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

OPEN ACCESS

Review Article

Ensuring Data Integrity In The Pharmaceutical Industry: Benefits, Challenges, Key Considerations And Best Practices

Akshaya U. Bhandarkar*[1], Balaji J.[2], Jasvanth R.[2], Dinesh Ragavendra S.[2]

[1]*Department of Pharmaceutics, Vidya Siri College of Pharmacy, Bangalore, Karnataka 560035, India.*
[2]*B Pharm students of Vidya Siri College of Pharmacy, Bangalore, Karnataka 560035, India.*

ARTICLE INFO

Published: 22 Oct 2024
Keywords:
Data Integrity, USFDA, EMA, WHO, ALCOA, ALCOA+, Data lifecycle, Warning letters, Product recalls, Block chain, Artificial Intelligence.
DOI:
10.5281/zenodo.13971925

ABSTRACT

Data integrity plays a critical role in the pharmaceutical industry, as it directly influences product quality, ensures patient safety, and maintains compliance with regulatory standards. It is the corner stone of the industry, ensuring the reliability of data used in drug development, manufacturing, testing, distribution and regulatory submissions. With increasing digitalization, maintaining data integrity has become more challenging, requiring robust systems and practices. Regulatory bodies such as the United States Food and Drug Administration [USFDA], European Medicines Agency [EMA] and World Health Organization [WHO], mandate strict adherence to data integrity principles, such as ALCOA and ALCOA+. These principles guide the entire data lifecycle, emphasizing the need for robust systems to maintain data accuracy, traceability and security. Compliance with these expectations is critical, as breaches can lead to severe consequences, including warning letters, product recalls, and legal action against the organization. The pharmaceutical industry is also adapting to emerging technologies like block chain and artificial intelligence, which offer potential enhancements in data integrity management. This review outlines the benefits, challenges and key considerations of data integrity for paper based and computerized records, strategies for strengthening data integrity and importance of quality culture in upholding data integrity. By understanding the complexities of data management, pharmaceutical companies can ensure regulatory compliance and enhance trust in their products.

INTRODUCTION

Data integrity is defined as "the extent to which the data are complete, consistent, accurate, trustworthy, and reliable and how these aspects of data are maintained throughout the lifecycle of the product" [1]. This is an important element and requirement of pharmaceutical quality system

***Corresponding Author:** Akshaya U. Bhandarkar
Address: *Department of Pharmaceutics, Vidya Siri College of Pharmacy, Bangalore, Karnataka 560035, India.*
Email ✉: akshaya@vidyasiricop.edu.in
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.

highlighting pharmaceutical industry's responsibility to ensure the safety, efficacy, and quality of the drugs, and FDA's role in safeguarding public health [1]. Inadequate data integrity practices and vulnerabilities compromise the quality of records and evidence, potentially jeopardizing the quality of medicinal products [2]. The principles of data management are applicable to paper-based and computerized records in the same way [3]. Integrity of data is essential in research and development, manufacturing, quality control and regulatory submissions. However, maintaining data integrity is fraught with challenges that can arise from various sources. Some of the key challenges are increasing digitalization, regulatory compliance, human error and legacy systems. ALCOA and ALCOA+ are acronyms introduced by USFDA to indicate data integrity in relation to pharmaceutical research, manufacturing, testing and supply chain as a best practice for ensuring data integrity and to overcome the challenges and data integrity breaches [4]. The term ALCOA was introduced to define its expectations of electronic data. Regulators adopted the term ALCOA Plus, emphasizing the consequences of not adhering to regulations and suggesting prevention methods such as simple checklists, self-audits, and self-inspections [2] ALCOA refers to the principles of data being Attributable, Legible, Contemporaneous, Original, and Accurate. ALCOA was further expanded to ALCOA Plus, Plus means Enduring, Available and Accessible, Complete, Consistent, Credible, and Corroborated [5].

