INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

OPEN ACCESS

Review Paper

Data Quality and Integrity Investigation in the Analytical Laboratory

Vairangi Patel*, Dhara Patel, Grishma Patel, Dhananjay Meshram

Department of Pharmaceutical Quality Assurance, Pioneer Pharmacy College, Vadodara - 390019, Gujarat, India

ARTICLE INFO

Published: 14 May 2026
Keywords:
Data Integrity; Data Quality; Analytical Laboratory; ALCOA Principle; ALCOA +; Regulatory Compliance; Audit Trail; Risk-Based Strategy; Good Manufacturing Practice (GMP); Quality Control (QC); FDA; EMA; MHRA; WHO; Data Integrity Breach; Life Cycle Management.
DOI:
10.5281/zenodo.20178235

ABSTRACT

Investigating the topic of data integrity in analytical labs may be considered one of the significant areas in pharmaceutical quality systems, which enables regulatory compliance and, ultimately, guarantees safety for patients. As the result of the development of Pharma 4.0 and the increasing digitization of the quality control process, managing electronic data has become one of the crucial aspects of lifecycle-based quality assurance. This review is focused on recent tendencies concerning data integrity standards and discusses the new guidelines suggested by FDA, EMA, MHRA, WHO, and EU GMP Annexes 11 and 22 on a global scale. Integrity of data can be defined using ALCOA and ALCOA+ attributes. They include attributability, legibility, contemporaneous, originality, accuracy, completeness, consistency, persistence, and accessibility. Examples of data integrity breach observed in analytical laboratories include data fabrication, data duplication, manipulation of data out of specification, backdating, lack of SOPs, shared passwords, insufficient raw data archiving, access control, and audit trail reviews. The need arises for risk-based strategies to identify any weaknesses in the data handling process and implement remedial action. Proper management of audit trail systems, computer validation, and cultivation of the quality culture within the laboratories becomes especially significant. The use of blockchain technology and artificial intelligence may prove beneficial in this regard. In general terms, it could be stated that a thorough investigation into data integrity helps guarantee laboratory data integrity.

INTRODUCTION

1.1 Background of Analytical Laboratories [1]

The majority of businesses have been audited and, when needed, they had to "defend" the work done in the analytical labs. In order to support the

***Corresponding Author:** Vairangi Patel
Address: *Department of Pharmaceutical Quality Assurance, Pioneer Pharmacy College, Vadodara - 390019, Gujarat, India*
Email ✉: patelvairangi2728@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

validity of the analytical results, laboratories have historically tended to offer information on the validation of their methods and procedures, the qualification and suitability of their analytical equipment, and information about the training of their laboratory workers. However, the FDA, the UK Medicines and Healthcare Products Regulatory Agency (MHRA), and other regulatory bodies' emphasis on data integrity during audits may mean that traditional methods of laboratory audit preparation and audit "defense" giving proof that the analytical results are not fake has taken preference over giving information related with technical reason and scientific rationale in a data integrity-focused audit. The recent announcement of MHRA and FDA warning letters has brought attention to the growing need of data integrity in the laboratories. Many laboratories currently have serious issues regarding data integrity. To exacerbate these worries, there are currently few reliable reference materials on the topic, and even the word "data integrity" might have wildly disparate definitions or interpretations.

1.2 Importance of Data Quality and Integrity [2][3]

Regulatory compliance: Data must be traceable, legible, contemporaneous, authentic, and accurate (ALCOA+) for laboratories to meet regulatory requirements put forward by entities such as the FDA and MHRA. There will be outcomes, including warning letters, penalties, or even product recalls, should there be a failure to adhere to these standards.

Accuracy and reliability: Research show that laboratory analysis is accurate and reliable in representing real experimentation rather than tampered data.

Scientific reproducibility: In scientific research, integrity in data is important since other scientists can reproduce experiments based on data presented.

Data Life Cycle Management: Data should be monitored through its entire life cycle for integrity and security.

Audit Trails: Examination of electronic audit trails helps in maintaining the integrity and security of the data.

Root Cause Analysis: Identifying the root cause of deviations helps prevent recurrence and guarantees good quality data in the future.

Out-of-Specification Results Management: Not conducting investigations for any out-of-specification results and continuing to test until passing instead.

1.3 Scope and Objectives of the Review [4]

This includes all elements that have anything to do with laboratory data management from receiving samples to archiving data. It would encompass computerized systems such as audit trails and access controls, the management of hard and soft copy records, reprocessing and deletion of chromatography and analysis data, and finally the training, ethics, and culture of data integrity of lab personnel. The primary purposes of such researches include detecting and preventing fraud, achieving regulatory compliance, checking the correctness of findings, securing the system from loss or tampering of data, ensuring data completeness, and improving laboratory processes through recognizing the vulnerable points.

2.Fundametals of Data Quality & Data Integrity

2.1 Definition of Data Quality [7]: Data quality means “with an emphasis on timeliness, correctness, and completeness, data quality evaluates how well data fulfils its intended function.”

2.**2 Definition of Data Integrity** [7]: Data Integrity, “which focuses on preventing corruption or illegal alteration, guarantees that data is

accurate, consistent, and safe throughout its whole lifecycle."

2.3 Key principles of ALCOA, ALCOA++ [1, 8-9]

When Stan Woollen was employed at the FDA, he used the acronym ALCOA to assist him remember compliance words related to data quality. Since then, the acronym has been widely associated with data integrity. Data integrity is covered in Appendix 3 of the good automated manufacturing practice (GAMP) good practice guide "A Risk-Based Approach to GxP Complaint Laboratory Computerized Systems". Because they include further terms based on the European Medicines Agency's concept paper on electronic data in clinical trials, the terms used in the appendix are commonly referred to as "ALCOA +." Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Durable, and Available are the terms related to ALCOA +.

