

Responsible AI Adoption in Clinical Trials

Best practices for
clinical data managers





03	Introduction
04	Common challenges facing clinical data managers
06	Overcoming challenges to technology adoption
08	Beneficial impact of AI implementation in clinical trials
09	Comparison of outcomes: AI-driven data quality tools versus traditional methods
10	AI implementation recommendations
14	Conclusion
15	Bibliography and About Medidata

Introduction

Industries across the globe are realizing the immense benefits that Artificial Intelligence (AI) powered solutions can deliver, and in Life Sciences, while some challenges remain, there is a clear and accelerating opportunity for transformation through responsible adoption of AI.

For Clinical Data Managers, time remains a critical operational constraint. Every delay in study build and each manual query incrementally extends the timeline to database lock (DBL), driving up operational costs and postponing the availability of clinically meaningful insights. This translates into an escalating workload driven by increasing protocol complexity and the expanding volume of data generated.

Implementing AI-powered solutions enables Clinical Data Managers to deliver faster, cleaner trial outcomes as trials become more complex, data rich, and time sensitive. They shorten study builds, harmonize data from diverse sources, and reduce manual site entry, delivering reliable data sooner. Real-time visibility scales to handle high-volume data, while resilient systems keep pace with rising velocity. These improvements eliminate bottlenecks, reduce operational burden, and enable timely, data-driven decisions.

This Best Practices Guide explores the unique challenges inherent to clinical trials and offers recommendations for improved Clinical Data Management through practical technology adoption.



Common challenges facing clinical data managers

Clinical Data Managers face mounting pressure as clinical trials grow more complex, data rich, and time sensitive. They must minimize study build timelines, reconcile fragmented data from diverse sources, and allow sufficient time for manual data entry by sites, while striving to deliver

clean, reliable data. Limited real-time visibility and the surge of high-volume, high-frequency data places a strain on traditional systems and processes. These challenges slow trial execution, increase operational burden, and hinder timely, data-driven decision making.

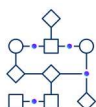


Common challenges facing clinical data managers (continued)



Delayed study build timelines

Lengthy and complex study build processes often delay clinical trials. Teams spend significant time designing and validating case report forms (CRFs), configuring databases and workflows, and aligning them with evolving protocols. Because traditional data collection systems lack the agility to support rapid study setup, these inefficiencies extend timelines and delay trial starts.



Data fragmentation

Clinical trials generate data from multiple sources, such as Electronic Health Records (EHRs), laboratory systems, wearables and sensors, as well as Electronic Data Capture (EDC), Imaging, and Electronic Clinical Outcome Assessments (eCOA), including Electronic Patient-Reported Outcome (ePRO) tools. Clinical data managers face difficulties integrating these diverse data streams into a unified view due to incompatible data formats, lack of interoperability, and the challenge of manual reconciliation. This fragmentation weakens scientific rigor, creates operational inefficiencies, and increases patient risk.



Comprised data quality and integrity

Many trials still rely on manual data entry, especially at sites with limited digital infrastructure. Manual processes have the potential to introduce human error, inconsistencies in data capture, and increased workloads for site staff. This compromises data integrity and increases the time and cost of data cleaning.



Delayed insights and decision-making

Timely access to trial data enables teams to monitor patient safety, identify risks to data quality, and ensure regulatory compliance. Because legacy systems lack real-time capabilities, they delay insights, obscure safety signals, and hinder efficient decision-making.



Difficulty analyzing immense data sets

Modern trials generate vast volumes of data. Managing and analyzing this data is resource-intensive due to limited automation, inadequate data visualization tools, and difficulty extracting actionable insights from the information collected. All of this can overwhelm clinical data management teams and delay trial outcomes.

Overcoming challenges to technology adoption

Technology platforms, such as those developed by Medidata, can be powerful tools to enhance study design, streamline trial processes, and improve overall trial outcomes. Yet some challenges to implementation remain.