KEY TERMINOLOGIES IN DATA INTEGRITY: DEFINITIONS3

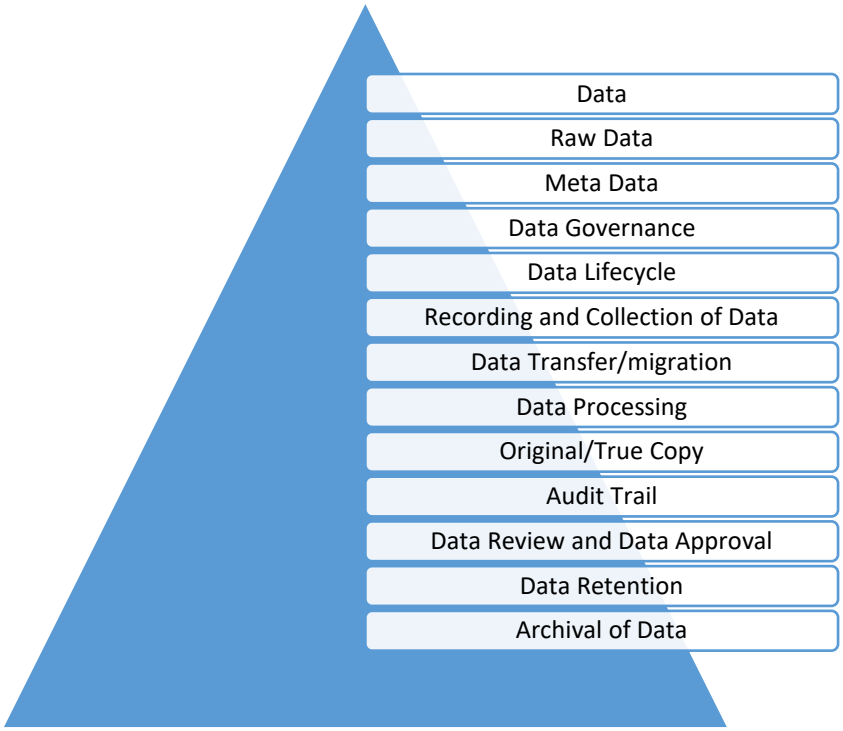

Figure 1: Key Terminologies

a. Data:

It includes facts, figures, and statistics, which are gathered for reference or analysis. This encompasses all original records and their true copies, such as source data and metadata, along with any subsequent transformations and reports. These records are generated or documented during the GXP activity, enabling a thorough reconstruction and evaluation of the activity.

- **Raw Data:**
  It refers to the initial record of information, captured for the first time, whether on paper or electronically. If the information is originally recorded in a dynamic state, it should remain accessible in that form. Raw data must allow for a complete reconstruction of activities.

b. Meta Data:
It refers to the data that describe the attributes of other data, providing context and meaning. They typically detail the structure, data elements, inter-relationships, and other characteristics, such as audit trails. They form an integral part of the original record, and without the context provided by metadata, the data lacks meaningful interpretation.

c. Data Governance:
It refers to the arrangements made to ensure that data, regardless of its format, is recorded, processed, retained, and utilized throughout its lifecycle. Data governance must establish clear data ownership and accountability, focusing on the design, operation, and monitoring of processes and systems to uphold data integrity principles, thereby ensuring that the data is readily traceable and directly accessible upon request from national competent authorities. Electronic data should be provided in a human-readable format.

d. Data Lifecycle:
It refers to all the phases in the life of a data, from generation and recording to processing [including analysis, transformation, or migration], use, retention, archiving/retrieval, and destruction. All these phases must be managed effectively. Data governance should be applied throughout the entire lifecycle to ensure data integrity.

e. Recording and Collection of Data:
Organizations should possess a thorough understanding of processes and technical knowledge of systems used for data collection and recording, including their capabilities, limitations, and vulnerabilities. The chosen method should ensure that data collected and retained is accurate, complete, and meaningful for its intended use.

f. Data Transfer/Migration:
Data transfer involves moving data between different storage types, formats, or computerized systems. Data migration, on the other hand, refers to relocating stored data from one durable storage location to another, which may include changing the data's format but not its content or meaning.

g. Data Processing:
It refers to a series of operations carried out on data to extract, present, or obtain information in a specified format.

h. Original/True Copy:
Original copy refers to the initial capture of data or information, such as an original paper record of a manual observation or an electronic raw data file from a computerized system, along with all subsequent data needed to fully reconstruct the GXP activity. True copy, on the other hand, refers to a copy of the original record, regardless of the media type, that has been verified [either by a dated signature or through a validated process] to contain the same information, including data describing the context, content, and structure, as the original.

i. Audit Trail:
It refers to a type of metadata that records actions related to the creation, modification, or deletion of GXP records. It securely logs life-cycle details such as creation, additions, deletions, or alterations of information in a record, whether on paper or electronically, without obscuring or overwriting the original record. An audit trail enables the reconstruction of event histories by detailing 'who, what, when, and why' of each action was performed.

j. Review and Approval of Data:

It refers to the method for reviewing specific record content, including critical data, metadata, cross-outs [in paper records], and audit trails [in electronic records], should comply with all applicable regulatory requirements and be based on risk. There should be a procedure outlining the process for data review and approval.

k. Data Retention:

It refers to archiving [protected data for long-term storage] or backup [data for disaster recovery purposes]. Arrangements for data and document retention should ensure records are protected from deliberate or accidental alteration or loss.

l. Archival of Data:

It refers to a designated secure area or facility for the long-term retention of data and metadata to verify processes or activities. Archived records may be the original record or a 'true copy' and should be safeguarded against alteration or deletion without detection, as well as protected from accidental damage such as fire or pests.

