Viewpoint

Data Challenges for Externally Controlled Trials: Viewpoint

Russanthy Ruthiran Velummailum[1*], MSc, MPH; Chelsea McKibbon[1*], MSc; Darren R Brenner[2], PhD; Elizabeth Ann Stringer[3], PhD; Leeland Ekstrom[3], PhD; Louis Dron[1], MSc

[1]Cytel, Inc, Vancouver, BC, Canada

[2]Department of Oncology, University of Calgary, Calgary, AB, Canada

[3]Nashville Biosciences, Nashville, TN, United States

[*]these authors contributed equally

Corresponding Author:
Louis Dron, MSc
Cytel, Inc
777 West Broadway
Suite 802
Vancouver, BC, V5Z1J5
Canada
Phone: 1 604 294 3823
Email: louis.dron@cytel.com

Abstract

The preferred evidence of a large randomized controlled trial is difficult to adopt in scenarios, such as rare conditions or clinical subgroups with high unmet needs, and evidence from external sources, including real-world data, is being increasingly considered by decision makers. Real-world data originate from many sources, and identifying suitable real-world data that can be used to contextualize a single-arm trial, as an external control arm, has several challenges. In this viewpoint article, we provide an overview of the technical challenges raised by regulatory and health reimbursement agencies when evaluating comparative efficacy, such as identification, outcome, and time selection challenges. By breaking down these challenges, we provide practical solutions for researchers to consider through the approaches of detailed planning, collection, and record linkage to analyze external data for comparative efficacy.

(J Med Internet Res 2023;25:e43484) doi: 10.2196/43484

KEYWORDS

external control arm; synthetic control arm; single-arm trial; real-world evidence; regulatory approval; data; clinical; decision-making; efficacy; rare conditions; trial

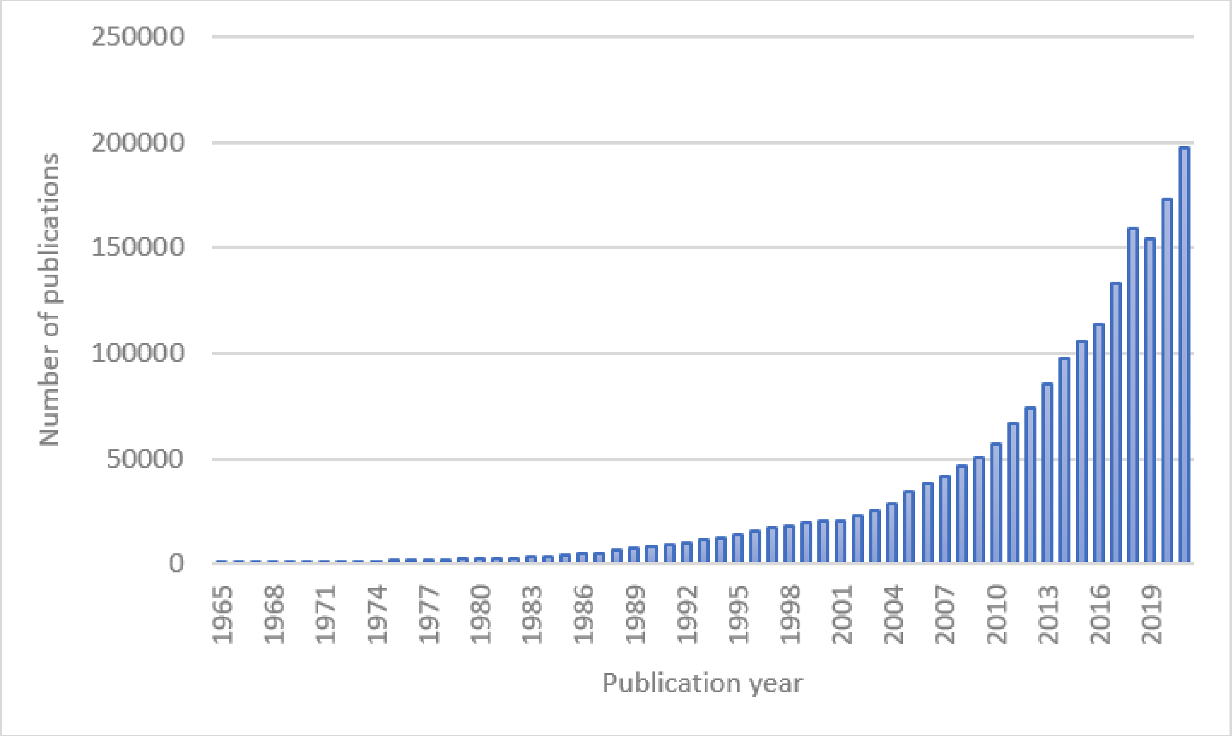
Introduction

Historically, real-world evidence from observational studies has had limited use for demonstrating therapeutic effectiveness for regulatory and reimbursement purposes. However, several recent developments, including the 21st Century Cure Act, increased accessibility to large-scale routinely collected health data, improved standardization of collection, and increased high-profile regulatory applications, and have resulted in an increased demand for these data and increased availability of these data [1,2]. This has changed the data landscape, with growing recognition of the value of using real-world evidence that is applicable for regulatory and reimbursement purposes. Frameworks from regulatory bodies, such as the United States Food and Drug Administration (FDA) and European Medicines Agency (EMA), and reimbursement agencies, such as the National Institute for Health and Care Excellence (NICE), specifically call out the use of external data sources in conjunction with single-arm evidence, particularly in rare diseases or clinical subgroups where there is a high unmet medical need and traditional randomized controlled trials (RCTs) may not be feasible [3-7].

Real-world evidence is derived from rigorous analyses of real-world data. Real-world data are data related to patient health status or delivery of health care outside of RCTs where sources commonly originate from electronic health records (EHRs), medical claims data, and product and disease registries [8]. In addition to what may be considered common forms, real-world data from outside of traditional medical charting, including data from mobile phones, wearables, and patient-reported outcomes, have provided an abundance of data, allowing comprehensive capture of the natural course of a disease from both the physician and patient perspectives [9-11]. Although outside of the scope of this viewpoint, data from historical clinical trials have been found to be influential in comparative efficacy analyses and have been found to have a role in supplementing an external control study through hybrid study designs [12,13].

Given the adoption and use of EHRs and other data content sources in general practice, real-world data and real-world evidence are being increasingly reported in publications in recent years (Figure 1) and are being increasingly used in health care decisions [14,15].

Figure 1. Number of real-world data and real-world evidence publications over time from 1965 to 2021. The following search terms have been considered in PubMed: “real-world” or “observational” or “nonrandomized” or “standard of care” or “external control” or “single-arm” or “historical-control” or “retrospective” or “noninterventional” or “case series” or “natural history” or “electronic health record” or “electronic medical record” or “claims”.



As the majority of criticism from regulatory and reimbursement agencies on applications using real-world data has so far focused predominantly on data features (type, quality, and frequency) and confounding and selection bias limitations, appropriate curation and evaluation of these sources are critical components of any exercise using external evidence [16]. In this paper, we provide an overview of the challenges in identifying data suitable for external control arms, including common pseudonyms, such as synthetic control arms and historical controls, when evaluating comparative efficacy, and provide solutions for researchers to consider.

