Rezzan Hekmat, MHS

Nawshin Kutub, PhD

Courtney VanHouten, MA

Ryad Ramda, MS

Van C. Willis, PhD

Gretchen P. Jackson, MD, PhD

Cindy Henderson

Dilhan Weeraratner, PhD

Robert A. DiCicco, PharmD

Jane L. Snowdon, PhD, FAMIA

Accelerating Clinical Trial Design and Operations

Fully-integrated, component-based CDMS offers flexibility, customization, and efficiency

Effective and efficient clinical data collection and management is one of the key factors affecting clinical trial success and is of heightened importance during the COVID-19 pandemic.[1] Not only have the scope and complexity of clinical trials continued to increase over the past decade, but the volume and diversity of clinical study data grows ever larger. As routine drug development cycle times continue to get longer, the challenges of attempting to accelerate study start up and database lock and reporting while maintaining quality has become harder.

Researchers have accelerated the development of vaccines and therapeutics for COVID-19 as evidenced by the 4,846 trials found on clinicaltrials.gov.[2] The COVID-19 pandemic presents a unique opportunity for understanding technologies that can enable trial data management and their effects on streamlining and expediting clinical trial design and implementation.

Goals and objectives

This study aims to evaluate how a clinical data management system (CDMS) supported COVID-19 vaccine and therapeutic drug clinical trials based on the experiences of a contract research organization (CRO) during the pandemic.

The study objectives were to evaluate how the CDMS was used during the conduct of two COVID-19 trials and identify any potential impact on clinical trial efficiency and/or CDMS usability and workflow. We examined data management processes and explored how the CDMS supported optimizing workflow in clinical trials during the COVID-19 pandemic and periods of remote work; identified perspectives regarding data management challenges either specific to or exacerbated by the COVID-19 pandemic; and measured usability, acceptability, and satisfaction.

Methodology and approach

We evaluated the two COVID-19 clinical trials conducted by the contract research organization (CRO), Veristat, as of the date of this research. IBM Clinical Development (ICD) was the data management system utilized. Project-level retrospective and prospective internal usage and administrative data, routinely collected to monitor utilization of the CDMS, were analyzed for each clinical trial. Internal data from usage reports included dates for key data management events such as database release and lock, module usage, changes to the database after its release (referred to as mid-study updates), and database downtime. The CRO provided additional operational and study management metrics including data management information, outlined in Table 1 on the facing page, to supplement CDMS usage data.

Specific measurements evaluated in this study were: database build efficiency, velocity of data collection, frequency of protocol amendments on the database, and the time impact of mid-study updates to the database. The cycle time to design and release the study database in the CDMS was used as an indicator for efficiency in database build. Specifically, this was defined as the number of weekdays between the date of the final approved protocol to the date of the first database "Go Live." The average cycle time to build these COVID-19 study databases was compared to both the CRO's and published industry baselines prior to the pandemic to establish study efficiency. Velocity of data collection was measured as the number of data points collected over time after the database was released.

Data managers completed a brief form on each protocol amendment and mid-study update to the database, providing the date and background information for the event, the reason for the protocol amendment, and an indication of whether a protocol amendment and mid-study update were related. Average database downtime was used to indicate impact of mid-study updates on the trial timeline. Database downtime refers to the number of days the database was "Locked" and no data collection could take place.

A survey was designed to investigate users' attitudes, experiences (e.g., impact on cycle times, efficiency and cost) and satisfaction with the CDMS's ability to meet the COVID-19 clinical trials' needs. A web-based survey composed of 37

STUDY DEFINITIONS

Milestones	Other major study and data management milestones or events not available in usage reports and their dates.
Study Design	Information on the trial design, including the scientific and operating characterisitics primarily found in protocal such as objectives, design, eligibility criteria, sites, procedures or other assessments, visits and schedules, and endpoints.
Database Design	Documentation of the database design for the study, including the study process, the casebook design, the individual page designs and system queries, and any report designed.
Protocol Amendments and Database Updates	A series of questions regarding each protocal amendment or database mid-study update, including classifying the cause of the change, whether the change could be avoided, and its impact.