BENEFITS OF DATA INTEGRITY

a. Data integrity ensures product quality, safety and patient compliance. In the pharmaceutical industry data integrity is essential to uphold the high standards of product quality and safety, there minimizing defects and reducing the risks to patients [6].

b. Data integrity enhances efficiency and productivity. Accurate data enhances efficiency, streamlines processes, and minimizes errors, enabling pharmaceutical companies to optimize production schedules, allocate resources effectively, and reduce downtime and waste.

c. Data integrity ensures regulatory compliance. It is crucial for regulatory compliance, as strict guidelines from FDA and EMA ensure pharmaceutical product safety, preventing costly fines and legal issues [6].

d. Data integrity helps to build trust with stakeholders. Transparent data practices demonstrate commitment to quality and safety, building trust with stakeholders like regulatory authorities, healthcare providers, and patients, fostering confidence in pharmaceutical companies' products and operations [6].

KEY CHALLENGES IN ENSURING DATA INTEGRITY

a. Increasing Digitalization

The pharmaceutical industry is rapidly adopting digital systems for data management. While this enhances efficiency, it also introduces vulnerabilities, such as cyber threats, unauthorized access, and software malfunctions, which can compromise data integrity [7].

b. Complex Regulatory Requirements

Pharmaceutical companies must comply with various regulations and guidelines, including those from the U.S. Food and Drug Administration [USFDA], European Medicines Agency [EMA], Medicines and Healthcare Products Regulatory Agency [MHRA] and other regulatory agencies. These regulations often have complex and evolving requirements, making compliance challenging [7].

c. Human Error

Maintaining data integrity is still greatly challenged by human error. Inaccuracies in data entry, documentation, and system operations can compromise the dependability of pharmaceutical products [7].

d. Legacy Systems

Many pharmaceutical companies still rely on outdated or legacy systems that lack robust data integrity controls. These systems can be prone to data loss, corruption, and unauthorized access, making them a significant risk to data integrity [8].

DI CONSIDERATIONS FOR PAPER BASED DOCUMENTS/RECORDS [2]

Effective management of paper-based documents is crucial for GMP/GDP compliance. Therefore, the documentation system should be structured in such a way that they fulfill these requirements, ensuring that the documents and records are effectively controlled and managed to maintain their integrity. Paper records should be managed to remain attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring [indelible/durable], and available [ALCOA+] throughout the data lifecycle.

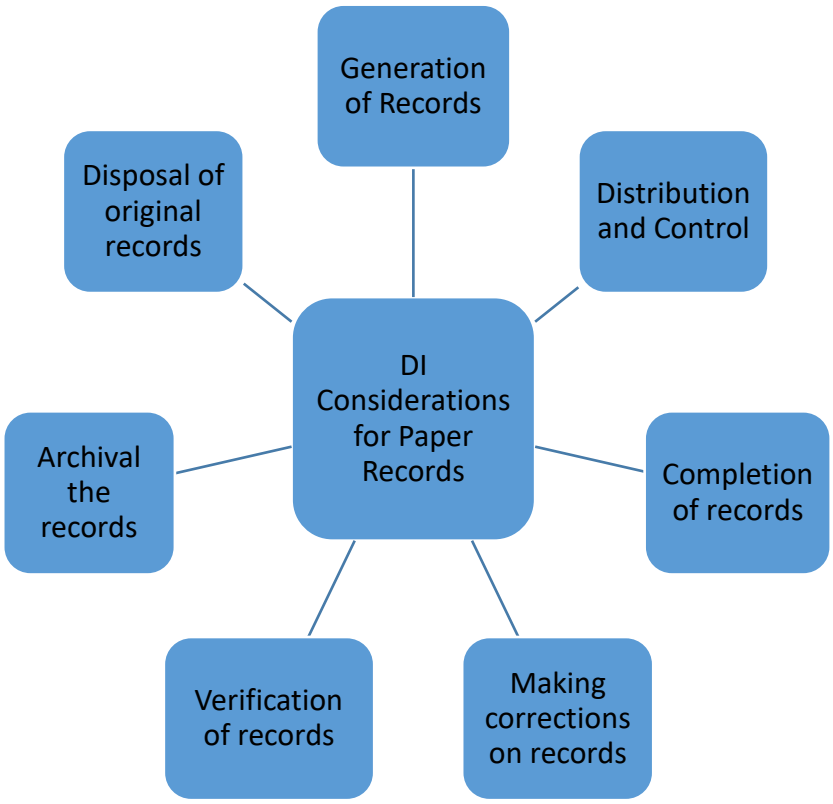

Figure 2: Considerations for DI of Paper Based Records

Expectations on Generation, Distribution and Control of Records [2]

