

# Impact of data management on clinical trials: new study

Recent research found a correlation between the upfront time to build and release a clinical database and its impact on downstream data management processes in conducting and completing trials. **Richard Young** reviews the study findings and offers guidance to improve the process.

Clinical trials continue to increase in complexity and scope. A typical phase 3 protocol, for example, now has many more endpoints, procedures, and data points collected compared to a decade ago.<sup>1</sup> At the same time, data management processes have become more complicated, as contract research organisations (CROs) and sponsors manage a variety of clinical trial data. Real-world evidence, electronic clinical outcome assessments, mobile device-driven data, social media communities, and electronic health and medical records are some of the new data sources now captured during clinical trials.

The volume and diversity of data presents integration, compatibility, and interoperability challenges that the pharma industry must address in order to optimise drug development.



A new survey we conducted with the US Tufts Center for the Study of Drug Development (CSDD) examines the state of clinical data management in life sciences and its impact on drug development. One of the largest, most in-depth studies of clinical data management professionals, the [2017 eClinical Landscape Study](#) found a correlation between the upfront time to build and release the clinical database and its impact on downstream data management processes in conducting and completing trials.

“The study results indicate that companies face a growing number of challenges in building and managing clinical study databases,” said Ken Getz, research associate professor and director at the Tufts CSDD.

## Database build delays have downstream impact

Typically there is a timescale of 60 days between submitting a protocol for finalisation and targeting first patient, first visit (FPFV). The new Tufts research shows that, on average, despite this significant window, the industry is not hitting that milestone. This is mainly because clinical trial database design is a slow, inflexible process, which has remained largely unchanged over the past 20 years.

### 2017 eClinical Landscape Study

One of the largest, most in-depth surveys of clinical digital management professionals shows that the time it takes companies to design and release clinical study databases is having a negative impact on conducting and completing trials. In addition to analysing data in the aggregate, Tufts broke the data down into several subgroups:

- Company Type
  - Sponsors (N=193)
  - CROs (N=56)
- Company Size (Total Annual Clinical Trial Volume)
  - Low: < 5 trials, median = 2 (N=84)
  - Medium: 5–15 trials, median = 8 (N=80)
  - High: > 15 trials, median = 50 (N=93)
- Primary EDC Provider
  - Industry Leaders (Medidata & Oracle) (N=144)
  - All Others (N=54)

Most (85%) life sciences organisations release some, if not all, of their clinical study databases after FPFV. These delays create issues themselves, but the research goes on to demonstrate that the challenges continue throughout the trial. Delays in releasing the study database are associated with downstream delays of up to a month in key data management activities, including patient data entry throughout the trial, up to last patient, last visit (LPLV) and database lock. Getz explains, “Our research shows that the release of the clinical study database after sites have begun enrolment is associated with longer downstream cycle times at the investigative site and at study close-out.”

Initial database delays also impact the time it takes sites to enter patient data into the electronic data capture (EDC) system. When the database is always released *before* FPFV, data-entry time is five

days. When the database is released *after* FPFV, ongoing patient data collection doubles to at least 10 days.

The impact of database-build delays is even greater by the time companies get to database lock. For companies that deliver the database *after* FPFV, it takes roughly 75% longer to lock the study database than for those that deliver the final database *before* FPFV. These timelines for data management activities are just as long today as they were 25 years ago, despite advances in technology (Figure 1).

Figure 1

Downstream Impact of EDC Release after First Patient, First Visit (FPFV)

Frequency	Percent	Time from Patient Visit to Data Entry	Time from LPLV to Database Lock
Never (N=39)	15%	5 Days	31 Days
Rarely (N=135)	53%	8 Days	34 Days
Often (N=75)	27%	10 Days	42 Days
Always (N=7)	3%	10 Days	54 Days

Source: 2017 eClinical Landscape Study, Tufts Center for the Study of Drug Development, September 2017

The survey finds several common causes for clinical database-build delays. Protocol changes are most often cited by respondents (45%), underscoring the challenge data management professionals have in dealing with changes as they are finalising the clinical trial database for the start of the trial. This highlights the need to optimise the database design process with standards and systems that support more flexible design and rapid development.

More than three-quarters of respondents (77%) also say they have issues loading data into their EDC applications, and that EDC system or integration issues are the primary reasons they are unable to load study data (65%).

