AUDIT TRAIL REVIEW:

STRATEGIES FOR IMPLEMENTING AUDIT TRAIL REVIEW

An Industry Position Paper

February 2026

Public Release Version 1.0

eClinical Forum

Authors: eClinical Forum (eCF)

Security: This is a public release. Non-eCF Members are welcome to use and share as per License in Appendix 1.

TABLE OF CONTENTS

TABLE OF CONTENTS ***1***

1. Abstract ***2***

2. Acknowledgements ***4***

3. Introduction ***5***

4. Clinical Relevance — Why audit trail review matters and where to focus ***6***

4.1 The Need for Audit Trail Review 6

4.2 Use Case Prioritization 6

4.3 Examples of Realized Risks 7

5. Data-Driven Analytics — How to access and analyze audit trail data ***8***

5.1 Data Access 8

5.2 Analytical Capability 10

5.3 Architecture Choices 14

6. Operational Integration — How to embed ATRA into routine clinical processes ***16***

6.1 Workflow 16

6.2 Maturity Assessment 20

6.3 Strategic Roadmap 23

7. Conclusion ***25***

Appendices ***26***

Appendix 1: DISCLAIMER and LICENSE for FAIR USE of eCLINICAL FORUM MATERIALS 26

Appendix 2: Definitions and Acronyms 27

Appendix 3: Use Cases Prioritized with Realized Risk Examples 29

Appendix 4: Step-by-Step Workflow for managing Audit Trail Reviews 36

Copyright eClinical Forum

1. Abstract

Since the publication of the 2021 position paper (Public Release: Audit Trail Review: Key Tool to Ensure Data Integrity Version 1.0 (eclinicalforum.org)), the eClinical Forum (eCF) has received requests for further guidance on implementing Audit Trail Review (ATR) strategies. ICH E6(R3) has called out the need for planned review of audit trail data.

> *"Procedures for review of trial-specific data, audit trails and other relevant metadata should be in place. It should be a planned activity, and the extent and nature should be risk-based, adapted to the individual trial and adjusted based on experience during the trial."*
>
> - Section 4.2.3 Review of Data and Metadata

In response, this paper offers a structured and actionable framework to help organizations prioritize, operationalize, and scale their audit trail review activities. The framework has three domains—**Clinical Relevance**, **Data-Enabled Analytics**, and **Operational Integration**—which cover the main capabilities for audit trail review per ICH E6(R3). Although these sections are presented sequentially, most organizations will approach this work as a continuous, iterative process. Reviewing the entire paper at a high level before focusing on individual sections can make it easier to read and understand.

Clinical Relevance — Why audit trail review matters and where to focus

1. **Need for Audit Trail Review:** The importance of audit trail review has increased with ICH E6(R3). While audit trail review was initially a focus point of European regulators, it is now globally seen as an important process to verify the integrity of a clinical study.
2. **Use Case Prioritization:** ICH E6(R3) emphasizes the need for proportional responses and a need to focus on data integrity to identify risks. We reviewed and ranked the use cases and risk scenarios from the 2021 position paper, resulting in an overall prioritization. Member companies and other organizations can use this prioritization as starting point for their own Audit Trail Review Analytics (ATRA) implementation
3. **Examples of identified risks and/or issues:** Member companies reported that audit trail review was perceived as a theoretical requirement and that the 2021 position paper lacked practical guidance. To address this, we collected anonymized, real-world examples of issues that occurred in clinical studies. These cases help illustrate the need for audit trail review and give organizations implementing ATRA insights into what kind of issues they may encounter.

Data-Driven Analytics — How to access and analyze audit trail data

1. **Data Access:** Reliable, automated access to audit trail data is foundational. This section outlines best practices for extracting, standardizing, and transforming data into analysis-ready formats, enabling consistent review across systems and vendors.
2. **Analytical Capability:** Legacy methods like listings and manual review are no longer sufficient. We introduce a maturity-based model that progresses from basic descriptive statistics to trend analysis, risk-based site ranking, and root cause investigation, helping teams prioritize signals and act with confidence.
3. **Architecture Choices:** Whether audit trail analysis is performed locally within source systems or centrally across platforms has significant implications for scalability and oversight. We explore the

Copyright eClinical Forum

trade-offs between both models and offer practical guidance for evolving towards more integrated, scalable approaches.

Operational Integration — How to embed ATRA into routine clinical processes

1. **RBQM Workflow:** This section describes guidelines for identifying, investigating, and tracking issues, leveraging existing RBQM workflows.
2. **Maturity Assessment:** A foundational assessment to measure current maturity across critical dimensions—data access, analytics, workflows, and skills—serving as a baseline for ATRA implementation planning.
3. **Strategic Roadmap:** A framework for aligning audit trail analytics implementation with the maturity and operational capabilities of clinical data organizations.

It is important to note that this paper is not a one-size fits all approach, and that any organization's needs will differ depending on the details of their trials and systems. It is the organization's responsibility to make sure that their implementation of Audit Trail Review Analytics is fit for purpose and proportional to the risks identified in their trials.

Copyright eClinical Forum

2. Acknowledgements

The information presented in these works draws upon the combined current understanding and knowledge of the eClinical Forum (eCF) and is provided as an aid to understanding the environment for electronic clinical research. Readers should assess the content and opinions in the light of their own knowledge, needs and experience as well as interpretation of relevant guidance and regulations.

This work is the property of the eClinical Forum and is released under a Creative Commons license for non-commercial use. For additional license information, see Appendix 1: DISCLAIMER and LICENSE for FAIR USE of eCLINICAL FORUM MATERIALS.

The eClinical Forum would also like to acknowledge the many volunteers who have participated in the eCF ATRA Working group/workshops and have contributed to forming the thoughts expressed in this position paper.

Lead Authors:

- Nechama Katan, Wicked Problem Wizards LLC, Chief Wizard
- Dan Cocozza, Boehringer Ingelheim, Director, Expert Capability Manager
- Jan Breemans, Strategic Clinical Advisor, Data Operations & Tech-enabled Clinical Trials
- Jun Zhao, Boehringer Ingelheim, Sr. AD, Sr. Prin. Clin Data Engineer
- Benjamin Shim, Eli Lilly, Clinical Data Strategist

Contributing and Reviewing eClinical Forum Member Companies

Astellas; BMS; Boehringer Ingelheim; BSI Life Sciences; CluePoints; CSL Behring; eClinical Solutions; Eli Lilly; Ferring; FGK Clinical Research; Johnson & Johnson Innovative Medicine; Medidata; Merck; Novartis; Novo Nordisk; Oracle; Oslo University Hospital; Pfizer; Signant Health; i2e Consulting; Viedoc

About the eClinical Forum:

Established in 2000, the eClinical Forum emerged as a response to the transformative impact of Electronic Data Capture on Clinical Trials. Initially centered on Data Management, we have since expanded to encompass all facets of Clinical Research and Development. Operating as a global not-for-profit organization, we unite over 65 member companies from pharmaceutical and biotech companies, CROs, technology providers, and healthcare organizations. As a member-driven network, we are committed to fostering dialogue and knowledge sharing among industry stakeholders. We organize working groups that regularly convene to explore regulatory changes and technological trends. This document is the result of such a working group and is part of our document repository (www.eclinicalforum.org).

Copyright eClinical Forum

3. Introduction

The integrity and reliability of clinical trial data are essential to ensuring patient safety, meeting regulatory expectations, and advancing medical knowledge. As the industry continues its transition toward electronic systems such as Electronic Data Capture (EDC), electronic Clinical Outcome Assessment (eCOA) and Interactive Response Technology (IRT), the role of audit trail data has grown in importance. Audit trails provide visibility into the history of data entry/changes, user activities, and system interactions, making them a foundational tool for transparency, traceability, and quality assurance in clinical research.

Regulatory authorities have increasingly emphasized the importance of metadata review, including audit trails, as a critical component of data integrity. This shift is particularly evident in ICH E6(R3), which underscores the need for risk-based approaches to monitoring and quality management. A key component of data quality oversight and monitoring includes implementing risk-based audit trail review in key systems where user actions and entries may compromise the integrity of clinical study data or the ability to recreate clinical study results. While the relevance of audit trail review (ATR) is now widely acknowledged, many organizations still struggle to move beyond viewing it as a regulatory checkbox. In practice, audit trail review often remains misunderstood, fragmented, inconsistently executed, or overly reliant on manual processes.