a. Every document should have a unique identification, including its version number, and should be reviewed, approved, signed, and dated. Local procedures should prohibit the use of uncontrolled documents and temporary recording practices as it increases the risk of omitting or losing critical data.
b. The documents should be structured to provide ample space for manual data entries and should also clearly specify what data/information is required. Handwritten data may become unclear and illegible if the provided spaces for data entry are too small. Document design should provide sufficient space for comments.
c. Documents should be stored to ensure proper version control. A well-designed document ensures that all critical data is recorded clearly, contemporaneously, and in a durable manner. Additionally, the document should be designed to record all the information sequentially according to the operational process and corresponding SOP, minimizing the risk of excluding any crucial data.
d. Master documents should have distinctive markings to differentiate them from copies,

such as using colored paper or ink to prevent inadvertent use.

Expectations on Distribution and Control of Records [2]

a. Revised versions of the documents and records should be distributed promptly. All the obsolete master documents and files should be archived in an effective manner and their access should be restricted. If the obsolete versions of the documents are available for use, it can lead to several issues such as noncompliance, quality control issues, data integrity risks, operational inefficiencies or product recalls.
b. Issuance of documents should be governed by written procedures or protocols that specify who issued the copies and the issuance dates. Clear methods should be implemented to distinguish approved document copies to ensure that only the current version is accessible for use. Each distributed copy should be numbered, and documents should be time-stamped upon generation. Failure to follow these methods may lead to the risk of data being rewritten or falsified by photocopying the template record.

Expectations on Completion of Records [2]

a. Only the individual who performed the task should make entries. Any unused or empty fields in the documents must be invalidated [e.g., crossed out], dated, and signed.
b. All the records related to manufacturing operations must be completed contemporaneously. Inspectors should anticipate that sequential recording can be conducted at the operational facility.
c. The records should be enduring which means that the entries noted down should be made in ink which is not erasable or should not fade away with time.
d. All the records generated should be signed and dated by an author who is using a unique identifying number which is attributable.

Expectations on Making Corrections of Records [2]

a. When changes in the entry need to be made, cross out with a single line on data which needs to be changed. Wherever necessary, the justification for the correction of any data should be explicitly documented and if crucial, verify the initial data and the change made. Overwriting is not permitted and the original data recorded should be readable and not obscured.
b. Corrections should be made in indelible ink which is long lasting.

Expectations on Verification of Records [2]

a. All the critical process steps should be recorded contemporaneously. Batch production records for critical process steps are typically reviewed by production personnel in accordance to an approved procedure.
b. Laboratory records for testing steps should also be reviewed by designated personnel [e.g.: second analysts] following completion of testing. All entries, critical calculations should be checked by the reviewer.
c. Data verification should be carried out following the completion of production-related tasks and activities. Data should be signed or initialed and dated by the authorized persons.

Expectations on How Records should be Verified [2]

a. Document should be verified, if every field has been accurately completed using the approved templates. The data noted should be compared to the acceptance criteria. The need for a secondary check of the data entered in the record is based on the principles of quality

risk management and also the criticality of the data generated.

b. In case of equipment/ instruments which produce direct printouts, they should be attached to batch processing or testing records.
c. The retention period for each type of record must comply with GMP/GDP requirements. Records can be stored internally or through an external storage service.

Expectations on Where and How Records should be Archived [2]

a. A system should be implemented to detail the steps for archiving records, including identifying archive boxes, listing records within each box, specifying retention periods, and designating archiving locations. Additionally, guidelines for storage controls, along with access and record recovery, should be implemented. The access to the archived documents should be restricted to authorized personnel to ensure integrity of the stored documents.
b. All hardcopy quality records should be archived in a secure location in a manner that is easy traceable and that ensures records are durable for their archival period. In case of an external storage service provider, the location must be audited and there should be a quality agreement in place.
c. The records must be protected from damage or destruction by fire, liquids, rodents, humidity or by any unauthorized persons who may destroy the records.

Disposal of Original Records or True Copies [2]

a. A documented procedure for disposal of records should be established to ensure that original records or true copies are properly disposed of after the specified retention period. The system must prevent accidental destruction of current records and ensure that historical records do not unintentionally re-enter the current record stream.
b. There should be register available to demonstrate appropriate and timely archiving or destruction of retired records in accordance with local policies.
c. Measures should be implemented to minimize the risk of deleting incorrect documents. Access rights for disposal of records should be controlled and restricted to a limited number of individuals.