Barriers to Overcome

Verification of accuracy

- AI tools, algorithms, and software must be rigorously tested to ensure accurate and reliable performance
- Sponsors, CROs, and vendors should collaborate with regulators and make AI systems explainable
- Inaccurate technology can compromise patient safety, corrupt data, or trigger regulatory rejection

Regulatory concerns

- Because clinical trials are tightly regulated, new technologies must comply with standards for data handling, audit trails, and software systems
- Many existing FDA and global regulations predate AI, creating compliance uncertainty

Data security and privacy

- Sensitive patient data must be safeguarded against unauthorized access, breaches, and misuse

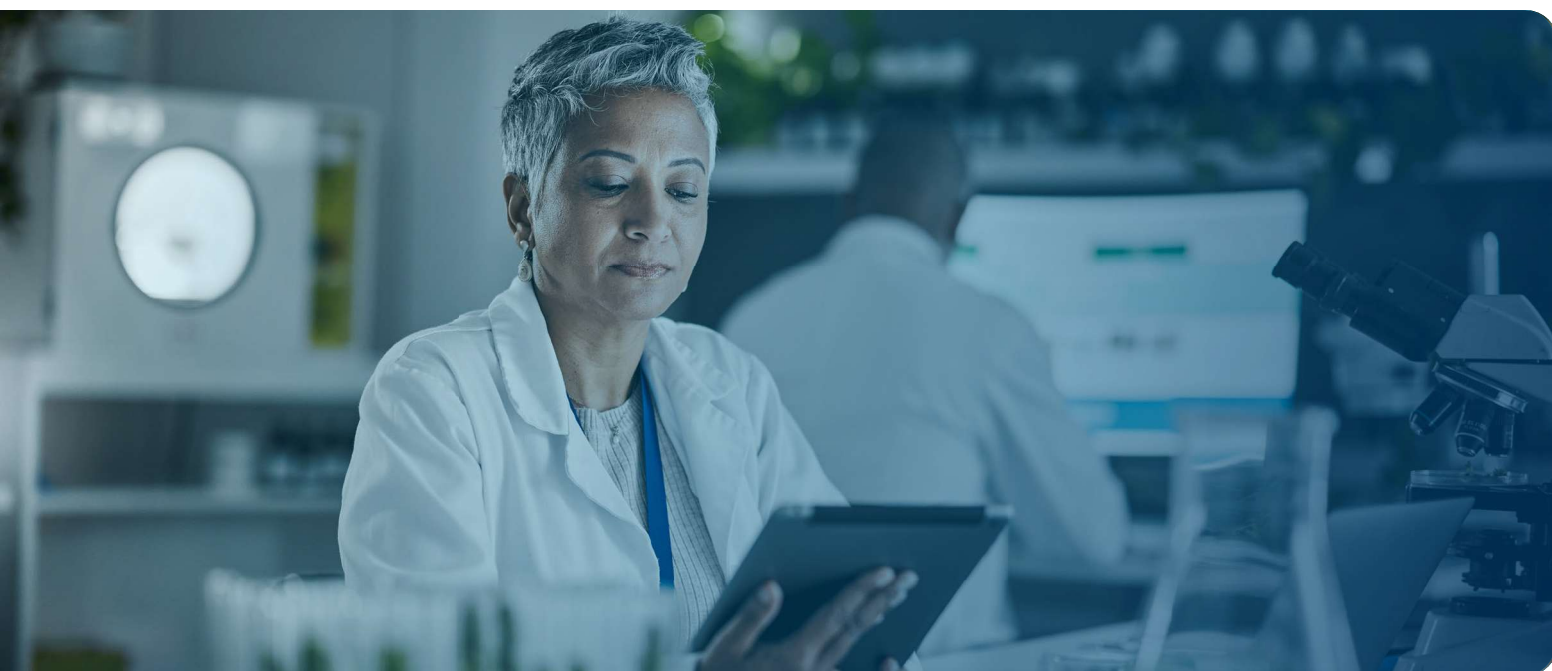
- Cloud platforms, third-party tools, and centralized data lakes increase the risk of security incidents and noncompliance with data protection laws

Transparency issues (ethics and consent)

- Clinical trial participants must understand how their data is used, ensuring informed consent and ethical oversight
- Because many AI systems lack transparency, patients may not grasp how automated decisions are made, raising ethical and legal concerns

Organizational readiness and capabilities

- Pharma, biotech, and CRO professionals often struggle to implement technologies like AI and real-time monitoring despite recognizing their value
- AI adoption requires process updates, revised SOPs, and governance changes, supported by targeted training



The evolving role of AI in clinical trials

The integration of AI into clinical research is advancing rapidly, offering significant opportunities to enhance data quality, operational efficiency, and scientific insight.

