REVIEW

Tactical Considerations for Designing Real-World Studies: Fit-for-Purpose Designs That Bridge Research and Practice

Nancy A Dreyer [1], Christina D Mack[2]

[1]Dreyer Strategies LLC, Newton, MA, USA; [2]IQVIA Real World Solutions, Research Triangle Park, NC, USA

Correspondence: Nancy A Dreyer, Dreyer Strategies LLC, 328 Country Club Road, Newton, MA, USA, Tel +1 617 733 9478, Email ndreyer@dreyerstrategies.com

Abstract: Real-world evidence (RWE) is being used to provide information on diverse groups of patients who may be highly impacted by disease but are not typically studied in traditional randomized clinical trials (RCT) and to obtain insights from everyday care settings and real-world adherence to inform clinical practice. RWE is derived from so-called real-world data (RWD), ie, information generated by clinicians in the course of everyday patient care, and is sometimes coupled with systematic input from patients in the form of patient-reported outcomes or from wearable biosensors. Studies using RWD are conducted to evaluate how well medical interventions, services, and diagnostics perform under conditions of real-world use, and may include long-term follow-up. Here, we describe the main types of studies used to generate RWE and offer pointers for clinicians interested in study design and execution. Our tactical guidance addresses (1) opportunistic study designs, (2) considerations about representativeness of study participants, (3) expectations for transparency about data provenance, handling and quality assessments, and (4) considerations for strengthening studies using record linkage and/or randomization in pragmatic clinical trials. We also discuss likely sources of bias and suggest mitigation strategies. We see a future where clinical records – patient-generated data and other RWD – are brought together and harnessed by robust study design with efficient data capture and strong data curation. Traditional RCT will remain the mainstay of drug development, but RWE will play a growing role in clinical, regulatory, and payer decision-making. The most meaningful RWE will come from collaboration with astute clinicians with deep practice experience and questioning minds working closely with patients and researchers experienced in the development of RWE.

Plain Language Summary: Diagnostics, medical interventions, and health services may not perform as well as expected when used in everyday care. The demand for real-world evidence (RWE) to support evidence-based medicine has been fueled by an explosion of accessible data from health encounters using information that clinicians record during everyday patient care, and from patients, caregivers, and biosensors worn by patients. Real-world data (RWD) is an all-encompassing term referring to data from clinical care and everyday life. RWE comes from coupling carefully curated RWD with strong study design and analytics. The primary use of RWE is to fill evidence gaps about real-world performance for people often excluded or underrepresented in clinical trials such as the elderly, those with co-morbidities, or those who use multiple medications. We provide examples of patient registries, longitudinal follow-up studies, evidence hubs with established data linkage and study-specific record linkage.

This paper offers tactical advice about the value of opportunistic study designs, what to plan for in terms of transparency in data generation and management, and how pharmacy claims are being linked with electronic health records and/or patient-generated health data. We also explain that treatment decisions may be made by statistical randomization, not by doctors and patients, but after randomization, naturalistic follow-up can be used, ie systematic data collection from patients as they present for care or using decentralized processes. The most meaningful RWE will come from a collaboration of astute clinicians working with patients and researchers experienced using RWD to support evidence-based clinical practice.

Keywords: real-world evidence, effectiveness, evidence-based clinical practice, methods and guidelines, master protocols

101

Introduction

Real-world evidence (RWE) that can help guide clinical practice is derived from real-world data (RWD), defined by the US Food and Drug Administration (FDA) as "data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources." The FDA defines RWE as "clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD."[1] Simply put, RWD is the collective term generally used to describe data collected outside of a traditional randomized clinical trial (RCT) research setting, often in patient-clinician encounters and/or from information generated by patients in terms of their perception of pain, function, quality of life, and so on, and/or from biosensors worn by study participants.[2]

Randomization may be introduced at baseline to facilitate balanced comparison groups for pragmatic clinical trials (PCT), which then assume many of the characteristics of RWE after randomization, such as studying treatments as used rather than according to intent-to-treat analyses and including more diverse populations.[3,4] PCT generally focus on health outcomes that inform a clinical or policy decision,[5,6] in contrast to intermediate and surrogate outcomes often studied in traditional RCT prepared for regulatory approval of new medical products and supplementary new indications, which may be less reflective of true clinical outcomes.[7]

RWE is becoming widely recognized as a complement to results from RCTs, especially as we recognize the importance of learning about how diverse patients respond to treatments administered outside of the clinical trial setting and the durability of any such benefits, especially for vulnerable patients who are often not included or underrepresented in traditional RCTs.[8] It is important to emphasize that RWE is not a replacement for traditional RCTs but rather an important supplement used to fill evidence gaps remaining after a product has received regulatory market authorization.[9,10] The evidence provided by RWE can inform medical practitioners about treatment effectiveness, safety and heterogeneity of treatment response, and can often be conducted more quickly than a traditional RCT and at lower cost. RWE is frequently the choice for studying the natural history of disease, evaluating care pathways,[11] developing clinical treatment guidelines and assessing their implementation,[12] and providing context for medical product marketing authorizations and label expansions in the US, Europe, and major markets in Asia.[13–17]

In real-world settings, patients present with clinical histories and risk factors that may be substantially different from those studied in most RCTs. Tan et al[18] recently reviewed all the RCTs listed on ClinicalTrials.gov and other public trial registration sites, recorded the indications under study and their inclusion and exclusion criteria (I/E). When these trial I/E exclusion criteria were applied to electronic medical record data from the UK National Health Service, more than 50% of patients with relatively prevalent and/or costly conditions would not have been eligible for these RCTs due to their age, co-morbidities, and/or use of prescription medication(s) for unrelated conditions. This demonstrates the impetus for the research community to use RWE to fill the evidence gaps about more diverse patients.