2.4 Life cycle of Data Analytics:[10]

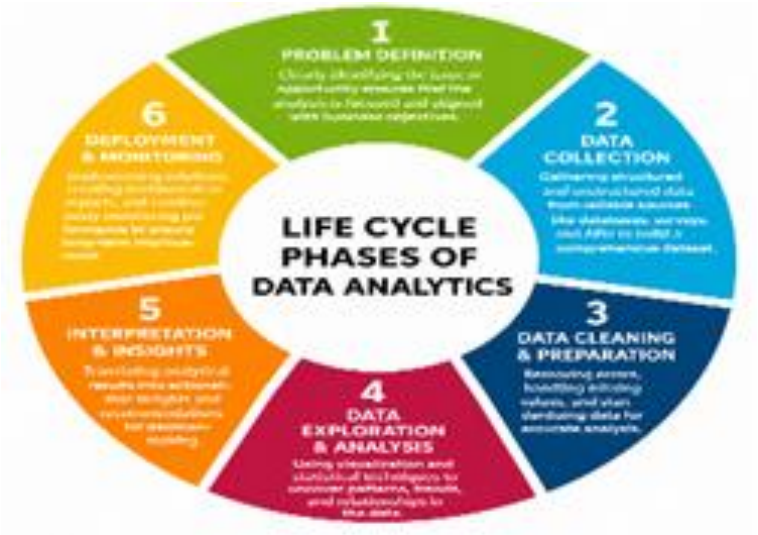

Figure 1: Life cycle of Data Analytics

2.5 Types of Laboratory Data: [2-3]

1. Raw Data: Original source records, consisting of information obtained using instruments like output from a spectrometer, data obtained from HPLC chromatography, and hand-written records in lab notebooks, which must be permanently documented when observed to properly reconstruct and evaluate a study.
2. Metadata: Additional information on the record that includes context, contents, structure, or other management characteristics necessary for monitoring changes in the record.
3. Instrument Data: System suitability test results, calibrations, detectors' settings (wavelengths, temperature), etc., constitute instrument data.
4. Sample and Method Data: Information about standard/reagent logs, weights (balance), and sample preparation.
5. Processed Results: Any kind of processing performed with external software (Excel, for example), entry into LIMS system, etc.
6. Electronic Records: Any document stored in a form readable by a computer (PDF reports, audit trail for full electronic systems and mixed computer/human systems, etc.).
7. Non-reportable Data: When there is a breach in data integrity, "trial" samples, test injections, or cancelled run are often omitted.

3.Regulatory framework and guideline

3.1 Overview of Global Regulatory Expectations [11][12]

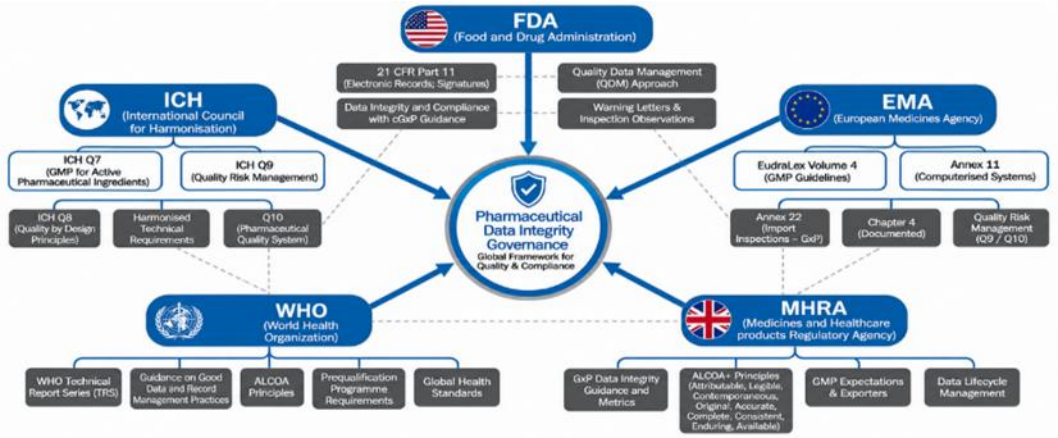

Figure 2: Overview of Global Regulatory Expectations

3.2 Good Laboratory Practices (GLP) [13]

The GLP assures proper personnel qualifications and their assignments in a study by using a Study Director. QA auditing ensures compliance with SOPs and procedures. Properly equipped laboratories with well-calibrated machines should be present for proper study conduct. Correct maintenance of test systems is needed for proper results. All activities must be conducted according to SOPs, while proper data archiving and reporting are vital. This process is regulated by the FDA 21 CFR Part 58 and OECD GLP Principles.

3.3 Good Manufacturing Practices (GMP) [11]

Consequently, Good Manufacturing Practices make sure that the food, medicine, and devices should be produced efficiently for ensuring quality, effectiveness, and safety. All of these include proper documentation (standard operating procedures), accredited plant and machinery, trained staff members, controlled production process, and quality control. Some of the most significant guidelines in this regard are FDA 21 CFR 210/211, WHO GMP, EU GMP, and ISO. Anomalies, avoiding contamination, documentations, continuous improvements (cGMP), and various other factors can be included in Good Manufacturing Practices.

3.4 FDA Data integrity Guidance [14]

The regulation by FDA on data integrity makes sure that all the data pertaining to pharmaceuticals is acceptable, accurate, and follows the rules of cGMP, and ALCOA+. This involved 21 CFR Parts 210, 211, 212, and Part 11. The process of verification must be undertaken for all systems, data must be accessed only by those authorized, and auditing trail must be kept. In addition to building a strong data integrity culture, documentation and investigation of incorrect data are key components.

3.5 WHO and MHRA Guidelines [15]

For pharmaceutical products, the international standards are set by the World Health Organization (WHO); however, legal standards for the same are adhered to by the UK-based Medicines and Healthcare Products Regulatory Agency (MHRA). The pharmaceutical industry regulation is hierarchical in its approach. WHO sets the international standards for manufacture, storage, and distribution of medicines that consider the elements of quality, safety, and efficacy; on the other hand, MHRA follows the guidelines set forth by agencies such as ICH under the "Orange Guide" guidelines.

