DIVERSITY, EQUITY, & INCLUSION in Digitized Clinical Trials

DIGITAL MEDICINE SOCIETY

Use Real-World Data and Real-World Evidence

What should I do?

Use Artificial Intelligence/Machine Learning (AI/ML) to interrogate Real-World Data/ Real-World Evidence (RWD/RWE) to identify populations at the community and individual levels that will benefit the most from your therapeutic area or clinical trial.

Why should I do it?

- Increase access and awareness for participants to clinical trial participation.
- Increase access and knowledge, for clinical trial designers and implementers, on diverse populations.
- Increase and improve patient engagement and retention.
- Reduce burden for participants and clinical trial teams by streamlining and creating efficiencies with awareness and outreach.
- Build trust by working directly with diverse communities.
- Accelerate recruitment timelines with fewer participants needed for analysis, decreased time lost to participant screening, and overall streamline recruitment processes for participants and clinical teams.
- Collect more relevant measures and non-clinical measures which can lead to improved outcomes and more robust data.

Ethical Considerations

AI/ML and RWD/RWE are trained on and applied to data that have been collected by systems that have historically excluded or underserved many populations based on race, ethnicity, gender, socioeconomic status and other diversity vectors. Therefore, the use of these technologies should be approached with an equitable lens and should only serve as a starting point. Multiple data sources should be used and any information extracted should be inspected for inclusion prior to implementation in clinical trials, and should be supplemented by additional measures for health equity.

Considerations for using RWD/RWE

Is the RWD/RWE source fit for purpose?

Determine which data are available in which RWD/RWE source, and for which populations. Identify the data types that are most relevant to the purpose of your clinical trial (e.g. diagnosis of diabetes in the past 10 years for a variety of populations (race, ethnicity, age, gender)).

Does your study design match your purpose for clinical endpoints?

Identify how endpoints were selected, defined, and validated for clinical and non-clinical measures. Learn how the source data have been documented, tested and results reported, including any use cases that may apply to your purpose for using this source.

How credible is your RWD/RWE source?

Obtain additional information on data quality including how the data were collected, cleaned and harmonized, where gaps exist, the time period for data collection, and last update of the data.

How many RWD/RWE sources do you need to use to fully capture all populations that can benefit from your clinical trial?

Use multiple sources to ensure that you are being inclusive, and that the data are generalizable or can be combined with other sources for a more comprehensive analysis.

How are you factoring in equity considerations with each RWD/RWE source?

Determine how different populations are represented in the dataset to ensure that health disparities are not reinforced or perpetuated. Confirm which populations are represented - e.g. in terms of race, ethnicity, age, socioeconomic status.

RWD/RWE Sources

- **Healthcare/Health system databases including Electronic Health Records** (EHRs) - Databases used by many health systems, where healthcare providers document and track routine clinical and laboratory data as standard of care.
    - Value provided
        - RWD on risks and benefits based on routine and standardized data collections on different medical treatments and effectiveness of those treatments.
        - Data are longitudinal.
    - Potential limitations
        - Data are not collected for research purposes.
        - May contain valid, inaccurate or incomplete data.
        - Quality and completeness of data varies within and among databases.
- **Patient Disease or Clinical Registries** - Databases that collect, analyze, and disseminate prospective, observational data on patients with a specific health condition
    - Value provided:
        - Disease-specific RWE that can estimate effectiveness with data on the natural history of a disease, clinical and comparative effectiveness of procedures, treatments, outcomes, safety of health technologies, and quality of care and life.
    - Potential limitations:
        - Quality of data and methods, lack of standards or uniformity across different registries.
        - May contain incorrect or missing data with patient-reported data.
        - Confounding and bias based on disease severity.
- **Medical Claims Data** - Data around billing codes that are collected by physicians, pharmacies, hospitals and other health care providers for payment.
    - Value provided:
        - Claims data follow a standard format and are standardized for pre-established codes that are specific to diagnoses, procedures, and drugs.

        - Claims data are a more reliable indication of treatment adherence, and provide information on the economic burden of diseases.
    - Potential limitations:
        - Limited interoperability and generalizability.
        - Use of non-prescribed treatment might not be reflected in the data.
        - Clinical information and history may be missing or incomplete.
        - Coding or other issues with data entry or sufficient details.
- **Wearables/Digital health products** - Patient-generated health data can be objective and continuous, thereby providing a wider lens to understand patients' lived experiences.
    - Value provided:
        - These health products collect a greater variety of real-time information on daily life that can be informative for safety, efficacy, and quality of life with a therapeutic.
    - Potential limitations:
        - Users of these products may not be an adequate or complete representation of the relevant patient population.
        - Privacy and security issues will have to be addressed.
        - Product developers may not provide full access to raw data or data standards to outside organizations.
- **Genomic Databases** - Data collected on genome-wide association and sequencing studies provide information on specific genetic variants within specific populations.
    - Value provided:
        - These databases contain genomic data on different populations from across research studies, health conditions, and populations with accompanying phenotype data.
    - Potential limitations:
        - The majority of data are from populations identifying as Europeans/White.
        - Data are cross-sectional and limited by information collected for specific studies.

