

Utilization of Real-World Data in Drug Development

Lu Hao

Bristol Myers Squibb, Shanghai 200040, China

*Corresponding author: Lu Hao, haolu916@hotmail.com

Copyright: © 2024 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: With the rapid development of modern science and technology, traditional randomized controlled trials have become insufficient to meet current scientific research needs, particularly in the field of clinical research. The emergence of real-world data studies, which align more closely with actual clinical evidence, has garnered significant attention in recent years. The following is a brief overview of the specific utilization of real-world data in drug development, which often involves large sample sizes and analyses covering a relatively diverse population without strict inclusion and exclusion criteria. Real-world data often reflects real clinical practice: treatment options are chosen according to the actual conditions and willingness of patients rather than through random assignment. Analysis based on real-world data also focuses on endpoints highly relevant to clinical benefits and the quality of life of patients. The booming big data technology supports the utilization of real-world data to accelerate new drug development, serving as an important supplement to traditional clinical trials.

Keywords: Real-world data; Drug development; Data mining

Online publication: June 19, 2024

1. Introduction

In clinical research, the classic method to evaluate the efficacy of a drug or treatment is the randomized controlled trial (RCT). Over the past few decades, RCTs have profoundly reshaped modern medical knowledge and practice strategies. The advantage of RCTs is their ability to significantly reduce bias in clinical trials. However, by design, RCTs also have certain limitations, and the clinical trial environment is often quite different from real medical practice. Therefore, the effectiveness of RCT results can be insufficient when extrapolated to the broader population or applied in real life. Due to these limitations of RCTs, real-world research, another clinical evaluation methodology, has been gaining increasing attention from researchers worldwide^[1].

Real-world studies (RWS) originated from practical randomized controlled trials, proposed as early as 1966, but gradually gained attention after being formally proposed by Professor Kaplan in his 1993

paper describing a prospective study of Ramipril's effects on hypertension. RWS are defined as providing interventions based on the actual conditions and wishes of patients, using a large and broadly representative sample size rather than random allocation, and conducting long-term follow-up evaluations of outcomes with broad clinical benefit to further evaluate the external effectiveness and safety of interventions. RWS can also be applied in various settings, including diagnosis, prognosis, etiology, and other aspects of certain diseases or medical conditions ^[2]. Comparative effectiveness studies, which involve patient-centered analyses of clinical data from different patient groups to compare the actual application of various interventions, also fall under the category of real-world research. These studies provide a better basis for clinical decision-making and health policy formulation.

Real-world data can be understood as clinical studies where data is derived from conventional medical environments. In this process, medical personnel aim to improve and safeguard the health status of patients, adopting appropriate methods and carrying out medical activities with patients at the core. Applying the concept of real-world data to drug development has both theoretical and practical advantages. There are many different ways to use real-world data to facilitate or accelerate drug development as a complement or replacement for RCTs. This paper reviews the latest trends in utilizing real-world research in clinical development and discusses data quality control in real-world research.

2. Features of real-world data

2.1. Nature of real-world data

Real-world data focuses on external authenticity and is primarily used to evaluate the effectiveness of a drug or treatment in a real clinical setting. Comparative effect studies in real clinical situations select rich samples to compare the advantages and disadvantages of different strategies in prevention, diagnosis, observation of disease conditions, or improvement of medical service levels. The research results help medical workers and policy implementers make decisions that improve clinical outcomes ^[3].

2.2. Design of studies based on real-world data

The authenticity of the clinical setting and the effectiveness of outcome evaluation requires observation and follow-up times in real-world studies to be generally longer. In designing real-world studies, medical personnel choose appropriate treatment plans according to the patient's conditions and wishes, determine research objectives, and then choose observational or analytical studies rather than simple randomization and blind trials using placebos to make the results more realistic. For example, in a multinational prospective observational study of acute coronary syndrome (ACS), 14,259 STEMI patients from multiple hospitals in several countries were studied. An observational method was applied to investigate the impact of thienopyridine and fibrinolytic treatments on STEMI (ST-elevation myocardial infarction) patients ^[4].

2.3. Real-world data inclusion and exclusion criteria and sample size

Strict inclusion and exclusion criteria often limit the broad generalization of a study's findings. To effectively reflect the true picture of the disease, real-world studies broaden inclusion criteria and narrow exclusion criteria, not excluding severe and complex cases with comorbidities. These approaches minimize the difference between the study population and the target population in real clinical practice, improve external authenticity, and reduce selection bias. Similarly, real-world data should take a holistic sample to include often overlooked populations, such as elderly patients, children, women, and minorities with special socioeconomic statuses. Patients with complex complications can be included in real-world studies, enhancing the applicability of conclusions to

healthcare workers. Of course, the sample size of real-world data is not always large, and sometimes a relatively small but uniform sample size can be chosen ^[5].

