be healthier. This difference, along with geographic and point-of-care differences, indicates that the patients from the two cohorts are sampled from different underlying

populations and may not be comparable. There were also significant concerns regarding differential misclassification of treatment assignment or incomplete capture of mortality data.

Overall, due to these limitations, the FDA was unable to make any definitive conclusion on the prognostic or predictive impact of FGFR alterations in advanced urothelial cancer, or use this data to support a claim of comparative treatment efficacy of erdafitinib as compared to treatment with anti-PD-L1/PD-1 therapy.[31] However, since responders in BLC2001 included patients with tumors that had previously not responded to anti-PD-L1/PD-1 therapy and the evidence supports treatment with erdafitinib regardless of prior treatment with these immune therapies, the drug was able to be approved.

External controls for isolation of treatment effect

Pembrolizumab and lenvatinib for patients with endometrial cancer. In September 2019, the FDA granted accelerated approval to pembrolizumab plus lenvatinib for the treatment of patients with advanced endometrial carcinoma that is not microsatellite instability-high (MSI-H) or mismatch repair deficient (dMMR), and who have disease progression following prior systemic therapy but are not candidates for curative surgery or radiation.[32]

The treatment effect of the combination of pembrolizumab and lenvatinib was estimated in Study E7080-A001-111/KEYNOTE-146, a single-arm, multicenter, open-label, multicohort trial which included patients who had previously treated endometrial carcinoma with tumors that were not MSI-H or dMMR (NCT02501096). The major efficacy outcome was ORR as per independent radiologic review committee using RECIST 1.1. Although the response rate and duration of response for the combination therapy were considered clinically meaningful, without contemporaneous controls randomized to each single agent, the contribution of each treatment component to the combination therapy was a key review issue.

Monotherapy data from three previously conducted clinical trials (Study204 for lenvatinib monotherapy and KEYNOTE-158 and KEYNOTE-028 for pembrolizumab monotherapy) were considered external control data to support contribution of treatment effect. Exploratory analyses included unadjusted cross-trial comparisons, which indicated a numerical improvement in ORR with the combination therapy as compared to each of the individual. The FDA also conducted exploratory adjusted analyses using propensity score methods to control for potential differences in baseline demographic and clinical characteristics across trials.[32] While the results of adjusted analyses were consistent with those seen in an unadjusted comparison, the analyses were limited by covariates that were measured in all four studies, and therefore could be subject to residual unmeasured confounding. Nonetheless, these results provided supportive evidence for the supplemental indication approval of the combination therapy.

Combination therapies for patients with previously untreated, locally advanced or metastatic renal cell carcinoma. Between April 2018 and January 2021, the FDA granted approval based on randomized controlled trials for four combination drug regimens in the treatment of patients with previously untreated, locally advanced or metastatic renal cell carcinoma (RCC): nivolumab in combination with ipilimumab (NCT02231749), pembrolizumab in combination with axitinib (NCT02853331), avelumab in combination with axitinib (NCT02684006), and nivolumab in combination with cabozantinib (NCT03141177). The four trial designs were similar with patients randomized to either the combination regimens or sunitinib. Because a factorial design was not used to isolate the effect of each agent in the combination, the FDA relied on supportive data from external clinical trials demonstrating the monotherapy activity of the drugs using objective response rate comparisons. In all cases, the safety profiles were well understood; the drugs had all been approved for other indications, and most of them had been previously approved for RCC. In some cases, *post hoc* propensity score matching was carried out to compare the combination therapies with the monotherapies in the matched patient populations. Ultimately, the results of all combination studies were strong with large magnitudes of benefit with respect to progression-free survival or OS, thus permitting the use of external controls to support the combination approvals.[33]

External controls for supportive comparative efficacy analyses

Blinatumomab for patients with precursor B-cell acute lymphoblastic leukemia in complete remission with detectable minimal residual disease. Blinatumomab was granted accelerated approval for the treatment of adults and children with B-cell precursor (BCP) acute lymphoblastic leukemia (ALL) in first or second complete remission with minimal residual disease (MRD) $\geq$0.1%. Limited data are available from randomized trials of treatment of patients with ALL and detectable MRD, and there are no randomized studies of chemotherapy that assess the long-term outcomes in adult or pediatric patients with BCP ALL who were treated after consolidation for MRD.[34]