By way of perspective, RWE is currently being used by regulators as context for single-arm studies, largely in oncology and rare diseases – situations where randomization may not be feasible or ethical.[9,19] Blinatumomab, for example, was approved for relapsing refractory acute lymphoblastic leukemia based on promising data from a Phase 2 trial.[20] The market authorization holder created an external control arm from pooled historical clinical data to support the interpretation of their single-arm trial and secured a label expansion before the Phase 3 RCT was completed.[21] Similarly, both the US FDA and the European Medicines Agency approved cerliponase alfa to slow the loss of ambulation based on a comparison of 23 treated symptomatic patients and 42 historical controls drawn from RWD.[22]

RWE has also been used to assess learning in risk management programs and to monitor adherence to guidelines, such as the "Get With The Guidelines" studies in coronary artery disease and heart failure.[23] Looking forward, machine learning models use RWD to understand disease progression, for example, by identifying rheumatoid arthritis patients who will switch from methotrexate to a biologic 30 days prior to occurrence of the medication change.[24]

Here, we describe four approaches used to generate RWE and provide pointers for clinicians interested in study design and execution. We offer tactical guidance about (1) opportunistic study designs, (2) considerations about representativeness of patients selected for study, (3) the need for transparency of data provenance, data handling, and data quality assessments, and (4) considerations for strengthening studies using record linkage and randomization in pragmatic clinical trials.

Study Designs for Real-World Data

Study design is foundational to developing reliable RWE. Much of the RWE that informs clinical practice comes from four broad types of study designs: (1) studies derived from patient registries, (2) follow-up after participation in traditional RCTs (sometimes known as "roll-over" studies) and other long-term safety studies, (3) pragmatic randomized trials, and (4) evidence hubs populated with RWD.

- Patient registries are used for many purposes including studying the natural history of disease, product safety, characterization of high-risk patients and identification of unmet needs. A strong advantage of patient registries is their adaptability. Broad data collection can be used to support a number of studies, and data collection may be adapted over time to include new exposures and outcomes of interest. See, for example, the registry maintained by the Cystic Fibrosis Foundation at www.cff.org. Registries have also been used to study physician decision-making, eg, the impact of PET and PET/CT scan results on intended management of various cancers.[25]
- Roll-over studies and long-term safety studies are follow-up studies designed to evaluate the durability of benefits and long-term safety, including the risks and benefits of various treatment sequences and combinations. In fact, post-market authorization long-term cohort (follow-up) studies are attracting renewed attention due to regulatory requirements to follow patients who receive cell and gene therapies for 5–15 years, as well requiring long-term studies for other treatments with limited follow-up data available at launch and/or for medical products approved using surrogate outcomes.
- Pragmatic randomized trials use baseline treatment randomization to achieve balanced comparison groups with regard to unmeasured confounders and then use naturalistic follow-up. These studies can also be used to understand surgical methods and patient management, not just treatments.
- Evidence hubs use multiple linked RWD sources, with data review and management (data curation) conducted in parallel with data collection. They are similar in concept and in benefit to RCT protocol approaches such as basket and umbrella trials and can be used to address questions as they arise, while also continuing to gather new data.[26]

Each study starts by crafting one or more research questions, identifying the target study population, exposures/ treatments, and outcomes of interest. The next step is to determine whether most or all necessary data elements are available in existing records like electronic health data or health insurance claims, or if additional data collection is needed. A variety of guidance documents are available to assist with study design and evaluation. User Guides for Patient Registries[27] and for Developing Protocols for Observational Comparative Effectiveness Research[28] provide detailed information to help with design and execution.

Feasibility and data quality assessments are critical, both at study conception and during design as needed. An emphasis on assessing data availability and completeness, alongside estimation of the size and distribution of a potential pool of study subjects, is key to understanding whether a research question can be reliably answered. That said, data completeness is often misunderstood by a lack of recognition that test data, for example, may appear to be missing because the test was not ordered and not because the test value was not recorded. While this information is still missing, it should not be attributed to faulty data quality but instead reflects the actual practice of clinical care. Also important, since many RWE studies are exploratory in nature, statistical considerations about multiple comparisons and reduction of alpha are not problematic; the goal of these studies is to estimate effect sizes, not test hypotheses.[29]

Tactical Guidance

In the spirit of pragmatism, we offer the following high-level guidance for study design:

Keep a Broad Eye Open to Naturalistic Opportunities for Study Design

Is there a natural experiment you could construct using existing data that would inform clinical decision-making by providing needed evidence? For example, hypersensitivity reactions following intravenous high-dose iron infusions with ferric carboxymaltose or isomaltoside 1000 were compared following switches in departmental purchasing over time.[30]