DI CONSIDERATIONS FOR COMPUTERIZED SYSTEMS [2]

Companies use various computerized systems, from simple standalone to complex integrated systems, impacting product quality. Regulated entities must evaluate and control these systems as per GMP and GDP requirements. Organizations should assess and document each system's use, function, and data integrity risks, focusing on criticality and product quality. Systems affecting product quality must be managed under a Pharmaceutical Quality System to prevent data manipulation. System design, evaluation, and selection should consider data management and integrity, ensuring vendors understand GMP/GDP and data integrity requirements. Legacy systems should meet these standards with additional controls if needed. A risk-based approach should be used to manage data risk and criticality, including metadata. Complete capture and retention of raw data and critical metadata are essential to reconstruct manufacturing events or analyses. Data vulnerability and risk should be assessed based on the computerized system's role in the business process, evaluating inherent data integrity controls, especially those vulnerable to exploits. During inspections, company expertise should be utilized for system access and navigation. These principles also apply to outsourced computerized systems, ensuring

compliance with GMP/GDP requirements and effective data management and integrity controls.

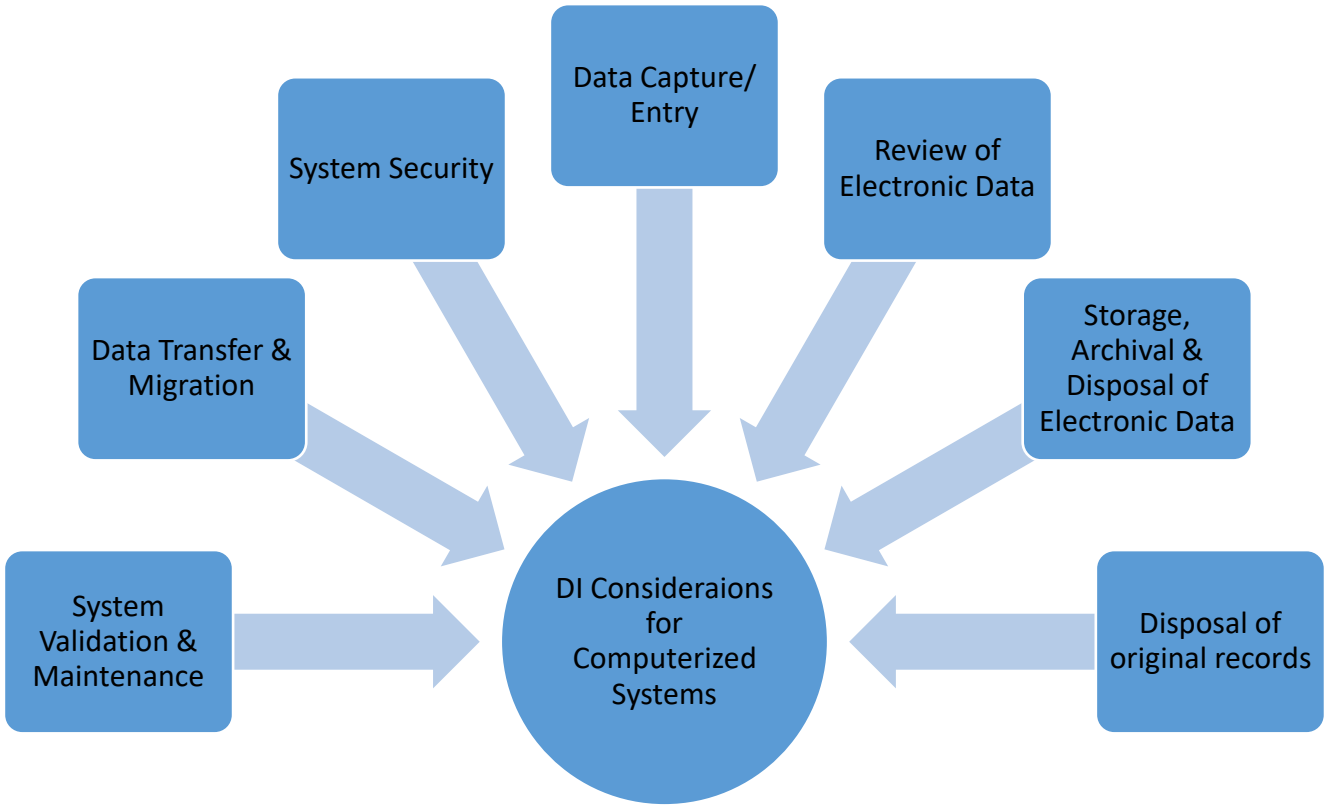

Figure 3: Considerations for DI for Computerized Records

Expectations on Validation and Maintenance of Computerized Systems [2]

Regulated companies must ensure data integrity from the start of system procurement through the entire lifecycle. Functional Specifications [FS] and User Requirement Specifications [URS] should address data integrity. Critical GMP/GDP equipment must be evaluated for data integrity controls before purchase.

a. Maintain an inventory of all computerized systems.
b. Include system name, location, function, criticality, and validation status.
c. Conduct risk assessments for data integrity controls based on system criticality and potential risk to product quality.