The primary study to establish the treatment effect of blinatumomab was the BLAST Study (NCT01207388), a single-arm, multicenter study that included adult patients with BCP ALL in complete remission with MRD at a level of $\geq$0.1%. The primary endpoint was undetectable MRD within one cycle of treatment, also referred to as complete MRD response, which is defined as the absence of detectable MRD confirmed in an assay with a minimum sensitivity of 0.01%. As a supportive analysis, the data from the single-arm trial were directly compared to an external control arm that was derived from Study 2120148, a retrospective cohort study of patients outside the United States with Philadelphia chromosome-negative BCP ALL in hematological complete remission with MRD. The initial selection of a comparable population for the external control and

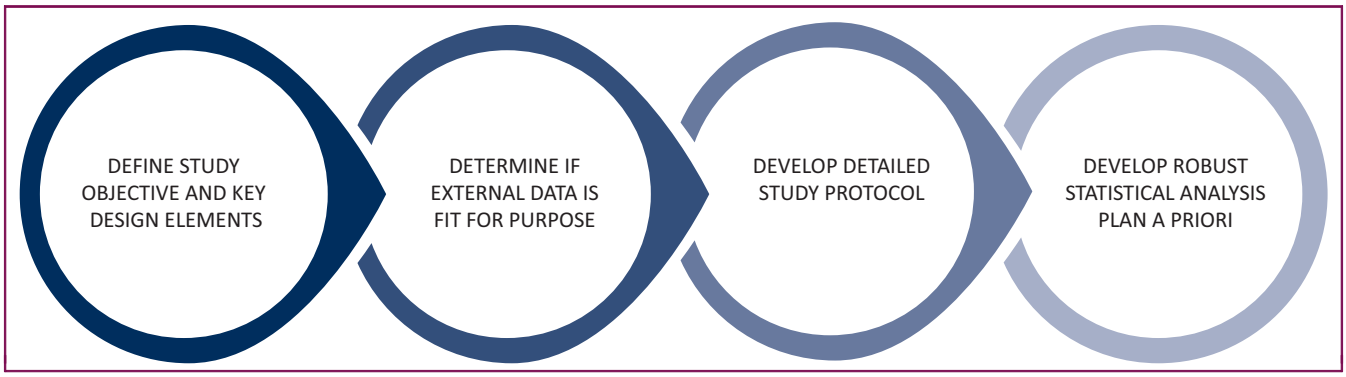

Figure 3. **A general approach to incorporating external control data.**

experimental arms included matching certain baseline patient clinical characteristics across studies, resulting in reduction of sample size in both patient cohorts. To ensure comparative populations, the two populations were matched by time from MRD measurement to start of therapy or relapse, which reduced the eligible population for a comparative analysis in the experimental arm by approximately one-third.[34]

After using stabilized inverse probability of treatment weights and adjustment for hematopoietic stem cell transplantation (HSCT), the two patient cohorts were compared for recurrence-free survival (RFS) and OS. These results were discussed at the March 2018 Oncologic Drugs Advisory Committee.[35] Though the results of the RFS analyses indicate a numerical advantage for blinatumomab, there were several limitations to the interpretation of the data analyses, including different rates of HSCT and subsequent treatment, resulting in residual confounding. Additionally, there were temporal differences in the data for the experimental and control arms, along with differential follow-up by arm. These results were considered exploratory and were considered supportive of the approval in this disease setting.

DISCUSSION AND FUTURE DIRECTIONS

The use of external control arms in oncology is an evolving area of methodological advancement and regulatory interest. Thus far, external control data have been limited to establishing natural history or providing supportive data analyses affording clinical context for an observed treatment effect. These data have not provided pivotal support as substantial evidence of treatment effectiveness in oncology, and therefore have not yet been included in the prescription drug labeling. In the future, it is possible that formal analyses for comparative treatment effect between an investigational therapy and an external control arm would provide the primary evidence of efficacy to support a regulatory approval in oncology, particularly for a disease type with a highly predictable natural history and the precisely measurable treatment effect.

