Readying your clinical trial for linkage with real-world data

How to Guide

datavant

When it comes to recruiting patients for clinical trials and collecting the necessary data to assess the impact and safety of novel therapies, trial sponsors and sites are facing greater challenges than ever before. The near ubiquity of real-world data (RWD) sources such as electronic health records, claims, and lab data has prompted sponsors to consider RWD's role in study programs, and regulatory bodies to issue guidance on their use in clinical development. While some types of RWD have been used widely for many years, RWD has remained isolated from the detailed study participant data collected during the course of a clinical trial.

Tokenization can significantly improve the way we run clinical trials.

Privacy Preserving Record Linkage (PPRL) or "tokenization" is the process by which we replace personally identifiable information (PII) with non-sensitive, encrypted codes (or tokens). The idea is to make it impossible for anyone – including your organization – to connect the token to the original identifiable information. Datavant's tokens are HIPAA-certified to ensure patient privacy, and expert determination services help provide further fit-for-purpose de-identification as needed to help ensure that any dataset resulting from linked clinical trial and real-world data preserves patient privacy. Using "tokens" we can link clinical trial data to RWD sources –and establish a more complete picture of each study participants' health journey.

Since launching, Datavant has tokenized over 100 trials and 150,000 study participants to support linkage between clinical trial data and the ecosystem of hundreds of real-world data sources.
To date, we have tokenized studies in a wide range of therapeutic areas and trial phases.

In this guide, we use our learnings from the field to examine in detail when a trial is ready or appropriate for tokenization.

Study flexibility and other considerations

Are there specific types of trials that are a better fit for tokenization than others?

Datavant has tokenized trials at all stages and across all major therapeutic areas. We are currently working with several sponsors and CROs on enterprise-wide tokenization strategies. When sponsors first engage in trial tokenization, one of the first questions is often where to start, or which study is the best fit.

The below parameters are study characteristics to keep in mind when looking to tokenize your first set of trials:

1. Are you running a decentralized or hybrid trial?

Siteless and hybrid studies are ripe for trial tokenization as they allow you to substantiate externally collected data with other clinical sources. Some sponsors are validating patient reported outcomes collected during a trial by using tokens to link patient-generated data to Electronic Health Records (EHR) data and monitoring patient encounters with other clinical sites that are not part of the study.

2. Does your protocol need data captured as standard-of-care elsewhere?

Does your trial require the collection of – or would it benefit from – clinical and patient data that could be found in RWD sources? Most studies would benefit from commonly available clinical data such as medication history or comorbidities, but certain studies may also benefit from the collection of data not easily captured in a site setting, such as extensive social determinants of health information.

3. Is your protocol introducing too much site or patient burden?

Does your study have a lengthy follow-up period that may place undue strain on patients and study sites? In such cases, patients can benefit from studies that instead use extended monitoring of trial patients in the real world, using claims and mortality data linkages.

4. Are you thinking about how to reduce costs of a future late phase trial?

Tokenization of existing protocols with appropriate patient consent can extend the effective monitoring period of current patient cohorts via RWD sources, providing insight into future patient clinical encounters, without introducing new site requests or additional protocols.

Patient consent and education

Do I need to capture patient consent for tokenization in my study?

Engaged clinical trial participants are the most crucial component of any clinical study. Datavant encourages sponsors to educate patients about the kinds of data being gathered about them while they are enrolled in the trial, including information that may become available as a result of linking their trial data to their RWD. Patients may specifically opt-in for tokenization in studies that have not yet begun or are in the process of accruing participants.

These studies can include tokenization in their consent collection process. When patients are given information about tokenization as part of their informed consent, Datavant has observed very high rates of patients explicitly opting in. Sites with on-going trials can introduce PPRL as Institutional Review Boards (IRBs) have determined that tokenization and linkage are considered non-human subject research because no new parties hold re-identifiable patient information.

Site education and operations

Do trial sites need to generate patient trial tokens?

Trial tokenization begins with the conversion of PII to encrypted tokens, necessitating a partnership between Datavant and sites, CROs, or technology partners who retain PII. Datavant has partnerships with many of the industry's leading CROs and eConsent providers so that trial tokens can be created with existing patient data and without additional site or sponsor effort - and without the need for new PII or data sharing.

For studies that do not use a CRO or eConsent platform, Datavant Trials has a web-based trial tokenization portal allowing sites to tokenize seamlessly and directly. This web-based portal alleviates burden on trial sites with limited capacity and allows them to generate tokens in minutes. Datavant has learned that after a short training session, site teams, including IRBs, technology teams, principal investigators, and research coordination personnel, are able to use the tokenization portal with ease.

The method used by Datavant to create tokens emphasizes and upholds the fundamentals of Good Clinical Practice and enables its partners to adhere to all pertinent laws. In accordance with essential laws like HIPAA and 21 CFR Part 11, creating Datavant tokens upholds requirements for accurate reporting, maintenance, and verification of electronic records, including safeguards for patient confidentiality.

Real-world data access and linkage

What data sources are available and how do I select the right one for my study?

Technically, any structured dataset utilizing Datavant tokens can be linked to clinical trial data. Oftentimes, the RWD of choice depends on the study or strategic questions that emerge during the trial.

Common sources of RWD include claims, EHR and mortality as they facilitate Phase IV studies, by reducing costs and patient burden. More sophisticated research studies are often building clinico-genomics datasets to inform study design and patient recruitment, and even adapt mid-study to better understand a cohort of interest, like patients with exceptional response to a therapy. Other sponsors can link to RWD during the trial to help with safety surveillance and medical chart retrieval during the experiment.

Some sponsors elect to tokenize their trials without a specific RWD source in mind – to support unforeseen data collection needs after the trial. For example, a regulatory concern may need to be addressed based on data that wasn't collected as part of the original protocol. Sponsors can connect to that data source post trial.

Regardless of whether your trial is still in the planning phase or has already concluded initial data collection, Datavant can serve as the independent and neutral party to help sponsors find the right dataset for their emerging research needs, at any point in the trial lifecycle.

Conclusion

At Datavant, we're excited by the rapid growth of trial tokenization and identification of even more use cases that include clinical trials. Trial tokenization is one example of the robust and exciting new technologies accelerating innovation in clinical trials and drug development.

As we experience high costs, patient burdens along with the need to rapidly bring life-saving therapies to patients, it is essential that we focus on the most impactful applications that can support patients, site researchers and sponsors in developing the next generation of safe and effective therapies.

Read our ebook of use case examples for connected trials.

Learn more at: datavant.com/product/datavant-trials