Technical Challenges

Challenges in Data Source Identification for Rare Conditions

RCTs have long been considered the gold standard approach to evaluate the comparative effectiveness of a drug or biological product due to the standardized methods to reduce bias, including randomization and blinding, as well as balancing known and unknown confounders. These gold standard trials, however, may not reflect the real-world setting where the patient is treated [17]. This is particularly notable for rare diseases, where the standard of care may be highly variable and disease definitions may change rapidly, than for nonrare indications. In the United States, a rare disease is defined as a condition that affects fewer than 200,000 individuals, with most diseases presenting in children or being oncology indications. Advancements in precision medicine have changed the rare disease paradigm, with many common diseases now being considered rare diseases by further splitting into subdiseases or considering stratification by mutated genes in common cancers [18]. Research focused on rare diseases is further challenged due to limited patient recruitment and siloed efforts focused on a singular therapeutic area [19,20]. Although common diseases and cancer indications are not safeguarded from similar complexities, they can often be more acute and rate limiting in the rare disease space. Regardless, both areas have their own unique technical challenges for researchers to broach throughout the life cycle of comparative efficacy analyses.

Computational models from machine learning and artificial intelligence have experienced success in early disease diagnosis and trial recruitment across several indications [21,22]. Challenges in trial recruitment have led researchers to generate computational models for patient identification, but they have faced difficulties due to lengthy diagnosis procedures, multiple physicians, and lack of gold standard–confirmed diagnoses to train models [23]. In scenarios where there is interest in rare diseases with a long disease course, these challenges can often be exacerbated due to the need for sufficient long-term follow-up data. Efforts to increase the use of data sharing in rare diseases, such as the use of FAIR (findable, accessible, interoperable, and reusable) Guiding Principles for implantation networks (eg, Global Open FAIR Implementation Network), aim to produce a conjoined effort to create data sources fit for translational research [24]. In many instances, patient identification within EHRs begins with the use of existing standard terminologies, such as the International Classification of Diseases (ICD) and the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT). For many common comorbidities, these coding strategies have been adopted into general administrative practice but have fallen short when classifying rare diseases. A study evaluating 6519 rare diseases considered the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM), International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM), and SNOMED CT, and found varying levels of coverage matching to a unique rare disease at 62%, 73%, and 85%, respectively [25]. With diseases being mapped to multiple ICD codes, a single representation may not constitute an accurate diagnosis. Increased coverage of rare diseases in coding languages has been acknowledged and incorporated in the International Classification of Diseases, 11th Revision (ICD-11) released in 2022, with 10 times more rare disease classifiers than the previous version [26]. Additional challenges related to executing machine learning and artificial intelligence initiatives arise due to several ethical and legal concerns, including consent for data use, transparency of algorithms, liability, and cyber security [27,28]. Given the adaptive nature, the FDA has released guidance and a proposed regulatory framework, mainly in relation to medical devices and decision support software, to enable development of these technologies while maintaining safety and oversight for transparency and performance [29,30]. Equivalent guidance has yet to be established for similar initiatives involving comparative effectiveness studies using real-world data. Similarly, advanced procedures for the diagnosis of rare diseases, such as whole genome sequencing (WGS) and advanced imaging, may not be sufficiently captured within EHRs. Initiatives surrounding WGS in particular are being developed to improve the storage, knowledge, and presentation of genomic information within EHR systems to increase use [31]. Despite the increase in coverage, the implementation of these coding strategies, conversion of previous ICD codes within EHR systems, and incorporation of advanced technologies will add an additional level of complexity for EHR-based studies.

Outside of traditional rare diseases, adoption of real-world data for comparative efficacy remains underutilized in certain categories, such as the incorporation of medications or surgical procedures in clinical trial conduct [32]. Similarly, the ability to replicate clinical trials based on real-world data alone continues to be an area to improve, as many replication failures can be attributed to analyzing endpoints not typically found in real-world data, requiring data that are unlikely to appear in a structured form, lack of complete medical history, or difficulties in closely emulating a placebo [33,34]. From herein, we focus on the following common challenges: outcome and covariate challenges, follow-up, time selection, and geography.

Outcome and Covariate Challenges

In a traditional clinical trial, there are frameworks present for outcome ascertainment and frequency through defined protocols for the monitoring and timing of key observations to evaluate clinical effectiveness. Given the structured nature of data from RCTs, analyses using these are seldom hampered by high levels of missingness or stochastic outcome measurements. In the context of real-world data and external control arms, a common challenge encountered is addressing the mindset of researchers to be open to the value of real-world data and the ability to develop suitable outcome proxies due to insufficient data availability.

EHR systems were not designed for use in the same structured trial environment for clinical effectiveness studies; therefore, additional steps are required to refine data sets for use. EHR data include routinely collected measurements from general practice, which may not exactly match defined clinical endpoints of interest or disease definitions as encountered in clinical trials. Accordingly, where analyses attempt to align these data for comparative efficacy, issues may arise. In these cases, plausible proxies need to be made. For example, a trial of interest in a hepatology setting might include an outcome for “ascites requiring treatment;” however, in EHR data sets, “ascites” might be captured, but the condition with the additional criteria of requiring treatment would not be preassembled and indexed. Indeed, it would be possible to generate the treatments for ascites and tie these back to an individual patient, but this may be a manual process. As such, estimating the total number of eligible patients with the outcome of interest may be subject to substantial manual data curation. Specifically, in the context of using EHRs for comparisons to data obtained from a prospective clinical trial, there may be an absence of overlapping outcomes. Clinical trials (randomized or single-arm) have been found to often capture data on outcomes not recorded routinely in clinical practice [34].

Similarly, the translation of many gold standard trial endpoints to real-world data initiatives may not always be a straightforward endeavor. These efforts require a series of methodological considerations during the planning, abstraction, and analysis phases to ensure regulatory-grade fit-for-purpose data [35]. Even where data may be recorded and decisions made consistent with clinical trials, there may be differences with respect to the timing and process of testing, which may influence the availability of testing. In clinical trials of solid tumors, the Response Evaluation Criteria in Solid Tumors (RECIST) are often used to define endpoints of progression-free survival (PFS) or objective response. However, the ability to operationalize the RECIST for retrospective analysis can be hindered due to lack of consistency in imaging reports in community practices [36]. The RECIST are partially qualitative measurements, and physicians or patients in clinical practice are not blinded to the treatment, providing another difference from a clinical trial environment. Where clinical trials may make assessments for progression on a scheduled basis, real-world settings often only do so when there is an indication of progression, generally resulting in a misalignment of timing when attempts are made to draw comparisons. The real-world equivalents of many clinical trial endpoints, such as PFS and objective response rate, may not be generated using the same level of standardization or may result in variations in assessment criteria between diagnosing physicians in comparison to when performed within a clinical trial setting. For indications with good survival, such as early stage breast cancer, surrogate endpoints have been influential in allowing for reduced development time, smaller patient populations, and shorter follow-up time [37]. Surrogate endpoints, such as pathologic complete response, would require the same level of attention as long-term outcomes to adequately define a real-world equivalent. As many surrogate endpoints for topics, such as oncology, remain controversial due to questions over their direct correlation to assess patient survival, their applications from real-world data are subject to similar criticism [38]. Real-world data–based surrogate endpoints require thorough validation and can often involve timely procedures to abstract adequate data for ascertainment [39].