As described in previous sections, the approach to incorporating external control data is a multistep process (see Figure 3), where the first efforts must be to define study purpose and determining whether an external control data source is fit for purpose. This process would include considerations of data quality, data comprehensiveness and completeness, and comparability to a potential experimental arm with respect to characteristics such as (but not limited to) underlying patient population, temporality, and key data ascertainment. Major barriers to establishing an appropriate external control data source that is fit for purpose in oncology often include challenges with non-comparable endpoints, differences in selection by or availability of information regarding biomarkers, and poor capture of information that may be generally prognostic such as standard of care or supportive treatment. Importantly, the granularity of the data with respect to measurement or categorization of important variables is not trivial, and must be similar between experimental arm and external control data to ensure comparative analyses.

Furthermore, the choice of external control comparator must be rigorously supported by the pre-specification of a detailed protocol and robust statistical analysis plan, including a comprehensive plan to address sources of bias. Bias, including confounding, selection bias, and survivor or lead-time bias, cannot be totally eliminated in any non-randomized comparison. However, as previously mentioned, study design elements can help to minimize bias, and statistical methods may address the influence of bias on the estimation of treatment effect. However, if there are major issues with the validity or accuracy of the data, such as inability to establish audit trails or concerns of measurement error and misclassification, or substantial concerns of bias, statistical methods cannot salvage the comparison. Development of additional methods and bias quantification are needed to appropriately characterize use of external control arm as substantial evidence in regulatory applications.

FUNDING

None declared.

DISCLOSURE

The authors have declared no conflicts of interest.

