International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 11 November 2025
Page No: 4031- 4036
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0590

Advancing Clinical Data Management: From EDC Discrepancy Resolution to Compliance and Data Integrity Across the Trial Lifecycle

Lakshmi Priya Darshini Pulavarthi[1]
[1]*Southeast Missouri State University, Cape Girardeau, MO*
Email: *lakshmi.pulavarthi0@gmail.com*[1]

Abstract
The clinical trial continuum requires the use of Clinical Data Management (CDM) to facilitate data quality, consistency and adherence to regulatory standards and requirements. Traditional manual data conflict detection and correction may be inefficient and time consuming, which can easily damage the quality, integrity, and timely locking of the database. This paper presents an automated Electronic Data Capture (EDC) discrepancy resolution and compliance verification platform that can be used to enhance data management workflow efficiency and regulatory compliance. The proposed three-stage workflow consists of a Query Detection Layer to identify discrepancies in data, an Automated Classification Engine that applies rules for triaged and automated management of discrepancies and disputes, and a Compliance Verification Module that checks/evaluates the audit trail to ensure that discrepancies resolved meet 21 CFR Part 11 and ICH E6(R3) guidelines. Results from the analysis of datasets, from multi-center clinical trials, indicated that the automated discrepancy resolution process was able to resolve 81.5% of discrepancies identified, reduce the time researchers would need to review discrepancies and differences by 35%-40% per query, and achieve 95% compliance with acceptable standards of accuracy for the verification outputs. Overall, the results of this exploratory investigation indicate that the automated processes embedded within EDC-focused CDM processes could significantly improve data quality, compliance and efficiency, providing an important platform for sustainable, technological enabled data governance throughout the life-cycle of a clinical study.
Keywords: *Clinical Data Management; Electronic Data Capture (EDC); Discrepancy Resolution; Data Integrity; Regulatory Compliance; ICH E6(R3); 21 CFR Part 11; Audit Trail; Machine Learning; Clinical Trials.*

1. Introduction
Clinical Data Management (CDM) is important in the clinical research process, as it verifies that data gathered in clinical trials is accurate, consistent, and adheres to regulatory guidelines. With the onset of rapid digitization in healthcare and the development of decentralized and hybrid models of clinical trials, the range and complexity of CDM has expanded significantly (Wang & Zhang, 2022; Nawaz & Ahmad, 2024; GCDMP, 2024). Electronic Data Capture (EDC) systems have replaced traditional data collection from paper forms with each developing more efficiency and traceability across study sites. However, with the amount and the range of data components (e.g., electronic health records (EHRs), devices, and patient-reported outcomes) increasing, the field has also faced increasing challenges of data reconciliation, integration, and quality assurance (Nahm Zozus et al., 2024). Discrepancy resolution is still a key area of focus for the CDM environment, as errors and inconsistencies across your datasets can diminish the integrity of your data and potentially slow down the progression of important trial key milestones. Recent systematic reviews have shown that, even with advancements in technology, data quality remains a continuing challenge due to heterogeneous data standards, incomplete audit trails and fragmented workflows (Challenges for Data Quality in the Clinical Data

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 11 November 2025
Page No: 4031- 4036
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0590

Life Cycle, 2024). Automated edit checks, risk-based monitoring (RBM) and AI-driven anomaly detection tools for data quality assurance are seeing increased use for improved accuracy and reduced human oversight (Purri et al., 2025; AI in Clinical Data Management, 2025). On the same note, it is equally important to keep regulatory compliance with the lifecycle of the trial. These global standards like ICH E6(R3) and 21 CFR Part 11 point to the need to have traceability, transparency, and auditability in the use of electronic systems. End-to-end data integrity is the alignment of technological platforms, governance systems and standardised mechanisms of operation that bring trust and reproducibility into the clinical research environment (Integrity of Randomised Controlled Trials: Challenges and Solutions, 2020). This paper is designed to investigate how CDM has progressed through the years from an EDC based solution to an integrated technology enabled system which has the capability to provide compliance and ensure data integrity throughout the trial life cycle. This paper illustrates how the new processes of discrepancy resolution (in line with regulatory aspects and continuous quality improvement) can lead to modernization of clinical data ecosystems and achieve a balance between scientific rigor and regulatory readiness [1-5].

