White Paper

# Control Your Data Flow, Control Your Trial

*The surge of healthcare data threatens to cripple healthcare organizations. But an effective data flow environment can get it under control*

**Tracy Mayer -** VP, CDM & Bios and Connected Devices
**Sabrina Steffen -** Sr Director Clinical Data Management
**David Jackson -** Director Clinical Data Management

# Table of contents

## Data overload

The flood of data pouring into the healthcare industry is creating productivity bottlenecks and introducing new risks that can delay projects and put patient safety at risk. To adapt, life sciences firms need to rethink their entire data environment, adding new technology and talent to transform all that raw data into usable insights.

***Here's how to do it.***

The volume of data pouring into the healthcare industry has the potential to inform every decision we make. But unless organizations figure out how to capture, integrate, analyze and share these data assets, it could also bring their clinical research to a screeching halt.

Consider COVID vaccine trials. During the pandemic sponsors were tasked with recruiting thousands of patients in a matter of weeks, monitoring their outcomes in near real time, and delivering a constant stream of results to regulators. It meant they needed to be able to capture and analyze billions of datapoints within extremely short timelines. If they hadn't had an effective system to capture and analyze all of that data, the trial workflow would have ground to a halt, adding months to the review process, and delaying access to life saving vaccines. It's one example of the constant pressure every clinical trial team faces.

IQVIA research shows that clinical studies now capture three times as much data as they did a decade ago, all of which comes in at a different cadence and in different forms[1]. It can include data from wearable devices, online diaries, labs results, home health nurses, and information captured during site visits. This can translate into thousands of data points per day from a single patient.

Advanced analytics capabilities that use artificial intelligence and machine learning (AI/ML), make it possible for researchers to find vital information within this sea of data about disease trends, response rates, and population outcomes that were previously impossible to uncover. But they can't do it unless they have the right talent, technology and data management practices embedded in the way they plan and deliver trials.

In today's competitive life sciences landscape, the data has to move as fast as the science — and in most organizations it can't. Unless they take the time now to build a comprehensive data flow system, and to add data strategists to every trial team to bring order to that system, they may find themselves falling behind.

Without this data flow infrastructure, life sciences companies should expect trials to experience new and longer delays, miss critical deadlines, and prevent stakeholders from accessing the information they need to make better decisions.

## Your current system doesn't work

The technology currently exists to build a seamless data flow system that can scale with the growth in data flow. But building and maintaining these systems is no small task.

While many life sciences companies now use analytics platforms to study data, the lack of a structured environment and automation limits the functionality of these tools. They acquire data sets with few rules for how it is purchased, stored and managed, with much of it coming in formats that can't be combined or analyzed in any meaningful way without tedious, time-wasting customization on every study.

As a result, most study teams preparing and utilizing the data still spend more than half of their time manually preparing and deploying data[2]. This prevents them from focusing on more value-added activities, while delaying time to analysis. By the time the data is captured, cleaned and combined with other data sets for analysis, its value has already diminished.

Users may also bring their own terminologies, priorities, and biases to the process, which can diminish the data's value. For example, biostatisticians and medical experts may have differing opinions on the same datapoints as they perform cleaning and review for their individual purposes; or are duplicating the review and queries for their individual needs. These kinds of disconnects drive confusion, waste effort, and clog the data flow.

## A single seamless system

To fix this patchwork approach, sponsors need to build a robust data flow environment that delivers consistent, repeatable, and scalable solutions for data management across the organization.

An ideal data flow environment replaces ad-hoc data acquisitions and siloed workflows with a single streamlined system. Any data that is created or acquired flows into a centralized data warehouse or data lake, where it is cleaned, refined, and harmonized. The data is hosted in a shared, access-controlled, and regulated environment where everyone can see all reviews along with all issues resulting from it. The environment adheres to common functional data model and standardization rules. End users can all work collaboratively on the same resource and see what changes are being made in real-time, which optimizes its use, and prevents the emergence of a disorganized data swamp filled with inaccurate or incomplete metadata.