3.6 OECD Principles [16]

Transparent, clear, and well-founded regulations are integral parts of the OECD principles. They pay much attention to the stakeholder involvement, cost-benefit assessment, burden reduction, risk proportionality, and clear objectives for regulation. The RIA, assessment of existing legislation, international cooperation in regulation, and good governance are significant principles. Taking everything into account, these principles encourage responsible approaches based on risk enforcement and development.

4. Sources of Data integrity Issues in Analytical Labs [1, 17-18]

4.1 Human Errors

The biggest reason why data integrity failures occur within the analytical laboratories is human error, which involves not only unintentional errors but intentional misconduct as well. Poor documentation, mis-handling of samples, conformity testing, and ignoring ALCOA+ are key the issues that needs to be addressed.

4.2 Instrumentation and Software Issues

Poor validation, shared logins, disabled audit trails, and antiquated systems all contribute to data integrity problems in analytical labs and raise the possibility of data tampering. Unsynchronized clocks, improper qualification, unsuccessful audit trails, and insufficient data storage are examples of instrument-related problems. Shared passwords, unauthorized user access, unverified spreadsheets, record deletion, and insecure legacy systems are examples of software problems. All things considered, these loopholes jeopardize traceability, data dependability, and adherence to ALCOA principles.

4.3 Documentation Deficiencies

There are many reasons for data integrity failures due to non-compliance with ALCOA+ requirements, which majorly include poor documentation, such as non-contemporaneous documentation, lack of metadata, and poor Documentation Practices. Such issues, includes both intentional data manipulation and transcription errors, which often lead to regulatory investigations, such as FDA Form 483s and warning letters.

4.4 Data Manipulation and Fraud

Intentional data manipulation and fraud, such as "testing into compliance," data falsification, and illegal manual integration, are common causes of data integrity problems in analytical laboratories that commonly result in FDA warning letters. Due to the things like insufficient audit trails, shared user access, and pressure to fulfil, these violations frequently include circumventing ALCOA+ standards, which can result in product recalls and harsh regulatory action.

4.5 Environmental and Systemic Factors

The environmental and system-related factors which mainly affect the precision and adherence of the process can be seen as the major source of the problem in terms of data integrity in analytical laboratories. Some examples of the environmental issues include changes in temperature and humidity, vibration, air contamination, and unreliable sources of energy. Systemic issues which have an adverse effect on the process include poor manual data management practices, inadequate controls regarding system access, lack of audit trails, out-of-date systems, and poor-quality systems.

5. Common Data Integrity Violations [19]

5.1 Backdating and Data Fabrication [12][20]

Data integrity infractions are rarely isolated procedural errors, according to an analysis of regulatory enforcement actions issued in 2024–2025; Instead, they often reveal structural flaws in

computerized system governance, organizational culture, and quality control. Inspection results from many countries show recurrent trends, especially in factory documentation procedures, third-party testing agreements, and quality control labs.

Table 1: Backdating and Data Fabrication

Violation Category	Specific Observation (2024-2025)	Regulatory Implication
Audit Trials	either not inspected during batch release, disabled, or improperly configured.	Reconstructing events is impossible; traceability is compromised
User Access	Uncontrolled system rights or shared passwords.	loss of accountability and attribution
Raw Data	Data Removal of experimental injections or unsuccessful outcomes.	breach of the standards of originality and completeness
Contemporaneousness	Timeliness using memory to retrospectively record data.	decreased dependability and higher fabrication risk.
Systems	Utilize unverified spreadsheets for computations.	Potential for undetected computational flaws to impact the quality of the final work.
Archival	Poor storage procedures; unreadable or unstable document.	Failure to meet enduring and legibility requirements.

5.2 Selective Reporting [21]

Selective reporting, also called as cherry-picking, is a major violation of the principle of data integrity because contradictory data is omitted while only data supporting the preconceived outcome is reported. Selective reporting leads to a major distortion of findings and is considered part of the "reproducibility crisis." It is also viewed as one form of falsification.

5.3 Unauthorized Data Deletion [22, 23]

Unauthorized deletion of data, data tampering, or accidently deleting data by employees, is a significant data integrity breach. It undermines accuracy, leads to loss of important of data, and creates issues of non-compliance. This issue tends to be linked to lack of proper audits or controls, and it is most times the focal point of regulatory warning letters.

5.4 Incomplete Audit Trails [24]

One of the major examples of integrity violation is the absence of audit trail, which is a requirement for the biotech industry, finance industry, and the pharmaceutical industry. Not being able to provide a proper sequence of all the events that take place during the life cycle of data does not comply with ALCOA+ requirements (Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Enduring, Available).

Problems: No audit trail, No date for change, No username, No justification, Inadequate logging review, No documentation of changes, Signing before changing the date, Password sharing, Weak system configuration.

5.5 Uncontrolled Spreadsheets [25]

Because of uncontrolled spreadsheets which lack security, validation, and audit records, they create a high danger to GMP.

Important issues: Data manipulation (adding, removing, and editing records), Absence of an audit trail (no change record), Unauthorized entry, Inconsistent or duplicate data, Inadequate version control, Errors in manual entry, Data loss risk

Impact on regulations: FDA non-adherence, Unverified computations, History has not changed, Inaccurate QC analysis.

6. Risk based approach to Data integrity

6.1 Risk assessment methodologies [9][26]

In a data integrity risk-based approach (DI-RBA), risk assessment entails methodically locating, evaluating, and reducing threats to data completeness and accuracy throughout its lifecycle (ALCOA+). ICH Q9 quality risk management concepts, 6M (Man, Machine, Material, Method, Measurement, Milieu), and FMEA (Failure Mode Effects Analysis) are important approaches. These techniques often prioritize controls by ranking risks according to their severity, probability, and detectability, usually in accordance with GMP rules. Finding computerized system configuration elements that affects the patient safety, product quality, and data integrity is the primary goal of the risk assessment and specially to find any high-risk locations that need further restrictions.