Another naturalistic experiment was created using RWD from a large health-care organization in Israel to evaluate the safety of messenger RNA-based vaccines against severe COVID-19, a study that has been described as "the most reliable scientific tool" the world had to evaluate the effectiveness of the [Pfizer] COVID vaccine and the impact of the vaccination program.[31,32]

The value of person-generated health data (PGHD) has also come to light for learning about people's perceptions of vaccine safety. PGHD, also known as patient-reported data, refers to health-related data created, recorded, and/or gathered from patients, family members, and/or caregivers without influence or intervention of clinical staff to help address a health concern.[33] During the pandemic, community-based volunteers were recruited online and asked if they had been vaccinated and if so, with which vaccine; if they noticed any side effects following vaccination; had sought medical attention for any such side effects or had tested positive for COVID-19 since vaccination. Follow-up over 30 days showed similar effectiveness and side effects for the three vaccines approved in the US at that time.[34–36]

Do Not Sacrifice Relevance for Representativeness

While studies are often criticized for not being geographically or demographically representative, this criticism does not hold up to scrutiny.[37] A study need not be broadly representative to be useful. Well-described study populations can provide important information unique to their patient characteristics, eg, race, ethnicity, and care settings.

Consider how large, closed-cohort occupational health programs for professional athletes were used to further understanding about COVID-19 transmission,[38] the effectiveness of COVID-19 diagnostic tests approved under Emergency Use Authorizations,[39,40] the impact of post-recovery viral shedding on test results and transmission,[41] and the ability of vaccine boosters to prevent incident infection.[42] This RWE was useful due to its timeliness, relevance, and on-going quality controls conducted as part of this evidence hub, including linked data from wearable devices, daily diagnostic results, contact tracing, and infection/vaccine history. Although these cohorts were largely composed of relatively healthy males, the evidence had broad relevance, with likely applicability to women, youth, elderly, and non-athletes.

Data Quality Matters. Documentation of Data Management, Staff Training, and an Audit Trail are Expected

It is important to understand why, where and how RWD were created and whether the data is likely to have been accurately and consistently recorded.[43] A clear description of the study population and data provenance, characterized in terms of person, place, and time, along with a description of data management and any data transformations will be expected.[44] The goal here is to provide enough information for reviewers to understand how the study population was recruited and the data were collected, handled, and analyzed, both to assist in interpretation and also so that others may replicate their methods, to the extent feasible, in different populations to evaluate whether the findings are broadly generalizable.

Most real-world studies use little, if any, source data verification. Instead, data are reviewed and curated, ideally on an ongoing basis rather than waiting until the end of the study. Established coding systems and algorithms should be used; exceptions should be justified. It is worth noting that roll-over studies and other safety studies may use MedDRA coding (Medical Dictionary for Regulatory Activities), which requires a potentially challenging mapping to ICD coding (International Classification of Diseases) used to classify disease, injuries, and other health conditions, as developed by the World Health Organization.

Purpose-driven data curation should focus on key variables and correction of data when possible and as needed. Supplementary data quality checks and/or validation for outcomes can be useful. For example, in creating the National Football League (NFL) Injury Analytics program, media reports of player injuries were compared with those reported through the NFL electronic medical record system to evaluate and enhance completeness of reporting, noting that injury descriptions from medical staff, not the media, were always considered the gold standard.[45]

There are a number of guidelines and frameworks available to guide quality review and data curation. For example, the Kahn Framework harmonizes a number of established data quality frameworks to define a comprehensive assessment

method for the quality of electronic health record data used in secondary settings.[46] This conceptual framework assesses quality of secondary EHR data across three dimensions of conformance, completeness, and plausibility, each of which is assessed through both the context of verification and the context of validation. Data quality frameworks such as the Kahn Framework are typically paired with standardized data structures, such as common data models (CDMs), to enable the standardization of collection, transformation, and review of data. Common data models can be designed with an eye toward data-types, such as the Observational Medical Outcomes Partnership (OMOP)[47] CDM are designed as a standard structure for capture of observational data, or they can be designed to standardize the capture of data unique to specific therapeutic areas, such as the mCODE data standard,[48] which defines the minimum common data elements needed for collection for oncology research.

The greater the potential impact of a decision, eg, approving a new indication for a marketed product, the more validations and data checks will be required.[49] If study results are intended for use in a regulatory submission, such as to support data on lack of effect in a non-treated group or natural history studies,[50] a detailed audit trail will be expected to ensure transparency in data provenance. Safety reporting is mandatory when funding is provided by a company that holds a market authorization for a product under study.[51]

Consider Strengthening a Study Using RWD Linkage, Patient-Generated Health Data Including Digital Health Technologies, and/or Treatment Randomization

Will new data be needed and from what sources? Long-term safety studies and roll-over studies often collect data directly from patients or their caregivers, only seeking clinical confirmation for events of special interest. The possibility of linking patient data to other medical information should be considered. Record linkage can help with long-term follow-up, historical patient information, clinical confirmation, or overall augmentation of foundational health records or registry data. The types of RWD that are often used for linkage include administrative health insurance claims data, pharmacy claims, and regional or national death records.[52–54] In a first-ever FDA acceptance of RWE as substantial evidence (in contrast to supporting evidence), transplant registry data were linked to US Social Security Administration death master files to support a label expansion for tacrolimus to prevent organ rejection in patients receiving lung transplantation.[13]