This paper was developed in response to feedback from member companies of the eClinical Forum, who expressed a need for more practical, detailed guidance on how to implement ATR in a way that is scalable, risk-aligned, and operationally sustainable. The goal is to shift ATR from a theoretical expectation into a proactive, value-generating activity that supports early detection of risks, informed decision-making, and continuous quality improvement.

Accordingly, the paper is organized into three integrated domains, each representing a critical pillar of effective audit trail review analytics (ATRA):

- **Clinical Relevance** – why audit trail review matters, guiding organizations on prioritizing high-risk use cases through real-world examples.
- **Data-Driven Analytics** – how to access, structure, and analyze audit trail data using a maturity model and scalable architecture.
- **Operational Integration** – how to embed audit trail review into RBQM workflows using defined processes, maturity assessments, and a strategic roadmap.

We advocate for the adoption of comprehensive ATRA processes that bring together the necessary people, processes, and technology. This includes using visualizations, exception reports, and prioritized checks, supported by automated data flows and RBQM-style oversight. As organizations mature in their audit trail capabilities, the goal is to create systems that not only detect anomalies but support continuous learning and improvement. As we embark on this journey, we recognize the importance of collaboration and dialogue among all stakeholders, including the eClinical Forum (eCF).

Ultimately, audit trail review is not just about meeting regulatory expectations but about building trust in clinical data, ensuring the robustness of trial results, and enabling earlier and more confident responses to emerging risks. This paper aims to help organizations of all sizes and levels of maturity in making meaningful steps toward implementing scalable, effective, and purpose-driven audit trail analytics.

Copyright eClinical Forum

4. Clinical Relevance: Why audit trail review matters and where to focus

Understanding the relevance of audit trail review is essential for determining where to focus implementation efforts. The release of ICH E6(R3) and the evolving role of Clinical Data Science have significantly elevated the importance audit trail review plays in supporting data integrity, risk-based monitoring, and operational oversight.

This section outlines the drivers behind ATR and highlights specific areas of focus, including use cases prioritized by eCF members and real-world examples showing meaningful risks. Together, these insights provide a practical foundation for implementing ATR as a process and aligning efforts with clinical and regulatory priorities.

4.1 The Need for Audit Trail Review

The integrity and reliability of clinical trial data are fundamental to the advancement of medical science and the protection of patient safety. As the industry increasingly adopts electronic data capture modalities, the role of audit trails in ensuring data integrity and transparency has become more critical than ever. Audit trails not only facilitate the reconstruction of data events but also provide valuable insights into site, CRO and sponsor processes, supporting more effective monitoring and oversight over the integrity of the study.

In the latest version of the ICH E6 guideline, Regulatory agencies have put explicit emphasis on metadata review (including but not limited to audit trail) as an essential component of data review. By implementing targeted, risk-based audit trail reviews, organizations can enhance their data governance and ensure compliance with regulatory requirements. This approach will enable sponsors, contract research organizations (CROs), and eClinical vendors to generate actionable insights, improve data quality, and maintain the highest standards of data integrity.

We advocate for the adoption of comprehensive ATR processes that encompass the necessary people, processes, and technology. This includes providing practical implementation guidance, utilizing visualizations and exception report listings, and fostering a culture of continuous learning and improvement. Additionally, we believe that robust ATR should incorporate prioritized checks, automated data flows and analysis, and a Risk-Based Quality Management (RBQM) type workflow with maturity models and roadmaps.

In conclusion, the routine review of audit trails is not just a regulatory expectations but a strategic imperative for the clinical research industry to increase efficiency and reduce timelines. By embracing this practice, we can ensure the integrity of our data, the reliability of our trial results, and ultimately, the advancement of medical science for patients.

4.2 Use Case Prioritization

ICH E6(R3) calls out the need for proportional responses to risks and observed issues. This comes with a need to understand the most critical risks and the ability to respond appropriately. The focus of Critical to Quality factors should be on protecting the integrity of the overall study rather than on single critical data points.

For example, the timely collection and entry of clinical study data are critical for the ongoing monitoring of protocol adherence and patient safety. It is therefore essential to monitor the timeliness of data entry across all data types, rather than limiting focus solely to endpoints or safety-related data.

Copyright eClinical Forum

We have grouped the different relevant use cases into three general categories:

- **Data Forensics**: This category centers on the audit trail of clinical data, offering a comprehensive record of initial data entries, modifications, and deletions. It enables the reconstruction of the data collection process, ensuring full traceability. This type of review is relatively straightforward, as it can be conducted using the audit trail from a single system, yet it often yields the most valuable insights. It also supports a wide range of in-depth analyses, including trend forecasting and signal detection. Due to its potential to uncover critical findings, this category should be considered the highest priority in audit trail reviews.
- **User Access & Activity:** This review involves comparing roles and qualifications of approved users with their actions within various systems such as eCOA, EDC, IRT and electronic Trial Master File (eTMF systems. It requires a thorough examination of both data and system audit trails, including user login logs, role assignments, and potentially even links to CTMS systems or company training records. While often critical, this review is more complex due to the need for data from multiple systems and sources.
- **System Integrity & Security**: This category addresses risks that compromise the integrity or security of clinical systems and their data, particularly those not covered by other review domains. It involves identifying unauthorized access attempts or system activities performed by users without appropriate privileges. While critical in principle, we consider it a lower priority in clinical trial data review practice, as most modern GxP-compliant systems incorporate robust access controls and audit mechanisms. Additionally, many system vendors actively monitor and mitigate such risks through built-in security features and ongoing surveillance.

An overview of these use cases and their realized risks can be found in *Appendix 3: Use Cases Prioritized with Realized Risk Examples.*

4.3 Examples of Realized Risks

Over the past years, audit trail review has often been discussed in abstract terms, resulting in a belief that it was a task that only needed to be done to satisfy a regulatory requirement.

When implemented properly, audit trail review will help to reveal real risks and threats to the outcome of a clinical study that may otherwise remain unnoticed or identified too late to mitigate. To demonstrate the value of ATR, we have consolidated several anonymized real-world examples drawn from the experiences of eClinical Forum members. These cases demonstrate how audit trail data can reveal operational issues, protocol deviations, or gaps in data integrity that might otherwise go undetected.

Each example is mapped to a defined use case and highlights which signals, such as unexpected data changes, delayed entry, or access anomalies that were identified and what impact these risks or issues might have had on the outcome of the study. For several of these cases, similar findings may emerge in other systems.

An overview of these use cases and their realized risks can be found in *Appendix 3: Use Cases Prioritized with Realized Risk Examples* These examples highlight the value of a systematic and proactive approach to audit trail review. They demonstrate how such practices, combined with related analytics, enable early detection of potential data integrity or compliance issues. This, in turn, supports improved study execution, strengthens regulatory compliance, and fosters continuous quality improvement. By sharing these case studies, the working group aims to offer practical reference points and guidance for organizations seeking to initiate or enhance their audit trail review processes.

Copyright eClinical Forum

5. Data-Driven Analytics: How to access and analyze audit trail data

Effective audit trail review cannot succeed without three core capabilities: access to well-structured data, an analytical capability that scales, and architectural design supporting consistent use. Many organizations encounter breakdowns at one or more of these layers:

- **Data Access**: Raw audit trail data often resides as large and complex data sets, often locked in disparate systems or delivered in formats that require extensive cleanup and transformation before use.
- **Analytical Capability**: Teams performing audit trail review may be limited to basic metrics and listings, lacking the statistical tools or confidence to perform trend detection, outlier analysis, or root cause investigation. A well-functioning audit trail review approach requires progression to more advanced methods and analytics.
- **Architecture Choices**: Analysis may occur either locally—within the source system—or centrally via data lakes or oversight platforms that consolidate audit trail data from different sources. While each model offers trade-offs in speed, cost, and integration, organizations may start with a local approach and move towards a more centralized approach as capabilities evolve.

To support scalable, risk-aligned audit trail review, all three components must evolve in parallel. This section outlines a practical approach for strengthening each area—starting with automation of data flow, maturing analytic methods, and aligning infrastructure decisions with regulatory and operational priorities.

5.1 Data Access

One of the primary barriers to effective audit trail analysis is timely and automated access to the data. Audit trail data is often large, complex, and system-generated, making manual review inconsistent and inefficient. To support reliable analysis, organizations need to move from ad hoc exports to a more structured workflow and automated ingestion of data into centralized repositories.