6.2 Critical Data Identification [26]

The identification of important data in case of a risk-oriented approach for data integrity would include the mapping of data lifecycles to identify areas that present risks to patient safety, quality of products, or regulatory decision-making. CPPs, CQAs, and data that impacts the batch release decision are among the important types of data identified using tools such as FMEA.

6.3 Impact analysis on Product quality [27]

To ensure both patient safety and product quality, the risk-based DI checks that all important data remains safe through its entire lifecycle. This reduces the number of product recalls, keeps patients safe, and helps with ensuring compliance (ICH Q10, FDA, EU GMP).

Critical Data: Identify key data affecting product quality, involved testing results, batch documentation, and stability tests.

Risk Assessment: Employing techniques such as FMEA for identifying and estimating risks such as data loss, interfering, or fabrication.

Effect on Quality: Poor DI practices may lead to improper decisions regarding the product quality and failure of the product.

Mitigation: Controls against issues with data include implementing audits, validations, and backup systems.

6.4 Risk Mitigation Strategies [26]

To ensure critical information through the use of both technical and procedural measures from when the information is first created until its disposal is the main objective of risk-based approaches to data integrity.

Risk assessment (FMEA): Assess and emphasize the risk according to the probability of being detected, happening, and the possible outcomes.

Technical controls: Introduce RBAC, audit trails, digital signatures, and encoded to ensure that no undesirable changes occur.

Procedural controls: Adhere to SOPs, conduct audit trails review, and maintain data integrity.

Data Life Cycle Management: Ensure monitoring and data integrity through the entire life cycle of data, from generation to destruction.

Vendor Management: To implement vendor compliance with DI through audits.

Education and culture: Conduct training for all employees regarding data integrity.

7.Investigation of Data quality Issues [28]

7.1 Triggers for Investigation

The failure of dashboards, the obsolescence of data, or unexpected changes in the volume, format, and distribution of data are some events that may either be actively monitored or passively reported as potential triggers for data quality investigations. Data discrepancies between systems, inaccurate data entry, improper transformation logic, and problems with ETL processes are critical triggers.

7.2 Investigation Workflow and Strategies

Technical data profiling, automated observability, and structured root-cause analysis are all part of a strong investigation workflow, which frequently makes use of frameworks like PICERL (Prepare, Identify, Contain, Eradicate, Recover, Learn).

7.3 Root Cause Analysis Tools [12]

According to ICH Q9(R1), quality risk management principles place a strong emphasis on the methodical identification and assessment of risk sources in pharmaceutical systems. Structured Root Cause Analysis (RCA) methods make it easier to differentiate between basic control failures and superficial symptoms when they are used in data integrity investigations.

Table 2: Root Cause analysis Tools

RCA Tool	Strategic application in Data Integrity	Outcome for Quality Assurance
5 Whys	Iterative inquiry to ascertain the reasons behind the retrospective entry of data or the failure to evaluate an audit trail.	Finds training shortages, ambiguous responsibility distribution, or impractical deadlines.
Fishbone (Ishikawa) Diagram	Grouping of contributing elements into the following categories: people, process, equipment, environment, and management.	Provide a comprehensive grasp of systemic weaknesses.
Failure Mode and Effect Analysis (FMEA)	A risk assessment of the potential risks involved in manipulating data within electronic systems.	Enables prioritization of the controls based on their severity, occurrence, and detection ability.
Fault Tree Analysis	Logical mapping of event combinations that compromise integrity.	Identifies organizational and technological contributors who interact.

7.4 Hypothesis Testing in Investigations [29]

A structured statistical procedure known as hypothesis testing is applied in the data quality assessments to determine if the observed anomalies in the data are statistically significant or merely due to chance. This approach takes the subjectivity out of scientific research, moving it from being based on one's intuition ("gut feeling") to objective evidence-based decision-making.

H0 (Null Hypothesis): There is no issue regarding data quality.

H1 (Alternative Hypothesis): An issue regarding data quality exists.

p Value: Where $p < \alpha$, there is an issue concerning the quality of data.

α Value (Level of Significance): Requirements for rejecting the null hypothesis (0.05).

7.5 Documentation of Investigation

Documenting the examination of data quality (DQ) problems is important for accountability, transparency, and averting such incidents in the

future. A centralized log, a thorough root cause analysis, and a record of corrective measures should all be part of a thorough documentation process.

8.Analytical Instrumentation and Data system [30]

8.1 Chromatographic system (HPLC, GC)

The methods of chromatography (GC and HPLC) are indispensable in the analysis of mixtures because they assist in the identification, quantification, and separation of the constituents. Although HPLC employs liquid as the mobile phase in the case of volatile and heat-sensitive chemicals, GC uses gaseous mobile phases for volatile and heat-stable chemicals. This occurs within the context of a Chromatography Data System (CDS).

The employment of HPLC is best suited for analysing biological substances and drugs because it applies pressure in pushing the liquid mobile phase through the solvent in controlled environments at room temperature (20-40°C). This can be achieved through pumps, injectors, columns, and certain detectors such as UV-Vis and PDA. During the process of gas chromatography, the sample substance (volatile chemical) is subjected to very high temperatures (150-300°C), causing it to evaporate. The gas pushes the evaporation through the column.

8.2 Spectroscopic Instruments

Spectroscopic devices are used for identifying the behaviour of light in relation to matter. Modern spectroscopic devices have been automated and are associated with the database management system.

Examples of some widely used spectroscopic devices are as follows:

1. UV-VIS spectrophotometry
2. FTIR spectrometry
3. Raman spectrometry
4. AA spectrometry
5. NMR spectrometry
6. EDX spectrometry
7. Laser induced breakdown

8.3 Laboratory Information Management System [LIMS] [31]

"An efficient computer system for managing laboratory data, samples and processes of today's age is called the Laboratory Information Management System (LIMS). LIMS maintains data integrity, automation of laboratory operations, and compliance with GMP, FDA 21 CFR Part 11 and other regulatory authorities. Sample tracking, interfacing of instruments, generation of reports and inventory control become crucial activities."