Other endpoints of interest, such as overall survival, may be clearly definable within a given data set, but are subject to limitations such as use restriction, delayed data availability, and missingness [40]. For example, in acute lymphoblastic leukemia, the concept of “fit-for-use” EHR data was evaluated by assessing the data suitability prior to examination of any survival-based outcomes [41]. Using a defined data quality assessment framework, data variables, such as diagnoses and demographics, were extracted and defined from EHRs over 70% of the time, but the approach fell short when categorizing laboratory values and death data, resulting in a high degree of missingness for these variables [41]. As such, when considering data missingness, it is important not to simply restrict the evaluation to availability versus nonavailability of data, but also consider the frequency of assessments and overall data suitability. While much of the current literature has focused on oncological examples and survival-based outcomes, similar considerations can be granted to other outcomes, such as adverse events, and assessment of the balance between efficacy and clinical benefit.

Follow-up

Real-world studies provide an opportunity to offer additional insights into the patient journey by providing a reflection of traditional care. In an RCT setting, long-term follow-up can be constrained by financial hurdles and trial logistics. To mitigate this, randomized trials have been extended by incorporating long-term follow-ups augmented using routinely collected health care data to investigate long-term outcomes [42]. In these contexts, there may be somewhat broader observation windows than in traditional RCTs, and there are still predefined study visits, follow-up criteria, and study investigators ensuring consistency of follow-up. In contrast, patients in general practice demonstrate higher variability with respect to their follow-up times and their frequency of follow-up visits. Patients’ frequency of visits may be heavily correlated with the severity of their condition or diagnosis, resulting in potential biases for follow-up times. Specifically considering this in the context of time-to-event outcomes, imbalances in follow-up frequency can result in biased estimations of relative treatment effects [43]. Empirically, this has been demonstrated with observational data, and showing the level of follow-up completeness has been influential in the accuracy of survival estimates [44]. In this example, simple reporting standardization of study follow-up, including in general practice, was found to be influential in outcome assessments, with a lack of systematic follow-up resulting in an underestimation of mortality [44]. Similarly, the frequency of assessment scans has been shown to be associated with higher median PFS for both treated and untreated populations [45]. As such, where assessments of PFS are made using real-world data with low or high frequency of assessments relative to the comparator data, the potential for bias may be high. In addition to visit frequency, the United States has additional hurdles contributing to the lack of sufficient follow-up data due to clinicians experiencing a higher administrative burden when using EHR systems in comparison to other countries, discontinuity resulting from out-of-system care, and a direct relationship with enrollment in health care insurance [46-48]. The value of incorporating real-world data into a framework to be considered as evidence for future extension studies has been recognized and endorsed by agencies such as the International Society for Pharmacoepidemiology [49].

Time Selection Challenges

Unlike in randomized controlled comparative efficacy studies, the definition of start date for measuring patient outcomes can be challenging in routinely collected health data. The concept of immortal time bias originated in the 1970s and is termed to account for the period of time within an observational or follow-up period of a cohort where the study outcome of interest cannot occur [50]. When unaccounted for, EHR studies can be prone to immortal time bias, and this phenomenon has been demonstrated in observational studies of long-term conditions [51]. For example, clinical trials may often be established to evaluate a therapy within a population having an existing therapeutic history, which is often referred to as a line of therapy. There may be limits with respect to the total lines of therapy, but they are otherwise permissive with respect to treatment lines or may even use treatment lines as a stratification factor. In routinely collected health data, this presents a challenge for patients with multiple lines of therapy, who are “multiply eligible” for the trial of interest. For a patient who has received 4 lines of therapy, which line of therapy should be considered their index date? Patients could be selected according to their most recent line of therapy or oldest line of therapy, or there could be a random selection of the therapy line. Each of these may have an influence on the associated outcomes, particularly those that may have a time-varying frequency. Hernan and Robins proposed several approaches within the context of target trial emulation. It has been proposed to use a single time zero, a randomly selected time zero, or all available time zeros for a given patient [52]. Within the context of a comparative efficacy analysis using external trial data, these approaches may not be as interchangeable. Indeed, simulation-based research in an oncology context identified that the use of either random line selection or the last line of therapy was subject to substantial inflation in type I error when compared to the use of all available lines of therapy [53]. As such, careful selection of the index date should be applied as decisions may substantially alter the associated inferences available.

Related to these issues are those representing general temporal biases associated with discordant timeframes of interest. While statistical methodologies may be used to minimize observable between-group differences, other aspects of care may vary over time in ways that cannot be accounted for in such a direct manner. For example, even where diagnostic criteria for a condition do not vary over the time period of interest, access to the requisite diagnostic technology may increase over time, leading to increased disease incidence outside of any other changes to practice. These patients may differ from patients identified at different times with respect to both measurable and unmeasurable characteristics, which may contribute to temporal bias [54]. This is separate from more clear-cut temporal bias, which may occur when diagnostic criteria change for patients, standards of care vary over time, or important prognostic factors of significance are identified, and it would need to be adequately accounted for. Finally, these changes in practice over time may result in essential covariates or populations of interest being simply unavailable owing to evolution of the standard of care.

Indeed, concerns regarding temporal bias have formed the basis of negative reception in regulatory submissions using external data for the FDA, the EMA, and Health Canada [55,56]. Statistical approaches exist to identify time-varying confounding [54], though these are predominantly for exploring the existence of such confounding, rather than to minimize their impact. As such, where possible, it is important to achieve close alignment with respect to the timeline of interest, and where this is not possible, it is important to identify covariates that may influence the differences in dates.

Geographical Context

Conditional on the circumstances associated with the application of an external control, geographic representation may be an important characteristic to consider. Regulatory and reimbursement bodies have often identified geographic discordance as part of the rationale behind negative endorsement of candidate products, particularly where these geographic differences are likely to translate to differences in prognostic factors, confounding, effect modifiers, and standards of care [57]. Management of these issues can be minimized through ensuring specificity of the geographic source of the data of interest to ensure consistency with the target trial of interest. This may not be possible in all instances. Real-world data may be abundant in high-income countries (eg, countries in North America and Europe) but can be significantly lacking in many low- and middle-income countries due to lack of EHRs, agreement between stakeholders, and regulation to support the generation and use of secondary data [58]. The commercialization of data from many high-income countries has made identification more attainable, but given that there is no organization dedicated to the tracking of sources in low- and middle-income countries, barriers remain for researchers to identify these sources with ease [59]. Further compounding these issues is the substantial variations with respect to the availability of data in certain geographies owing to the varying legal frameworks associated with the use of patient data and differences in language. Efforts to address the ability to generalize study inferences outside of the intended study population of a geography to another is an emerging topic of interest known as transportability, and it has shown success under special circumstances [60,61]. Despite advancement, geography remains a concern for researchers to contend with when wanting to have sufficient population coverage across high- and low-income countries.

Solutions: How Can Challenges Be Minimized?