The implementation will look different depending on the organization's needs and risks. However, there are general best practices that apply to managing this type of data. A general methodology is described below. Any given organization will want to make sure that their systems are fit for purpose. The goal is to develop a solution that provides analysis-ready datasets in a robust and efficient way.

A three-tier data architecture with an integrated presentation layer is typically recommended to support audit trail analysis and ensure data integrity throughout clinical trials. This approach —commonly known as the medallion architecture— provides a clear pathway from raw audit data to useful insights. Clinical teams can maintain oversight without needing deep technical knowledge.

Bronze Layer: Foundational Data Collection (Raw Audit Trail Data):

The bronze layer serves as the initial storage area for all audit trail data collected from clinical systems including EDC, ePRO, IRT, and other study platforms. Data is stored in its original format with minimal changes. Complete audit history is preserved, and no information is lost during collection.

Key characteristics of the bronze layer:

- Original data formats and structures from source systems are maintained
- Complete audit metadata and system timestamps are preserved

Copyright eClinical Forum

- The authoritative historical record for regulatory compliance is provided
- Data tracking back to source systems is enabled

Silver Layer: Standardized and Cleansed Data (Reporting-Ready Dataset):

Raw audit data is transformed into a standardized, quality-assured format in the silver layer. Business rules, data validation, and standardization processes are applied to create a unified view across all clinical systems.

Audit trail data is structured so each record captures a single data transaction with standardized fields including:

- **Date and time tracking:** Timestamps of when changes occurred are standardized and normalized across different time zones and study sites
- **Data lineage:** Field name, form, and dataset context are recorded
- **Change tracking**: Previous and new values are preserved with data type information
- **User attribution:** User ID and role for both initial entry and subsequent modifications are captured, with separate records created for each user interaction
- **Study context:** Site and subject identifiers are included with appropriate masking
- **Change justification:** Change type classification (new record, edit, deletion), coded reason categories, and free-text explanations are captured where available

This standardized structure enables consistent analysis across different systems and study phases, creating a reliable foundation for quality monitoring activities.

Gold Layer: Analytics and Insights (Analysis-Ready Datasets and Presentation)

The gold layer represents purpose-built datasets optimized for specific audit trail analysis use cases. Silver layer data is aggregated and enriched with calculated metrics, trend analysis, and contextual information to support strategic decision-making

- These curated datasets are designed to address specific sets of use cases such as those identified in Appendix 3.
- Identifying sites with unusual change frequencies or patterns
- Examining relationships between data modifications and query generation cycles
- Site behavior evolution and improvement over time is tracked
- Early warning signals for potential data quality issues.

The gold layer enables rapid exploration of complex relationships within the audit data. Quality managers and data managers can identify patterns requiring intervention.

Presentation Layer: Visualization and Monitoring Tools

The presentation layer delivers audit trail insights through intuitive interfaces that support both routine monitoring and investigative analysis. This includes integration with RBQM platforms, executive dashboards, and specialized audit review tools that consume data from the gold layer. Depending on the organization, this layer might be in the analytics tools discussed in the next section.

Copyright eClinical Forum

Effective presentation tools should provide:

- Clear drill-down capability from overall trends to individual transactions
- Proactive identification of anomalies requiring attention
- Benchmarking across sites, studies, and therapeutic areas
- Standardized outputs for inspection readiness

Benefits of the Structured Approach

This layered architecture provides several advantages for clinical data integrity programs:

- Growing data volumes and new source systems are accommodated without redesigning analysis workflows
- Diverse analytical requirements are supported while data consistency is maintained
- Complete data tracking from source to insight is preserved for regulatory inspection
- Time-to-insight is reduced by pre-computing common analytical patterns
- Enables controlled access and data quality management at each layer.

By implementing this structured approach, clinical teams can maintain robust oversight of data integrity while enabling sophisticated analysis that supports proactive quality management and regulatory compliance.

5.2 Analytical Capability

Once data access is in place, the next challenge is making sense of it. To help organizations scale their review efforts, we propose a **four-level analytics maturity model**. The model progresses from basic descriptive statistics to advanced and exploratory techniques, enabling organizations to detect meaningful patterns of behavior and take proportionate action. While the maturity levels are presented sequentially, they are not strictly linear. In practice, analytics capabilities are often developed iteratively: a solution is built, tested, refined, and expanded over time as teams learn which signals matter most. In the maturity assessment an additional layer of manual analysis is added. This is for those organizations without robust analytical capability, where all analysis is ad hoc and manual.

The four levels of analytic capability include:

1. **Descriptive Statistics** – Quantify volume, frequency, and distribution of actions, flag outliers
2. **Time-Trend and Rule-Based Monitoring** – Detect unexpected activity patterns using control charts, threshold rules, or predictive models
3. **Advanced Analytics** – Signal prioritization and intelligent risk scoring to identify atypical behaviors
4. **Root Cause Analysis and Feedback Loops** – Support investigation, resolution tracking, and continuous improvement

This framework helps shift oversight from reactive to proactive, offering a practical foundation for scalable, risk-aligned audit trail review, and is designed to support timely decision-making. Effective audit trail analytics help teams answer key oversight questions: Which sites, users, or systems are behaving in unexpected ways? Are issues isolated or recurring? Have prior concerns been resolved or are they resurfacing? And what operational or systemic factors might explain these patterns?

Copyright eClinical Forum

To deliver useful answers, the methodology must be flexible, and capable of shifting focus depending on the context. Whether the analysis centers on a site, user, patient, study, program, or data system, the goal remains the same: to generate relevant, timely insights that drive confident, data-informed action.

Level 1: Descriptive Statistics

Descriptive statistics form the foundation of audit trail analytics. They offer basic metrics such as the mean, count, and standard deviation of audit trail events, summarized by site, user, role, or field. These metrics help establish operational baselines and are most effective for identifying gross outliers: sites or users whose audit trail behavior significantly deviates from the norm.

Key measures at this level may include but not be limited to:

- Total number of audit trail events (e.g., data changes, queries)
- Mean and variability of events per site, user, or patient
- Frequency of specific event types (e.g., updates to critical fields, deletions)
- Aggregated counts by site, patient, user ID, or role—optionally filtered by risk category
- Ratios such as queries per change or form-level change density

These summary views provide a useful first look at where unusual patterns might exist. For example, a site generating five times more data changes than its peers may warrant deeper review, even if the context for those changes is not yet clear.

Level 2: Time-Based Trends and Thresholds

This level adds a simple but powerful question to audit trail review: When did the issue start—and is it getting better or worse? Looking at audit trail activity over time helps teams spot patterns that one-time summaries can miss.

Plotting changes, queries, or updates over time can show whether a site is stabilizing, trending toward trouble, or responding to earlier feedback. This is where you move from “what happened” to “what’s happening now.”

To make the trends meaningful, setting thresholds can be helpful, creating lines on a chart that show what’s typical. These can be:

- Fixed thresholds, based on similar studies or historical data
- Rolling thresholds, which adjust as more data comes in
- Context-aware markers, such as study start-up or protocol changes

These thresholds and limits help focus on what is actually statistically unusual. For example, sites often show more changes early in a study, which is normal. A spike in late-stage activity, however, might be worth a closer look.

It’s also important to normalize adjusting for factors such as number of patients or forms. A high-volume site will naturally generate more changes— that does not mean it is a problem.

This level helps you to:

- Spot problems that are just starting—or getting worse

Copyright eClinical Forum

- See if early issues have improved after action was taken
- Avoid overreacting to normal activity during startup or protocol changes

Challenges include:

- Deciding what "normal" looks like for each phase of a study
- Reviewing thousands of data points across large trials
- Communicating what you see in a way monitors and study teams can act on

With the increasing amount of data, simply monitoring trends is no longer sufficient. That is where the next level—ranking and prioritization—helps you decide where to focus.

Level 3: Advanced Analytics

At this stage, the challenge is not getting data, it is knowing what to do with it. Most teams already have more metrics, graphs, and dashboards than they can reasonably review, and adding more audit trail signals without a way to prioritize just adds noise. Effective scaling analytics in an organization requires the ability to focus analytics to the highest priority issues and remove noise from systems. Level 3 is about cutting through that noise and surfacing what matters most.