8.4 Electronic Lab Notebook (ELN) [32]

An Electronic Lab Notebook (ELN) is a software system that serves as an alternative for laboratory notebooks containing original data obtained from experiments conducted, protocols of these experiments, as well as characteristics of the instruments utilized. The ELN serves as the data management system through which the collected data is stored efficiently as searchable records in one place, interacting with analytical instruments.

8.5 Audit Trails and System Controls [33]

With controlled conditions such as GMP and GLP, audit trails and system control play the role of being the foundational basis of data within analytical instruments and systems, ensuring that ALCOA+ principles are met in data management. Questions of "who, what, when, and why" data management remain one of the major concerns under regulation in 2026, specifically in computerized systems such as CDS and LIMS.

9.Audit Trails and Electronic Records Review [12]

9.1 Importance of Audit Trails

The term audit trail refers to a reliable computer-generated and timestamped document that allows for the reconstruction of events relating to the generation, modification, or deletion of electronic records. The audit trail is now a proactive compliance tool and not merely a passive archival feature in modern regulation. In routine quality assurance procedures, it is becoming more common to see the documented assessment of audit trail information regarding data that impact critical quality parameters or batch decisions**.**

9.2 Types of Audit Trails

By chronologically documenting system actions, audit trails in electronic records management guarantees data integrity, accountability, and regulatory compliance. System-level (logins, configuration changes), application-level (software activities), user activity (individual actions), data modification (field-level alterations), financial (transaction logs), and electronic signature trails are important categories.

9.3 Review Procedures [34]

The audit trail and review process of the electronic record maintains integrity through systematic verification of the legitimacy and accuracy of electronic records by establishing risk-based scope, reviewing system logs for changes (CREDO – created, read, edited, deleted, occurred), recording reasons for changes, and the access rights of the users.

Steps Involved in Audit Trail Review Process:

1.Scope and Frequency: Scope and perform audits (such as before batch release).

2.Critical Data Identification: Concentration is on GMP/GLP data including laboratory results and system changes.

3.Audit Trail Review: Detect suspicious transactions in the logs (CREDO – created, read, edited, deleted, occurred).

4.Issues Evaluation: Detection and document any unauthorized or missing changes.

5.Documentation: Using electronic signatures to document the name of the reviewer, date, and findings.

6.CAPA: Prevention and Corrective Actions if required.

9.4 Common Findings in Audit Trails Review [35]

Data integrity issues, such as unauthorized data modifications, missing entries, a lack of user-specific logins, and insufficient periodic reviews, are frequently the focus of audit trail and electronic records reviews. These results frequently show unapproved alterations, falsification, and inadequate documentation practices, such as omitting change justifications, which are against FDA and GMP requirements.

9.5 Data Security and Access control [36]

Data integrity, tracking, and compliance (FDA 21 CFR Part 11, GDPR, HIPAA) can be guaranteed by data security and access control by auditing the logs and checking electronic documents. Role-based access, data storage security, and evaluation for any data alteration are essential. System log monitoring for any anomaly and identification of who, what, and when is very important.

10. Case studies of Data Integrity Failures

10.1 Pharmaceutical Industry Cases [37]

In many cases, the pharmaceutical company has been seen manipulating, altering, or fabricating data in laboratory tests like HPLC chromatograms to meet the batch specifications. These practices result in warnings being issued, as well as monetary penalties. Examples include: Wockhardt Ltd. (poor quality control and alteration/destruction of data); and Ranbaxy Laboratories ($500 million fine for fabricating data).

Ranbaxy Laboratories Inc. (2013): After the revelation that the company had falsified data regarding the quality of their drugs, the FDA imposed a ban on imports from the company, resulting in a fine of $500 million.

Wockhardt Ltd. (2013): The company was given warning letters by the FDA due to violations including destruction of data, failure to prove the integrity of the aseptic manufacturing process, and destruction of laboratory data to conceal negative results.

Johnson & Johnson – McNeil Consumer Healthcare (2011): The FDA observed numerous violations of the companies in terms of poor maintenance of equipment and accuracy of data.

Generic Manufacturers/Lab Fraud (2020): There was an FDA discovery of violations related to ALCOA+ criteria (Attributable, Legible, Contemporaneous, Original and Accurate) since HPLC tests lacked all results with no explanations from the companies.

10.2 Regulatory Warning Letter Analysis [38]

Any laboratory or factory setting must have data integrity, but this is especially true in the biotechnology and pharmaceutical industries. Data integrity violations frequently result in warning letters from regulatory agencies like the FDA, according to recent trends. This paper provides a thorough, methodical approach to comprehending these breaches, the ensuing legal ramifications, and the efficient implementation of data integrity compliance services.

10.3 Lessons Learned from Real Incident

Any data manipulation, accidental or intentional, that causes incorrectness or unavailability of the information due to corruption, loss, or changes in the data can be classified as a failure in data integrity. Incidents have taught us that the risk is often intrinsic to culture, poor validation, and access management.

11.Corrective Actions and Preventive Actions [39]

11.1 CAPA Framework

There are many tools that can be used to ensure that there are quality and continuous improvements in any products and services. This includes CAPA. There is assurance that the product or service meets all of the requirements of quality and regulations. CAPA is an important tool that is used to manage quality risks. Therefore, implementing CAPA in industries will enable the cause of non-compliance to be identified in the future.

11.2 Immediate vs Long-term Actions

Corrective and Preventive Action (CAPA) is intended to resolve any quality issues in two different phases, namely corrective action (short term) and preventive action (long term). The intention is to reduce damage by taking immediate action to fix the current nonconformity before finding the root cause that prevents its recurrence.