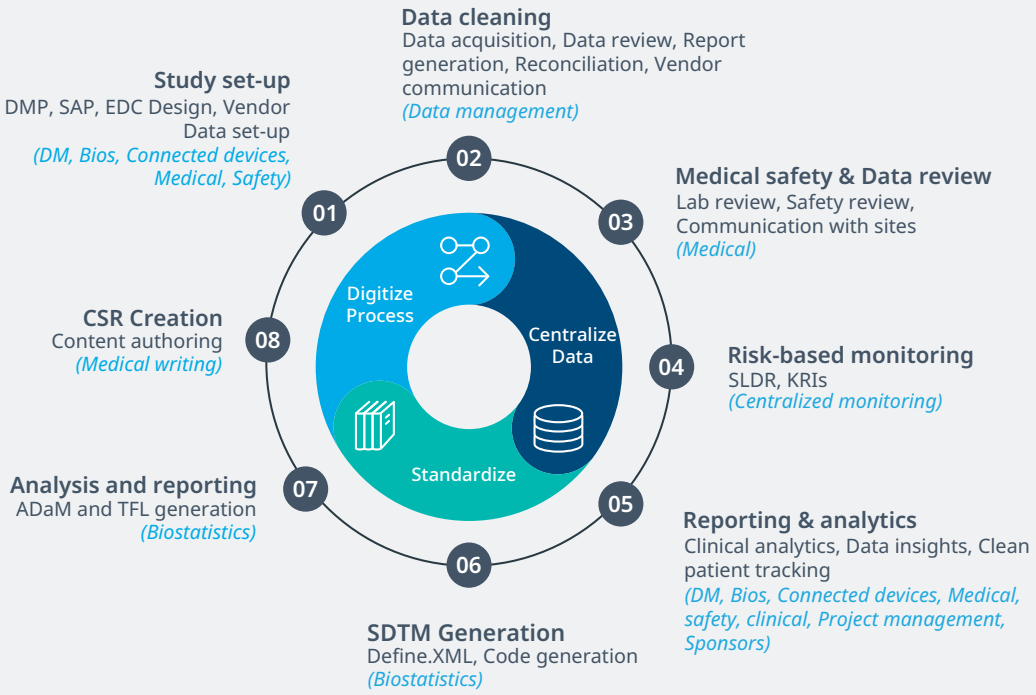
The analysis environment can be further optimized with automated analytic capabilities, and easily shared via user-friendly dashboards that inspire better, faster decision making. Automation reduces variability and rework, and takes over manual data management tasks, freeing staff to spend more time on patient care.

This robust, flexible ecosystem provides a single model to rapidly support new trial types, including adaptive designs and decentralized clinical trials (DCTs) with a data environment built for speed, transparency, and regulatory compliance.

The addition of data strategists, and implementation of complex data flow technology systems is still new in the life sciences industry. But for those leading edge organizations, we are already seeing measurable benefits. In one example, the sponsors saw **a 50% decrease in time to clean data; less than 28 days to database lock, and final study data tabulation model generation within 2-3 days,** all while decreasing burden on site staff.

**IQVIA DATA ECOSYSTEM**

**The data is what connects us — a single ecosystem for all data-centric activities**




**Robust, flexible ecosystem provides a single model to support new trial types focused on speed and data transparency**

- Complex / Adaptive Study Designs
- Decentralized Clinical Trials

## Digitize, centralize, standardize, automate

The data flow transformation requires investment and change but will be necessary if organizations want to generate the best returns from their clinical research. The two biggest drivers of value for a clinical asset is time to market and probability of getting to market. So, the faster companies can turn data into insights, the more value they will generate from their efforts.

The companies that are willing to rethink their data management process, and invest in the talent and technology required to create end-to-end data flow environment, stand to gain significant advantages now and in the future.

IQVIA has optimized clinical data flow environments to the benefit of sponsors, sites and patients. They have developed a four-step innovation process to streamline the connected data flow environment.

**Digitalization.** Many clinical data assets still exist as Word documents, spreadsheets, emails, handwritten notes, and natural language narratives. Before any data can be part of the analysis environment, it must be turned into digital asset that is ready to be formatted, cleaned, and merged with the rest of the data environment.

**Centralization.** After it is digitized, all data and issues be stored in a central warehouse or data lake that acts as a single repository for all data assets. Keeping data in a central location prevents duplicate acquisitions, ensures rigor in the way data is cleaned and handled, and makes it easy for teams across the organization to leverage the same resources in their decision making process.

**Standardization.** Every asset is created in different formats, many of which are unstructured, e.g. physician's notes, customer service requests, images and audio files. They also use different rules, codes, and naming conventions, which are relevant in isolation but create challenges for deeper analysis. Before these assets can be combined for analysis, they need to be adapted to a single set of business rules, including deidentification. These adaptations should occur as close to source in the data flow chain as possible and connected to standards along the data flow to create a consistent end to-end-model. A metadata repository makes it possible to govern the end-to-end standards in a quality streamlined way. When defining these rules, it is important to include end user feedback about how the data will be used, to ensure it maximizes value without compromising patient privacy.