Other RWD can also be linked, and studies based on these linked data illustrate the utility of PGHD, genomic data, laboratory results, and direct-to-patient mobile data as part of an evidence hub. This approach is useful across clinical and occupational health settings. For example, some elite sports organizations were early adopters of injury surveillance programs linked with other RWD to improve athlete health.[45,55] These evidence hubs integrate information from electronic medical records (EMR) with player participation, game statistics, sideline clinician reports, equipment data, and wearable devices, and utilize regular quality reviews to support the reliability of ongoing research findings.[56,57] During the COVID-19 pandemic, these hubs were recrafted to provide a clinical foundation and longitudinal diagnostic data from frequent COVID-19 surveillance testing within a closed cohort,[38–40] serological testing for SARS-CoV-2 antibodies,[41] and genomic sequencing performed for all infections to determine the SARS-CoV-2 variant and vaccination history.[42] These RWD formed the basis for timely reliable evidence, including demonstrating that recovered individuals who continued to test positive for SARS-CoV-2 following discontinuation of isolation were not infectious to others,[41] understanding the viral trajectory of illness,[58,59] and demonstrating that booster vaccinations were associated with a significant reduction in incident infections during the Omicron wave.[42] Studies of the NFL occupational cohorts used data from wearables worn by players and staff while in their facilities and during travel to drive insights on disease transmission, showing that COVID-19 could be transmitted with less than 15 min of contact[38] and that there was great variability in the accuracy of testing from newly approved diagnostics.[39]

Treatment randomizations may be useful in helping avoid key sources of bias, such as selection bias and imbalance in baseline risk factors, by randomly allocating patients to treatment or control, a technique that lends itself to use in registries and health systems or as a stand-alone approach.[60] The TASTE (Thrombus Aspiration in ST-Elevation Myocardial Infarction) randomized registry trial, for example, randomized patients from an existing registry, substantially lowering the cost of patient recruitment and speeding time to completion since the data of interest were already being collected as part of the registry protocol.[61] Similarly, the Diuretic Comparison Project by the US Veterans

Administration is a switching study where participants who were using hydrochlorothiazide diuretics, 25 or 50 mg daily, were randomized at the point of care (here, the pharmacy) to either stay on their current regimen or switch to chlorthalidone at a suggested equipotent dose. Patients enroll in the study with permission from their primary care provider and informed consent is obtained by telephone.[62]

Randomization adds confounding control and simplifies the interpretation of study findings, which can be particularly useful to differentiate treatments in a crowded market. By way of example, the PRIDE (Palmitate Research in Demonstrating Effectiveness) study was conducted to study real-world outcomes in schizophrenics who had recently been incarcerated, hospitalized for an episode of psychosis, attempted suicide, and/or engaged in high-risk behaviors.[63] This study group was chosen since it is easier to detect a true effect (benefit or harm), should one exist, in a group at high risk for the events of interest. Participants were followed in a time-to-relapse study after being randomized to a monthly injection of paliperidone palmitate or their choice of one of seven daily oral medications. This practical comparison yielded RWE describing the comparative effectiveness, which was then included in the Patient Exposure section of the label (section 6.1).

Discussion

Clinicians are the bedrock of data creation and data validation. Community settings are integral to understanding treatment heterogeneity among diverse populations, how marketed medical products are used in everyday life, and which patients are most likely to benefit or be harmed and under what circumstances. The systematic use of RWE will not only contribute to evidence-based clinical practice and decisions by regulators and payors around product approval, effectiveness, and safety,[64] but may also facilitate clinician-led discovery of off-label drug therapies.[65]

Expertise specific to the design and conduct of studies that use RWD cannot be overlooked, as these methods are complex. High-quality studies that use RWD generally start with feasibility assessments to assure that critical exposure(s), outcome(s) and covariates are recorded or able to be collected[66] and require study designs that are appropriate for the study goal (often termed "fit for purpose") and are transparent about study methods, data collection, curation, and analysis.[67] Emulating a target trial design[68] in RWE is helpful for guiding most study designs – be they studies derived from patient registries or follow-up (cohort) studies using clinical or PGHD or RWD available through evidence hubs.

Systematic error (bias) is always a concern, especially as it relates to recruitment and loss-to-follow-up. People willing and able to participate in research studies can be hard to find and may differ from the target population of interest. Studies that rely on volunteer input must keep in mind that, once recruited, participants may soon tire of participation. Losses to follow-up can introduce bias since dropouts may be severely ill or dead. Successful follow-up requires investment and attention to participant engagement, and the impact of misclassification or missing data must be assessed to avoid misinterpretation. A salient example is the potential late mortality risk class signal identified after the approval of paclitaxel-coated balloons and paclitaxel-eluting stents for peripheral artery disease. Interpretation of the increased mortality observed 2 years after device implantation was complicated by missing follow-up information in the data sources used, leading the FDA to convene an advisory panel to evaluate the mortality signal and re-focus efforts on signal detection methods.[69]

Mitigation strategies include blinding outcome assessors, using appropriate statistical analyses to account for measured and unmeasured confounders, quantitative estimation of bias, and small incentives to increase patient participation and retention. That said, it is important to publish study results including detailed descriptions of how patients were recruited and studied, losses to follow-up, etc., so that we may continue to fill in evidence gaps, with a mind to looking both for consistencies and inconsistencies that could be attributed to bias or to differences in baseline study population characteristics. Even imperfect studies can be useful, especially for identifying potential safety signals and in situations where little if any quantitative information is available,[27] eg, early studies of the impact of using hand-held mobile phones on the risk of brain tumors.[70] Such publications may stimulate other descriptive and exploratory research. Consistency in the directionality and relative impact of treatment benefits and risks should increase trust and clinical acceptance.