Solution 1: Transparent Prespecified Description of Data Element Definitions and a Detailed Data Analysis Plan

The abundance of real-world data certainly presents a vast set of challenges, including “data dredging” where the hoped-for result may stem from ad hoc data mining or from selection of one of the many data sources with limited characteristics for adjustment [62-65]. With the large quantities of available real-world data across the United States and Europe, data sets can conceivably be stratified and matched in ways that could provide favorable results in an opaque manner. To mitigate this, solutions have been presented by regulatory bodies and academic groups for improving transparency for comparative efficacy exercises using real-world data. FDA guidance specifically calls out that study design elements, including data source definitions of all data elements and analyses, should be prespecified prior to analysis, and encourages groups submitting to partake in preanalysis discussions early in the drug development program about whether conducting an externally controlled trial instead of an RCT is reasonable [7]. Academic groups, such as the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) task force, also have guidance on good practices for real-world data studies of treatment effectiveness [66]. The provided advice is to have a transparent and systematic review process of combing through the data source availability and matching to trial characteristics.

The intended steps of using real-world data to evaluate clinical efficacy should be discussed or described in a protocol and statistical analysis plan (SAP) with the associated regulatory or reimbursement body ahead of initiating the externally controlled trial, in an effort to not “cherry pick” the results. Here, sponsors should include a justification for selecting or excluding relevant data sources and should demonstrate that the choice of the final data source for the control arm best answers the research question of interest [7]. Further, it would be imperative to describe how the relevant EHR data were extracted and imported into the sponsor’s electronic system, and how the data obtained from EHRs are consistent with the data collection specified in the clinical trial protocol.

FDA guidance specifies that the protocol and SAP that will be submitted following initial discussions should describe the data provenance curation and transformation procedures in the final study-specific analytic data set and describe how processes could affect data integrity (consistency of data and completeness) and overall study validity [3]. Given the origin of real-world data, SAPs should also specify all proposed sensitivity analyses or quantitative bias analyses as suggested by FDA guidance documentation to address the influence of outcome misclassification and unmeasured confounders in order to ensure appropriate conclusions are drawn [3,7]. Taken together, prespecifying data source collection methods, data element availability, and data integrity to regulatory authorities before conducting any analyses can provide transparency regarding the upcoming efficacy analyses and provide a solution to data mining.

One approach to doing this is to define the patient population for the external control arm, specify the outcome of interest, identify prognostic factors associated with the outcome of interest, and specify the control therapy. During the data source identification phase of the study, it may be common to undertake a scoping exercise to understand the breadth of data available and decipher top-level patient counts. From these selection criteria, data sets that do not have available data for elements of interest can be excluded. The derived top-level counts will reduce as further matching with trial selection criteria is applied. It is an iterative process to assess the best data source that will have sufficient patient counts and covariates to confirm a match, by assessing the strengths and limitations of each of these data sources simultaneously. Since the identification of a suitable real-world data database is an iterative process when proposing the external control arm to regulators at the time of filing for regulatory discussions, a prespecified description of the database is suggested to inform regulators, but a major limitation is that some data element criteria might be unknown until further exploration [7]. Among included data sets, assessments may be made for sample size estimates of the closest resemblance to the effective population and data missingness. The data sets with the highest effective sample size and better coverage of data elements are considered for a feasibility assessment with a deeper look at the data. Through tokenization or other methodologies, multiple data sources could also be used to get more complete patient health care data, follow-up information, and geographic coverage.

Solution 2: Data Collection Leveraging Real-World Data

Situations where existing data are not available or suitable for an indication, a necessary exposure, an outcome, or a key covariate to measure confounding from sources, can pose a major limitation to conducting comparative efficacy studies to support regulatory decisions. As other researchers have stated, a common external control arm critique is about the mitigation of confounding [16]. Unfortunately, there are limits to ascertain real-world data proxies for variables in historically collected databases. This could also be problematic when a group may wish to minimize temporal bias due to changes in the standard of care when using older data, by restricting the dates of eligible patients. This in turn could result in low sample sizes of well-matched control patients.

A possible solution for these issues is to conduct de novo retrospective data collection where there would be manual review of patient charts, in conjunction with pre-existing data. Here, clinicians and key experts for the indication create a customized data collection form (also known as an electronic case record form) that can be standardized across multiple sites and gather useful details. In these scenarios, assessments can be obtained and linked directly to events of interest, such as a new medication or an annual physical examination. As EHR databases can provide access to longitudinal data, further evaluation of measures both before and after the event window can be performed. Drawing upon these aspects further emphasizes the value of real-world data to contribute to both predictive analytics and assessment of long-term outcomes in comparative effectiveness analyses.

De novo data capture will not fix challenges in relation to the frequency of visits or collection of variables known to exist exclusively within trial settings, which is a limitation of this solution. However, qualitative assessments and clinical narratives can contain valuable insights absent from the structured data fields of EHRs. Improvements in advanced analytics and natural language processing have provided increased automation to abstract valuable clinical information from unstructured fields, which is traditionally time consuming. A direct comparison between machine learning review and chart-based review in the diagnosis of rheumatoid arthritis demonstrated the ability of the algorithm to identify patients with a strong overlap in disease classification criteria and baseline disease characteristics [67]. On the contrary, missing data and poor data quality can introduce bias in automated methods, and variability in models can limit comparisons [68,69]. Commercial data providers have realized the potential of custom abstraction projects incorporating both manual and automated abstraction to increase the availability of data outside of structured fields, as well as algorithms to calculate real-world endpoints in product offerings.

Solution 3: Record Linkage/Tokenization

The ability to characterize the entire patient journey is critical to perform clinical effectiveness analyses using real-world data, though a complete picture of a patient's medical health is often not available within a single data source. Patients who move in and out of health care networks could have their data omitted from analyses of single data sets. If this movement is related to the patient's standard of care or access to treatment, the process can result in unintended biases in an analysis that is unable to account for these patient movements and missing data. A solution to the fragmentation of the patient journey and perspective could be to engage with data providers who tokenize their data set. In tokenized data sets, patient "tokens" or unique identifiers are assigned to link a patient within a given data source to assist identification in other separate real-world data sources, avoiding duplication while protecting patient privacy at the same time [70]. Tokenization and record linkage can aid in this endeavor and have been recognized by regulators through inclusion in draft guidelines to provide considerations when performing data linkage to ensure a comprehensive data set, quantification of errors, and resolution of discrepancies [3]. Similarly, initiatives have been developed to combine machine learning methodologies and tokenization to expand upon use for EHR analysis, while drawing attention to areas needing improvement in patient phenotyping or patient identification for research purposes [71].

Examples of the role of tokenization have been identified for broadening data sets to improve associations of testing with outcomes in diseases such as COVID-19 [72,73]. While these approaches may be applicable in select cases, it is important to note that not all data sources are capable of linking data due to privacy concerns regarding indirect personal data and data ownership, and additional challenges due to differences in language. These themes highlight the potential limitations that can arise when considering linkage and tokenization to aid in downstream analyses. In turn, this necessitates a flexible approach to data identification and consideration of which approaches of further data acquisition are available for a given project.

Summary of Real-World Data Identification Challenges and Case Studies

Every scenario in identifying a suitable real-world data source as an external control arm for comparative efficacy is unique. The solution of tokenization, for example, could work in some instances, but might not be appropriate in cases where complete coverage of data is available, though the definitions in the data set are poorly defined in the external data source. Table 1 lists the challenges that frequently arise in these data landscaping exercises and summarizes some examples of solutions or case studies for further reading.