Immediate actions (Correction & Containment): This refers to fast, reactive measures taken as soon as a nonconformity occurs in order to curb the issue from getting any worse or spreading to other areas of the client's business.

Long-term Actions (CAPA): After conducting an extensive root cause analysis (using "5 Whys" or "Fishbone analysis"), these systematic measures are taken to prevent recurrence.

11.3 Effectiveness checks

Another crucial component of a CAPA action is verification that the measures taken have been successful. An evaluation must be performed to ensure that the underlying reason for the problem has been tackled, and all the consequences have been sorted out, with appropriate controls put in place and monitoring of the issue conducted. The evaluation can include assessing whether any new

negative impacts arise as a result of the measures. It is necessary to monitor the results of such investigations and assessments. Documentation of the entire process undertaken in a corrective or preventive action from the onset of the problem to its resolution is critical, although necessary to meet regulations.

11.4 Continuous Improvement

A CAPA continuous improvement initiative will transform a quality system's approach from being solely based on error correction to a proactive and strategic process that drives growth. In order to avoid any recurrence of the issues and to improve efficiency by 20%, RCA, risk assessment, and action efficacy need to be considered.

12. Role of Quality Assurance (QA)

12.1 QA Oversight in Laboratories [40]

The complete procedure that ensures that the lab's results are as precise as possible is referred to as quality assurance (QA). Material analysis, evaluation of transcriptional processes, use of the most reliable methods, and verification of results. Apart from promoting quality laboratory practices, a QA system ensures that the testing facility conducts itself in such a way as to ensure that the HIV tests conducted for surveillance are as precise and accurate as possible. Quality and accurate laboratory results are obtained through the use of standardized test algorithms together with the right SOPs and quality control mechanisms. Procedure manuals should incorporate vital elements of QA.

12.2 Internal Audits and Self-Inspections [41]

Some of the sections of the GMP, GMP API, GDP, GDP API guidelines, other best practices adapted in the EU, and the PIC/S specify the broad requirements and principles regarding the procedure of internal audits. These internal audits serve the purpose of assessing pharmaceutical enterprises' compliance with the principles of PQS. One of the most critical pre-requisites for internal audits is that they should be conducted according to a well-established plan.

12.3 Training and Competency Programs [42]

Competency and quality assurance training processes are systematic and scientific ways devised to ensure that workers have the required skills and competencies that will help them carry out their duties effectively. These processes are important since they lead to fewer mistakes, increased efficiency, and compliance with set standards.

12.4 Data Governance Policies [40]

In quality assurance (QA), data governance policies are explicit guidelines that specify how data is categorized, protected, and handled to guarantee correctness, compliance, and usefulness. They provide as a framework for handling data as a strategic asset, frequently addressing security, ownership, and quality requirements for trustworthy decision-making. Establishing positions such as data stewards for daily quality management is necessary for effective policies.

13. Digitalization and Emerging Technologies

13.1 Automation in Analytical Labs [43]

Accuracy and efficiency form major drivers for automation and computerization of laboratory tests. Through uniformity in operation, automation technologies reduce mistakes and guarantee correctness in the output. Additionally, they can work continuously without problems like fatigue. Automation improves efficiency by doing away with labour-intensive activities like recording information manually, mixing reactants, and sample preparation.

13.2 Artificial Intelligence in Data Monitoring [12]

The use of artificial intelligence (AI) technology is becoming common practice in laboratory activities such as visual inspection, process optimization, and analysis. The visibility, repeatability, and explain ability of AI tools are key criteria required by regulatory authorities. AI systems should always deliver results that can be verified and operate within specified limits. AI can increase efficiency and support quality management initiatives, although it also requires risk assessment, validation, regular monitoring, and human intervention.

13.3 Blockchain for Data Integrity [12]

It is anticipated that the application of blockchain technology for producing tamper-proof documents through time-stamping and sharing among all parties involved will ensure greater traceability and security. In this manner, issues related to batch history and the transparency of the materials used can be improved. Nevertheless, some of the key obstacles include validating the data, scalability, regulation, privacy, and the compatibility with GMP regulations.

13.4 Challenges with Digital Transformation [44]

All businesses in the world value digital transformation (DT) highly; however, those that don't adapt themselves face the phenomenon called "Digital Darwinism," where only the most adaptable will make it through. Despite the fact that DT is currently being embraced by many organizations (87% of CEOs regard it as an essential strategy), it remains very complicated and often conflicts with existing processes and operating models. The success rate is also comparatively low, ranging below 30%, and even lower for traditional sectors such as the pharmaceutical industry, at 4-11%.

14. Best Practices for Ensuring Data Integrity

14.1 Good Documentation Practices (GDP) [45]

Using ALCOA+, GDP ensures that data remains accurate throughout its life cycle. Examples of such critical tasks include using non-erasable ink, signing and dating of entries when conducting actions, keeping electronic records with an audit trail, and training staff.

14.2 Standard Operating Procedures (SOPs) [19]

It is essential for SOPs to be developed and adhered to so that workers are able to understand how to record and retain data. The SOPs lay out the process of how data should be collected, stored, retrieved, and retained.

14.3 Periodic Data Review [46]

Data integrity can be achieved by following certain best practices, one of which is the process of data review. Data review is carried out regularly to ensure accuracy and consistency of the data in compliance with the ALCOA+ concept, which means that the data should be attributable, legible, contemporaneous, original, and accurate. Data reviews usually take place on a quarterly basis. Some of the mistakes made during data review are validation mistakes and audit trail mistakes.

14.4 Vendor and Software Qualification [45]

To make sure that data integrity complies with ALCOA+ standards (Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Enduring, and Available) for the whole software lifetime, the optimal practices of vendor and software qualification, which can be found in articles devoted to GxP compliance, utilize risk assessment methods.

15. Challenges and Future Perspectives

15.1 Regulatory Challenges [47]

By 2026, the issues of integrity and data quality will extend beyond compliance. There is a growing concern for data integrity and quality

among regulators like the FDA, MHRA, and WHO as the usage of AI and digitization continues, particularly in computerized systems and AI-driven research.

