

White Paper

Unlocking the Potential of RWD in **Clinical Trial Recruitment**

April 2025



Introduction

Despite a robust clinical trial landscape and the increasing popularity of a decentralized trial model designed to make participation in clinical trials easier, drug developers continue to struggle with recruiting patients for their clinical research. Failure to achieve recruitment goals puts sponsors at risk for costly setbacks, invalidated findings, or trial termination, delaying the delivery of potentially life-changing treatments to patients.

The use of real-world data (RWD) in clinical trial recruitment and retention offers sponsors a solution for moving forward. Technological advancements such as artificial intelligence (AI) and machine learning (ML) have recently enabled the instantaneous processing of vast amounts of RWD, enabling new insights to streamline and perfect previously tedious and time-consuming processes.



Despite growth, recruiting challenges remain

Clinical trial activity, which dipped during the global COVID-19 pandemic, has recently shown a remarkable rebound. [Trialtrove data](#) show that during the 2023 calendar year, 9,959 Phase I–III clinical trials investigating at least one drug were initiated, a significant 9.4% increase over the previous year. The industry continues to demonstrate a robust interest in developing new and innovative medications, with data showing the initiation of Phase I trials growing by 3.3% from 2022 to 2023.

While the [present and future look bright](#) for clinical trial growth, patient recruitment remains a persistent challenge for study sponsors. One study showed¹ that over roughly the past decade, as studies have become larger (as measured by the number of clinical sites registered), the number of patient participants has significantly decreased while the duration of the recruitment process has increased. Other studies suggest² that approximately 85% of clinical trials fail to meet their enrollment goals and nearly one-third of Phase III trials fail due to enrollment issues. Poor recruitment has been cited as the most prevalent cause³ of discontinued trials.

In recent years, new patient pool considerations⁴ and stringent trial inclusion/exclusion requirements have increased the complexity of clinical trial patient recruitment. In June 2024, draft guidance from the US

FDA (since removed⁵) called for sponsors of investigational new medicines or biologics to submit with their application a Diversity Action Plan (DAP) outlining their strategy for expanding inclusion in their clinical trial, justifying their patient enrollment goals along demographic lines, and explaining how they intend to reach their goals for diverse patient enrollment. While diversifying patient enrollment is essential to ensure study findings accurately reflect the efficacy and safety of a drug for the broader population, recruiting and enrolling traditionally underrepresented groups has long been a challenge due to barriers including access to healthcare, language differences, and mistrust⁶.

Sponsors, especially [rare disease](#) drug developers, also face difficulties in patient enrollment due to the constraints inherent in the nature of their clinical trials: Often, inclusion/exclusion criteria for this type of research are extremely specific, shrinking already limited potential patient pools even further.

Once patients are located and enrolled, retaining them for the duration of the trial becomes the next hurdle to successfully ushering the trial to completion. Targeting highly specific populations to include in clinical research, recruiting underrepresented populations for these clinical trials, and retaining those patients can be new and challenging frontiers even for seasoned drug developers.

Powering patient recruitment with RWD

Sponsors continue to struggle to locate, engage, and enroll the patient populations needed for their clinical trials, but the strategic use of RWD offers a powerful solution for these challenges. Drawn from a wide range of resources, RWD can include electronic health records (EHR), electronic medical records (EMR), patient registries, claims data from insurance companies, wearable health device data, social determinants of health (SDOH) data, lab data, and public and proprietary data drawn from other clinical trials. Using the right datasets in the right combinations, sponsors can leverage RWD to gain a highly accurate view of patients within their disease areas. From there, they can easily determine disease prevalence as it relates to patient demographics and to the demands of the trial. RWD can return highly specific results,

down to pinpointing patients with the necessary biomarkers⁷ for a rare disease trial, for example.

Insights derived from RWD can be used not only to identify potential patients for participation in clinical research, but also to inform the overall patient engagement and retention strategy of a clinical trial. By leveraging actionable insights from combined datasets, sponsors can enhance and refine their approaches to patient identification and engagement, dynamically adjusting their outreach strategies⁸ based on real-time patient engagement metrics. The information surfaced by RWD provides the basis for data-driven recruitment and retention and can also inform other critical clinical trial factors such as trial design, inclusion/exclusion criteria, and broader patient-centric trial considerations.

Not all RWD are created equal

The quality of RWD insights largely depends upon the variety, scope, and quality of the data sources informing them. Sources such as open claims data can provide vital insights to help sponsors diversify patient pools or satisfy the narrow parameters of their trial protocols. Adding additional relevant data sources further strengthens RWD, revealing new information that enables even faster and more agile decision making.