TABLE 1. OPERATIONAL AND STUDY MANAGEMENT METRICS

SOURCE: IBM, Veristat

questions divided into seven sections was used to collect survey data. These sections included questions covering the System Usability Scale (SUS), a Net Promoter Score (NPS), general clinical trial data management, general use of the CDMS before and during the COVID-19 pandemic, specific use related to the COVID-19 trial being run in the CDMS, and demographics, such as years of experience in the field and years of experience using CDMS tools. SUS is a simple, ten-item attitude Likert scale that provides a subjective assessment of usability.[3] NPS is a well-established measure to assess customer satisfaction on a scale from 1 (not at all likely) to 10 (extremely likely) by asking the single question "How likely is it that you would recommend our system to a friend or colleague?"[4] The survey was distributed to the CRO's data managers and was open from July 27 to August 7, 2020. Descriptive methods for analyzing the objective data were a five-point Likert scale and cause-and-effect relationships.

Interviews were conducted with the lead data managers and CDMS users from each clinical trial in order to highlight unique challenges or approaches for their specific COVID-19 clinical trial. Interviews were conducted via a secure online video communication instrument, used an approved interview guide, and each lasted approximately 60 minutes. Using an open coding approach, interviews were recorded, transcribed, and analyzed. Thematic analysis was defined by a combination of deductive review of the direct responses to interview questions and inductive analysis to identify emerging patterns. The interview guide is given in the Supplement. The study was submitted to the Western Institutional Review Board for regulatory review and received an exempt status determination. Participants provided informed consent for both the survey and interview.

Results

Two COVID-19 clinical trials were included in this study: a therapeutic drug and a vaccine. The vaccine trial is a Phase I open-label, non-randomized study evaluating the safety, tolerability, and immunogenicity of the intervention in 120 healthy volunteers across three US sites. The therapeutic drug trial is a Phase II open label, randomized control, proof-of-concept study enrolling 50 patients hospitalized with COVID-19 across eight US sites.

Figure 1 on the following page depicts key clinical data management milestones and the volume of data over time for the two COVID-19 clinical trials. The average time to design and release the study database for a COVID-19 trial was 11 business days. This was a 63.3-72.5% reduction compared to the CRO's typical 30-40 day timeline and an 83.9% reduction compared to an industry benchmark of 68.3 days for a non-COVID trial.[5] Mid-study updates to the database after first release were primarily a result of protocol amendments. Both trials had three mid-study updates during the period evaluated. On average, the database was only locked for less than one hour per mid-study update in each trial. Data managers classified all protocol amendments and mid-study updates as "completely unavoidable."

Four (100%) clinical data managers and data management professionals were contacted, and all completed the survey; 75% were 25 to 34 years old. Each participant managed three to four trials simultaneously and used the CDMS on average 25.75 hours / week (range 20 – 30). Years
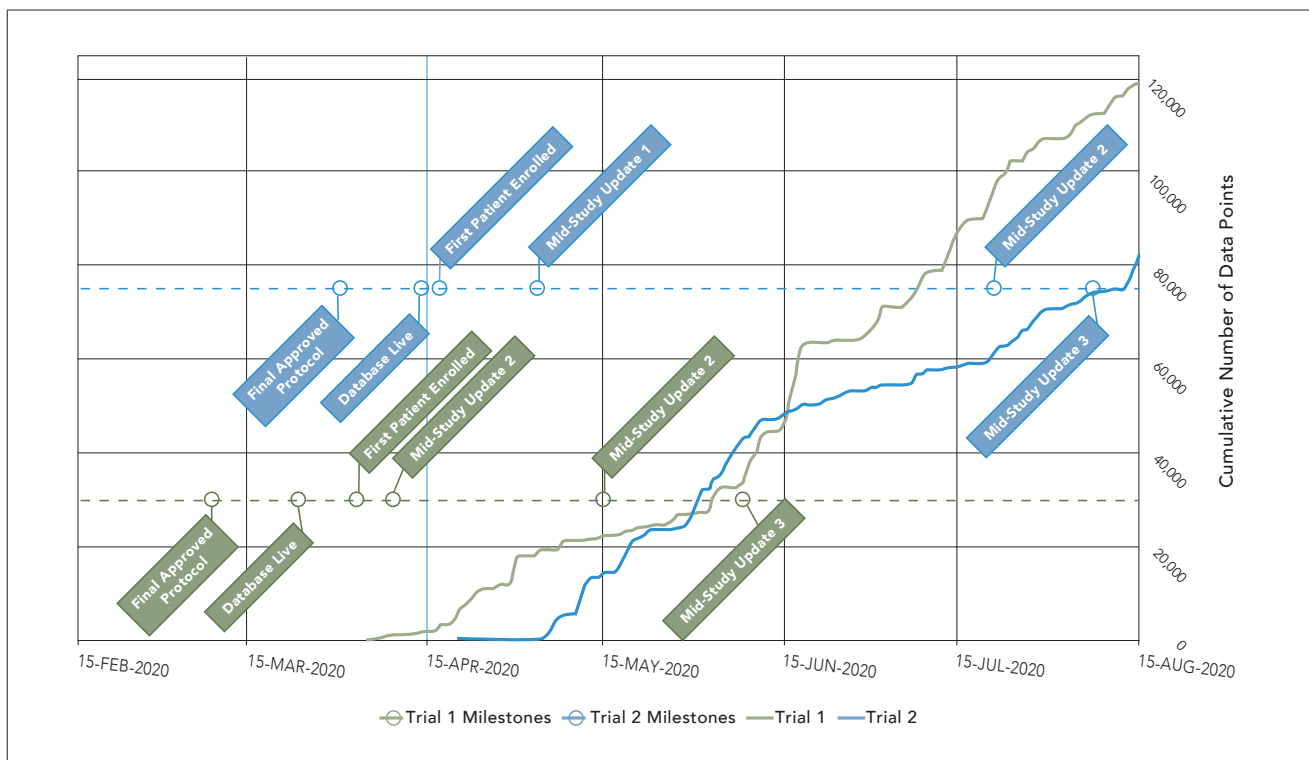
TRIAL MILESTONES