Table 1. Summary of real-world data identification for comparative efficacy using externally controlled trial challenges, examples, and application of solutions.

Challenges	Examples	Application of solutions
Data source identification of rare conditions		
Indecision gaps due to abundance of real-world data	It is difficult to parse out important data sources, rare disease candidates, and data linkage options.	Machine learning applications can improve accuracy and quality (type and frequency) in data source selection and patient selection [23,67,71]
Outcome and covariate		
Poorly defined variables or inconsistent definitions from clinical trial to real-world data for limited comparability of real-world data	The conceptual definition of a data element does not align with the operational definition.	• De novo data collection [74] • Automated electronic health record (EHR) abstraction [69] • Characterization of real-world variables or surrogate endpoints [75] • Prespecify sensitivity analyses, including quantitative bias analyses [76], in the statistical analysis plan
Medical claims data might have limited use to support regulatory-grade decision-making	Claims data have limited clinical outcome data.	Combine with EHRs to expand the applicability, coverage, and depth of data [77]
Follow-up		
Difficult to capture continuity of care in a single data source	Diagnosis is spread across multiple physicians; if the patient moves and seeks care outside of the care network, follow-up data will be lost.	• Tokenization/data linkage and advanced analytics with EHR data for capturing a more complete patient journey (particularly helpful for rare conditions where the sample size would be low) [23,44,71] • Analytical approaches (ie, imputation) for missing data [23,44,71,78]
Time selection		
Timing of therapy	Patient has multiple lines of treatment; what should be considered the index date?	Define a proper index date or “time zero” following the target trial emulation framework [52]
Timing of data collection – inconsistent standard of care over time	Data may be present, but are not current enough to provide a reasonable comparison to the current standard of care.	• De novo data collection [55] • Tokenization/data linkage [78]
Geography		
External control arm nongeneralizable to clinical practice	Geographic representation where the main external control arm data source is from outside of the country of interest. Select two unlinked data sources with available data to obtain a sufficient sample size. However, it is unclear if patients overlap in care networks.	• Tokenization/data linkage, which improves patient counts with geographic representation while accounting for duplicates [79] • Transportability [60]
Analysis phase		
Data loss or insufficient sample size to detect power	In the analysis phase, during matching, the power to detect an effect is reduced.	• De novo data collection [55] • Tokenization/data linkage [23,44] • Analytical approaches (ie, imputation) for missing data [78]
Avoid the appearance of the analysis as post-hoc or cherry picked	Data dredging/post-hoc analysis (eg, regulators can assume the most appealing analysis was conducted).	Transparent prespecified description of data selection, data provenance, and the statistical analysis plan [3]

Conclusion

Given the considerable costs of novel drug development pipelines and the increasing stratification of diseases and subtypes, real-world data will become increasingly relevant for regulatory and reimbursement discussions in the decades to come. Exclusive reliance on an RCT framework as the entire evidence generation plan does not adequately acknowledge the shortcomings of a singular research strategy, despite the advantages for comparative efficacy research. Simultaneously, the field of biostatistics has expanded to focus on data missingness to allow effectiveness analyses when a subject presents with incomplete data, while taking the necessary precautions and accounting for bias.

The versatility of the use of real-world data extends beyond its use in comparative efficacy analyses. There are additional concerns with safety, which have postmarket authorization requirement differences. Similarly, important topics outside the scope of this paper include the many challenges associated with patient privacy and reidentification risk, and the necessary consideration needed when performing these analyses. EHR-based real-world data were the focus of this paper, but additional sources have been successful in comparative efficacy analyses. Claims, like other sources of real-world data, have their own unique challenges for consideration [80], but have demonstrated success in clinical effectiveness studies and in rare disease studies [77,81,82]. Conjoined efforts, such as those between the ISPOR and the International Society for Pharmaceutical Engineering, have provided recommendations to highlight the need for transparency in planning and reporting of observational real-world evidence studies and comparative effectiveness studies [66]. Although real-world evidence based on real-world data studies is critical to the operation of providing timely insights into what works for patients and when, the identification and evaluation of real-world data sources for comparative effectiveness studies have many challenges. The solutions suggested in this paper could minimize these challenges; however, the selection and evaluation of a good real-world data source is not as straightforward as it may appear.

Conflicts of Interest

RRV, CM, and LD work for Cytel Inc (paid employment). RRV and LD have shares of Cytel Inc. EAS and LE work for Nashville Biosciences (paid employment). DRB has no conflicts to declare.

References

1. Mahendraratnam N, Mercon K, Gill M, Benzing L, McClellan MB. Understanding Use of Real-World Data and Real-World Evidence to Support Regulatory Decisions on Medical Product Effectiveness. Clin Pharmacol Ther 2022 Jan;111(1):150-154. [doi: 10.1002/cpt.2272] [Medline: 33891318]
2. Gabay M. 21st Century Cures Act. Hosp Pharm 2017 Apr;52(4):264-265 [FREE Full text] [doi: 10.1310/hpj5204-264] [Medline: 28515504]
3. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products - Draft Guidance for Industry. Food and Drug Administration. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory [accessed 2023-02-27]
4. CHMP Guideline on registry-based studies. European Medicines Agency. URL: https://www.encepp.eu/publications/documents/3_KellyPlueschke_CHMPGuidelineonregistry-basedstudies.pdf [accessed 2023-02-27]
5. NICE real-world evidence framework. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/corporate/ecd9/chapter/overview [accessed 2023-02-27]
6. E10 Choice of Control Group and Related Issues in Clinical Trials. ICH Expert Working Group. 2000 Jul 20. URL: https://database.ich.org/sites/default/files/E10_Guideline.pdf [accessed 2023-02-27]
7. Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products. Food and Drug Administration. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-design-and-conduct-externally-controlled-trials-drug-and-biological-products [accessed 2023-02-27]
8. Use of Electronic Health Record Data in Clinical Investigations Guidance for Industry. Food and Drug Administration. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/use-electronic-health-record-data-clinical-investigations-guidance-industry [accessed 2023-02-27]
9. Maruszczyk K, Aiyegbusi O, Cardoso V, Gkoutos G, Slater L, Collis P, et al. Implementation of patient-reported outcome measures in real-world evidence studies: Analysis of ClinicalTrials.gov records (1999-2021). Contemp Clin Trials 2022 Sep;120:106882 [FREE Full text] [doi: 10.1016/j.cct.2022.106882] [Medline: 35973663]
10. Huhn S, Matzke I, Koch M, Gunga H, Maggioni MA, Sié A, et al. Using wearable devices to generate real-world, individual-level data in rural, low-resource contexts in Burkina Faso, Africa: A case study. Front Public Health 2022 Sep 30;10:972177 [FREE Full text] [doi: 10.3389/fpubh.2022.972177] [Medline: 36249225]
11. Liu J, Barrett J, Leonardi E, Lee L, Roychoudhury S, Chen Y, et al. Natural History and Real-World Data in Rare Diseases: Applications, Limitations, and Future Perspectives. J Clin Pharmacol 2022 Dec;62 Suppl 2:S38-S55 [FREE Full text] [doi: 10.1002/jcph.2134] [Medline: 36461748]
12. Baumfeld Andre E, Reynolds R, Caubel P, Azoulay L, Dreyer NA. Trial designs using real-world data: The changing landscape of the regulatory approval process. Pharmacoepidemiol Drug Saf 2020 Oct;29(10):1201-1212 [FREE Full text] [doi: 10.1002/pds.4932] [Medline: 31823482]
13. Ghadessi M, Tang R, Zhou J, Liu R, Wang C, Toyoizumi K, et al. A roadmap to using historical controls in clinical trials - by Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG). Orphanet J Rare Dis 2020 Mar 12;15(1):69 [FREE Full text] [doi: 10.1186/s13023-020-1332-x] [Medline: 32164754]
14. Evans RS. Electronic Health Records: Then, Now, and in the Future. Yearb Med Inform 2016 May 20;Suppl 1(Suppl 1):S48-S61 [FREE Full text] [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]