Integrated datasets, like NorstellLinQ, available from Citeline's parent company Norstell, combine typical RWD data sources with additional forecasting, clinical, regulatory, payer, and commercial intelligence from their

proprietary databases, enabling sponsors to meet recruitment challenges at scale. NorstellLinQ incorporates data from the Citeline suite of pharmaceutical intelligence solutions. This includes the RWD-driven Citeline PatientMatch tool as well as Trialtrove, which provides detailed benchmarking, enrollment, patient population, and geographic data. It also includes Sitetrove, providing granular data on half a million investigators and nearly 200,000 clinical trial sites in 185 countries. These vast stores of data, along with other RWD — EHR, patient registries, etc. — can be instantly combed and analyzed using Citeline PatientMatch's AI capabilities and used to refine potential targets.

Real-world RWD applications

The use of integrated datasets for patient recruitment and retention can take different forms. For example, [one rare disease trial sponsor](#) successfully added lab data and information from structured and unstructured EMR to the claims and registry datasets it already had, swiftly expanding its study cohort from less than 200 patients to more than 3,000 patients — a remarkable 1,500% increase.

Another rare disease research sponsor needed to locate participants for its Phase II clinical trial on treatment for extensive-stage small cell lung cancer and previously treated small cell

lung cancer. The sponsor was having difficulty pinpointing patients who met the study's stringent eligibility criteria. The organization used RWD from Citeline PatientMatch to find additional eligible patients via a weekly alert system, notifying the sponsor when new patient matches were found.

These types of applications enable study sponsors to use RWD to target the right patients for their clinical trials, capture a holistic view of the patient population, and quickly streamline the recruitment process, allowing for increased and ongoing site efficiencies.

References

1. Mette Brøgger-Mikkelsen, John Robert Zibert, Anders Daniel Andersen, et al. (2022), Changes in key recruitment performance metrics from 2008–2019 in industry-sponsored phase III clinical trials registered at ClinicalTrials.gov. Available from <https://pmc.ncbi.nlm.nih.gov/articles/PMC9321424/>
2. Amy Kasahara, Jennifer Mitchell, and Tim K. Mackey (2024), Digital technologies used in clinical trial recruitment and enrollment including application to trial diversity and inclusion: A systematic review. Available from <https://pmc.ncbi.nlm.nih.gov/articles/PMC10981266/>
3. Benjamin Kasenda, Erik von Elm, John You, et al. (2014), Prevalence, characteristics, and publication of discontinued randomized trials. Available from <https://pubmed.ncbi.nlm.nih.gov/24618966/>
4. Michelle D Kelsey, Bray Patrick-Lake, Raolat Abdulai, et al. (2022), Inclusion and Diversity in Clinical Trials: Actionable Steps to Drive Lasting Change. Available from <https://pmc.ncbi.nlm.nih.gov/articles/PMC9133187/>
5. Giuliana Grossi (2025), FDA Quietly Removes Draft Guidance on Diversity in Clinical Trials Following Executive Order on DEI. Available from <https://www.ajmc.com/view/fda-quietly-removes-draft-guidance-on-diversity-in-clinical-trials-following-executive-order-on-dei>
6. Alice McCarthy (2023), Embracing Diversity: The Imperative for Inclusive Clinical Trials. Available from <https://postgraduateeducation.hms.harvard.edu/trends-medicine/embracing-diversity-imperative-inclusive-clinical-trials>
7. Chiara Pizzamiglio, Hilary J Vernon, Michael G Hanna, et al. (2022), Designing clinical trials for rare diseases: unique challenges and opportunities. Available from <https://pmc.ncbi.nlm.nih.gov/articles/PMC7613711/>
8. Kelly J Craig, Yanrong (Jerry) Ji, Yuxin (Chloe) Zhang, et al. (2024), Real-world Application of Racial and Ethnic Imputation and Cohort Balancing Techniques to Deliver Equitable Clinical Trial Recruitment. Available from <https://pmc.ncbi.nlm.nih.gov/articles/PMC10785904/>



Citeline, a [Norstella](#) company, powers a full suite of complementary business intelligence offerings to meet the evolving needs of life science professionals to accelerate the connection of treatments to patients and patients to treatments. These patient-focused solutions and services deliver and analyze data used to drive clinical, commercial, and regulatory-related decisions and create real-world opportunities for growth.

Citeline's global teams of analysts, journalists, and consultants keep their fingers on the pulse of the pharmaceutical, biomedical, and medtech industries, covering them all with expert insights: key diseases, clinical trials, drug R&D and approvals, market forecasts, and more. For more information on one of the world's most trusted health science partners, visit [Citeline](#) and follow on [LinkedIn](#) and [X](#).

Copyright © 2025 Citeline, a Norstella company.

Pharma Intelligence UK Limited is a company registered in England and Wales with company number 13787459 whose registered office is 3 More London Riverside, London SE1 2AQ.