A number of guidance documents for study design, conduct, and evaluation have been developed that may be of interest to research-minded clinicians, both those that are broadly applicable,[71,72] and those developed with a focus on

specific therapeutic areas such as asthma[73] and to medical devices.[74] For those in search of high-level guidance about comparative effectiveness, the GRACE (Good Research for Comparative Effectiveness) checklist for RWE may also be of interest, a stand-out due to its validation rather than having been formulated solely by consensus and its broad applicability.[75]

Conclusion

RCT remains a mainstay for drug development and is increasingly being complemented with RWE to fill in evidence gaps about medical products, services, and interventions as used in real-world settings, including risks and benefits among diverse patients and situations not typically studied in RCT. RWE will play an increasingly important role in clinical, regulatory, and payer decision-making by evaluating the experience of diverse patients and care settings, including quantitative evaluation of long-term benefits, risks, and risk mitigation activities, but the studies need careful design and execution, all tempered by the main study goals. We expect to see more use of linked RWD and supplementation with clinical outcome assessments, wearables, and other PGHD. Success will depend on the contributions of astute clinicians empowered by efficient data capture tools and information provided directly by patients about their experiences, coupled with researchers expert in the design, analysis, and interpretation of RWE.

Abbreviations

FDA, US Food and Drug Administration; CDM, common data models; COVID-19, Coronavirus disease. The number 19 refers to the fact that the disease was first detected in 2019; GRACE, Good Research for Comparative Effectiveness; ICD, International Classification of Diseases; I/E, Inclusion and exclusion criteria; MedDRA, Medical Dictionary for Regulatory Activities; NBA, (US) National Basketball Association; NFL, (US) National Football League; OMOP, Observational Medical Outcomes Partnership; PET, Positive emission tomography; PET/CT, Positive emission tomography/computed tomography; PRIDE, Paliperidone Palmitate Research in Demonstrating Effectiveness; RWD, Real-world data; RWE, Real-world evidence; PCT, Pragmatic clinical trial; PGHD, person (or patient) generated health data; RCT, randomized clinical trial; TASTE, Thrombus Aspiration in ST-Elevation Myocardial Infarction.

Acknowledgments

Author Contributions

Both authors made a significant contribution to the work reported, whether that is in the conception, execution, acquisition of data, analysis, and interpretation, or in all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Disclosure

Nancy A. Dreyer is affiliated with Dreyer Strategies LLC. Christina D Mack is a full-time employee of IQVIA. The authors report no other conflicts of interest in this work.

References

1. US Food and Drug Administration. Framework for FDA's real-world evidence program; 2018. Available from: https://www.fda.gov/media/120060/download. Accessed September 20, 2023.
2. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-world evidence — what is it and what can it tell us? *N Engl J Med*. 2016;375(23):2293–2297. doi:10.1056/NEJMsb1609216
3. Brass EP. The gap between clinical trials and clinical practice: the use of pragmatic clinical trials to inform regulatory decision making. *Clin Pharmacol Ther*. 2010;87(3):351–355. doi:10.1038/clpt.2009.218

4. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol.* 2009;62(5):464–475. doi:10.1016/j.jclinepi.2008.12.011

5. Ford I, Norrie J, Drazen JM. Pragmatic Trials. *N Engl J Med.* 2016;375(5):454–463. doi:10.1056/NEJMra1510059

6. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis.* 1967;20(8):637–648. doi:10.1016/0021-9681(67)90041-0

7. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med.* 1989;8(4):431–440. doi:10.1002/sim.4780080407

8. US Food and Drug Administration. Diversity plans to improve enrollment of participants from underrepresented racial and ethnic populations in Clinical Trials Guidance for Industry; 2022. Available from: https://www.fda.gov/media/157635/download. Accessed September 20, 2023.

9. Dreyer NA. Strengthening evidence-based medicine with real-world evidence. *Lancet Healthy Longev.* 2022;3(10):e641–e642. doi:10.1016/S2666-7568(22)00214-8

10. Marc BMO, Daniel G, Frank K, et al. A framework for regulatory use of real-world evidence. White Paper. Duke Margolis Center for Health Policy; 2017.Available from: https://healthpolicy.duke.edu/publications/framework-regulatory-use-real-world-evidence. Accessed September 20, 2023.