15. Rahal RM, Mercer J, Kuziemsky C, Yaya S. Factors affecting the mature use of electronic medical records by primary care physicians: a systematic review. BMC Med Inform Decis Mak 2021 Feb 19;21(1):67 [FREE Full text] [doi: 10.1186/s12911-021-01434-9] [Medline: 33607986]
16. Jaksa A, Louder A, Maksymiuk C, Vondeling GT, Martin L, Gatto N, et al. A Comparison of Seven Oncology External Control Arm Case Studies: Critiques From Regulatory and Health Technology Assessment Agencies. Value Health 2022 Dec;25(12):1967-1976 [FREE Full text] [doi: 10.1016/j.jval.2022.05.016] [Medline: 35760714]
17. Jones R, Jones RO, McCowan C, Montgomery AA, Fahey T. The external validity of published randomized controlled trials in primary care. BMC Fam Pract 2009 Jan 19;10:5 [FREE Full text] [doi: 10.1186/1471-2296-10-5] [Medline: 19152681]
18. Chung BHY, Chau JFT, Wong GK. Rare versus common diseases: a false dichotomy in precision medicine. NPJ Genom Med 2021 Feb 24;6(1):19 [FREE Full text] [doi: 10.1038/s41525-021-00176-x] [Medline: 33627657]
19. Halley MC, Smith HS, Ashley EA, Goldenberg AJ, Tabor HK. A call for an integrated approach to improve efficiency, equity and sustainability in rare disease research in the United States. Nat Genet 2022 Mar 07;54(3):219-222 [FREE Full text] [doi: 10.1038/s41588-022-01027-w] [Medline: 35256804]
20. Kempf L, Goldsmith J, Temple R. Challenges of developing and conducting clinical trials in rare disorders. Am J Med Genet A 2018 Apr;176(4):773-783 [FREE Full text] [doi: 10.1002/ajmg.a.38413] [Medline: 28815894]
21. Chen Z, Liu X, Hogan W, Shenkman E, Bian J. Applications of artificial intelligence in drug development using real-world data. Drug Discov Today 2021 May;26(5):1256-1264 [FREE Full text] [doi: 10.1016/j.drudis.2020.12.013] [Medline: 33358699]
22. Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput 2022 Jan 13:1-28 [FREE Full text] [doi: 10.1007/s12652-021-03612-z] [Medline: 35039756]
23. Colbaugh R, Glass K, Rudolf C, Tremblay Volv Global Lausanne Switzerland M. Learning to Identify Rare Disease Patients from Electronic Health Records. AMIA Annu Symp Proc 2018;2018:340-347 [FREE Full text] [Medline: 30815073]
24. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3(1):160018 [FREE Full text] [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
25. Fung KW, Richesson R, Bodenreider O. Coverage of rare disease names in standard terminologies and implications for patients, providers, and research. AMIA Annu Symp Proc 2014;2014:564-572 [FREE Full text] [Medline: 25954361]
26. Aymé S, Bellet B, Rath A. Rare diseases in ICD11: making rare diseases visible in health information systems through appropriate coding. Orphanet J Rare Dis 2015 Mar 26;10:35 [FREE Full text] [doi: 10.1186/s13023-015-0251-8] [Medline: 25887186]
27. Gerke S, Minssen T, Cohen G. Chapter 12 - Ethical and legal challenges of artificial intelligence-driven healthcare. In: Bohr A, Memarzadeh K, editors. Artificial Intelligence in Healthcare. Cambridge, MA: Academic Press; 2020:295-336.
28. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36 [FREE Full text] [doi: 10.1038/s41591-018-0307-0] [Medline: 30617336]
29. Clinical Decision Support Software - Guidance for Industry and Food and Drug Administration Staff. Food and Drug Administration. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software [accessed 2023-02-27]
30. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback. Food and Drug Administration. URL: https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf [accessed 2023-02-27]
31. Williams MS, Taylor CO, Walton NA, Goehringer SR, Aronson S, Freimuth RR, et al. Genomic Information for Clinicians in the Electronic Health Record: Lessons Learned From the Clinical Genome Resource Project and the Electronic Medical Records and Genomics Network. Front Genet 2019 Oct 29;10:1059 [FREE Full text] [doi: 10.3389/fgene.2019.01059] [Medline: 31737042]
32. Rogers JR, Lee J, Zhou Z, Cheung YK, Hripcsak G, Weng C. Contemporary use of real-world data for clinical trial conduct in the United States: a scoping review. J Am Med Inform Assoc 2021 Jan 15;28(1):144-154 [FREE Full text] [doi: 10.1093/jamia/ocaa224] [Medline: 33164065]
33. Bartlett VL, Dhruva SS, Shah ND, Ryan P, Ross JS. Feasibility of Using Real-World Data to Replicate Clinical Trial Evidence. JAMA Netw Open 2019 Oct 02;2(10):e1912869 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.12869] [Medline: 31596493]
34. Zarbin M. Real Life Outcomes vs. Clinical Trial Results. J Ophthalmic Vis Res 2019;14(1):88-92 [FREE Full text] [doi: 10.4103/jovr.jovr_279_18] [Medline: 30820292]
35. Miksad RA, Abernethy AP. Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. Clin Pharmacol Ther 2018 Feb;103(2):202-205 [FREE Full text] [doi: 10.1002/cpt.946] [Medline: 29214638]
36. Griffith SD, Tucker M, Bowser B, Calkins G, Chang CJ, Guardino E, et al. Generating Real-World Tumor Burden Endpoints from Electronic Health Record Data: Comparison of RECIST, Radiology-Anchored, and Clinician-Anchored Approaches

for Abstracting Real-World Progression in Non-Small Cell Lung Cancer. Adv Ther 2019 Aug;36(8):2122-2136 [FREE Full text] [doi: 10.1007/s12325-019-00970-1] [Medline: 31140124]