**Automation.** Automation driven by AI/ML will streamline the entire data flow process, from acquisition through final analysis. This makes it easy for end users to query data assets asking detailed questions that generate real-time results. Automation also enables them to dig deeper, asking follow-up questions, and updating results when new data flows in, all in near real time. Ideally this automation will also include graphical dashboards that make the results easy to interpret for all users (not just data scientists).

Once all of these elements are in place, companies can propagate the data to teams, providing easy access with the latest updates so they generate insights that are relevant to each group. This flow of data offers the opportunity to identify trends, make decisions faster, and to support more agile development strategies. This includes adaptive trial designs, that allow study planners to test progressive treatment modifications, and variation in eligibility while continuing to meet safety and quality data requirements.

## Data strategist to the rescue

Achieving this seamless data flow environment requires more than just technology investments. Any organization can buy the best solution in the world, but if no-one on the trial team knows how to manage it, it won't add value.

If sponsors want to streamline the way, they capture and analyze trial data they need to add data strategy experts to the clinical trial team.

These experts aren't there to provide help desk support, or to monitor data. They develop a consistent, repeatable, and scalable strategy for the entire end-to-end data journey. Initially, they will lead a team to define, acquire, and assemble the layers of data technology along with the standards and processes needed to create this seamless data flow environment. Sponsors are enabled to rapidly capture disparate data sets in a central location for cleaning and analysis, while adhering to regulations that govern its use and capitalizing on standardization downstream of those capture systems.

They then function as the data curator and coach through the entire study conduct, ensuring a collaborative and efficient data experience is maintained for all end users, so they in turn can access relevant insights while meeting all regulatory data requirements.

In this role, they ensure the data plan is effectively executed, and provide guidance throughout the trial, so the rest of the team can focus on the downstream operational use of the data (ex: data cleaning, medical monitoring, statistical analysis, etc.).

In the past, data flow management was an added task that fell to the project leader. But the quantity and complexity of data that clinical trials now gather, the standards and processes required to drive value and increased efficiency from this volume of data, and the pressure they face to manage it faster while maintaining quality, has created the need for a new dedicated data management professional to join the team.

For smaller organizations, a single data strategist can function in an oversight role, managing the data streams for every trial. However, for complex studies with many participants and sources of data, having a dedicated data strategist is essential in avoiding data-related pitfalls that can add costly delays to the process.

**NOW HIRING: CLINICAL TRIAL DATA STRATEGISTS**

In the future, data strategists will become a core part of the clinical trial team. Here's what the job description might include:

- Provide oversight of the set-up and integration of individual systems (EDC, IRT, connected devices, non-EDC vendor data, etc), ensuring data standards are followed leading to successful integration of systems to IQVIA CDR and connected services
- Liaise with IT to build custom data integration solutions
- Establish agreements with vendors on data flow technologies, formats, structures, and content
- Ensure the data collected and integrations can be successfully utilized in all downstream systems by First Patient In
- Work with downstream teams and Project Management throughout the entire study conduct, ensuring data integration allows for continuous data flow while informing all connected teams of change impact via the metadata repository tools
- Champion people, process and technology improvements and standards to enhance quality, increase productivity, and reduce data latency within clinical trial systems
- Proactively provide education and training to ensure adherence of the relevant start-up SOPs and best practices during the RFI and post-award phases over data flow strategy

The data strategist will guide the development of the entire end-to-end data ecosystem, ensuring it provides a robust, agile network through which the data can easily flow and to which users consistently find value and insights with ease.

## No time to waste

The deluge of data has the potential to drive innovation in the life sciences space. When drug developers have the power to combine and analyze massive healthcare data sets, researchers can uncover insights that can speed development of the right therapies to save lives, improve the patient experience, and drive new sources of revenue for the industry.

It can be overwhelming to think about how to transition the data flow from a siloed environment, with dozens of disparate resources that all have infrequent updates, to a real time, streamlined shared experience where you can turn those billions of data points into real insights. But it can be accomplished when companies work with the right vendors who have industry knowledge, global healthcare data assets, and technical expertise to build and deploy comprehensive data management systems that accommodate the needs of every user.

*To learn more about IQVIA's data flow process, and how our experts can help you transform your data management into a streamlined collaborative environment,* www.iqvia.com/datamanagement.

[1] Tufts Center for the Study of Drug Development. "Rising Protocol Design Complexity Is Driving Rapid Growth in Clinical Trial Data Volume" Tufts CSDD Impact Report Jan/Feb 2021. https://csdd.tufts.edu/

[2] https://blogs.idc.com/2021/06/23/the-data-dilemma-and-its-impact-on-ai-in-healthcare-and-life-sciences/

**CONTACT US**

**iqvia.com**