- **Patient Lifestyle Information** - Social or other patient/participant online networks provide access to qualitative information on treatment preferences and adherence, adverse reactions, and quality of life.
    - Value provided:
        - Qualitative information from patient perspectives can be used in estimating effectiveness and safety.
        - These platforms allow for efficient collection of patient perspectives, from large geographic areas.
        - Information on new or existing adverse events can be collected, closer to real time.
        - These platforms may be more familiar to and convenient for patients.
    - Potential limitations:
        - No standard methodology for collecting information.
        - Authenticity cannot be verified; the person discussing the condition may not be the one with the condition.
        - The information may not be a good or complete representation of the relevant patient population.
        - Privacy and security issues will have to be addressed.

RWD/RWE Databases

Clinical Measures

Digital Medicine Society Library of Digital Endpoints

International Consortium for Health Outcomes Measurement

Population Data

Agency for Healthcare Research and Quality: Social Determinants of Health

Center for Disease Control and Prevention

Digital Equity

US Census Bureau

Biomedical Databases

All of Us Research Program

Database of Genotype and Phenotypes

International HapMap Project

The International Genome Sample Resource

UK Biobank

References & Resources

1. Advancing digital health applications: priorities for innovation in real-world evidence generation. Findings from discussions with the aims of (1) accelerating and stimulating innovative approaches to digital medical product evaluation, and (2) promoting international harmonization of best evidentiary practices.
2. All of Us: Release of Nearly 100,000 Whole Genome Sequences Sets Stage for New Discoveries. This historic program is building a vital research community within the United States of at least 1 million participant partners from all backgrounds.
3. A Street-by-Street View of Digital Inequity in the United States. A new Digital Equity Data Dashboard to help create better understanding of the economic opportunity gaps in towns, cities and neighborhoods across the United States.
4. Best Practices For Evaluating The Relevance And Quality Of RWD. The value of RWE derived RWD has been demonstrated in a broad range of biopharmaceutical applications including marketing strategy, clinical trial optimization, label expansion, and payer approval.
5. CTTI Recommendations: Use of Real-World Data to Plan Eligibility Criteria and Enhance Recruitment. RWD and RWE collected through the routine delivery of health care, are potentially powerful tools for enhancing the quality and efficiency of clinical trials.
6. Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. Standardized criteria facilitate the evaluation of RWD and data analysis in order to establish the confidence level in an RWE result.
7. How Can I Use Real-World Data In Clinical Trials? Real-world data describes what is happening in routine everyday clinical practice, outside of a clinical trial setting.
8. How Data Science Is Ushering in a New Era of Modern Medicine. Algorithms, artificial intelligence, machine learning and other technologies are transforming the way physicians identify, treat and manage diseases.
9. Injecting real-world data into every day clinical trials – RWE and RWD. Where trials fall short is in a lack of access, representation, and diversity, and data from the real world may be the key to picking up the slack.
10. Opportunities and challenges in using real-world data for health care. Although the best-known use case of RWD has been in drug regulation, RWD is being generated and used by many other parties, including bio-pharmaceutical companies, payors, clinical researchers, providers, and patients.
11. Real World Data & Real World Evidence In Clinical Trials. The clinical research landscape is shifting drastically due to an increase in the availability of electronic data allowing researchers to make assessments based on available RWE obtained from the analysis of RWD.

12. Real-world data: a brief review of the methods, applications, challenges and opportunities. We provide a brief overview on the type and sources of real-world data and the common models and approaches to utilize and analyze real-world data.
13. Real-world data quality: What are the opportunities and challenges? Growth in the availability and variety of real-world data, including non-health sources of data, opens up new opportunities, as well as challenges, in their application to real-world evidence and improving health outcomes.
14. RWE Navigator - Sources of real-world data. Real-world data (RWD) is an overarching term for data on the effects of health interventions (such as benefits, risks or resource use) that are not collected in the context of conventional randomized controlled trials (RCTs).
15. Toward a better understanding about real world evidence. RWE is derived from a spectrum of data generated from the real-world setting, using two broad study designs including observational studies and pragmatic clinical trials.
16. Understanding Use of Real-World Data and Real-World Evidence to Support Regulatory Decisions on Medical Product Effectiveness. RWE is derived from a spectrum of data generated from the real-world setting, using two broad study designs including observational studies and pragmatic clinical trials.