37. Gion M, Pérez-García JM, Llombart-Cussac A, Sampayo-Cordero M, Cortés J, Malfettone A. Surrogate endpoints for early-stage breast cancer: a review of the state of the art, controversies, and future prospects. Ther Adv Med Oncol 2021;13:17588359211059587 [FREE Full text] [doi: 10.1177/17588359211059587] [Medline: 34868353]
38. Korn EL, Sachs MC, McShane LM. Statistical controversies in clinical research: assessing pathologic complete response as a trial-level surrogate end point for early-stage breast cancer. Ann Oncol 2016 Jan;27(1):10-15 [FREE Full text] [doi: 10.1093/annonc/mdv507] [Medline: 26489443]
39. Kehl KL, Riely GJ, Lepisto EM, Lavery JA, Warner JL, LeNoue-Newton ML, American Association of Cancer Research (AACR) Project Genomics Evidence Neoplasia Information Exchange (GENIE) Consortium. Correlation Between Surrogate End Points and Overall Survival in a Multi-institutional Clinicogenomic Cohort of Patients With Non-Small Cell Lung or Colorectal Cancer. JAMA Netw Open 2021 Jul 01;4(7):e2117547 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.17547] [Medline: 34309669]
40. Zhang Q, Gossai A, Monroe S, Nussbaum NC, Parrinello CM. Validation analysis of a composite real-world mortality endpoint for patients with cancer in the United States. Health Serv Res 2021 Dec;56(6):1281-1287 [FREE Full text] [doi: 10.1111/1475-6773.13669] [Medline: 33998685]
41. Ngo V, Keegan TH, Jonas BA, Hogarth M, Kim KK. Assessing the Quality of Electronic Data for 'Fit-for-Purpose' by Utilizing Data Profiling Techniques Prior to Conducting a Survival Analysis for Adults with Acute Lymphoblastic Leukemia. AMIA Annu Symp Proc 2020;2020:915-924 [FREE Full text] [Medline: 33936467]
42. Katkade VB, Sanders KN, Zou KH. Real world data: an opportunity to supplement existing evidence for the use of long-established medicines in health care decision making. J Multidiscip Healthc 2018;11:295-304 [FREE Full text] [doi: 10.2147/JMDH.S160029] [Medline: 29997436]
43. Zeng L, Cook RJ, Wen L, Boruvka A. Bias in progression-free survival analysis due to intermittent assessment of progression. Stat Med 2015 Oct 30;34(24):3181-3193 [FREE Full text] [doi: 10.1002/sim.6529] [Medline: 26011411]
44. von Allmen RS, Weiss S, Tevaearai HT, Kuemmerli C, Tinner C, Carrel TP, et al. Completeness of Follow-Up Determines Validity of Study Findings: Results of a Prospective Repeated Measures Cohort Study. PLoS One 2015;10(10):e0140817 [FREE Full text] [doi: 10.1371/journal.pone.0140817] [Medline: 26469346]
45. Haslam A, Gill J, Prasad V. The frequency of assessment of progression in randomized oncology clinical trials. Cancer Rep (Hoboken) 2022 Jul;5(7):e1527 [FREE Full text] [doi: 10.1002/cnr2.1527] [Medline: 34821077]
46. Holmgren AJ, Downing NL, Bates DW, Shanafelt TD, Milstein A, Sharp CD, et al. Assessment of Electronic Health Record Use Between US and Non-US Health Systems. JAMA Intern Med 2021 Feb 01;181(2):251-259 [FREE Full text] [doi: 10.1001/jamainternmed.2020.7071] [Medline: 33315048]
47. Lin KJ, Glynn RJ, Singer DE, Murphy SN, Lii J, Schneeweiss S. Out-of-system Care and Recording of Patient Characteristics Critical for Comparative Effectiveness Research. Epidemiology 2018 May;29(3):356-363 [FREE Full text] [doi: 10.1097/EDE.0000000000000794] [Medline: 29283893]
48. Hatch B, Tillotson C, Angier H, Marino M, Hoopes M, Huguet N, et al. Using the electronic health record for assessment of health insurance in community health centers. J Am Med Inform Assoc 2016 Sep;23(5):984-990 [FREE Full text] [doi: 10.1093/jamia/ocv179] [Medline: 26911812]
49. Burcu M, Manzano-Salgado CB, Butler AM, Christian JB. A Framework for Extension Studies Using Real-World Data to Examine Long-Term Safety and Effectiveness. Ther Innov Regul Sci 2022 Jan;56(1):15-22 [FREE Full text] [doi: 10.1007/s43441-021-00322-8] [Medline: 34251656]
50. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol 2008 Feb 15;167(4):492-499. [doi: 10.1093/aje/kwm324] [Medline: 18056625]
51. Tyrer F, Bhaskaran K, Rutherford MJ. Immortal time bias for life-long conditions in retrospective observational studies using electronic health records. BMC Med Res Methodol 2022 Mar 27;22(1):86 [FREE Full text] [doi: 10.1186/s12874-022-01581-1] [Medline: 35350993]
52. Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol 2016 Apr 15;183(8):758-764 [FREE Full text] [doi: 10.1093/aje/kwv254] [Medline: 26994063]
53. Backenroth D. How to choose a time zero for patients in external control arms. Pharm Stat 2021 Jul;20(4):783-792. [doi: 10.1002/pst.2107] [Medline: 33655598]
54. Jackson JW. Diagnostics for Confounding of Time-varying and Other Joint Exposures. Epidemiology 2016 Nov;27(6):859-869 [FREE Full text] [doi: 10.1097/EDE.0000000000000547] [Medline: 27479649]
55. 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 [FREE Full text] [doi: 10.2147/CLEP.S242097] [Medline: 32440224]
56. Lau C, Jamali F, Loebenberg R. Health Canada Usage of Real World Evidence (RWE) in Regulatory Decision Making compared with FDA/EMA usage based on publicly available information. J Pharm Pharm Sci 2022;25:227-236. [doi: 10.18433/jpps32715] [Medline: 35760071]
57. FDA Acceptance of Foreign Clinical Studies Not Conducted Under an IND: Frequently Asked Questions: Guidance for Industry and FDA Staff. Food and Drug Administration. URL: https://www.fda.gov/regulatory-information/

search-fda-guidance-documents/fda-acceptance-foreign-clinical-studies-not-conducted-under-ind-frequently-asked-questions [accessed 2023-02-27]