REFERENCES

1. ICH Harmonised Tripartite Guideline: E10 Choice of Control Group and Related Issues in Clinical Trials. May 2001. Available at https://www.fda.gov/media/71349/download.
2. U.S. Food and Drug Administration. FDA Guidance for Industry: Use of Electronic Health Record Data in Clinical Investigation. July 2018. Available at https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM501068.pdf.
3. U.S. Food and Drug Administration. FDA Guidance for Industry: Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drugs and Biologics. May 2019. Available at https://www.fda.gov/media/124795/download.
4. Thorlund K, Dron L, Park JJH, Mills EJ. Synthetic and external controls in clinical trials - a primer for researchers. *Clin Epidemiol*. 2020;12:457-467.
5. Seeger JD, Davis KJ, Iannacone MR, et al. Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiol Drug Saf*. 2020;29(11):1382-1392.
6. U.S. Food and Drug Administration. FDA Guidance for Industry: Rare Diseases: Natural History Studies for Drug Development. March 2019. Available at https://www.fda.gov/media/122425/download.
7. Burcu M, Dreyer NA, Franklin JM, et al. Real-world evidence to support regulatory decision-making for medicines: considerations for external control arms. *Pharmacoepidemiol Drug Saf*. 2020;29:1228-1235.
8. Friends of Cancer Research Whitepaper "Characterizing the use of external controls for augmenting randomized control arms and confirming benefit". Available at https://www.focr.org/sites/default/files/Panel-1_External_Control_Arms2019AM.pdf.
9. Amiri-Kordestani L, Xie D, Tolaney SM, et al. A Food and Drug Administration analysis of survival outcomes comparing the adjuvant paclitaxel and trastuzumab trial with an external control from historical clinical trials. *Ann Oncol*. 2020;31(12):1704-1708.
10. Sun Z, Niman SM, Pagani O, et al. Estimation of historical control rate for a single arm de-escalation study - application to the POSITIVE trial. *Breast*. 2020;53:1-7.
11. Ventz S, Lai A, Cloughesy TF, et al. Design and evaluation of an external control arm using prior clinical trials and real-world data. *Clin Cancer Res*. 2019;25(16):4993-5001.
12. Carrigan G, Whipple S, Capra WB, et al. Using electronic health records to derive control arms. for early phase single-arm lung cancer trials: proof-of-concept in randomized controlled trials. *Clin Pharmacol Ther*. 2020;107(2):369-377.
13. Friends of Cancer Research. 2019 Annual Meeting Panel Presentation. Characterizing the use of external controls for augmenting randomized control arms and confirming benefit. Available at https://friendsofcancerresearch.org/sites/default/files/Panel_1_Slide_Deck_AM19.pdf.
14. U.S. Food and Drug Administration. Framework for FDA's Real-World Evidence Program. December 2018. Available at https://www.fda.gov/media/120060/download.
15. U.S. Food and Drug Administration. FDA Guidance for Industry: Demonstrating Substantial Evidence of Effectiveness for Human Drug and Biological Products. December 2019. Available at https://www.fda.gov/media/133660/download.
16. Webster-Clark M, Jonsson Funk M, Stürmer T. Single-arm trials with external comparators and confounder misclassification: how adjustment can fail. *Med Care*. 2020;58(12):1116-1121.
17. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424.
18. Stuart EA, Lee BK, Leacy FP. Prognostic score—based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *J Clin Epidemiol*. 2013;66(80):S84-S90.
19. Lim J, Walley R, Yuan J, et al. Minimizing patient burden through the use of historical subject-level data in innovative confirmatory clinical trials: review of methods and opportunities. *Ther Innov Regul Sci*. 2018;52:546-559.
20. Hobbs BP, Carlin BP, Sargent DJ. Adaptive adjustment of the randomization ratio using historical control data. *Clin Trials*. 2013;10(3):430-440.
21. Dron L, Golchi S, Hsu G, et al. Minimizing control group allocation in randomized trials using dynamic borrowing of external control data - an application to second line therapy for non-small cell lung cancer. *Contemp Clin Trials Commun*. 2019;16:100446.
22. Ye J, Reaman G, De Claro RA, Sridhara R. A Bayesian approach in design and analysis of pediatric cancer clinical trials. *Pharm Stat*. 2020;19:814-826.
23. Ye J, Travis J. A Bayesian approach to incorporating adult clinical data into pediatric clinical trials. Paper presented at Pediatric Trial Design and Modeling: Moving into the Next Decade, FDA Public Workshop. FDA Public Workshop. 2017; US FDA, White Oak, MD. Available at https://www.fda.gov/drugs/news-events-human-drugs/pediatric-trial-design-and-modeling-moving-next-decade.
24. Ferner RE, Huson SM, Thomas N, et al. Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. *J Med Genet*. 2007;44(2):81-88.
25. Korf BR. Plexiform neurofibromas. *Am J Med Genet*. 1999;89(1):31-37.
26. Mautner VF, Asuagbor FA, Dombi E, et al. Assessment of benign tumor burden by whole-body MRI in patients with neurofibromatosis 1. *Neuro Oncol*. 2008;10(4):593-598.
27. U.S. Food and Drug Administration. FDA "NDA Multi-disciplinary Review and Evaluation of KOSELUGO (selumetinib)". May 2020. Available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/213756Orig1s000MultidisciplineR.pdf.
28. Casey D, Demko S, Sinha A, et al. FDA approval summary: selumetinib for plexiform neurofibroma. *Clin Cancer Res*. 2021;27(15):4142-4146.
29. Gross AM, Singh G, Akshintala S, et al. Association of plexiform neurofibroma volume changes and development of clinical morbidities in neurofibromatosis 1. *Neuro Oncol*. 2018;20(12):1643-1651.
30. Widemann BC, Dombi E, Gillespie A, et al. Phase 2 randomized, flexible crossover, double-blinded, placebo-controlled trial of the farnesyltransferase inhibitor tipifarnib in children and young adults with neurofibromatosis type 1 and progressive plexiform neurofibromas. *Neuro Oncol*. 2014;16(5):707-718.
31. U.S. Food and Drug Administration. FDA "NDA Multi-disciplinary Review and Evaluation of BALVERSA (erdafitinib)". 2019. Available at https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/212018Orig1s000MultidisciplineR.pdf.
32. Arora S, Balasubramaniam S, Zhang W, et al. FDA approval summary: pembrolizumab plus lenvatinib for endometrial carcinoma, a collaborative international review under project orbis. *Clin Cancer Res*. 2020;26(19):5062-5067.
33. Brewer JR, Chang E, Agrawal S, et al. Regulatory considerations for contribution of effect of drugs used in combination regimens: renal cell cancer case studies. *Clin Cancer Res*. 2020;26(24):6406-6411.
34. Jen E, Xu Q, Schetter A, et al. FDA approval: blinatumomab for patients with B-cell precursor acute lymphoblastic leukemia in morphologic remission with minimal residual disease. *Clin Cancer Res*. 2019;25(2):473-477.
35. U.S. Food and Drug Administration. FDA Briefing Information for the March 7, 2018 Meeting of the Oncologic Drugs Advisory Committee. Available at https://www.fda.gov/media/111622/download.