2.4. Choice of interventions and evaluation criteria

Studies based on real-world data do not need to use special criteria to intervene in the treatment of patients. Instead, they only collect data according to the actual clinical indications of patients, their diagnosis and treatment processes, and the long-term outcomes of such treatments. Evaluation of results is also not limited to one or a few specific clinical indicators but includes dose frequency, side effects, improvement rate, mortality rate, recurrence rate, disability conditions, and other indicators with clinical practical value ^[6].

2.5. Requirements and statistical methods for real-world data ethics

The ethical issues of real-world data are mainly distributed across three stages of research: data collection, data analysis, and results reporting, each involving different ethical issues. However, real-world data do not involve special or ineffective intervention measures for the target population but only observation and follow-up of the treatment process. Thus, it can minimize the ethical harm to study subjects. The statistical methods used in real-world studies are essentially the same as those used in other clinical studies (randomized controlled trials, single-case randomized controlled trials, etc.) (chi-square test, Fisher test, log-rank test, ROC curve, etc.). However, due to the many factors that influence real-world research, more comprehensive and advanced statistical methods (multi-factor analysis, instrumental variables, propensity scores, etc.) are needed to analyze the results more accurately.

3. Cases of real-world data application in drug development

3.1. Application in drug life cycle management

Drug life cycle management refers to the comprehensive oversight of drugs throughout their development stages, including design, clinical trials, marketing, and market withdrawal. Real-world data (RWD) plays a crucial role in this process, helping researchers understand drug safety, efficacy, and patient usage to develop better treatment options ^[7].

For example, post-marketing authorization of a drug, RWD can help monitor its safety, identify potential adverse reactions, and implement timely measures to ensure patient safety. Additionally, RWD helps understand patient responses and the effectiveness of different drugs, facilitating the creation of more personalized treatment plans. It also informs the development of digital or portable devices (e.g., mobile apps, wearables) that help understand patient behavior and treatment adherence, thereby enabling the creation of effective digital interventions.

3.2. Application in drug clinical trials

Clinical trials are essential for measuring the efficacy and safety of drugs. However, due to limited sample sizes and strict inclusion/exclusion criteria, clinical trials may not fully reflect real-world outcomes ^[8]. Therefore, drug evaluations relying solely on clinical trial data have limitations. RWD, through large-scale research designs, can provide a more comprehensive description of real-world case characteristics and treatment practices, summarizing treatment effects and patient responses to better assist in clinical trial design and further drug evaluation.

RWD is increasingly important in pharmaceutical drug development. Well-designed clinical studies using RWD analyze sources such as electronic medical records and wearables. Unlike controlled trials, these studies

offer insights into patients' daily lives regarding safety and efficacy. Researchers can use the results to identify potential patients and optimize criteria for clinical trials. For example, during the pandemic, RWD played a key role in assessing the efficacy and safety of COVID-19 vaccines. It helped evaluate the effectiveness of mRNA vaccines for healthcare workers and critical personnel ^[9]. A recent trial based on RWD also demonstrated the efficacy of AstraZeneca's COVID-19 vaccine against certain critical medical conditions and hospitalizations.

RWD is used not only pre-approval but also post-approval by health authorities. It answers remaining questions about vaccines after clinical trials, conducted within a short period, helping understand patient characteristics and behavior, predict disease progression, and evaluate vaccine response.

The use of RWD applies to both vaccines and drugs. It provides insights into actual outcomes, including adverse events, treatment adherence, and long-term effects. By analyzing data from diverse patient populations, RWD can complement clinical trial results. In summary, leveraging data from diverse sources, RWD enhances drug development for decision-making in business, regulatory, and medical contexts, improving the understanding of drug safety, efficacy, and real-world impact.

4. The future of real-world data in drug development

4.1. Real-world data in clinical trials

In the long-term battle against diseases, clinical trials have adhered to unique methodologies and practices, with a people-centric concept being the most prominent characteristic. Guided by the International Council for Harmonisation (ICH) Good Clinical Practice (GCP), clinical studies should not only focus on disease management but also on improving patients' quality of life, subjective feelings, and ability to perform daily activities. Randomized controlled trials (RCTs) require simple and clear interventions, effective control measures (such as placebos), and highly homogeneous study populations, which may no longer fully reflect the needs of modern drug development. Therefore, clinical trials should aim to establish homeostasis in the body, maintain normal physiological functions, and improve adaptability to the external environment. This holistic approach determines the complexity of drug interventions, relying on clinically meaningful endpoints beyond morbidity and mortality ^[10]. Real-world data (RWD) encompasses the impact of different interventions under unique patient conditions, combined with social and environmental factors. Integrating these considerations into traditional clinical trial models through RWD enables more comprehensive and fit-for-purpose efficacy evaluations.

4.2. Big data technology supporting real data research

The 21st century has witnessed rapid advancements in internet technology, cloud computing, mobile terminals, and data storage methods, heralding the era of "big data." Concurrently, data mining technology has emerged, defined as the process of extracting hidden information and knowledge from large, incomplete, noisy, fuzzy, and random databases. Big data technology emphasizes the comprehensive collection of data, prioritizing the approximate grasp of large datasets over the precise accuracy of a few individuals. This aligns with the overall concept of drug development, focusing on practical results, making it well-suited for clinical research supporting drug development.