Applications such as pattern recognition and anomaly detection are demonstrating tangible value in supporting more robust and timely decision-making across the clinical trial lifecycle.

As adoption accelerates, the focus is shifting toward ensuring that AI-driven methodologies meet the rigorous standards of reliability, reproducibility, and transparency required in a regulated environment.

Continued collaboration among technology providers, sponsors, and regulatory authorities will be essential to establish clear frameworks for quality assurance, governance, and ethical deployment.

By advancing these principles, the life sciences community can foster trust in AI-enabled innovation, enabling clinical trials that are not only more efficient and informative, but also compliant, transparent, and patient-centered.





Beneficial impact of AI implementation in clinical trials

The Medidata Data Experience unifies the entire clinical data lifecycle—from study build to data capture and database lock—on a single connected platform, alongside the Patient and Study Experiences. By harnessing AI and automation, it streamlines data workflows, minimizes manual effort, and boosts data quality.

AI-driven anomaly detection, data reconciliation, and audit trail review identify issues early, such as outliers and inconsistencies, strengthening data integrity and reducing risk. Adopting a technology-based approach

helps teams focus on the most critical data challenges, ensuring smarter resource allocation and proactive quality management.

With real-time visibility into clean, reconciled data, teams gain faster insights and make more informed decisions. By turning raw data into actionable intelligence, the [Medidata Data Experience](#) accelerates trial execution—enabling safer, more efficient, and higher-quality clinical trials.



Comparison of outcomes: AI-driven data quality tools versus traditional methods

The table below actively compares the most likely outcomes of using AI-driven data quality tools versus traditional manual or legacy methods.

Comparison criteria	AI-driven data quality tools	Traditional methods
Speed and throughput	Faster	Slower, with manual review bottlenecks
Errors and variability	Lower error rates and greater consistency	Higher variability, due to more human interpretation noise
Scalability	Easily handles large and complex datasets	Harder to scale without substantial resource increase
Early-issue detection and real-time monitoring	Real-time anomaly detection and proactive flagging	Issues may be identified later, requiring more rework, or at worst, derailing the trial
Regulatory trust and audit readiness	Improving, however concerns remain regarding transparency; requires external verification	High regulatory familiarity due to its long track record
Human oversight and exception handling	Positive impact when combined with human oversight	Strong human judgment, flexible in exceptions
Resource utilization	Elevating and supporting the workforce for improved performance and greater efficiency	High workforce and time cost

AI tools deliver significant benefits in data review and monitoring—enhancing speed, reducing errors, improving scalability, and enabling earlier detection of issues when applied effectively. Yet, for tasks that demand deep human judgment, contextual understanding, ethical reasoning, or interpretation of novel endpoints, manual methods still excel.

Traditional approaches remain essential for oversight, exception handling, and regulatory compliance. Their reliability and trustworthiness make them indispensable,

though they struggle to keep pace with expanding data volumes and increasing trial complexity.

A hybrid model offers the strongest value: using AI to automate and accelerate data quality assurance while empowering human experts to focus on critical decisions, rare or unexpected events, and regulatory safety. This combination maximizes efficiency, accuracy, and compliance throughout the clinical trial process.

AI implementation recommendations

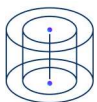
Adopting AI in clinical trial data management drives major efficiencies, delivers earlier insights, accelerates data clean-up, and enables smarter monitoring—ultimately improving data quality and decision-making. However, it also brings new governance, compliance, and operational challenges that must be managed proactively.