As audit trail capabilities mature, advanced analytics techniques can help streamline review. These tools do not replace humans but instead help teams focus their time where it counts.

Examples include:

- Combining multiple metrics into a single site-level risk score
- Grouping sites into "typical" vs. "atypical" profiles based on activity
- Using machine learning to flag unusual patterns or predict risk based on past behavior

These models help uncover weak signals that do not trigger a single alert but may still point to a problem when viewed together. They also enable reviewers to shift from reactive review to proactive triage without drowning in graphs.

One simplistic but practical approach would be to rank sites based on threshold crossings. This means counting how often a site's behavior falls outside the expected range, whether it is a sudden spike in queries, a string of repeated changes, or data that looks a little too clean. You can refine this by:

- Giving more weight to recent behavior
- Counting both high and low-end outliers (e.g., zero queries on complex forms)
- Looking at how many different metrics a site is triggering—not just one

The result is a prioritized list—an at-a-glance view of which sites need follow-up and why. Instead of scrolling through thousands of graphs, your team sees a shortlist of potential risk areas, organized by urgency.

Level 4: Root Cause Analysis and Feedback Loops

Finding a signal is only the beginning. Whether it comes from a trend, a threshold breach, or a risk score, the next step is the most important: figuring out why it happened. Level 4 is about investigation, context, and follow-through, linking signals to causes and causes to action.

Copyright eClinical Forum

Many tools can tell you where something looks off. But understanding the why takes more than a dashboard. It requires combining data with operational context and engaging the study team to validate what is really happening.

A typical investigation might unfold in two parts:

Step 1: Look across the data
Start by asking the basics. Is this isolated to one patient? One site? One point in time?
Join datasets to explore broader explanations "going wide". For example, are AE underreporting patterns linked to patient age? Site location? Time since site initiation?

Step 2: Talk to people
Once a pattern is found, engage the study team, CRAs, or even the site directly. Ask for an explanation on how a process is implemented, for example how AE reporting is handled at that site. This helps confirm if the signal reflects a data issue, a training gap, or a systemic process problem.

A good investigation links signals to potential sources such as:

- Site training history or documentation practices
- Protocol complexity or recent amendments
- System usability or access issues
- Broader operational trends across studies or programs

What sets this level apart is the feedback loop. Root cause findings should not live in a spreadsheet but instead should flow into:

- Issue tracking systems
- Risk libraries
- Ongoing site oversight plans

This closes the loop between detection and action and sets the stage for continuous improvement across studies. It also makes sure that the same issue does not quietly reappear elsewhere.

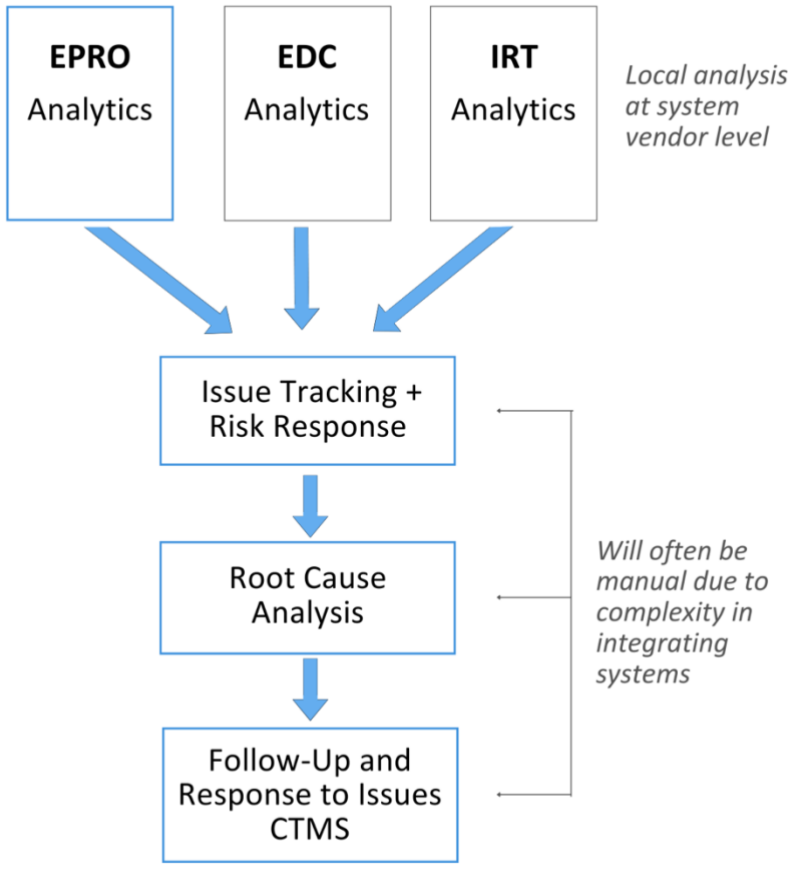
Copyright eClinical Forum

5.3 Architecture Choices

Audit trail analysis can be performed locally within systems like EDC, ePRO or eTMF, or centrally by consolidating data into a central repository. This directly impacts how well teams can detect issues, connect the dots, and take meaningful action. Organizations must choose between analyzing data within each system—like EDC or eCOA—or consolidating it into a centralized environment. Each approach comes with trade-offs in speed, effort, and visibility.

Local or system-level analysis means reviewing audit trail data within the tools where it originates. This approach is fast to implement and requires less upfront infrastructure, making it a reasonable starting point for many sponsor and vendor teams. But it comes with real limitations:

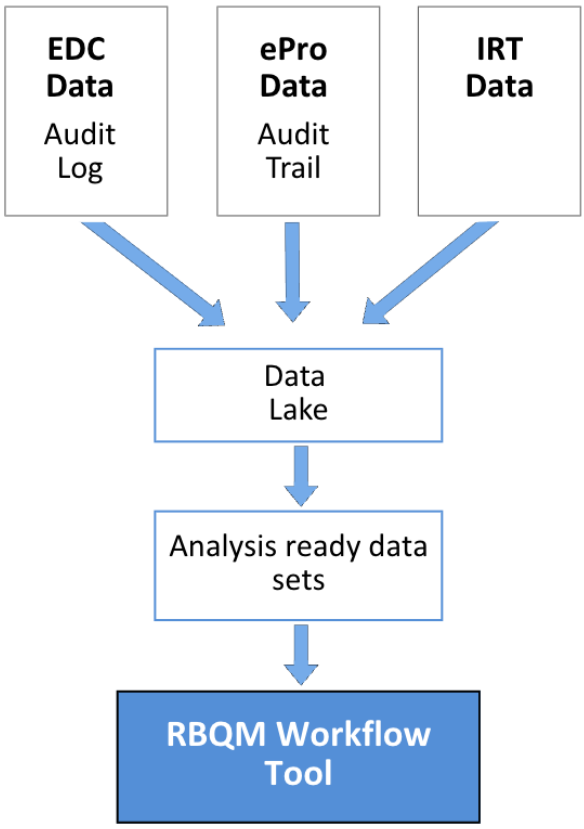
1. **Fragmented visibility:** Issues in one system may appear isolated, even if they are part of a broader site-level pattern.
2. **Traceability gaps:** There is often no shared record of what was reviewed, what was investigated, or what conclusions were reached, making oversight inconsistent and follow-up difficult.
3. **Manual coordination burden:** Teams must rely on people to transfer findings between systems, spreadsheets, trackers, and RBQM tools. This increases risk, slows response time, and reduces transparency



Graph 1: Local or System-Level Analysis and Follow-Up

Copyri

Centralized analysis, by contrast, moves audit trail data into a common structure across systems. This supports cross-functional review, automation, and traceable workflows. Findings can be tied to issue tracking systems, risk libraries, and monitoring plans, helping teams work from a single source of truth.



Graph 2: Centralized Analysis and Follow-Up

The trade-off is effort: centralized models require investment in data pipelines, governance, and validation.

In practice, sponsors typically don't transition to a fully centralized model immediately—and there isn't always a need to do so. Over time, however, centralizing analyses can lead to more comprehensive results.

Local analysis can serve as a practical steppingstone, especially when paired with a plan to consolidate findings over time. As technology and industry expectations evolve, the shift toward centralized, structured analytics will become more natural and more necessary.

Ultimately, your architecture model—where data lives, how it flows, and where insights are generated—forms the backbone of ATRA. If those elements are not connected, even great analytics will not scale or stick.