Traditional clinical trial design and conduct have evolved through inherited and repeated practices, often leading to challenges in addressing diverse clinical questions, lengthy completion times, and frequent criticisms ^[11]. Big data techniques such as association rule cluster analysis, complex networks, decision trees, and hidden Markov models can objectively present clinical practice experiences using real-world data generated in clinical settings, providing more objective and representative insights. Additionally, the proliferation of electronic medical

record systems and the exponential growth of medical information make big data research in the medical field inevitable. Sharing and integrating big data in clinical drug development, and analyzing it, can significantly enhance the efficiency of clinical research. This approach provides opportunities to discover new clinical practice patterns and features, promoting the innovation of basic theories into practical methods for guiding patient selection, treatment choices, and evaluation criteria within clinical development. Ultimately, big data technology will facilitate the inheritance and evolution of clinical development practices.

5. Conclusion

In summary, real-world data, with its origins in clinical application, is closely intertwined with people-oriented clinical development practices and evaluations. It places significant emphasis on practical value and holds promise for further application in the future. However, real-world data is still in its infancy and presents some challenges. For instance, quality control in real-world research requires collective efforts from researchers to develop appropriate quality control strategies. Additionally, the large sample sizes in real-world studies, the extensive and complex databases generated, and the difficulties in statistical analysis pose new challenges for data management and statistical evaluation of efficacy. Opportunities and challenges coexist, offering a unique chance to drive future developments in the field of drug development.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Dagenais S, Russo L, Madsen A, et al., 2022, Use of Real-World Evidence to Drive Drug Development Strategy and Inform Clinical Trial Design. *Clin Pharmacol Ther*, 111(1): 77–89. <https://doi.org/10.1002/cpt.2480>
- [2] Comi E, 2022, Real-World Evidence in Clinical Research: Challenges and Opportunities. *Chimica Oggi – Chemistry Today*, 40(5), viewed Jan 1, 2024, https://www.teknoscienze.com/tks_article/real-world-evidence-in-clinical-research-challenges-and-opportunities/
- [3] Wang SV, Sreedhara SK, Schneeweiss S, et al., 2022, Reproducibility of Real-World Evidence Studies Using Clinical Practice Data to Inform Regulatory and Coverage Decisions. *Nat Commun*, 13(1): 5126. <https://doi.org/10.1038/s41467-022-32310-3>
- [4] Barrett JS, Azer K, 2023, Opportunities for Systems Biology and Quantitative Systems Pharmacology to Address Knowledge Gaps for Drug Development in Pregnancy. *J Clin Pharmacol*, 63 Suppl 1: S96–S105. <https://doi.org/10.1002/jcph.2265>
- [5] Bastarache L, Brown JS, Cimino JJ, et al., 2021, Developing Real-World Evidence from Real-World Data: Transforming Raw Data into Analytical Datasets. *Learn Health Syst*, 6(1): e10293. <https://doi.org/10.1002/lrh2.10293>
- [6] Volansky R, 2022, Real-World Data Confirms Safety, Efficacy of Guselkumab in Plaque Psoriasis. *Healio Psoriatic Disease*, viewed Jan 1, 2024, <https://www.healio.com/news/dermatology/20221019/realworld-data-confirms-safety-efficacy-of-guselkumab-in-plaque-psoriasis>
- [7] Huang CW, Wu BCY, Nguyen PA, et al., 2023, Emotion Recognition in Doctor-Patient Interactions From Real-World Clinical Video Database: Initial Development of Artificial Empathy. *Comput Methods Programs Biomed*, 233: 107480. <https://doi.org/10.1016/j.cmpb.2023.107480>
- [8] Wee LE, Tay AT, Chiew C, et al., 2023, Real-World Effectiveness of Nirmatrelvir/Ritonavir Against COVID-19

- Hospitalizations and Severe COVID-19 in Community-Dwelling Elderly Singaporeans During Omicron BA.2, BA.4/5, and XBB Transmission. *Clin Microbiol Infect*, 29(10): 1328–1333. <https://doi.org/10.1016/j.cmi.2023.06.016>
- [9] Bertsimas D, Li ML, Liu X, et al., 2023, Data-Driven COVID-19 Vaccine Development for Janssen. *INFORMS Journal on Applied Analytics*, 53(1): 1–95. <https://doi.org/10.1287/inte.2022.1150>
- [10] Kouki Y, Okada N, Saga K, et al., 2023, Disproportionality Analysis on Hypothyroidism With Roxadustat Using the Japanese Adverse Drug Event Database. *J Clin Pharmacol*, 63(10): 1141–1146. <https://doi.org/10.1002/jcph.2300>
- [11] Beck D, Winzenborg I, Gao W, et al., 2022, Interdisciplinary Model-Informed Drug Development for Extending Duration of Elagolix Treatment in Patients with Uterine Fibroids. *Br J Clin Pharmacol*, 88(12): 5257–5268. <https://doi.org/10.1111/bcp.15440>

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.