To help pharmaceutical and biotechnology companies and CROs capture AI's full value while maintaining rigorous standards, the following best-practice recommendations span people, processes, data, technology, and regulatory domains.



Strategic alignment

- **Build a roadmap**
Prioritize use cases by complexity and risk—start with low-risk tasks like automated anomaly detection, then expand to advanced applications such as generative AI for audit trail review.
- **Gain cross-functional involvement**
Ensure leadership in clinical data management, clinical operations, medical/safety monitoring, biostatistics, IT regulatory compliance, and the CRO (if applicable) are all engaged up front.
- **Define scope clearly**
Decide whether your AI data management initiative is limited to data cleaning and QA, or encompasses broader trial operational applications.
- **Partner strategically**
Select vendors with proven clinical trial expertise, strong interoperability and governance capabilities, and high-quality historical trial data to train AI models.



Data and infrastructure

- **Data quality and standards**
Ensure high-quality metadata, standardized definitions, and consistent data across studies. AI performs only as well as the data behind it.
- **Data integration and cohesion**
Data sources must be fully integrated to support successful AI adoption and increase cohesiveness.
- **Appropriate data sources**
Include both structured and unstructured data, and potentially real-world data (RWD), and ensure provenance, relevance, and representativeness.

AI implementation recommendations (continued)

- **Privacy, security, and de-identification**
Sensitive patient data mandates strong encryption, access controls, anonymization, and adherence to all relevant regulatory requirements.
- **Governance of training datasets**
To effectively train AI models, keep clear records of datasets, versions, labeling rules, sources, site origin, and external versus internal data.

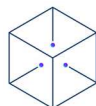


Process integration and operational adoption

- **Embed into existing workflows**
Integrate AI with current data management systems and workflows rather than as a standalone, add-on product.
- **Change management and training**
Successful adoption depends on the user acceptance; data managers, monitors, CRAs, and statisticians all should be trained on how the models work, their limitations, and how to interpret results.
- **Pilot and iterate**
Before full-scale roll-out, run proof-of-concepts on defined use cases, measure user acceptance, model performance, and workflow fit.
- **Scalability and infrastructure readiness**
Ensure your infrastructure can support growing volumes, real-time analytics, and multiple concurrent use cases.



AI implementation recommendations (continued)



Regulatory compliance and ethical considerations

- **Regulatory alignment**
Engage regulatory affairs early to ensure AI use in clinical trials fully complies with all applicable regulations and patient privacy laws.
- **Explainability and auditability**
Design models to produce transparent, traceable, and defensible outputs—especially when supporting regulatory submissions.
- **Address bias and fairness**
Train models on representative datasets to prevent algorithmic bias and avoid excluding or misrepresenting patient populations.
- **Data sovereignty and IP**
Define data ownership, storage location, access rights, and intellectual property terms clearly when using vendor platforms.
- **Ethical oversight**
Conduct ethics reviews for all AI applications that affect patients, maintain transparency, obtain consent when required, and establish strong governance for AI-driven decisions.
- **Privacy-preserving techniques**
Apply robust privacy methods when sharing data across sites or partners to safeguard patient identities.



Readiness and capabilities

- **Foster data science and domain expertise collaboration**
Enable data scientists and clinical experts to work closely together to ensure AI models reflect real-world context and deliver meaningful insights.
- **Promote transparency and accountability**
Assign clear roles and responsibilities for model ownership, data stewardship, validation, and operations to ensure responsible AI governance.
- **Data integration across sources**
Unify EDC, ePRO/eCOA, lab, imaging, and real-world data into a shared analytics layer to power AI models with deeper, cross-domain insights.

Read our eBook, [**5Es for Operationalizing AI**](#), to learn our perspective on how to strategically leverage AI for the benefit of your organization.



Conclusion

Capturing the AI advantage in clinical research

AI empowers the clinical research industry to cut costs, manage growing complexity, and transform the drug development process. By streamlining patient recruitment, automating data monitoring, and accelerating study completion, AI boosts efficiency and enhances data quality across every stage of research.