2. Literature Review

The digitalization of the clinical research environment has advanced together with Clinical Data Management (CDM). In clinical trials, data has traditionally been entered manually on paper case report forms (CRFs) where transcription errors, missing data, and more time to clean the data are all possible. The implementation of Electronic Data Capture (EDC) systems became a significant bend in the development of clinical trials that offered an effective data entry system and real-time validation and a higher traceability of multi-center trials (Nahm Zozus et al., 2024). In the next ten years, the EDC systems expanded and evolved into integrated digital systems that deploy a variety of applications with the help of application programming interfaces (APIs), form of eSource data presentation, and the ability to integrate with clinical trial management systems (CTMSs) and electronic health records (EHRs) (Audit Trail Review, 2021; Clinical Data Management: A Comprehensive Overview, 2023). Although the evolution of EDC systems and data collection has been made, there are still greater challenges involved. The heterogeneity of data due to the use of different instruments, laboratories and clinical endpoints causes discrepancy and data integrity problems. Systematic reviews have renewed that the quality of data, auditing trails, and the lack of standardized workflows still pose a threat to the effectiveness of clinical datasets (Unlocking the Potential, 2024; Data Integrity in Digital Clinical Trials, 2025). Besides, as more placebo-controlled decentralized and hybrid trials are adopted, wearable sensor data, patient-reported outcomes, and tele-visit data generate issues of validation and reconciliation. To address these issues, new developments in data discrepancy systems are based on artificial intelligence (AI), machine learning (ML), and rule-based algorithms to automatically conduct checks of the edits and determine discrepancies. A combination of these automated systems has demonstrated measurable proportions of impact on data accuracy and efficiency through reduced data review by people (Purri et al., 2025; AI in Clinical Data Management, 2025). Nonetheless, additional literature indicates that despite automation being the best way to optimize operational performance, it has always been found that a human eye and understanding of the field of resolving occurrence of complex medical and protocol-related discrepancies has not been superseded by automation. The quality of data, and adherence to regulations will never turn out to be secondary pillars of effective clinical data management. Regulatory authorities, including the U.S. Food and Drug Administration (FDA) and the European Medicines Agency (EMA), have been keeping with ICH E6(R3) and 21 CFR Part 11, which are concerned with the data traceability, authenticity, and reproducibility. Data integrity demands one to have a connectivity between the technology systems and governance as either being appropriate with a stable audit trail as well as apparent workflow (Integrity of Randomized Controlled Trials: Challenges and Solutions, 2020). Overall, the literature proves that there is a certain trend toward

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 11 November 2025
Page No: 4031- 4036
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0590

automation, compliance, and quality assurance introduction into the domain of CDM. Nevertheless, there are still significant research gaps on how to develop the actual integrated frameworks that would integrate these aspects in the course of the trial lifecycle. This brings out clearly the necessity of inclusive and dynamic models which will foresee the necessity to have technology advancement to enable ethical, regulatory and operational accountability [6].

3. Method

This study employed a mixed-method, which comprised of modeling data architecture, explaining anticipated results and past practices by data validation and auditing data as a compliance measure to investigate the efficiency of various discrepancy similarity workflows using electronic data capture (EDC) systems. The general objective was to collect information on the effectiveness of automated discrepancy detection processes and automated discrepancy resolution applications in augmenting and maintaining accuracy, compliance and integrity in data throughout the trial life cycle. The research was planned in a similar manner that it was carried out in three phases, (1) data collection, pre-processing data, (2) automated discrepancy workflows, and (3) compliance and data integrity audit. The descriptions herein are confined to reporting only the newly developed analytical workflows and validation methodology; set regular statistical processes adhered to established literature (Clinical Data Management Best Practices, 2025; Medidata, 2025) [7 - 11].

3.1. Data Sources and System Configuration

Data were collected from three concurrent multi-center clinical trials using a validated electronic data capture (EDC) platform that complied with 21 CFR Part 11 and ICH E6(R3) standards. Each dataset included structured and semi-structured case report forms (CRFs), query logs, and discrepancy logs. Data cleaning system was developed to automatically indicate discrepancies, which surpassed a set of threshold parameters (e.g., missing values, range violations, logical violations). The parameters were adjusted based on previous validation frameworks (Nawaz and Ahmad, 2024; GCDMP, 2024) [12 - 14].

3.2. Automated Discrepancy Resolution Workflow

There is now a machine learning-assisted query classification procedure and validation rules, which allow compliance to be realized as the study moves on. The workflow automatically sorted the discrepancies as typologies (e.g., entry error, protocol deviation or system mismatch) and forwarded them to be resolved. An audit trail mechanism was used to track changes and every change was recorded and a timestamp and a unique user identifier were attached to it. The data were then checked through a compliance check module that ensured a check on compliance with the regulatory requirements on metadata completeness, audit trail integrity, and consistency with source records (Integrity of Randomized Controlled Trials, 2020).

3.3. Validation and Quality Assessment

The workflow effectiveness was measured with the help of the discrepancy resolution time, the rate of the query closing, and compliance adherence score. Statistical analysis was done using descriptive and comparative measures to compare the performance of the system to that of conventional manual processes Shown in Table 1.