Expectations on Data Transfer and Migration [2]

a. Interfaces should be validated to ensure accurate data transfer, with built-in checks like secure transfer, encryption, and checksums to minimize data integrity risks. Automated GMP/GDP data transfer should be designed and qualified.
b. When installing or updating system software, ensure existing and archived data is readable by the new software. If conversion to the new format isn't possible, maintain the old software as a backup. Data migration should be controlled and verified according to documented protocols.
c. When legacy software is unsupported, maintain it for data access, possibly in a virtual environment. Migrate data to a new format if needed, balancing long-term access with functionality. Assess risks and document controls to prevent unauthorized changes. Verify the effectiveness of these controls.

Expectations on System Security for Computerized Systems [2]

a. User access controls must prevent unauthorized access, changes, and deletion of data, tailored to the system's criticality. Individual login IDs must be provided and only authorized personnel should input or change data. Administrator access rights must be strictly controlled. Systems should have an inactivity logout to prevent unauthorized access, requiring re-authentication after inactivity.
b. Computerized systems must be protected from accidental changes or manipulation mainly by restricting access to hardware, servers, and media. Also, periodic security reviews must be conducted and updated.

Expectations on Data Capture or Entry in Computerized Systems [2]

c. Systems must be designed to accurately capture data, regardless of whether it is obtained manually or automatically.
d. All required data modifications must be authorized and managed following approved procedures.

Expectations on Data capture and Entry [2]

a. Systems should accurately capture data, whether collected manually or automatically.
b. Any required data changes must be authorized and controlled according to approved procedures.

Expectations on Review of Data within Computerized Systems [2]

a. Regulated users must assess risks to identify GMP/GDP relevant electronic data and its criticality. Critical data should be audited and verified for accuracy, with all changes authorized.
b. The quality unit should schedule ongoing audit trail reviews based on criticality and system complexity, incorporating them into the self-inspection program. Procedures must address and investigate discrepancies, with escalation to senior management and authorities as needed.

Storage, Archival, and Disposal of Electronic Data [2]

a. Data storage must encompass the complete original data and all relevant metadata, including audit trails, using a secure and validated process.
b. Data storage must include the full original data and all relevant metadata, such as audit trails, using a secure and validated process.
c. Data should be periodically backed up and archived as per written procedures. Archive copies must be securely stored in a separate, remote location from the original and backup data.
d. It should be possible to print a clear and meaningful record of all data, including metadata, generated by a computerized system. Any changes to records should also be printable, showing when and how the original data was altered.

BEST PRACTICES/ STRATEGIES FOR ENSURING DATA INTEGRITY

a. Implementation of ALCOA and ALCOA+ principles [Refer Figure 5a and Figure 5b] provide a framework for ensuring data integrity. These principles should be integrated into every aspect of data management in the pharmaceutical industry [6].
b. Regular audits and inspections help identify and mitigate risks to data integrity. Audits should focus on evaluating the effectiveness of data management systems, adherence to procedures, and compliance with regulatory requirements [6].
c. The adoption of advanced technologies, such as AI and machine learning, can enhance data integrity by automating processes, reducing human error, and providing real-time monitoring and analysis of data [6].
d. To ensure data accuracy and consistency, pharmaceutical companies should also implement ongoing monitoring and validation processes. Regular audits and checks are crucial in identifying and correcting discrepancies early. This proactive approach helps maintain the integrity of data, ensuring it remains a trustworthy asset for the company [6].
e. Maintaining data integrity also relies heavily on the continuous education and training of staff. Employees must be well-informed about the latest data management practices, regulatory requirements, and the critical importance of accurate and dependable data. Regular training sessions help instill a culture of data integrity within the organization, ensuring that all staff members are equipped to manage data effectively [6].