6. Operational Integration: How to embed ATRA into routine clinical processes

Embedding audit trail review analytics into clinical trials requires more than analytical tools, it demands alignment with existing workflows which include (risk frameworks and operational processes) and organizational maturity. This section explores how ATRA can be integrated into Risk-Based Quality Management (RBQM) workflows, offering a practical framework for issue identification, investigation, and resolution. This framework will touch on other sections of the paper at a high-level as it describes the overall process. There is a diagram of this process in Appendix 4. It also introduces a maturity assessment model to evaluate current-state capabilities and outlines a strategic roadmap to guide phased implementation across sponsors, CROs, and technology partners.

6.1 Workflow

ATRA is inherently aligned with the ICH E6 (R3) principles on quality management and adoption of risk-based approaches to ensure that data are reliable and have been processed correctly. Successfully integrating ATRA into an RBQM framework is not only a technical task but also a significant operational and cultural undertaking. It requires breaking down historical silos between functions like Clinical Operations and Data Management and establishing new, collaborative workflows. Recognizing that organizations exist on a spectrum of RBQM adoption and technological sophistication, a phased, maturity-based approach is essential for successful implementation. This section provides a practical roadmap for operationalizing the integrated workflow, from defining roles and responsibilities to outlining a clear path for organizational development.

Just like with RBQM, the implementation of an audit trail review strategy requires two core pillars: (1) **risk management,** encompassing the identification, evaluation, control, and reporting of risks, and (2) **signal detection**, where analytical review identifies anomalies that may indicate underlying process issues. A systematic evaluation of the data collection system and data flow for the trial will be an important step in determining the risks related to the study. This includes an evaluation of the critical data and critical processes by identifying where audit trial records (both user activity and actual data) can pose risks to data integrity.

It is important to note that while we are suggesting the use of the RBQM workflow we are not necessarily suggesting that ATRA analysis and workflow must be implemented in the RBQM tool your organization uses. The choice of where ATRA analysis will occur and be tracked will depend on the maturity of the organization and the tools that they use.

There are also other trial level governance activities underway where real-time audit trail review may be included as part of an existing activity and does not always have to be a "unique" exercise. For example, high risk scenarios with a new technology or eCOA may warrant the audit trail to be reviewed as part of the monitoring visit in conjunction with SDR and other evaluations to confirm the contemporaneous entry and/or entry by someone other than the authorized user. The workflow below is focused on how to address analytics on these data sets, not all possible review of the audit trail data.

As sponsors commence with their audit trail review activities it is advisable to start with the use cases listed in Appendix 3: Use Cases Prioritized with Realized Risk Examples. As they progress, they may build up a library with ATR templates and lists of audit trail checks for series of trials.

Copyright eClinical Forum

Step-by-Step Workflow Description for Managing Audit Trail Reviews

An overview of a step-by-step workflow diagram can be found in *Appendix 4: Step-by-Step Workflow for managing Audit Trail Reviews*

Develop a Strategy for Managing Audit Trail Reviews

- Evaluate
    - Evaluate the computerized system to identify whether it is a high-risk system (e.g. EDC, eCOA, and IRT). High risk may be determined by, but not limited to, patient data, critical data, control of critical process and or safety reporting, and ensure that data integrity measures are in place.
    - If it is a high-risk system, develop a strategy for integrating an audit trail data review (ATR) into the existing clinical data review process.
- Identify
    - Map out key trial activities where audit trail data can assist in detecting potential risks:
    - Activity Audit Trails focus on the actions performed within a system. (e.g. Critical Processes, Access Management, Data entry)
    - Data Entry primarily focuses on the changes made to data within a system (Critical Data, Randomization Blinding, Investigational Product Management.
- Define
    - Establish standard templates for the ATRs:
    - Create a central repository or ATR Library for the types of ATRs to conduct on the targeted system.
    - Refer to *Appendix 3: Use Cases Prioritized with Realized Risk Examples* to identify which ATRs can be added to the targeted system central repository or ATR library and where ATRA can support risk monitoring.

Integrate Audit Trail Reviews into existing Trial Level Risk Management Processes

- Access
    - Evaluate the ATR risks related to the trial during Trial Level Risk assessment.
    - Select the ATRs from a central repository or ATR Library that may affect critical data or processes.
    - Customize and validate according to the trial needs.
- Plan
    - Determine the frequency of ATR, assign responsibility, and required follow-up actions.
    - Note: If no automated ATR tool is being used, limit the scope of ATR reviews to ad hoc manual ATRs which will limit your ability to implement ICH E6(R3) 4.2.3 which calls for planned analysis.

Copyright eClinical Forum

- Document
    - Document the identified risks in the appropriate trial-level plans, including the Risk Management Plan, Central Monitoring Plan, Data Management Plan, and Data Validation Plan.
- Program
    - Set-up programs to support ATR reviews. Customize and validate the ATR reporting to visualize data according to trial needs.
    - Define thresholds for flagging anomalies (e.g., excessive data edits, late entries).

Conduct Data Audit Trail Reviews

- Execute
    - ATR analytics are reviewed as scheduled by assigned role, through automated tracking or manual ad hoc audit trail reviews (ATRs) prompted by signals that necessitate a review. Actions taken include:
    - Prevent: Minimize the risks of non-compliance by establishing robust preventive controls.
    - Monitor: Ensure the timely detection and correction of issues identified by audit trail.
    - Respond: Address and resolve discrepancies effectively, maintaining data quality and regulatory compliance.
- Report
    - When high-risk signals, trends, or anomalies requiring actions are detected, the related issues should be investigated further and communicated through the appropriate channels by the designated roles. All analysis resulting in the issue needs to be reproducible and repeatable and captured.
    - This role, in collaboration with the study team, identifies which risks need to be monitored closely, escalated, or resolved.
- Manage
    - ATR issues are managed through the Trial Issue Management process and assigned to the appropriate roles. Review flagged audit trail entries to determine the root cause of the issue.
    - As part of ongoing risk evaluation- pre-defined ATR risks are re-evaluated regularly for latest trends and insights
    - Document findings and retain according to organizational retention guidelines.
    - If necessary, update the Risk Management Plan to reflect new ATR risk findings

Integrate Audit Trail Reviews into existing Trial Level Risk Management Processes

- Using FMEA (Failure Mode Effect Analysis) to rank risk → 1x3x3 = 9
    - Likelihood = 1: Rare
    - Detectability = 3: High with an ongoing check of data changes at a site.

Copyright eClinical Forum

- Impact to Patient Safety= 3: High, if sites are changing primary endpoint data and/or safety data there is a concern about the conduct of the trial.

Conduct Data Audit Trail Reviews

- Execute
    - Monitor the number of data changes at a site (on critical data and non-critical data). Normalize those for changes that were requested (queries) and those not requested. Also look at the number of changes per data point entered (a site with many changes but even larger number of datapoints entered is not the same as a site with a smaller number of changes of very little data).
    - Compare the number of changes against other sites in that study or other data in the therapeutic area to identify outlier behavior.
    - Review sites with an unexpectedly large or small number of data changes (normalized correctly).
- Report
    - Perform additional analysis or investigation to understand possible causes of many data changes. It is important to ask questions about why these changes might have happened.

For example: What does it mean if a participant always enters data in sequential order according to subject number? What does it mean if an investigator suddenly changes endpoint-related data for certain participants after gaining access to specific lab values?

These are the kinds of questions that highlight the importance of not just detecting anomalies but also interpreting and acting on them.

- Potential actions for addressing anomalies
    - New employee is making mistakes.
      Action: Have a conversation with the employee to understand why they are making mistakes, then possibly retrain the employee in proper data entry.
    - A confusing CRF page (many sites having a similar problem can point to a database problem). **Action:** Make the CRF page clearer and retrain
    - A lack of understanding of the protocol causes the PI to enter and then delete information that they feel might no longer be relevant.
      Action: First understand if this is a site problem or a study problem (unclear protocol) and address the issue.
    - Misconduct at a site; a PI removed medical history data to make a patient pass inclusion/exclusion criterion.
      Action: Address site misconduct, shut down site, etc.
    - Where these are core GCP compliance aspects, these should be properly linked and referenced.

Copyright eClinical Forum

- There are times when data review uncovers an unexpected issue that then needs to be explored. If that issue turns out to be a missing risk, the organization needs to add it to the risk library.
- Note, while some of this might be statistical in nature, they might also require a site visit, a site audit or other non-quantitative review to determine root cause. Statistical correlation does not imply causation. A complete analysis of what is happening will almost always require contact with the individuals involved at the site.