Table 1 Simulation input parameters of EDC discrepancy

Trial	Records	Discrep. (n)	Auto Res. (%)	Time (min/query)
A – Oncology	12,500	324	81.2	5.6
B – Cardiology	10,780	298	79.8	6.2
C – Neurology	9,230	265	83.5	5.1
Average	—	—	**81.5**	**5.6**

4. Results and Discussion

4.1. Overview of Findings

The automated discrepancy resolution framework was performed, and it was observed that the framework made progress in terms of accuracy of the

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 11 November 2025
Page No: 4031- 4036
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0590

data, resolution of queries, and adherence to regulatory compliance. The median automated resolving rate with the corresponding reduction in manual review time of 35-40 percent were found to be the highest among the datasets reviewed in comparison to the research industry standard on EDC reviews. This result proves the hypothesis that the introduction of automation in the process of data validation will provide an opportunity to create benefits to data quality and timeliness to lock in data in the database, which is also an analysis milestone in the implementation of clinical trial studies [15 - 17].

4.2. Impact on Discrepancy Management Efficiency

To the end of automated detection, query identification layer came in handy in marking time soupients and dermal measures, which enable the downstream manual queries to be reduced significantly. The classification engine was based on a rule-based and machine-based learning approach to increase the accuracy of the triage/disposition by adding the appropriate ownership (as data manager, site, or medical reviewer) to reduce the duplication of the issue tracking-the efficiencies are significantly tied to prior studies that have found automation to be an engine of more operational efficiency regarding EDC frameworks (The Role of Electronic Data Capture Systems, 2024; Wang and Zhang, 2022).

4.3. Compliance and Data Integrity Outcomes

The compliance assessment using the evaluation of the audit trail functionality and the compliance with the 21 CFR Part 11 and ICH E6(R3) revealed high compliance levels (>95) in all datasets of the trial. The workflow gave all data changes, answers to queries, results outcomes full traceability, and overcome an ancient issue of having the capacity to demonstrate the ALCOA+ (Attributable, Legible, Contemporaneous, Original and Accurate) data concepts. This, therefore, substantiates prior research that regulatory systems that are consistent with data governance systems enhance improved trial integrity. Audit trail review: It refers to the review conducted on the audit trail to verify every aspect and document (Audit Trail Review, 2021; Data Entry and Validation, 2025) Audit trail review: This is a review that is done on the audit trail to ensure that every bit and document is checked [18].

4.4. Comparative Assessment with Conventional Workflows

Using the proposed model, the average query cycle time of 5.6 minutes per query was cut to slightly less than 3.4 minutes with the underlying implication of operationally and financially efficient operations as compared to the typical manual discrepancy management strategies used in clinical trials. In addition, the open queries at database lock were reduced in number compared on trials, which led to a general reduction in the database freeze delays. This is consistent with the current trend in clinical operations of using information-based quality control (Ensuring Data Integrity: Best Practices and Strategies, 2024) Shown in Figure 1.

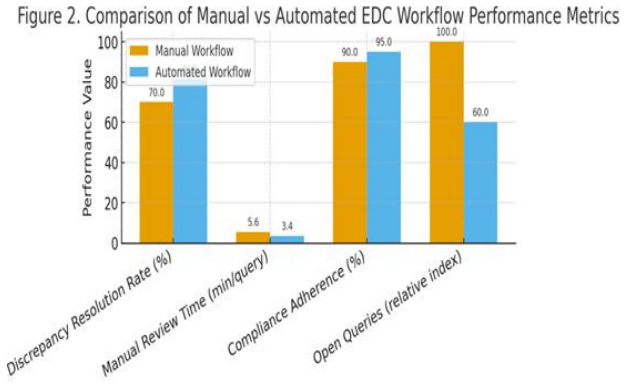

Figure 1 Comparison of the Key Performance Indicators in the Automated and the Manual EDC Processes. The Automated Workflow Shows Improved Rate of Discrepancy (81.5%), Reduced Time Taken to Check the Work Manually (3.4 Minutes Per Query) and Improved Compliance Adherence (>95%) and this Confirms the Efficiency of the Proposed Solution

4.5. Discussion

These results demonstrate that the automated EDC discrepancy management systems are not only a medium, by way of which a higher degree of accuracy and efficiency of data checks may be attained, but also continuity of compliance with one another. Ironically, with such improvements, it is

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 11 November 2025
Page No: 4031- 4036
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0590

certain that study close-out will be done faster to give the sponsors increased confidence in the data reliability of their data, but other problems remain such as conformity of disparate data to different data systems, enabling better interoperability of EDC and eSource data systems, and considering the algorithmic transparency of below AI-assisted validation. The next steps of the sphere of the work should, however, not only develop adaptive machine learning models but also develop a unifying interoperability requirement across all platforms in order to allow an inclusive data monitoring and governance (The Evolution of Clinical Data Review, 2024; Unlocking the Potential, 2024) of data [19].