Compliance Fatigue & Fragmentation: With the emergence of new regulatory obligations, documentation activities continue to escalate, leading to the inability to pay attention to security. Risks related to integrity in the healthcare industry are brought about by fragmented systems and manual entry of data.

Governance Issues in AI: Transparency is required by regulatory organizations regarding AI decision-making process. Proofing AI decisions is difficult when it comes to drug discovery, particularly from the AI 'black box'.

Data Quality Problems: Inefficiencies caused by fragmented data pipeline across multiple platforms result in poor data quality that causes loss of context, information, and traceability.

Integrity in Electronic Systems: Some of the issues with integrity in an electronic system include shared logins, weak audit trails, inappropriate alterations, and lack of preservation of raw data.

15.2 Technological Limitations [47]

No Automation: As no software exists to ensure data quality, everything remains manually oriented.

Data Degradation and Speed: In particular, real-time systems face rapid data degradation issues.

Increase in Unstructured Data: Much of the data produced tends to be unstructured data.

Compatibility Problems: Changes in schemas and integration of various data sources can result in errors.

AI Risks: Data drift and data poisoning make AI less reliable.

IoT Risks: IoT devices lacking proper security measures can cause access to be compromised.

15.3 Cultural and Organizational Barriers [48]

The main obstacles to obtaining good data quality and integrity are commonly identified as organizational and cultural hurdles, which frequently surpass technical constraints. Although data problems are expected to cause 60% of AI initiatives to fail in 2026, successful companies are moving away from considering data as a solely IT issue and toward fostering a "Data Quality Culture."

15.4 Future Trends in Data Integrity [49]

Data integrity issues for the future would concentrate more on AI-based observability, edge security, and ALCOA+. The difficulties related to the management of data, protection from cyberattacks, and assurance of quality in a legacy system are some of the major concerns. Opportunities for the future in this context are enhanced data literacy and decentralized governance.

CONCLUSION

16.1 Summary of key findings

In most cases, any investigation that is carried out to look into the data integrity of lab data identifies major system weaknesses. Most of these problems stem from the improper management of electronic record keeping, inadequate supervision, and improper handling of failed tests. The key findings from such an investigation will mostly deal with data manipulation, violation of ALCOA standards, and inadequate training.

16.2 Importance of a Quality Culture

The basis of scientific validity, compliance, and safety for patients is dependent on the quality culture of the laboratory with regards to data integrity. In accordance with FDA research, the establishment of a quality culture helps in making sure that employees realize the importance of data integrity as an essential business aspect, which means they will feel empowered to raise any

concerns. This means the data is ALCOA+ due to the quality culture established.

16.3 Final Recommendations

In a laboratory setting, more particularly in highly regulated industries such as those in pharmaceutical, biotechnology, and clinical diagnostics, the conclusions that will arise from a study regarding data quality and integrity assessment would primarily emphasize how to prevent future violations through a risk-based approach rather than a reactive strategy. These recommendations would include strengthening data governance policies, setting up strict SOPs, improving employee education regarding data integrity (ALCOA+) principles, and conducting adequate documentation. While identifying the potential errors that could exist in the laboratory, a large budget should be set aside to validate the computer systems with an audit trail system, put access controls in place, and conduct risk assessments and audits. It is only through continuous monitoring and supervision that we can ensure data integrity and reliability.

REFERENCES

1. Smith P. Data integrity in the analytical laboratory. *Pharmaceutical Technology*. 2014 May 2;38(5).
2. Madhanraj S, Smith AA. Data integrity in pharmaceuticals. *Int J Drug Regul Aff.* 2025 Dec 15 ;13(4):26.
3. Ahmad S, Kumar A, Hafeez A. Importance of data integrity & its regulation in pharmaceutical industry. *Authorea Preprints.* 2022 Sep 8;8(1):306–313.
4. https://www.chromatographyonline.com/view/data-integrity-focus-part-1-understanding-scope-data-integrity.
5. Chakraborty T. Data integrity principles ALCOA+++ contents. *pharma VA*. 2025 Jul.
6. Choudhary A. Importance of data integrity for pharmaceutical regulatory agencies. *Pharma guideline*. 2025 Dec 8.
7. https://www.ibm.com/think/topics/data-integrity-vs-dataquality?utm_source=chatgpt.com
8. Woollen SW. Data quality and the origin of ALCOA. *Newsletter of the Southern Regional Chapter Society for Quality Assurance*. 2010.
9. International Society for Pharmaceutical Engineering (ISPE). *GAMP good practice guide: a risk-based approach to GxP compliant laboratory computerized systems*. 2nd ed.2012. 978.
10. https://www.3ritechnologies.com/6-life-cycle-phases-of-data-analytics/.
11. Chen Q, Cheng HJ, Bian ZX, Lyu AP, Yang Y, Chan KW. Regulatory frameworks and evidence requirements for traditional, complementary and integrative medicines. *J Ethnopharmacol*. 2025 Sep 3;103(11):696–707.
12. Salunkhe P, Mane C, Kumbhar S, Battalwar V, Kadam K, Salunkhe S. Data integrity in pharmaceutical manufacturing: evolving global regulatory frameworks, enforcement trends, and digital governance strategies. *Int J Pharm Sci.* 2026;4(2):3799–3818.
13. Chen Z. Good laboratory practice (GLP) 101 – regulations and basic studies. CDER inspections of good laboratory practice, animal rule, and bioavailability/bioequivalence study sites; 2022 Jul 19.
14. https://www.fda.gov/inspections-complianceenforcement-and-criminalinvestigations/compliance-actions-andactivities/warning-letters.
15. https://cdn.who.int/media/docs/defaultsource/medicines/norms-andstandards/guidelines/inspections/trs1033-annex4-guideline-on-data-integrity.pdf.