11. Curtis JP. Association of physician certification and outcomes among patients receiving an implantable cardioverter-defibrillator. *JAMA.* 2009;301 (16):1661. doi:10.1001/jama.2009.547

12. Heidenreich PA, Hernandez AF, Yancy CW, Liang L, Peterson ED, Fonarow GC. Get with the guidelines program participation, process of care, and outcome for medicare patients hospitalized with heart failure. *Circ Cardiovasc Qual Outcomes.* 2012;5(1):37–43. doi:10.1161/CIRCOUTCOMES.110.959122

13. Concato J, Corrigan-Curay J. Real-world evidence — where are we now? *N Engl J Med.* 2022;386(18):1680–1682. doi:10.1056/NEJMp2200089

14. Purpura CA, Garry EM, Honig N, Case A, Rassen JA. The role of real-world evidence in FDA-approved new drug and biologics license applications. *Clin Pharmacol Ther.* 2022;111(1):135–144. doi:10.1002/cpt.2474

15. Bakker E, Plueschke K, Jonker CJ, Kurz X, Starokozhko V, Mol PGM. Contribution of real-world evidence in European Medicines Agency's regulatory decision making. *Clin Pharmacol Ther.* 2023;113(1):135–151. doi:10.1002/cpt.2766

16. Li M, Chen S, Lai Y, et al. Integrating real-world evidence in the regulatory decision-making process: a systematic analysis of experiences in the US, EU, and China using a logic model. *Front Med.* 2021;8:669509. doi:10.3389/fmed.2021.669509

17. Storm NE, Chang W, Lin T-C, et al. A novel case study of the use of real-world evidence to support the registration of an osteoporosis product in China. *Ther Innov Regul Sci.* 2022;56(1):137–144. doi:10.1007/s43441-021-00342-4

18. Tan YY, Papez V, Chang WH, Mueller SH, Denaxas S, Lai AG. Comparing clinical trial population representativeness to real-world populations: an external validity analysis encompassing 43 895 trials and 5 685 738 individuals across 989 unique drugs and 286 conditions in England. *Lancet Healthy Longev.* 2022;3(10):e674–e689. doi:10.1016/S2666-7568(22)00186-6

19. Simon R, Blumenthal G, Rothenberg M, et al. The role of nonrandomized trials in the evaluation of oncology drugs. *Clin Pharmacol Ther.* 2015;97 (5):502–507. doi:10.1002/cpt.86

20. Topp MS, Gökbuget N, Zugmaier G, et al. Long-term survival of patients with relapsed/refractory acute lymphoblastic leukemia treated with blinatumomab. *Cancer.* 2021;127(4):554–559. doi:10.1002/cncr.33298

21. Gökbuget N, Kelsh M, Chia V, et al. Blinatumomab vs historical standard therapy of adult relapsed/refractory acute lymphoblastic leukemia. *Blood Cancer J.* 2016;6(9):e473–e473. doi:10.1038/bcj.2016.84

22. European Medicines Agency. Product information – Brineura 150 mg solution for infusion; 2017. Available from: https://www.ema.europa.eu/en/documents/product-information/brineura-epar-product-information_en.pdf. Accessed 20 September, 2023.

23. American Heart Association. Get with the guidelines - coronary artery disease. Available from: https://www.heart.org/en/professional/quality-improvement/get-with-the-guidelines/get-with-the-guidelines-coronary-artery-disease. Accessed September 20, 2023.

24. Shankar R, Poole L, Halmos T, et al. Using AI to support evidence & market access strategy development. Presentation presented at: ISPOR; May 7–10; 2023; Boston, MA, USA.

25. Hillner BE, Siegel BA, Shields AF, et al. Relationship between cancer type and impact of PET and PET/CT on intended management: findings of the National Oncologic PET Registry. *J Nucl Med.* 2008;49(12):1928–1935. doi:10.2967/jnumed.108.056713

26. Park JJH, Siden E, Zoratti MJ, et al. Systematic review of basket trials, umbrella trials, and platform trials: a landscape analysis of master protocols. *Trials.* 2019;20(1):572. doi:10.1186/s13063-019-3664-1

27. Gliklich RE, Dreyer NA, Leavy MB. Registries for evaluating patient outcomes: a user's guide. AHRQ Publication No. 13(14)-EHC111 ; 2014.

28. Velentgas PDN, Nourjah P, Smith SR, Torchia MM. *Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide.* Agency for Health Care Research and Quality; 2013:EHC099.

29. Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology.* 1990;1(1):43–46. doi:10.1097/00001648-199001000-00010

30. Bager P, Hvas CL, Dahlerup JF. Drug-specific hypophosphatemia and hypersensitivity reactions following different intravenous iron infusions. *Br J Clin Pharmacol.* 2017;83(5):1118–1125. doi:10.1111/bcp.13189

31. Barda N, Dagan N, Ben-Shlomo Y, et al. Safety of the BNT162b2 mRNA Covid-19 vaccine in a nationwide setting. *N Engl J Med.* 2021;385 (12):1078–1090. doi:10.1056/NEJMoa2110475

32. Bourla A. *Moonshot: Inside Pfizer's Nine-Month Race to Make the Impossible Possible.* New York, NY: Harper Business; 2022.

33. Patient-generated health data. Available from: https://www.healthit.gov/topic/scientific-initiatives/pcor/patient-generated-health-data-pghd. Accessed September 20, 2023.