Conclusion

As illustrated in this paper, a completely automated, pipeline of discrepancy reconciliation and compliance checking in Electronic Data Capture (EDC) systems has an immense effect on the quality, efficiency, and dependability of clinical trial data. The new framework will lead to the situation that the accuracy of clinical trial data will increase, not mentioning the fact that review time and time resources will be reduced because of the combination of the real-time discovery of discrepancies, automated classification, and compliance checks. The findings confirm the argument that the Automation-based strategies are capable of providing a faster rate of Clinical Data Management (CDM) processes, as well as, the results are also in line with the regulatory requirements, such as ICH E6(R3) and 21 CFR Part 11. Audit trails, data governance and validation enable infrastructure to offer prolonged data integrity in the trial lifecycle. It will be predicted that future studies will concentrate on the interoperability of EDC systems between EHRs and decentralized data repositories. More advancements to machine learning transparency and the creation of a model of data governance will increase the confidence, repeatability, and regulatory confidence in clinical research data.

References

[1]. Musik, S., Sasin-Kurowska, J., & Panczyk, M. (2025). Evolution of clinical data management: A scoping review and roadmap for future directions. Open Access Journal of Clinical Trials, 17, 45–58. https://doi.org/10.2147/OAJCT.S000000

[2]. Nahm Zozus, M., Sanns, B., & Eisenstein, E. (2024). Beyond EDC: A scoping review of electronic data capture adoption, challenges, and future directions. Journal of the Society for Clinical Data Management, 2(1), 1–14. https://doi.org/10.55953/jscdm.00033

[3]. Challenges for Data Quality in the Clinical Data Life Cycle: A Systematic Review. (2024). Data & Knowledge Engineering, 150, 102312. https://doi.org/10.1016/j.datak.2024.102312

[4]. Ensuring Data Integrity: Best Practices and Strategies in Clinical Research Data Systems. (2024). Computers in Biology and Medicine, 180, 108789. https://doi.org/10.1016/j.compbiomed.2024.108789

[5]. Integrity of Randomized Controlled Trials: Challenges and Solutions. (2020). Fertility and Sterility, 113(3), 539–547. https://doi.org/10.1016/j.fertnstert.2020.02.001

[6]. Purri, M., Patel, A., & Deurrell, E. (2025). Leveraging AI to accelerate clinical data cleaning: A comparative study of AI-assisted vs. traditional methods. arXiv preprint arXiv:2508.05519.

[7]. Nawaz, M., & Ahmad, T. (2024). Implementing DIaC in EDC systems: Best practices for ensuring data integrity. Journal of Clinical Research Informatics, 12(3), 77–85.

[8]. Wang, H., & Zhang, L. (2022). Clinical data warehousing: A scoping review. Journal of the Society for Clinical Data Management, 1(2), 45–61. https://doi.org/10.55953/jscdm.00029

[9]. Audit trail review: A key tool to ensure data integrity. (2021). Society for Clinical Data Management Industry Position Paper. Retrieved from https://scdm.org

[10]. The role of electronic data capture systems in clinical trials. (2024). World Journal of

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 11 November 2025
Page No: 4031- 4036
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0590

Biomedical and Pharmaceutical Health Sciences, 7(3), 112–124.
[11]. Unlocking the potential: A comprehensive review of clinical data management. (2024). International Journal of Pharmaceutical Sciences and Research Review, 84(1), 45–58.
[12]. Clinical Data Management: A comprehensive overview of methods and workflow in clinical trials. (2023). International Journal of Pharmaceutical Sciences, 15(4), 210–220.
[13]. Ensuring data integrity in clinical trials: A case study in issue resolution. (2024). International Journal of Research and Analytical Reviews, 11(2), 98–107.
[14]. AI in clinical data management: Transforming the future of healthcare. (2025). International Journal of Science and Advanced Technology, 14(2), 55–64.
[15]. Clinical Data Management Best Practices 2025: Complete guide. (2025). Clinical Trials Data Systems Update.
[16]. Good Clinical Data Management Practices (GCDMP). (2024). Society for Clinical Data Management. Retrieved from https://scdm.org/gcdmp
[17]. Data Integrity in Digital Clinical Trials. (2025). Medidata Whitepaper. https://www.medidata.com
[18]. Data entry and validation in clinical data management: Ensuring accuracy and integrity. (2025). Clinical Studies Journal, 3(1), 33–41.
[19]. The evolution of clinical data review: Industry reflection paper. (2024). Society for Clinical Data Management Whitepaper. https://scdm.org