- Manage
  - Depending on the root cause of the issue, apply corrective action. This can be anything from site retraining, to addressing issues in study instructions, etc.
  - Continue to monitor metrics to see that corrective action fixed the process problem and the metric returns to typical behavior (generally determined by a statistical measure).

6.2 Maturity Assessment

As clinical trials become more complex and data-driven, implementing Audit Trail Review Analytics (ATRA) requires more than technical tools. It demands coordination across systems, processes, and teams. To support scalable, high-quality analytics, organizations need a structured understanding of where they are today and what is needed to advance.

This section introduces a capability-based maturity model to help sponsors, CROs, and solution providers assess their readiness and define a roadmap for ATRA implementation. The model aligns with ICH E6(R3) principles of data integrity, proportionality, and quality-by-design, and includes four core domains:

1. Data Access
2. Analytical Capabilities
3. Workflow Integration
4. Organizational Attributes

To form a complete score each of the areas will be scored from 1-5, 1 being the lowest level and 5 being the highest level. Not all scores will be strictly linear. This score would then be used to form the basis of a roadmap, and for vendor selection.

Data Access

Reliable audit trail analytics start with timely, structured, and automated access to system-generated metadata. Without stable data pipelines, organizations cannot scale analytics, ensure consistency, or support risk-based quality management across studies. This domain assesses the maturity of the audit trail data pipeline—not just the type of data accessed, but how it is accessed and prepared for analysis.

Each system might have its own score. Key systems include but are not limited to: EDC platforms, ePRO/eCOA systems, eTMF, IRT and operational metadata sources, System logs and similar information.

SCORE	TITLE	DESCRIPTION
1	Manual	Data is accessed through manual exports or ad hoc file pulls, with no automation. Often stored in spreadsheets or individual flat files.

Copyright eClinical Forum

SCORE	TITLE	DESCRIPTION
2	Semi Manual	Data is pulled into some centralized locations but there are still many manual steps.
3	Local Automation	Systems support automated data access, but pipelines are system-specific or limited to a local environment. Data is not reusable across systems.
4	Centralized	Data is extracted automatically across systems into a centralized repository (e.g., data lake or analytics workspace).
5	Reusable	Data Pipelines exist which are validated, reusable, and support multiple use cases.

This maturity scoring helps teams evaluate whether current data infrastructure supports scalable ATRA and where investment is needed to move toward a harmonized, cross-study, and cross-functional analytics model.

Analytical Capability

Analytics maturity reflects an organization's ability to define key metrics, apply statistical thresholds, and scale insight generation. The level of analytical maturity might vary from data source and use case to use case. For example, you might have high maturity for ePRO (vendor provided) but lower for EDC audit trail. This should be done utilizing the different types of risk scenarios outlined in *Appendix 3: Use cases Prioritized with Realized Risk Examples*. While this is a ranked score, tools that specialize in one area but skip steps should be called out. There are tools on the market that can aggregate but do not provide time trend analysis, and others that can perform root cause analysis but do not aggregate. When evaluating your system call out what is included and what is not, using real-world examples.

SCORE	TITLE	DESCRIPTION
1	Manual Analysis	Analysis is completed manually on an ad hoc/as needed basis.
2	Descriptive Analysis	Quantify volume, frequency, and distribution of actions, flag outliers.
3	Time-Trend and Rule-Based Monitoring	Detect unexpected activity patterns using control charts, threshold rules, or predictive models.
4	Advanced Analytics	The ability to group data across domains and generate an aggregate score for a Study/Site/PT/ Domain.
5	Root Cause Analysis and Feedback Loops	The ability to perform and track root cause analysis, completed and then integrate back into regular checks.

Copyright eClinical Forum

Workflow Integration

Analytics without a workflow and process are rarely effective and should therefore be part of the ongoing data cleaning and data quality process.

SCORE	*NAME*	*DESCRIPTION*
1	**Ad Hoc**	ATRA activities are informal, manual, and disconnected. Audit trail data is rarely used; roles unclear.
2	**Initial**	Some structured processes exist. Basic data review and triage occur manually with limited consistency.
3	**Defined**	Formal workflows are in place. Audit trail analytics support some triage; roles partially defined.
4	**Managed**	Integrated systems automate signal detection and routing. ATRA informs decisions; clear RACI is in place.
5	**Optimized**	Proactive, enterprise-wide RBQM. Real-time ATRA supports continuous improvement; cross-functional ownership is embedded.

Organizational: People Skills, Roles and Responsibilities

ATRA spans clinical, data, and quality teams. Clear ownership and governance, especially across sponsors and solution providers, ensure findings are acted upon.

SCORE	STAGE	DESCRIPTION
1	**Unaware**	Teams have little or no exposure to audit trail data. ATRA is not recognized as relevant to daily work. No training or curiosity-building in place.
2	**Exposed**	Individuals are introduced to audit trail concepts. Early engagement begins via presentations or passive consumption. Curiosity may be emerging.
3	**Learning**	Teams begin hands-on work with ATRA. Training is available. People experiment with basic data exploration. Recognition of its potential is growing.
4	**Skilled**	Individuals across clinical, data, and quality functions can independently explore and interpret audit trail data. Curiosity is becoming cultural.
5	**Integrated & Curious**	ATRA skills are embedded in team capabilities. Continuous learning is supported. Data curiosity and exploration are core to how problems are solved.

Copyright eClinical Forum

6.3 Strategic Roadmap

The maturity assessment provides a snapshot of your organization's current capabilities across four ATRA domains: **Data Access, Analytical Capability, Workflow Integration,** and **Organizational Aspects**. However, knowing your current level is only the starting point.

The next step is to define how to move from **where you are** to **where you need to be**, based on the **risk scenario** or **use case** in focus. Different use cases, such as Data Forensics or User Access and Activity, may require different levels of maturity. The roadmap below supports flexible progression, helping organizations design and implement ATRA in a structured, scalable way.

Progressive stages include foundational steps such as automating data access and initiating basic analytics, followed by more advanced capabilities like cross-system integration, predictive modeling, and embedded issue management workflows. Each stage should be aligned with use cases, tooling requirements, and ownership roles to help teams prioritize actions, allocate resources, and track progress toward a scalable, risk-proportionate ATRA program.

Key Principles

- **ATRA maturity is not one-size-fits-all**: It varies by system, risk type, and organizational readiness.
- **Progress requires alignment**: Tools, vendors, processes, and people must evolve together.
- **The roadmap is modular**: You can start at different points depending on current capabilities and advance iteratively.

Roadmap Phases: From Discovery to Optimization

PHASE	FOCUS	KEY ACTIONS
1. DISCOVER	*Understand where you are*	- Identify audit trail data sources (EDC, ePRO, IRT, eTMF, etc.) - Assess maturity across ATRA domains using the scoring model - Begin stakeholder education and build awareness - Understand variations across use cases
2. DESIGN	*Define where you're going*	- Prioritize high-impact use cases (e.g., informed consent errors, site-level change patterns) - Set target maturity levels for each domain - Align on goals with internal stakeholders and vendors
3. BUILD	*Implement localized solutions*	- Automate audit trail data access for select systems (start with high-priority use cases) - Develop analysis-ready datasets - Apply descriptive and rule-based analytics - Log findings in structured trackers
4. SCALE	*Expand across systems and teams*	- Build cross-system data pipelines and normalize audit trail structures - Integrate ATRA findings into operational workflows (e.g., issue management or monitoring plans) - Conduct root cause investigations

eClinical Forum

PHASE	FOCUS	KEY ACTIONS
5. OPTIMIZE	*Embed continuous improvement*	- Deploy advanced analytics (e.g., machine learning, clustering) - Enable real-time alerts and dashboards for proactive site and study oversight - Institutionalize upskilling and governance practices

Putting It into Practice

Few organizations progress along this path in a straight line. For example, a sponsor might already have advanced analytics capabilities for ePRO data but still be exploring options for EDC. Similarly, a CRO could depend on vendors for its technical infrastructure while focusing significant resources on developing its staff.