34. Dreyer N, Reynolds MW, Albert L, et al. How frequent are acute reactions to COVID-19 vaccination and who is at risk? *Vaccine.* 2022;40 (12):1904–1912. doi:10.1016/j.vaccine.2021.12.072

35. Reynolds MW, Xie Y, Knuth KB, et al. COVID-19 vaccination breakthrough infections in a real-world setting: using community reporters to evaluate vaccine effectiveness. *Infect Drug Resist.* 2022;15:5167–5182. doi:10.2147/IDR.S373183

36. Reynolds MW, Secora A, Joules A, et al. Evaluating real-world COVID-19 vaccine effectiveness using a test-negative case–control design. *J Comp Eff Res.* 2022;11(16):1161–1172. doi:10.2217/cer-2022-0069

37. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol.* 2013;42(4):1012–1014. doi:10.1093/ije/dys223

38. Mack CD, Wasserman EB, Perrine CG, et al. Implementation and evolution of mitigation measures, testing, and contact tracing in the National Football League, August 9–November 21, 2020. *MMWR Morb Mortal Wkly Rep.* 2021;70(4):130–135. doi:10.15585/mmwr.mm7004e2

39. Mack CD, Osterholm M, Wasserman EB, et al. Optimizing SARS-CoV-2 surveillance in the United States: insights from The National Football League occupational health program. *Ann Intern Med.* 2021;174(8):1081–1089. doi:10.7326/M21-0319

40. Mack CD, Wasserman EB, Hostler CJ, et al. Effectiveness and use of reverse transcriptase polymerase chain reaction point of care testing in a large-scale COVID -19 surveillance system. *Pharmacoepidemiol Drug Saf*. 2022;31(5):511–518. doi:10.1002/pds.5424

41. Mack CD, DiFiori J, Tai CG, et al. SARS-CoV-2 transmission risk among national basketball association players, staff, and vendors exposed to individuals with positive test results after COVID-19 recovery during the 2020 regular and postseason. *JAMA Intern Med*. 2021;181(7):960. doi:10.1001/jamainternmed.2021.2114

42. Tai CG, Maragakis LL, Connolly S, et al. Association between COVID-19 booster vaccination and omicron infection in a highly vaccinated cohort of players and staff in the National Basketball Association. *JAMA*. 2022;328(2):209. doi:10.1001/jama.2022.9479

43. Simon GE, Bindman AB, Dreyer NA, et al. When can we trust real-world data to evaluate new medical treatments? *Clin Pharmacol Ther*. 2022;111 (1):24–29. doi:10.1002/cpt.2252

44. Nicholas MNH, Silcox C, Aten A, et al. Determining real-world data's fitness for use and the role of reliability. White Paper; 2019 https://healthpolicy.duke.edu/sites/default/files/2019-11/rwd_reliability.pdf.

45. Dreyer NA, Mack CD, Anderson RB, Wojtys EM, Hershman EB, Sills A. Lessons on Data collection and curation from the NFL injury surveillance program. *Sports Health*. 2019;11(5):440–445. doi:10.1177/1941738119854759

46. Kahn MG, Callahan TJ, Barnard J, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. *eGEMs*. 2016;4(1):18. doi:10.13063/2327-9214.1244

47. Informatics OHDSa. Standardized data: the OMOP common data model. Available from: https://www.ohdsi.org/data-standardization/#:~:text=The %20Observational%20Medical%20Outcomes%20Partnership,that%20can%20produce%20reliable%20evidence. Accessed September 20, 2023.

48. Osterman TJ, Terry M, Miller RS. Improving cancer data interoperability: the promise of the Minimal Common Oncology Data Elements (mCODE) initiative. *JCO Clin Cancer Inform*. 2020;4(4):993–1001. doi:10.1200/CCI.20.00059

49. Daniel G, Silcox C, Bryan J, et al. Characterizing RWD quality and relevancy for regulatory purposes; 2018 https://healthpolicy.duke.edu/sites/default/files/2020-08/Characterizing%20RWD%20for%20Regulatory%20Use.pdf.

50. US Food and Drug Administration. Demonstrating substantial evidence of effectiveness for human drug and biological products; 2019. Available from: https://www.fda.gov/media/133660/download. Accessed September 20, 2023.

51. Dreyer NA, Sheth N, Trontell A, Gliklich RE. Good practices for handling adverse events detected through patient registries. *Drug Inf J*. 2008;42 (5):421–428. doi:10.1177/009286150804200502

52. Niu X, Divino V, Sharma S, Dekoven M, Anupindi VR, Dembek C. Healthcare resource utilization and exacerbations in patients with chronic obstructive pulmonary disease treated with nebulized glycopyrrolate in the USA: a real-world data analysis. *J Med Econ*. 2021;24(1):1–9. doi:10.1080/13696998.2020.1845185

53. Rivera DR, Gokhale MN, Reynolds MW, et al. Linking electronic health data in pharmacoepidemiology: appropriateness and feasibility. *Pharmacoepidemiol Drug Saf*. 2020;29(1):18–29. doi:10.1002/pds.4918

54. Pratt NL, Mack CD, Meyer AM, et al. Data linkage in pharmacoepidemiology: a call for rigorous evaluation and reporting. *Pharmacoepidemiol Drug Saf*. 2020;29(1):9–17. doi:10.1002/pds.4924