58. McNair D, Lumpkin M, Kern S, Hartman D. Use of RWE to Inform Regulatory, Public Health Policy, and Intervention Priorities for the Developing World. Clin Pharmacol Ther 2022 Jan;111(1):44-51 [FREE Full text] [doi: 10.1002/cpt.2449] [Medline: 34655224]
59. Barrett JS, Heaton PM. Real-World Data: An Unrealized Opportunity in Global Health? Clin Pharmacol Ther 2019 Jul;106(1):57-59 [FREE Full text] [doi: 10.1002/cpt.1476] [Medline: 31188467]
60. Ramagopalan SV, Popat S, Gupta A, Boyne DJ, Lockhart A, Hsu G, et al. Transportability of Overall Survival Estimates From US to Canadian Patients With Advanced Non-Small Cell Lung Cancer With Implications for Regulatory and Health Technology Assessment. JAMA Netw Open 2022 Nov 01;5(11):e2239874 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.39874] [Medline: 36326765]
61. Degtiar I, Rose S. A Review of Generalizability and Transportability. Annu. Rev. Stat. Appl 2022 Oct 19;10(1):103837. [doi: 10.1146/annurev-statistics-042522-103837]
62. Brixner DI, Holtorf A, Neumann PJ, Malone DC, Watkins JB. Standardizing quality assessment of observational studies for decision making in health care. J Manag Care Pharm 2009 Apr;15(3):275-283. [doi: 10.18553/jmcp.2009.15.3.275] [Medline: 19326959]
63. PLOS Medicine Editors. Observational studies: getting clear about transparency. PLoS Med 2014 Aug;11(8):e1001711 [FREE Full text] [doi: 10.1371/journal.pmed.1001711] [Medline: 25158064]
64. Segal JB, Kallich JD, Oppenheim ER, Garrison LP, Iqbal SU, Kessler M, et al. Using Certification to Promote Uptake of Real-World Evidence by Payers. J Manag Care Spec Pharm 2016 Mar;22(3):191-196. [doi: 10.18553/jmcp.2016.22.3.191] [Medline: 27003547]
65. Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets. Food and Drug Administration. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic [accessed 2023-02-27]
66. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler H, Goettsch W, et al. Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. Pharmacoepidemiol Drug Saf 2017 Sep 15;26(9):1033-1039 [FREE Full text] [doi: 10.1002/pds.4297] [Medline: 28913966]
67. Maarseveen TD, Maurits MP, Niemantsverdriet E, van der Helm-van Mil AHM, Huizinga TWJ, Knevel R. Handwork vs machine: a comparison of rheumatoid arthritis patient populations as identified from EHR free-text by diagnosis extraction through machine-learning or traditional criteria-based chart review. Arthritis Res Ther 2021 Jun 22;23(1):174 [FREE Full text] [doi: 10.1186/s13075-021-02553-4] [Medline: 34158089]
68. Ayala Solares JR, Diletta Raimondi FE, Zhu Y, Rahimian F, Canoy D, Tran J, et al. Deep learning for electronic health records: A comparative review of multiple deep neural architectures. J Biomed Inform 2020 Jan;101:103337 [FREE Full text] [doi: 10.1016/j.jbi.2019.103337] [Medline: 31916973]
69. Brazeal JG, Alekseyenko AV, Li H, Fugal M, Kirchoff K, Marsh C, et al. Assessing quality and agreement of structured data in automatic versus manual abstraction of the electronic health record for a clinical epidemiology study. Research Methods in Medicine & Health Sciences 2021 Nov 22;2(4):168-178. [doi: 10.1177/26320843211061287]
70. Dagenais S, Russo L, Madsen A, Webster J, Becnel L. Use of Real-World Evidence to Drive Drug Development Strategy and Inform Clinical Trial Design. Clin Pharmacol Ther 2022 Jan;111(1):77-89 [FREE Full text] [doi: 10.1002/cpt.2480] [Medline: 34839524]
71. Yang Z, Dehmer M, Yli-Harja O, Emmert-Streib F. Combining deep learning with token selection for patient phenotyping from electronic health records. Sci Rep 2020 Jan 29;10(1):1432 [FREE Full text] [doi: 10.1038/s41598-020-58178-1] [Medline: 31996705]
72. Harvey RA, Rassen JA, Kabelac CA, Turenne W, Leonard S, Klesh R, et al. Association of SARS-CoV-2 Seropositive Antibody Test With Risk of Future Infection. JAMA Intern Med 2021 May 01;181(5):672-679 [FREE Full text] [doi: 10.1001/jamainternmed.2021.0366] [Medline: 33625463]
73. Dron L, Kalatharan V, Gupta A, Haggstrom J, Zariffa N, Morris AD, et al. Data capture and sharing in the COVID-19 pandemic: a cause for concern. Lancet Digit Health 2022 Oct;4(10):e748-e756 [FREE Full text] [doi: 10.1016/S2589-7500(22)00147-9] [Medline: 36150783]
74. Alzu'bi AA, Watzlaf VJM, Sheridan P. Electronic Health Record (EHR) Abstraction. Perspect Health Inf Manag 2021;18(Spring):1g [FREE Full text] [Medline: 34035788]
75. Ma X, Bellomo L, Magee K, Bennette CS, Tymejczyk O, Samant M, et al. Characterization of a Real-World Response Variable and Comparison with RECIST-Based Response Rates from Clinical Trials in Advanced NSCLC. Adv Ther 2021 Apr;38(4):1843-1859 [FREE Full text] [doi: 10.1007/s12325-021-01659-0] [Medline: 33674928]
76. Popat S, Liu SV, Scheuer N, Hsu GG, Lockhart A, Ramagopalan SV, et al. Addressing challenges with real-world synthetic control arms to demonstrate the comparative effectiveness of Pralsetinib in non-small cell lung cancer. Nat Commun 2022 Jun 17;13(1):3500 [FREE Full text] [doi: 10.1038/s41467-022-30908-1] [Medline: 35715405]

77. Bennett KJ, Mann J, Ouyang L. Utilizing Combined Claims and Clinical Datasets for Research Among Potential Cases of Rare Diseases. Int J Healthc Inf Syst Inform 2018;13(2):1-12 [FREE Full text] [doi: 10.4018/ijhisi.2018040101] [Medline: 32913425]
78. Wells BJ, Chagin KM, Nowacki AS, Kattan MW. Strategies for handling missing data in electronic health record derived data. EGEMS (Wash DC) 2013;1(3):1035 [FREE Full text] [doi: 10.13063/2327-9214.1035] [Medline: 25848578]
79. Just BH, Marc D, Munns M, Sandefer R. Why Patient Matching Is a Challenge: Research on Master Patient Index (MPI) Data Discrepancies in Key Identifying Fields. Perspect Health Inf Manag 2016;13(Spring):1e [FREE Full text] [Medline: 27134610]
80. Johnson EK, Nelson CP. Values and pitfalls of the use of administrative databases for outcomes assessment. J Urol 2013 Jul;190(1):17-18 [FREE Full text] [doi: 10.1016/j.juro.2013.04.048] [Medline: 23608038]
81. Strom JB, Faridi KF, Butala NM, Zhao Y, Tamez H, Valsdottir LR, et al. Use of Administrative Claims to Assess Outcomes and Treatment Effect in Randomized Clinical Trials for Transcatheter Aortic Valve Replacement. Circulation 2020 Jul 21;142(3):203-213. [doi: 10.1161/circulationaha.120.046159]
82. Curtis JR, Baddley JW, Yang S, Patkar N, Chen L, Delzell E, et al. Derivation and preliminary validation of an administrative claims-based algorithm for the effectiveness of medications for rheumatoid arthritis. Arthritis Res Ther 2011;13(5):R155 [FREE Full text] [doi: 10.1186/ar3471] [Medline: 21933396]

Abbreviations

EHR: electronic health record
EMA: European Medicines Agency
FAIR: findable, accessible, interoperable, and reusable
FDA: Food and Drug Administration
ICD: International Classification of Diseases
ISPOR: International Society for Pharmacoeconomics and Outcomes Research
PFS: progression-free survival
RCT: randomized controlled trial
RECIST: Response Evaluation Criteria in Solid Tumors
SAP: statistical analysis plan
SNOMED CT: Systematized Nomenclature of Medicine – Clinical Terms
WGS: whole genome sequencing

Edited by A Mavragani; submitted 12.10.22; peer-reviewed by P Kumwichar, N Dreyer; comments to author 11.01.23; revised version received 01.02.23; accepted 19.02.23; published 05.04.23

Please cite as:
Velummailum RR, McKibbon C, Brenner DR, Stringer EA, Ekstrom L, Dron L
Data Challenges for Externally Controlled Trials: Viewpoint
J Med Internet Res 2023;25:e43484
URL: https://www.jmir.org/2023/1/e43484
doi: 10.2196/43484
PMID: 37018021