What matters is that organizations:

- Anchor the roadmap in real use cases
- **Use the maturity model** to identify domain-specific gaps
- **Create a plan to close those gaps** across tools, vendors, skills, and processes

Example Path: From Manual Change Logs to Proactive Site Oversight

Use Case	Current Maturity	Target Maturity	Actions Needed
High frequency of manual data changes	Data Access: 2 Analytics: 2	Data Access: 4 Analytics: 5	Automate EDC audit trail extraction, apply root cause analysis, and integrate alerts into oversight
Informed consent workflow anomalies	People: 1 Workflow: 2	People: 4 Workflow: 4	Train study teams on audit trail review, embed findings into issue triage process

From Compliance to Capability

Audit trail data is often seen as a regulatory requirement, but it needs to become a strategic capability. With a clear maturity baseline, a use-case-driven roadmap, and investment in the right people and partnerships, ATRA can move from being reactive and manual to being a proactive driver of quality, insight, and oversight.

Copyright eClinical Forum

7. Conclusion

The evolving regulatory landscape, most notably the release of ICH E6(R3), has placed increased emphasis on proactive, risk-based oversight of clinical trials, including the effective use of audit trail data. This paper has presented a practical framework for organizations seeking to implement audit trail analytics as a core component of Risk-Based Quality Management (RBQM).

Throughout the document, we explored how audit trail data can be used to surface meaningful insights about site behavior, data integrity risks, and protocol compliance. We introduced a maturity model for analytics, described foundational data access requirements, and emphasized the importance of system integration, signal detection, and actionable workflows. We also provided a readiness assessment to help organizations evaluate their current state and a roadmap to guide the development of scalable, efficient audit trail analytics programs.

The key takeaway is that audit trail analysis is not simply a technical exercise but an operational, analytical, and strategic discipline. Implementing ATRA requires collaboration across sponsors, solution providers, and technology vendors; it requires thoughtful alignment of data, tools, workflows, and people.

For all organizations this will be a need for continual improvement as data access and tools mature. But the value is clear: by operationalizing audit trail analytics, we can more effectively detect and mitigate risks, improve data quality, and build a stronger foundation of trust and integrity in the conduct of clinical research.

The eClinical Forum encourages all stakeholders—data managers, auditors, quality teams, and solution providers—to build upon the foundations laid out in this paper and to continue advancing the science and practice of audit trail analytics in support of safer, smarter, and more reliable clinical trials.

Copyright eClinical Forum

Appendices

Appendix 1: DISCLAIMER and LICENSE for FAIR USE of eCLINICAL FORUM MATERIALS

This work is the property of the eClinical Forum and is released under Creative Commons – Attribution-Non-Commercial-No Derivatives 4.0. Under the terms of the license, you are free to:

- **Share,** copy and redistribute the material in any medium or format

4. The licensor cannot revoke these freedoms if you follow the license terms.

Under the following terms:

- **Attribution:** You must give appropriate credit, provide a link to the license, and indicate if changes are made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- **Non-Commercial:** You may not use the material for commercial purposes.
- **No Derivatives**: If you remix, transform, or build upon the material, you may not distribute the modified material.

5. **No additional restrictions:** You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

- You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.
- No warranties are given. The license may not give you all the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Copyright eClinical Forum

Appendix 2: Definitions and Acronyms

AE	Adverse Event
ATR	Audit Trail Review
ATRA	Audit Trail Review Analytic(s)
BYOD	Bring Your Own Device
CRF	Case Report Form
CRO	Contract research organization
CTMS	Clinical Trial Management System
eCF	eClinical Forum
eCOA	Electronic Clinical Outcome Assessment
eCRF	Electronic Case Report Form
EDC	Electronic data capture
EMA	European Medicines Agency
ePRO	Electronic Patient Reported Outcome
eTMF	electronic Trial Master File
FDA	Food and Drug Administration
FMEA	Failure Mode and Effects Analysis
I/E	Inclusion/Exclusion
ICH	International Council for Harmonization
ICH E6(R3)	ICH E6(R3) integrated addendum
ID	identification
IDMS	Investigational Drug Management System
IRT	interactive response technologies
IT	Information Technology
PI	Principal Investigator
PII	Personally Identifiable Information
QbD	Quality by Design
RACI	Responsible, Accountable, Consulted, Informed

RBQM	Risk Based Quality Management
SAE	Serious Adverse Event
SCDM	Society for Clinical Data Management
SD	Standard deviation
SDTM	Study Data Tabulation Model
TA	Therapeutic Area

Copyright eClinical Forum

Appendix 3: Use Cases Prioritized with Realized Risk Examples

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
Data Forensics	1.1: Incomplete data- data deleted	Inadvertent deletion of critical study data (e.g. safety or endpoint data) may result in incomplete datasets and thus impact analysis results. This may include a few different scenarios e.g. a) A site user is accidentally / inadvertently deleting data (especially log form data such as adverse events & conc. medications) b) A system administrator is accidentally deleting data records during e.g. system maintenance	High	EDC, IRT, eCOA	The IT admin super user accidentally deletes a large number of objects from the system.	This resulted in loss of data which required a recovery from the backups
Data Forensics	1.2: Changes to inclusion/exclusion (I/E) criteria, primary efficacy, key secondary data, safety data and/or other critical data as defined by risk-based monitoring	a) Inadvertent changes to inclusion / exclusion criteria or eligibility data may result in enrollment of ineligible patients and as such affect the outcome of the study b) Inadvertent changes to relevant analysis data (primary & secondary efficacy & safety) or other critical data types, thereby affecting outcome of the study	High	EDC, eCOA	Sites initiating DCFs to update baseline (pre-randomization) pain scores recorded by the patient in the e-diary. Sites who (frequently) requested those changes became subject to increased monitoring and potentially being audited. Note: Depending on the type of data, you may accept a low amount of data change requests; but the highlighted case was not just an inclusion criterion, but also a baseline score. Every data change request was critically verified. Critical EDC data were changed multiple times, back and forth query by site users	Average baseline pain (over selected days) was used as inclusion criterion. Simply accepting data changes without revalidation (and check against potential source data) would have led to potential inclusion of ineligible patients. The fact that data was changed multiple times back and forth resulted in suspicions that the site was either not understanding the CRF or trying to tamper data

Copyright eClinical Forum

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
Data Forensics	1.3: Inconsistent data- critical data changes after key timepoints such as data locks or subject's study disposition status is marked as complete	In selected studies there is a risk that critical data are changed after key study timepoints have been met e.g. after a safety or analysis cutoff (e.g. soft lock). Such changes - if undetected - may invalidate the previously reported safety or other analysis data This risk and its related impact is dependent on the specific study parameters (e.g. are changes to the reported data expected; are data protected from updates in the system) and can thus not be considered as standard checks	Medium	EDC, eCOA IRT (if using to collect CRF-like data)	NA	NA
Data forensics	1.4: Timing of changes after initial data entry or source is obtained is longer than expected	**EDC:** Timing of data changes after initial data entry or assessment is longer than expected **eCOA:** Timing of initiating (requesting) data changes after initial entry or assessment is longer than expected	Low	EDC, IRT, eCOA	Sites answered queries and / or initiated data changes very late in the process (compared to the timepoint of collecting the data)	This suggested that sites were not verifying their data in timely manner and resulted in questions on integrity and reliability of the reported data
Data Forensics	1.5: Data collected at suspicious timing or not contemporaneous to event	Data collected at suspicious timings or not contemporaneous to event e.g. entry times that are implausible for the time zone of the site or patient. This may suggest that patient-reported data or other electronic source data is initially collected on paper and retrospectively entered in the system	High	eTMF requiring contemporaneous filing IRT & eCOA when collecting patient-reported data EDC only	eTMF audit trail indicating that filing was not contemporaneous, but rather just before the announcement of the inspection eCOA assessments for multiple patients were performed at the same time resulting in patients going through the study in synchronous way	the TMF was not deemed readily available for inspections / the TMF did not always contain the essential documents this was picked up during inspections re. post submission and lead to major regulatory concerns. Sponsor had to justify and re-analyzed the data with and without the site in question Ultimately it turned out that the way the site used flexibility in the protocol to put patients on a synchronous pattern. While suspicious, it was formally not