55. Mack CD, Meisel P, Herzog MM, et al. The establishment and refinement of the national basketball association player injury and illness database. *J Athl Train*. 2019;54(5):466–471. doi:10.4085/1062-6050-18-19

56. Mack C, Sendor RR, Solomon G, et al. Enhancing concussion management in the national football league: evolution and initial results of the unaffiliated neurotrauma consultants program, 2012–2017. *Neurosurgery*. 2020;87(2):312–319. doi:10.1093/neuros/nyz481

57. Mack C, Myers E, Barnes R, Solomon G, Sills A. Engaging athletic trainers in concussion detection: overview of the National Football League ATC spotter program, 2011–2017. *J Athl Train*. 2019;54(8):852–857. doi:10.4085/1062-6050-181-19

58. Hay JA, Kissler SM, Fauver JR, et al. Quantifying the impact of immune history and variant on SARS-CoV-2 viral kinetics and infection rebound: a retrospective cohort study. *Elife*. 2022;11. doi:10.7554/eLife.81849

59. Kissler SM, Fauver JR, Mack C, et al. Viral dynamics of acute SARS-CoV-2 infection and applications to diagnostic and public health strategies. *PLoS Biol*. 2021;19(7):e3001333. doi:10.1371/journal.pbio.3001333

60. Ferreira JC, Patino CM. Choosing wisely between randomized controlled trials and observational designs in studies about interventions. *J Brasileiro de Pneumologia*. 2016;42(3):165. doi:10.1590/S1806-37562016000000152

61. Lauer MS, D'Agostino RB. The randomized registry trial — the next disruptive technology in clinical research? *N Engl J Med*. 2013;369 (17):1579–1581. doi:10.1056/NEJMp1310102

62. US Department of Veterans Affairs. VA cooperative studies program; 2018. Available from: https://www.vacsp.research.va.gov/CSP_597/Information_for_VA_Providers.asp#:~:text=The%20Diuretic%20Comparison%20Project%20%28DCP%29%20is%20the%20first,or%20more%20approved%20treatments%20when%20clinical%20equipoise%20exists. Accessed September 20, 2023.

63. Alphs L, Benson C, Cheshire-Kinney K, et al. Real-world outcomes of paliperidone palmitate compared to daily oral antipsychotic therapy in schizophrenia: a randomized, open-label, review board–blinded 15-month study. *J Clin Psychiatry*. 2015;76(5):554–561. doi:10.4088/JCP.14m09584

64. 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;29(10):1201–1212. doi:10.1002/pds.4932

65. DeMonaco HJ, Ali A, Hippel EV. The major role of clinicians in the discovery of off-label drug therapies. *Pharmacotherapy*. 2006;26(3):323–332. doi:10.1592/phco.26.3.323

66. Gatto NM, Campbell UB, Rubinstein E, et al. The structured process to identify fit-for-purpose data: a data feasibility assessment framework. *Clin Pharmacol Ther*. 2022;111(1):122–134. doi:10.1002/cpt.2466

67. Hall GC, Sauer B, Bourke A, Brown JS, Reynolds MW, Casale RL. Guidelines for good database selection and use in pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf*. 2012;21(1):1–10. doi:10.1002/pds.2229

68. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *JAMA*. 2022;328(24):2446. doi:10.1001/jama.2022.21383

69. US Food and Drug Administration. Paclitaxel-coated balloons and stents for peripheral arterial disease. Available from: https://www.fda.gov/medical-devices/cardiovascular-devices/paclitaxel-coated-balloons-and-stents-peripheral-arterial-disease. Accessed September 20, 2023.

70. Rothman KJ, Loughlin JE, Funch DP, Dreyer NA. Overall mortality of cellular telephone customers. *Epidemiology*. 1996;7(3):303–305. doi:10.1097/00001648-199605000-00015

71. Miksad RA, Abernethy AP. Harnessing the power of Real-World Evidence (RWE): a checklist to ensure regulatory-grade data quality. *Clin Pharmacol Ther*. 2018;103(2):202–205. doi:10.1002/cpt.946

72. Vandenbroucke JP. Strega, Strobe, Stard, Squire, Moose, Prisma, Gnosis, Trend, Orion, Coreq, Quorom, Remark... and Consort: for whom does the guideline toll? *J Clin Epidemiol*. 2009;62(6):594–596. doi:10.1016/j.jclinepi.2008.12.003

73. Roche N, Campbell JD, Krishnan JA, et al. Quality standards in respiratory real-life effectiveness research: the REal Life EVidence AssessmeNt Tool (RELEVANT): report from the Respiratory Effectiveness Group—European Academy of Allergy and Clinical Immunology Task Force. *Clin Transl Allergy*. 2019;9(1):20. doi:10.1186/s13601-019-0255-x

74. Allen A, Patrick H, Ruof J, et al. Development and pilot test of the registry evaluation and quality standards tool: an information technology–based tool to support and review registries. *Value Health*. 2022;25(8):1390–1398. doi:10.1016/j.jval.2021.12.018

75. Dreyer NA, Bryant A, Velentgas P. The GRACE checklist: a validated assessment tool for high quality observational studies of comparative effectiveness. *J Manag Care Special Pharm*. 2016;22(10):1107–1113. doi:10.18553/jmcp.2016.22.10.1107