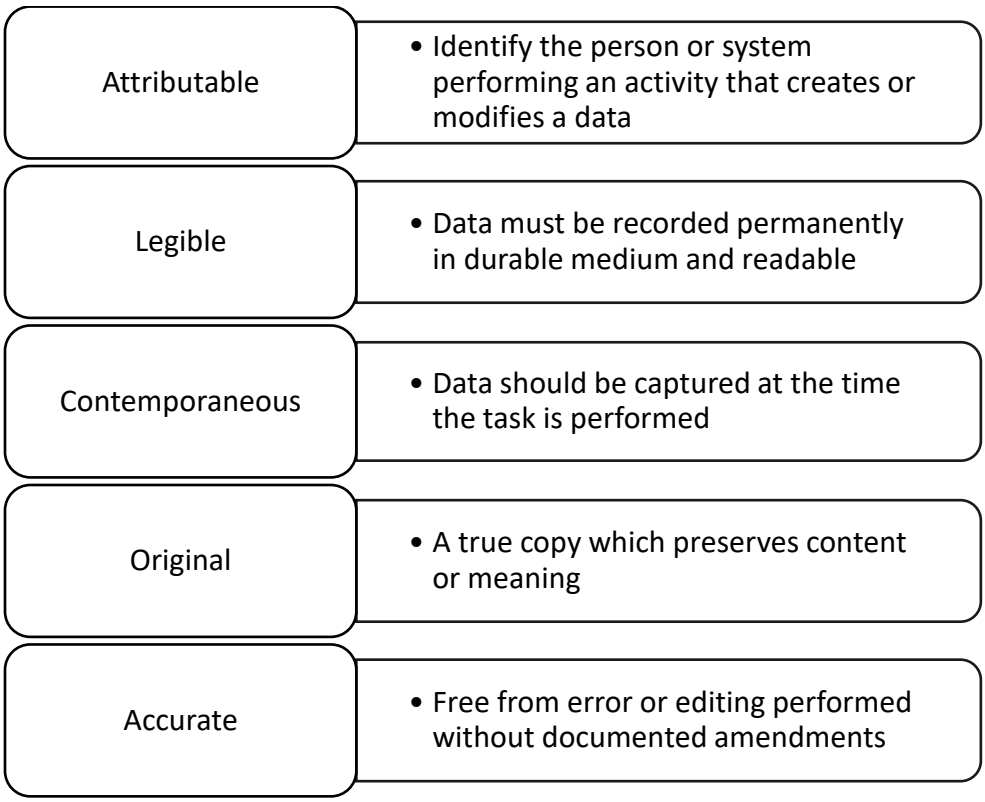

Figure 5a: ALCOA

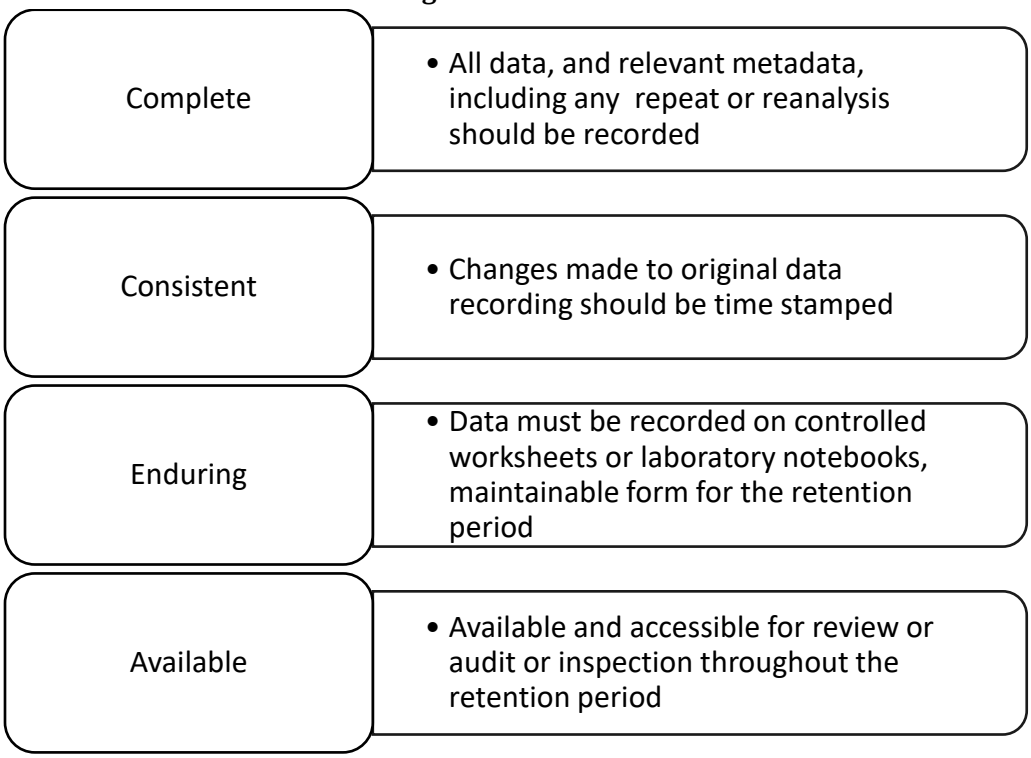

Figure 5b: ALCOA Plus

DATA INTEGRITY VIOLATIONS

Data integrity violations in cGMP have led to regulatory actions like warning letters, import alerts, and seizures. Numerous serious DI issues have been uncovered, posing long-term risks to companies and affecting their culture. Managing