Copyright eClinical Forum

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
				when real-time data entry is required or applied	Identification of 2 subjects with almost the same diary timings over a period of several weeks Unusual patterns of data, especially in the ePROs (eg. Weekends, Low mean duration to complete) Nearly 70% of daily ePRO diary entries at a site occurred in same hour of day (6 pm local time) across all patients	against the protocol Likelihood that this occurs is extremely low thus triggering suspicion of fraud 4th highest enrolling center of study was audited & major findings identified: gaps in PI oversight, gaps in training to patients, gaps in record-keeping It turned out that devices were never distributed to patients and that site staff entered data at close of business every day. Sponsor closed the site and removed patients from the analysis
Data Forensics	1.6: Data not collected per protocol timing	Data were not collected or entered in line with protocol requirements	High for critical data Medium to Low for other data	EDC, eCOA, IRT (if using to collect CRF-like data)	Site with substantial amount of sleep diaries not completed within 60 minutes from final awakening as required by the protocol. Timepoints for data collection / data entry in eCRF for one site very different compared to others based on a plot of collection / entry timepoints	This was not deemed as a fraud issue but led to low validity of the PRO responses and exclusion of patients from the analyses, thus reducing power This suggested that the site was not collecting or entering the data in appropriate timeframe following the visit and resulted in questions on reliability / usability of the data
Data Forensics	1.7: Varying completed durations	Data entry duration is either very short or very long indicating that there are potential issues with the completion of a scale, assessment or instrument. E.g. a) Very long duration may indicate that	High	eCOA, IRT when collecting patient-reported data	For a newly added site in the study, the mean questionnaire completion was significantly shorter than for other sites	Initial suspicion of fraud / misconduct, although may also point to lack of training, sloppiness, etc 4th highest enrolling center of

Copyright eClinical

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
		assessment or scale is not fully understood b) Very short duration may indicate that the scale or assessment is not performed adequately (just randomly filling in results) or that results are first captured on paper (underlying source)		EDC either for assessments that require real-time data entry or for which a start & end-time is provided on the CRF	Unusual patterns of data especially in the ePROs (eg. Weekends, Low mean duration to complete) At one site (20+ patients), the average duration to complete a key efficacy assessment related to psoriasis (using eCOA) was very short: 4 minutes vs. a study average of 15 minutes	study was audited & major findings identified: gaps in PI oversight, gaps in training to patients, gaps in record-keeping This finding led to further investigations which revealed further significant issues at the site
Data Forensics	1.8: Reporting time is outside of required timing	Data entry date is too late after data collection dates thereby not meeting established entry criteria for timeliness. Delayed entry of data prevents timely review / monitoring of critical study aspects (e.g. safety; efficacy) and thus delayed identification of issues that may impact outcome of the study **Note:** Focus is here on delayed data entry (considering EDC). Late SAE reporting (to sites / ethics is normally part of pharmacovigilance and constantly monitored)	Medium	EDC	NA	NA
Data Forensics	1.9: Missing data	a) Missing (critical) data points may indicate poor conduct by the site or patient (in case of an eCOA) and may impact the outcome of the study b) Missing PI signatures indicate lack of review and endorsement of the reported data by the PI which may lead to issues with acceptance of the data	High	EDC, eCOA Note that investigator signatures may not be captures	PIs were not signing the eCRF pages in a timely manner	This resulted in an internal audit finding and was deemed a non-compliance issue. As a mitigation, it led the team to regularly review the signature audit trail report to monitor timely signatures and ensure that potential safety issues would not go unnoticed

Copyright eClinical Forum

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
		c) Late signatures are a sign that investigators do not review and endorse data in a timely manner. While looking for missing signatures may help in immediately mitigating the gap, checking for late signatures allows for identification of sites that have a persistent problem with signing in timely manner (note: this could also be considered under 2.3)		in all systems		
Data Forensics	1.10 Patient making multiple changes to questionnaires or diaries prior to submission of responses	Excessive amount of data changes to a diary or questionnaire prior to submission may indicate a lack of understanding of questions or poor design of the diary	Medium	eCOA	NA	NA
User Access & Activity	2.1: Unauthorized user-access / Changes by right role, right person	a) A user is assigned incorrect access rights / user role and is thus able to perform or performing actions in the system that they are not supposed to perform e.g. - study coordinator is able to sign off on data instead of the investigator (EDC) - study coordinator is able to enter data for patients at another site - sponsor role (CRA; CDM) enters or correct data instead of site staff (EDC) or for a patient (eCOA) - site users able to mark data as source verified b) A user is no longer working on a study or system but has not been deactivated in time and can thus continue to perform actions in the system c) Study site personnel has shared password	High	EDC, IRT, eCOA	In an eDiary dataset, a patient in the trial was assigned 'Clinician' as user role. This was the only patient with such result, all other patients had Subject as user role Data changes made a sponsor / CRO role (project manager in this case) while not permitted for eCOA data Study site personnel shared passwords with other staff members not working on the study / authorized to access the clinical system and	Issue turned out to be a programmatic error and required updated programming and correction of all previously reported entries NA This may be seen as misconduct on site and result in major / critical audit or inspection findings, potentially resulting in rejection of (part of) the study data

Copyright eClinical Forum

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
		with other staff members not authorized on the study / to access the clinical system and enabled them to enter / update data			enabled them to enter / update data in the eCRF	
User Access & Activity	2.2: Access without training	a) PI completed scales prior to rater training b) Sponsors or sites getting access without training c) Patient did not receive tool training	Medium	EDC, eCOA	NA	NA
User Access & Activity	2.3: User login and activity	Users are not accessing the clinical trial system routinely as per protocol requirements or study expectations, thereby impacting data quality and/or study oversight. Examples include: a) Investigators not reviewing patient data in timely manner (EDC, eCOA) b) CRAs not performing site monitoring in timely manner c) Data managers not reviewing data in timely manner (EDC when using review flags) e) Study documents are not uploaded in a timely manner (eTMF)	Medium	EDC, IRT, eCOA, eTMF	PI's did not sign CRF pages because they did not have access to the eCRF system Sponsor only verified CRO's eTMF at selected (critical) timepoints (e.g. right after announcement of inspection or other critical timepoint in the study)	As this was only discovered at the end of the study it resulted in delaying database lock and considered a compliance issue due to lack of PI oversight Sponsor did not have sufficient and ongoing overview over the quality of the CRO's TMF filing, thus not executing their sponsor oversight obligations.
User Access & Activity	2.4: Lack of PI access at site	PI cannot review or sign-off data or cannot unblind in an emergency situation (e.g., due to PI was never granted access by vendor, PI access was revoked due to inactivity, or PI never took required training)	High	EDC, IRT, eCOA	PI did not sign CRF pages because they did not have access to the eCRF system	As this was only discovered at the end of the study it resulted in delaying database lock and considered a compliance issue due to lack of PI oversight
System Integrity & Security	3.1: Suspicious account details	Password or PIN too simple / obvious and therefore prone to misuse or lack of protection of data e.g. a) PIN or password with same character	Low	eCOA	Blinded vendor analyses revealed that at certain study sites patients had the exact same or similar PIN number. E.g. PIN number corresponds	If identified in an audit or inspection it may trigger suspicions of fraud (i.e. inspectors wondering whether data is entered by site

Category	Use Cases: Data integrity risk	Risk scenarios / descriptions	Overall ranking	Applicable system(s)	Example (description of the case / circumstances)	Potential root cause & impact (how it hurts)
		sequences used across subjects at the same site b) PIN or password based on the site number c) PIN or password is based on the participant's DoB			with DoB, with site ID or same sequential numbers (1234)	staff instead of patients themselves)
System Integrity & Security	3.2: Duplicate datasets	Duplicate data (both individual records & complete data sets) may be available for the same timepoint, thereby causing uncertainty which set of data is the latest / valid that needs inclusion in the analysis	Low	Any system	NA	NA
System Integrity & Security	3.3: Date / time-stamp inaccurate	Inaccurate date/timestamps may result in implausible data. These may be caused by e.g. a) Dead battery or other technical malfunctioning / synchronization issues b) Internal clock manipulation resulting in either past or future in-device timestamps c) Time zone adjustments	Low	eCOA (app-based), handhelds	Due to battery issues or synchronization issues, the internal clock of an eCOA device got reset to an implausible value, resulting in study measurements that were not matching with the study / visit period	Due to this issue the study measurements were no longer matching with the study / visit period and required manual updates of the data via the eCOA supplier. If missed, the data might have to be fixed after database lock or excluded from the study

Copyright eClinical

Appendix 4: Step-by-Step Workflow for Managing Audit Trail Reviews