## More data sources, more challenges

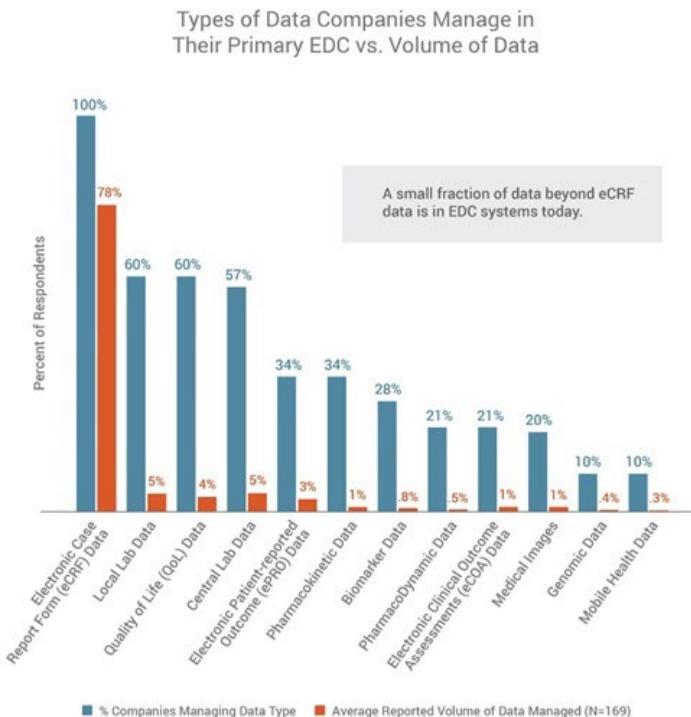
The survey found that EDC is the most widely-adopted clinical application, used by all respondents (100%), followed by randomisation and trial-supply management (77%), electronic trial master file (eTMF) (70%), and safety systems (70%). All (100%) CROs and sponsors confirmed that their organisations were using their primary EDC systems to manage electronic case report form (eCRF) data, followed by local lab and quality-of-life data (60% each).

The research also confirmed that companies are collecting diverse, heterogeneous sets of data elements throughout trials – from eCRF data to social media and online community information, real-world data, mHealth data, imaging, lab data, and data from other mobile applications.

However, when it comes to *volume* of data managed in a company’s EDC, the EDC is used predominantly to manage eCRF data, while just a small portion use their EDC to manage non-eCRF data as well. In fact, eCRF data represents 78% of the data managed in a company’s EDC. All other data types represent 5% or less each of the total data volume, including mobile-health data

(Figure 2). This demonstrates the need for processes and systems to better support the industry's vision to have complete study data in EDC.

Figure 2



"EDC has been widely adopted to accommodate the management of electronic clinical data, but this landscape is changing dramatically," added Getz. "EDC systems have been handling increasingly complex global clinical trials requiring the collection and management of very large volumes of data. And data is coming from a wide variety of sources, including not only case-report forms, but also mobile devices, biomarker and genetic data, and social media. Our research is examining this changing landscape and the projected impact that this will have on data management practices."

### A way forward for faster trials

Clinical data is highly fragmented and often managed in system silos. As study designs become more complex, clinical data management will play a critical role in trial success, especially in dealing with protocol changes when finalising the clinical trial database for the start of the trial. While protocol design changes are the most frequent event to impact timelines, they tend to have the least-observed impact on the cycle time from LPLV to database lock. Meanwhile, changes to the database design itself and its functionality have a much more significant impact on cycle time at the end of the study. The survey qualifies and quantifies the differences in the impact that these various changes or causes might have on cycle times (Figure 3).

Figure 3

Top Causes of Database Build Delays

	Percent of Total (N=257)	Time from LPLV Database Lock
Protocol Changes	45%	32 Days
User Acceptance Testing (includes review and approvals)	17%	34 Days
Database Design Functionality	15%	50 Days
Study Database Move from Development into Production	8%	39 Days
Standards Management	4%	38 Days
Ethics Approval Delays/Changes	1%	33 Days

Source: 2017 eClinical Landscape Study, Tufts Center for the Study of Drug Development, September 2017

Reducing the time it takes to build and release the clinical database can have a positive impact on subsequent trial timelines and, ultimately, lead to developing treatments more quickly, effectively, and safely. The Tufts research also provides an industry benchmark of more than 250 life sciences companies. But, possibly most significant is that the research can also serve as reference data for the industry to set new, far more aggressive, goals in the future.

Overall, a modern approach is needed in clinical data management to address the growing complexity and volume of data in clinical trials today. Better data management will lead to smarter, real-time decision making during trials – not after they are completed. Next-generation strategies and solutions will give clinical data management teams the flexibility to design complex studies faster and handle protocol changes with little to no downtime. Ultimately, the best approach will bring together data from every source – from patient to regulator – and in real time, ensuring that every observation, result, and event is captured as it occurs.

### Reference

['Getz: Site Activations Hurt By Commodity Mentality](#), Clinical Leader, 16 May 2016.

**Reprinted with permission from**  
[www.pharmaphorum.com](http://www.pharmaphorum.com)

### About the author



**Richard Young** is vice president for Veeva Vault EDC. With almost 25 years of experience in life sciences, he is known for his experience in data management, eClinical solutions, and advanced clinical strategies.

Most recently, Richard served as vice president of global consulting partners at Medidata Solutions. He consults on adaptive trials, risk-based monitoring, big data, mobile health, and other major data strategies.

[Veeva Systems](#) is a leader in cloud-based software for the global life sciences industry and has more than 550 customers, ranging from the world's largest pharmaceutical companies to emerging biotechs. The [Tufts CSDD](#) at Tufts University provides strategic information to help drug developers, regulators, and policy makers improve the quality and efficiency of pharmaceutical development, review, and utilisation.