Fact Sheet

Managing Sample Data from Vendors

Best practices for clinical trials

Most clinical trials involve a variety of laboratory vendors and CROs to provide services during the course of the study. Disparity in how each vendor handles, tracks and reports data presents a challenge.

Each vendor functions relatively independently. They have their own data systems in place to meet their internal needs. Whether it's processing or storage of the samples or shipping and receiving, their data systems are designed to make their own work the most efficient. Although they're providing data to the sponsors as part of their service offering and the sponsor typically has requirements on the content, each vendor is free to reference the data in their systems as they need it.

Various aspects of the data, such as the name of the study, the sample type reference, the name of the visit or test type associated with the sample, can take on different forms for different vendors. Later, when the data needs to come together for tracking or submission, the differences in the data need to be reconciled so everything is in the same language. This takes resources from the study side for data management and the analytics team to piece the data together, often manually, which could mean hard coding data from vendors, just to get alignment. In the ideal case, there would be global standards in place for vendor system setup, including sponsor standards for nomenclature and data structure. This would greatly reduce the effort and resources required for data harmonization.

Common components within sample data are present in most vendor systems. Shipping and receiving details, requisition or manifest numbers, sample and subject identifiers — the actual barcode on the tube — can be extracted and sent back to the sponsor.

However, it's the data values themselves that are potentially most problematic. Most challenging, for example, is the application of a vendor's ID to the actual sample tube with a new sticker containing their own barcode rather than tracking the original sample ID that arrived on the tube. This situation poses a tremendous challenge when trying to create chain of custody data for end-to-end tracking for all aspects of the trial. One can't make the alignment happen if the original ID is unknown.

Less painful are variations in visit and time point naming — cycle one, day one, pre-dose, for example, or using abbreviations or even full text strings. This type of difference can be resolved more easily, but it is still a challenge to harmonize when pulling the data together across different entities for a study. It can be a waste of time and resources.

It's not safe to make assumptions. To avoid that, one must work with a vendor to understand what their data contains, what they call the visit, what they call the tube type, what they call the analysis. Do they have multiple records for a given sample in their data or just one?

Approaching vendors of different capabilities

Not every vendor has the capability to track every aspect of the sample data. Smaller, specialty labs may not have a sophisticated data architecture or enterprise in place for managing sample tracking data. University labs, for example, may have a few staff members to manage all aspects of receiving and analyzing the data. Their focus is getting the analytics and the assay right, less so on the chain of custody. Their approach to sample tracking may not be as robust as what bigger vendors have available.

It's better to approach these smaller vendors with the awareness that resources and data availability can be an issue. One might come to the table with a minimum set of criteria that allows more flexibility on the side receiving the data to accommodate any issues.

If they can extract their data and format it in a way that is more aligned with the data for the rest of the study, the *bare minimum* is still requested to reduce resources.

When working with a larger, more sophisticated lab that has a laboratory management system in place to house their data, it's possible to be more specific about the data points requested, and even more restrictive as to the structure and delivery of the data file. Typically, these vendors have more flexibility in how the data are pulled out of their system, and even some flexibility in how the data are delivered, whether it's by email or transferring data to a secure FTP or loading it into a vendor portal.

For these larger labs, one might start with a more formal set of data transfer specifications that aim to meet the sponsor's data requirements for whatever system the data will ultimately end up in. There should be negotiation on the data format up front.

For example:

- Which columns are allowed to have missing values?
- How are duplicate records handled?
- How often are files transferred to the sponsor?
- How will the vendor manage any queries or questions about the data in terms of data quality or errors — things that must happen at the vendor side because they're the authoritative source?

When is the best time to engage a vendor?

The answer to this is always as soon as possible. Questions about data handling are often overlooked early in the contracting process because the emphasis is on the logistics of testing and storage. Frequently, contracts will specify at most that a data manifest will be provided. But much more detail is needed. It's a good idea to ask a lot of questions at this point. What will be in that manifest? How often will it be delivered?

Sample handling

Early conversations should cover: What sample data does the lab store? For example, visit and time point, receipt date. Who shipped the sample to whom? Is there a destruction date or does the vendor plan to ship any sample remnants back to a central lab or to the sponsor? If so, can they provide shipping and destination information for those samples?

Data delivery

How will sample and subject identifier be captured in the data that they're storing? It may be different for each study. It may be different for different systems that they're using internally. Is it possible to get clarification on that early on? Does the lab have export capabilities? Or is it a manual process for them to pull this type of data together? Can they perform data corrections in their system and how quickly?

Frequency

How often can the data be provided? We want to track samples as close to real time as possible. If a vendor is receiving samples every week, can the data be provided every week? Or if we are only testing twice a year, once we have six patients and all the samples associated with them, can we receive those data feeds then twice a year?

Deciding on these items in advance minimizes any downstream issues with data transfers and can avoid contract changes once the data is needed. There's nothing worse than going back to amend the contract once it's already been set and signed. Updates can sometimes take months to put into place with justifications and budget adjustments. That can ultimately impact the data analysis timelines, with the study team waiting for these changes.

A success story

IQVIA Laboratories Digital Technologies has been working with a large vendor that provides a lot of translational work for one sponsor across multiple studies. Working through the data transfer design and specifications on a study-by-study basis with this vendor was reinventing the wheel every time. Collectively, with the sponsor and the vendor, we worked to create a single data transfer specification — an agreement on exactly how all the data would be formatted and sent. Now that information works for every study from that sponsor. It's very efficient. The vendor was able to tailor their sample tracking exports to align with these specifications and do it in a way that was partially automated and required minimal resourcing.

Now data arrives much faster. Also negotiated was a way to receive all possible data values for each study set up in their system long before samples are received so that IQVIA Laboratories Digital Technologies could have all of the pieces into place as the study begins. Knowing all the values is important so that translation mentioned above can take place and the receiving system can be ready when that first data feed arrives. This would be the ideal case for vendors of any size. It's a tall order and requires some resources upfront but it's well worth it.

IQVIA LABORATORIES

CONTACT US
Toll free: +1 855.277.9929
Direct: +1 919.998.7000
International: +44 (0) 1506 814000
labs.iqvia.com