



ABSTRACT

With the constant rush in clinical trials to "clean" data, CROs are looking for smart ways to "clean" data. When digitized, centralized data flows and standards connect to provide an integrated, streamlined, and automated data management process, it evolves into a tech-enabled process called Real-Time Data Cleaning. Continuous data ingestion and automated cleaning, results in shortened database locks, issue detection and mitigation. The exponential increase in the type of data sources together with the ever-increasing complexities of trial designs require a new methodology to meet these challenges efficiently while maintaining high quality, without raising costs. The obvious solution is automation. Automation requires electronic repeatability, digitized consistency, standards and compliance. Digitized processes leverage connected data and standards to drive data management, reducing manual processes, ensuring consistency, and increasing efficiency leading to faster identification of issues to enable companies to make faster decisions in clinical trials, and enable data insights.

RTDC ESSENTIALS

Real-Time Data Cleaning (RTDC) requires the use of inhouse tools for automated data cleaning. The centralized data flow strategy leverages the experience and technology to drive data standards, manage data as an asset and facilitate greater data transparency to help the trial gain improved analytics and insight and make study critical decisions faster. With the complexity of clinical trials, it becomes imperative to work with accurate and error free data in real time which is crucial to ensure business continuity. RTDC is transparent and in real-time, where all team members across functions work on the same data and see all issues at the same time. RTDC cleans data daily, combining standards, data, and operations to provide a more integrated and streamlined data cleaning process thus optimizing cycle times, and increases accountability.

END-TO-END DATA ECOSYSTEM

100% Data Ingestion Automation **Data-Driven Standardized Process** **Real-Time Data Cleaning** **Real-Time Flow & Access**

WHAT IS REAL TIME DATA CLEANING?

Real-Time Data Cleaning (RTDC) is an intelligent, data-driven, continuous, automated check of process flows to track data sources across all steps of the data flow eco system. For sponsors, it means reduced data errors, greater efficiencies, increased patient safety, and powerful oversight of data flow.

	Traditional	RTDC
Data Acquired (EDC, Vendors)	Manual data ingestion 30 Days	Automated data ingestion* 1 Day
Issues Generated	Manual data ingestion 2 Days	Automated generation of data check programs 0 Day
Issues Reviewed & Queries Opened	Review 30 days worth of data issues 14 Days	Review 1 day worth of data issues 1-3 Days
Query Answered by Site	NA – Site activity 10 Days	NA – Site activity 10 Days
Vendor Issues Escalated	Collate 1x/month in excel 14 Days	Escalate individual issues as sites respond 1-3 Days
Data Issues Addressed by Vendor	Process 1x/month in excel 14 Days	Vendor access to workflow system, address issues as escalated 1-5 Days
Total Data Cleaning Cycle Time	84 Days	14 - 25 Days

* Assumes daily vendor transfers

Continuous cleaning facilitates accelerated preparation for IAs **Database Lock Decreased from 60+ Days to <28 Days (all phases)**

Measuring the benefits of Real-Time Data Cleaning process

100% Data ingestion automation	\$0 Budget to increase to daily data transfers	20% Decrease in programming cycle	0 Omission of error detection	50% Efficiency to manage data issues
--	--	---	---	--

REFERENCES

<https://www.iqvia.com/>