FIGURE 1. KEY CLINICAL DATA MANAGEMENT MILESTONES AND THE VOLUME OF DATA OVER TIME FOR TWO COVID-19 CLINICAL TRIALS

SOURCE: IBM, Veristat

of experience using the CDMS were one (50%), two (25%), and five (25%). The average SUS score was 80.6 (range 77.5 to 85.0 and considered excellent. NPS was 75% with 0% detractors, which is also considered excellent.

Table 2 on the following page reports how much time the CDMS saves performing various tasks compared to other solutions used where 100% of participants believed the CDMS either somewhat or very much saved time for study startup, data collection, study closeout, and overall. All (100%) survey respondents agree or strongly agree that the CDMS integrates with workflow, increases productivity, is interoperable, and research studies are easy to set up and start up in the CDMS, while 75% think the CDMS is flexible and customizable and reduces manual errors. Of the survey respondents, 100% thought the CDMS somewhat or quite a bit helped to avoid cost, reduce manual errors, and decrease delays from mid-study updates. The most commonly reported causes for delays when building a clinical trial database were protocol changes

(50%) and user acceptance testing (50%).

Interviews highlighted two specific functions as the main factors contributing to the rapid launch of study databases. First, the system easily allows reuse of code for database creation. The users created and routinely consulted a template library to generate new customized databases more efficiently. Bypassing de novo coding for both common and specialized database creation allowed users to focus efforts on the specific customizations required by the study sponsor.

Second, the ability to execute mid-study updates or post-go live changes with minimal to no downtime (< one hour) allowed the users and sponsors to work in parallel rather than in serial fashion, speeding up trial start dates and implementation of protocol amendments, and accommodating adaptive COVID-19 trial design. The ability to copy portions or the entirety of a database in a single environment permits users to 'split' the release of the trial database and allow the sponsor to initiate trial startup (e.g., screening and randomization) while the user completed

CDMS UTILIZATION

Overall, how much time, if any, do you believe ICD saves you during the following tasks compared to other solutions you use?	1 - Not applicable	2 - Not at All	3 - A little	4 - Somewhat	5 - Very Much
Study Startup				3 (75%)	1 (25%)
Data Collection				4 (100%)	
Data Processing and Cleaning			1 (25%)	2 (50%)	1 (25%)
Database Mid-Study Updates			1 (25%)	3 (75%)	
Study Closeout				3 (75%)	1 (25%)
Overall from Startup to Closeout				4 (100%)	
Indicate how strongly you agree or disagree with the statement	**Strongly Disagree**	**Disagree**	**Agree**	**Strongly Agree**	
I feel ICD accomodates individual needs (flexible and customizable)		1 (25%)	2 (50%)	1 (25%)	
I feel ICD integrates with my workflow			3 (75%)	1 (25%)	
I think ICD increases my productivity			3 (75%)	1 (25%)	
I think research studies are easy to setup and startup in ICD			2 (50%)	2 (50%)	
I think ICD has the ability to exchange with and make use of information from other tools, i.e. interoperable			4 (100%)		
I think ICD reduces manual errors		1 (25%)	3 (75%)		
Overall, how much impact, if any, do you believer mid-study updates on ICD have on the following?	**1 - Not at All**	**2 - A little**	**3 - Somewhat**	**4 - Quite a Bit**	**5 - Very Much**
Cost Avoidance			2 (50%)	2 (50%)	
Reduction of Manual Errors			2 (50%)	2 (50%)	
Reduction in Delays			1 (25%)	3 (75%)	

TABLE 2. CDMS USABILITY, TIME SAVINGS, AND EFFECTS ON MID-STUDY UPDATES (N=4)

the database without risk of error or corruption of data already collected. This is a direct attribute of having a single codebase for all applications of the CDMS.