Organizations that integrate AI successfully do so by:

- Piloting technologies in controlled, auditable settings
- Maintaining human oversight over AI outputs
- Collaborating with experienced, verified partners
- Ensuring alignment with regulatory expectations and data standards

Medidata harnesses AI and advanced analytics to transform clinical trials

By leveraging vast historical trial data gathered from more than 38,000 trials, 12+ million patients, and harnessing 70B+ data points, Medidata empowers research organizations to optimize design, streamline execution, reduce costs, and accelerate breakthroughs that improve patient outcomes.

Medidata uses advanced analytics and machine learning on historical trial data to enhance design, execution, and data quality. Our digital solutions streamline clinical trials, reduce costs, accelerate therapy development, and improve patients' lives.

Our platform-enabled experiences were purpose-built to manage clinical trial complexity:

- **Patient Experience** enhances engagement, simplifies participation, and improves trial outcomes
- **Study Experience** optimizes design, planning, startup, and execution for smarter, faster clinical trials
- **Data Experience** mobilizes AI technology to simplify data workflows from study build to database lock

By partnering early and strategically with Medidata, pharmaceutical and biotechnology companies and CROs can future-proof their development pipelines. Ultimately, AI can be leveraged as a scientific enabler, enhancing decision-making, minimizing uncertainty, and speeding the delivery of innovative therapies to patients.

Contact an expert to learn more.

Bibliography

1. Mao, X., Zeng, D., Wang, X. et al. Digital health technology use in clinical trials of rare diseases: a systematic review. *Commun Med* 5, 449 (2025).
2. NajafZadeh M, Fernández Oromendia A, Burcu M, et al. Linkage of Clinical Trial Data to Routinely Collected Data Sources: A Scoping Review. *JAMA Netw Open*. 2025; 8(4):e257797.
3. Copland RR, Hanke S, Rogers A, Mpaltadoros L, Lazarou I, Zeltsi A, Nikolopoulos S, MacDonald TM, Mackenzie IS. The Digital Platform and Its Emerging Role in Decentralized Clinical Trials *J Med Internet Res* 2024; 26:e47882.
4. Krishnankutty B, Bellary S, Kumar NB, Moodahadu L. Data management in clinical research: An overview. *Indian J Pharmacol*. 2012;44(2):168-172.
5. Clinical trial data management and quality metrics system. *Clin Trials Qual Metric Syst*. 2016; PubMed abstract.
6. A review of clinical data management systems used in clinical trials. *J Clin Trials Clin Data Manag Syst*. 2018; PubMed abstract.
7. Eade D, Pestronk M, Russo R, et al. Electronic Data Capture (EDC) study implementation and start-up. *J Soc Clin Data Manag*. 2021;1(1):1-24.
8. Izmailova ES, Middleton D, Yunis R, et al. Implementing sensor-based digital health technologies in clinical trials: Key considerations from the eCOA Consortium. *Clin Transl Sci*. 2024;17(11):e70054.
9. Evaluating the Impact of Electronic Health Record to Electronic Data Capture Technology on Workflow Efficiency: a Site Perspective. *J Am Med Inform Assoc*. 2025; PubMed abstract.
10. Medidata eBook: The “Five Es” of Operationalizing AI in Clinical Data Management.
11. Medidata white paper: Navigating data capture complexity in early- and late-phase clinical trials.

About Medidata

Medidata is powering smarter treatments and healthier people through digital solutions to support clinical trials. Celebrating 25 years of ground-breaking technological innovation across more than 38,000 trials and 12+ million patients, Medidata offers industry-leading expertise, analytics-powered insights, and one of the largest clinical trial data sets in the industry. More than 1 million registered users across approximately 2,400 customers trust Medidata’s seamless, end-to-end platform to improve patient experiences, accelerate clinical breakthroughs, and bring therapies to market faster. A Dassault Systèmes brand (Euronext Paris: FR0014003TT8, DSY.PA), Medidata is headquartered in New York City and has been recognized as a Leader by Everest Group and IDC.