16. https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/09/glp-data-integrity_c2f067ec/45779212-en.pdf.
17. Choudhary A. FDA's top data integrity issues found during inspections. *Pharma guideline*. 2024 Apr 17.
18. https://www.pharmtech.com/view/data-integrity-analytical-laboratory.
19. Ingale MH, Tayade MC, Patil YP, Salunkhe R. Data integrity violations in the pharmaceutical industry and regulatory measures. *Int J Pharm Sci Rev Res*. 2023 Jun 25.
20. Sandle T. Data integrity and pharmaceutical quality systems. *J Validation Technol.* 2016;22(4):15–21.
21. https://editorresources.taylorandfrancis.com/publishing-ethics-for-editors/research-integrity-and-selection-bias/.
22. Borchers AT, Hagie F, Keen CL, Gershwin ME. The history and contemporary challenges of the US Food and Drug Administration. *Clin Ther.* 2007;29(1):1–16.
23. Charoo NA, Khan MA, Rahman Z. Data integrity issues in pharmaceutical industry: common observations, challenges and mitigation strategies. *Int J Pharm.* 2023;631.
24. Rattan AK. Data integrity: history, issues, and remediation of issues. *PDA J Pharm Sci Technol.* 2018;72(2):105–116.
25. De la Torre BG, Albericio F. The pharmaceutical industry: an analysis of FDA drug approvals from the perspective of molecules. *Molecules.* 2020;25(3):745.
26. Jain SK**.** Strategy to avoid data integrity issues in pharmaceutical industry. *The Pharma Innovation.* 2017 Feb 1;6(2 Pt B):110.
27. Ayillath Kezhadath A, Amarapalli L. Ensuring data integrity in pharmaceutical quality systems: a risk-based approach. *J AI-Powered Med Innov.* 2024;1(1).
28. Rao D, Gudivada V, Raghavan VV. Data quality issues in big data. In: *Proceedings of the IEEE International Conference on Big Data*; 2015 Oct 29.
29. https://www.dasca.org/newsroom/hypothesis-testing-in-data-science-validating-decisions-with-data.
30. Ahmed S, Hossain MA, Islam S, Islam MM. Review article on introduction of analytical instruments analysis in pharmaceutical industry according to pharmacopoeia. *Int J Pharm Med Sci.* 2019;8.
31. Raviteja M N, Gupta N V. A review on electronic data management in pharmaceutical industry. *Asian J Pharm Clin Res*. 2013 Apr 1;6(2):38-42.
32. https://www.yokogawa.com/library/documents-downloads/technical-information/what-is-eln/.
33. Dougherty M. How legal is your EHR: identifying key functions that support a legal record .*J ahima*. 2008;79(2):24–30.
34. McCartney P. HIPAA and electronic health information security. *MCN Am J Matern Child Nurs*. 2003; 28.
35. Khandelwal M. The importance of audit trails: Enhancing transparency and accountability. *The Compass – A Weekly Newsletter*. 2023.
36. Sonkhiya B, Verma N. Audit trail: Impact analysis & reporting responsibility of auditor. *Taxmann*. 2023.
37. Nainita MK. Case studies on data integrity in pharma. Raghavendra Institute of Pharmaceutical Education and Research; 2020.
38. https://www.pharmaregulatory.in/case-study-warning-letters-for-breaches-in-data-integrity/.
39. Jain SK, Jain RK. Investigations and CAPA: Quality system for continual improvement in pharmaceutical industry. *Int J Res Pharm Sci.* 2017;2(6):47–54.

40. Arnold JE, Camus MS, Freeman KP, Giori L, Hooijberg EH, Jeffery U, et al. ASVCP guidelines: principles of quality assurance and standards for veterinary clinical pathology developed by the American Society for Veterinary Clinical Pathology's (ASVCP) Quality Assurance and Laboratory Standards (QALS) Committee. *Vet Clin Pathol.* 2019;48(4):542–618.
41. Nikityuk V, Karamavrova T, Lebedynets V. The self-inspections (internal audits) process as a part of the pharmaceutical quality system: formation of a risk-based approach to internal audits planning. *J Pharm Pharmacol.* 2019; 7:385–397.
42. Vatharkar P, Ghanwat A. A comprehensive literature review on competency-based training systems in manufacturing. *Int J Sci Res Eng Manag.* 2026;10(1):1–9.
43. Alzahrani ASA, Alharbi HAA, Alkhawlani HAA, Orepi AA, Alamri AA, Alzahrani AA, et al. Automation and digitalization in laboratory testing: revolutionizing accuracy and efficiency. *Review of Contemporary Philosophy*. 2023;22(1):2252–2266.
44. Maksimenko I, Vashko T, Zdrestova-Zakharenkova S. Digital transformation and its challenges to the strategic management system. *MIT Sloan Management Review*. 2017;58(2):17.
45. https://www.sciencedirect.com/science/article/pii/S2949866X24001060.
46. https://www.ijpsjournal.com/article/Ensuring-Data-Integrity-In-The-Pharmaceutical-Industry%3A-Benefits%2C-Challenges%2C-Key-Considerations-And-Best-Practices.
47. https://escalon.services/blog/life-sciences/preparing-for-2026-regulatory-data-integrity-and-compliance-trends-life-sciences-leaders-must-address-in-q1.
48. Ghafoori A, Gupta M, Merhi MI, Gupta S. Toward the role of organizational culture in data-driven digital transformation. *Int J Prod Econ*. 2024 Mar;271.
49. Munagandla VB, Dandyala SSV, Vadde BC. The future of data analytics: trends, challenges, and opportunities. *Rev Intell Artif Med*. 2022;13(1).

HOW TO CITE: Vairangi Patel, Dhara Patel, Grishma Patel, Dhananjay Meshram, Data Quality and Integrity Investigation in the Analytical Laboratory, Int. J. of Pharm. Sci., 2026, Vol 4, Issue 5, 3287-3303, https://doi.org/10.5281/zenodo.20178235