Furthermore, the mid-study update feature continued to be critical to the flexibility of the database after study startup. Participants reported decreased downtime for mid study updates. In fact, usage data revealed that databases in these trials were only locked on average for about one hour in order to execute a mid-study update.

Conclusions

The sudden advent of the COVID-19 pandemic caused unprecedented stress to global health systems due to the novelty of the disease, rapid societal transmission, and absence of established therapies or vaccines. This global exigency prompted pharmaceutical companies and spon-

sors to fast-track the clinical trial clock to expediently bring effective therapies and vaccines to patients. One integral component of accelerating the clinical trial timeline is deployment of an efficient data management system. Utilization of an end-to-end solution with multiple embedded functionalities built on a single code base, including randomization, endpoint adjudication, electronic clinical outcome assessment, inventory management and medical coding within one platform was one supporting factor to the rapid execution of the trials in this study. In addition, the flexibility of the EDC platform facilitated efficient mid-study updates and post-go-live protocol amendments which are inherent in an adaptive clinical trial. The CDMS also enabled seamless transition to remote work during the pandemic without disruption to data entry and patient enrollment. Users highlighted that CDMS compartmentalization facilitated quick study launch and enabled database splitting, which also allowed parallel processing of downstream workflows to further accelerate study start-up.

This study is limited by the small sample size of survey respondents and clinical trials assessed. Many other factors in addition to the specific data management functionalities and practices utilized in the trials could have contributed to compressing COVID-19 clinical trial timelines; these include speed up in regulatory review, efficient protocols, reallocation of organization priorities and resources, exceptional efforts from the research community, and extraordinary awareness and participation from subjects. Evaluation of additional COVID-19 trials and CDMS users is needed to clarify the contribution of each of these factors.

References

1. World Health Organization. Coronavirus disease (COVID-19) situation reports. https://covid19.who.int/
2. NIH U. S. National Library of Medicine, https://clinicaltrials.gov/ct2/results?cond=COVID-19, accessed February 23, 2021.
3. Brooke J. SUS: a "quick and dirty" usability scale. In P. W. Jordan; B. Thomas; B. A. Weerdmeester; A. L. McClelland (eds.). Usability Evaluation in Industry. London: Taylor and Francis. 1986.
4. Reichheld F.F. The one number you need to grow. Harvard Business Review. 2003;81(12):46–55.
5. Wilkinson M, Young R, Harper B, Machion B, Getz K. Baseline Assessment of the Evolving 2017 eClinical Landscape. Ther Innov Regul Sci. 2019 Jan;53(1):71-80. doi: 10.1177/2168479018769292. Epub 2018 Apr 26. PMID: 29714600.

About the authors

Rezzan Hekmat, MHS is a Biomedical Program Manager in the Center for Artificial Intelligence, Research and Evaluation at IBM Watson Health, Nawshin Kutub, PhD, Chief of Staff and Center Program Manager, Center for Artificial Intelligence, Research and Evaluation, IBM Watson Health, Courtney VanHouten, MA, Medical Ethnographer, Center for Artificial Intelligence, Research and Evaluation, IBM Watson Health, Ryad Ramda, MS, is Associate Director Data Management, Veristat, Van C. Willis, PhD is a Senior Biomedical Writer in the Center for Artificial Intelligence, Research and Evaluation at IBM Watson Health, Gretchen P. Jackson, MD, PhD is the Chief Data Scientist and Vice President in the Center for Artificial Intelligence, Research and Evaluation at IBM Watson Health, Cindy Henderson is Executive Vice President, Strategic Development, at Veristat, Dilhan Weeraratne, PhD is a Genomics Expert in the Center for Artificial Intelligence, Research and Evaluation at IBM Watson Health, Robert A. DiCicco, PharmD, is a Deputy Chief Health Officer at IBM Watson Health, Jane L. Snowdon, PhD, FAMIA is Associate Chief Health Officer in the Center for Artificial Intelligence, Research and Evaluation at IBM Watson Health and an IBM Industry Academy Member.