DI is particularly challenging in the pharmaceutical industry due to rapidly increasing data volumes. Poor data quality can damage an organization's reputation. Implementing data controls without understanding regulatory and business processes can lead to questionable data validity and regulatory action [9]. DI is crucial for ensuring data accuracy and consistency throughout its lifecycle. Without proper measures, there's a high risk of corrupted results. DI errors often stem from human error, inadequate procedures, data transfers, software defects, and physical damage. Maintaining DI is essential for accountability, ensuring the safety, effectiveness, and quality of drug products, and is a critical aspect of regulatory compliance [10].

ABBREVIATIONS

a. USFDA – United States Food and Drug Administration
b. EMA – European Medicines Agency
c. WHO – World Health Organization
d. ALCOA – Attributable, Legible, Contemporaneous, Original, Accurate
e. ALCOA+ – Attributable, Legible, Contemporaneous, Original, Accurate + Complete, Consistent, Enduring, Available
f. FDA – Food and Drug Administration
g. GMP – Good Manufacturing Practice
h. cGMP – Current Good Manufacturing Practice
i. MHRA – Medicines and Healthcare Products Regulatory Agency
j. GDP – Good Distribution Practice
k. DI – Data Integrity
l. AI–Artificial intelligence
m. SOP–Standard Operating Procedure
n. GXP–good practice guidelines and regulations in the life sciences industry, including good clinical, laboratory, manufacturing, and other practices

CONCLUSION

Ensuring data integrity in the pharmaceutical sector is paramount to safeguarding public health and maintaining the credibility of the industry. The benefits of upholding data integrity include enhanced product quality, compliance with regulatory standards, and fostering trust among the stakeholders. However, the sector faces key challenges such as human error, inadequate training, outdated systems, and intentional data manipulation. Addressing these key challenges requires a strategic approach, involving robust data management systems, continuous employee training, and the adoption of modern technologies like blockchain and AI for real-time monitoring and verification. Furthermore, compliance with regulatory expectations from regulatory bodies such as FDA, EMA and others, is critical for preventing potential violations and ensuring patient safety. In conclusion, maintaining data integrity is not only a regulatory requirement but a fundamental responsibility for pharmaceutical companies. A proactive, technology driven strategy aligned with regulatory standards will enhance transparency, ensure compliance, and ultimately contribute to the delivery of safe and effective pharmaceutical products.

REFERENCES

1. Food and Drug Administration. Data Integrity and Compliance with Drug CGMP Questions and Answers Guidance for Industry: December 2018.
2. PIC/S Pharmaceutical Inspection Convention. Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments: July 2021.
3. MHRA GXP Data Integrity Guidance and Definitions; Revision 1: March 2018.
4. ISPE-FDA 3rd Annual CGMP Conference Baltimore, MD Current Inspectional and ISPE-FDA 3 rd. Annual CGMP Conference Baltimore, MD What is Data Integrity? Why is Data Integrity Important? June 2014.

5. Anil K. Rattan. PDA Journal of Pharmaceutical Science and Technology March 2018, 72 [2] 105-116; DOI: https://doi.org/10.5731/pdajpst.2017.007765
6. Pharma 4.0: the importance of data integrity,eubusiness.com/focus/pharma-4-0-data-integrity : August 2024.
7. Parenteral Drug Association[PDA], Elements of a Code of Conduct for Data Integrity in the Pharmaceutical Industry | PDA : 2016.
8. Documentation and Data Integrity in Pharmaceutical Industry | SpringerLink: March 2024.
9. Bhatia K. Current scenario and future challenges in pharma segment, pharmabiz.com; 2014.
10. International Journal of Pharmaceutics sciencedirect.com/science/article/abs/pii/S0378517322010584 : January 2023; Volume 631, 122503

HOW TO CITE: Akshaya U. Bhandarkar, Balaji J., Jasvanth R., Dinesh Ragavendra S. , Ensuring Data Integrity In The Pharmaceutical Industry: Benefits, Challenges, Key Considerations And Best Practices, Int. J. of Pharm. Sci., 2024, Vol 2, Issue 10, 1198-1210. https://doi.org/10.5281/zenodo.13971925