DATA INTEGRITY ISSUES, STATISTICAL TRENDS AND UTILIZATION OF ALCOA+ IN PHARMACEUTICAL INDUSTRY

Abstract

Background: Data integrity is critical to regulatory compliance, and the fundamental reason for 21 CFR Part 11 published by the U.S. Food and Drug Administration (FDA). Good documentation frames are an essential aspect of good manufacturing practises (GMP) in pharmaceutical enterprises because they play a vital part in the sanity of data preservation and its development, certifications, registrations, commodification, and life-cycle governance for pharmaceutical goods. Data integrity is a key concern in the regulated pharmaceutical sector due to observations and results due to inappropriate record keeping practices or data manipulation.

Main Body of the Abstract: In the present review article, data is collected from various online sources which includes articles, CFR and WHO guidelines and statistical data from Blogs. The articles have been thoroughly gathered and reviewed from a variety of databases and indexed journals.

Short Conclusion: An upsurge in data integrity breaches in 2019 led regulatory organisations such as the WHO and FDA to issue guidance on effective documentation practices. The present review article will help researchers to understand what is data integrity? Its importance in pharmaceutical industry, ways to reduce the data integrity issues by utilizing concept if ALCOA+ and recent WHO guidelines for data integrity issues.

Keywords: Data Integrity, 21 CFR Good Documentation Practices, Regulatory Bodies, ALCOA +, WHO

Authors

Mrs. Nisha S. Shirkoli
Department of Pharmaceutical Quality Assurance,
KLE College of Pharmacy, Belagavi.
KLE Academy of Higher Education and Research
Belagavi, Karnataka, India
nishashirkoli@klepharm.edu

Mrs. Kishori P. Sutar
Department of Pharmaceutics,
KLE College of Pharmacy, Belagavi.
KLE Academy of Higher Education and Research
Belagavi, Karnataka, India
kishorisutar@klepharm.edu

Mr. Veerkumar P. Japti
Department of Pharmaceutical Quality Assurance,
KLE College of Pharmacy, Belagavi.
KLE Academy of Higher Education and Research
Belagavi, Karnataka, India
veerkumarjapti@klepharm.edu

I. INTRODUCTION

Data has always been an important part of in pharmaceutical manufacturing and research activities. The importance of data is growing tremendously in range of different factors in pharmaceutical manufacturing.During recent cGMP inspections FDA has observed numerous data integrity associated violations. Ensuring data integrity is vital component in pharma sector as it directly impacts the efficacy, safety and quality of drugs. Increasing data integrity issues, questions FDA's ability to protect public health. These data integrity issues have led regulatory bosies to carry out stringent actions which include warning letters, import alerts, form 483's and mutual agreement on termination[1]

Data integrity is a serious concern in the regulated pharmaceutical sector as a result of observations and outcomes of inadequate record keeping practises or data manipulation. This has led in multiple FDA warning letters along with data integrity guidelines provided by regulatory authorities like as the MHRA, WHO, FDA, and PIC/S, as well as industry associations.[2]

Every healthcare industry is concerned with Good Documentation Practises (GDP) within the organisation in order to lead the global marketplace of pharmaceuticals or medical devices. The procedures also involve extensive documentation to ensure that they are audit-ready and have traceable records. [5,6]

In recent years, healthcare sectors, particularly in India, have repeatedly received warning letters for data integrity flaws that have resulted in massive losses, for example [7]:

1. Regulations imposed by regulatory authorities such as the WHO, FDA, and others.
2. Due to lack of traceability, organisations are unable to respond to agencies
3. Constrained reputation in the market
4. Loss of client trust results in loss of business.
5. Questions on the marketed product regarding authenticity
6. Major deprivation in trades and share market

False and unethical pharmaceutical data integrity practises have resulted in significant regulatory and financial ramifications for an extensive number of companies.
Major issues for why we are dealing with data integrity are:

1. Senior management who doesn't themselves learn and don't support data integrity. They don't promote support or guide the employees for following correct procedures.
2. Employees who lack technical and regulatory knowledge of process or product hence cannot perform their jobs adequately and accurately.
3. People don't understand importance of cybersecurity and data completeness[8]

The U.S. Food and Drug Administration's (FDA) 21 CFR Part 11 was established for the main objective that data integrity is essential to regulatory compliance. Since the FDA launched its first guideline in 1963, it has collaborated with the European Union to develop numerous guidelines on a variety of topics related to data integrity for pharmaceutical companies. Data integrity is security of data from unintentional changes to information protecting data from unauthorised parties.[9]

Integrity stems from the Latin word 'integer' which means whole and complete. So, integrity requires an inner sense of 'wholeness' and 'consistency' of character.Integrity is not possible without compassion and makes it clear that doing the right thing includes doing it for the right reason[10]

Integrity, ethically can be defined as "the honesty and truthfulness or accuracy of one's actions and a concept of consistency of actions, values, methods, measures, principles, expectations, and outcomes.[11]

Data can be defined as attributes or details, usually numeric, collected as a result of various observations. Data in more technical sense may be defined as collection of qualitative or quantitative variables about one or more persons or objects. In complete sense data integrity is characteristics or information which is complete, accurate and reliable.
"Integrity is doing wright thing when no one is watching"

- C. S. Lewis

II. DEFINITIONS

1. **Meta Data:** Metadata is a contextual data that depends on the preceding or following parts of a text to clarify meaning. In short it is a set of data that describes and gives information about other data. It helps organize electronic resources, provide digital identification, and archive and preserve resources. For example, to denote file size number 23 is meaning less without metadata i.e., the unit KB. Other examples of metadata may be ID given to a person, time or date of activity performed. Many distinct types of metadata exist, including descriptive metadata, structural metadata, administrative metadata, reference metadata and statistical metadata. [12]

2. **Raw Data:** Raw data, also known as primary data, are data (e.g., numbers, instrument readings, figures, etc.) collected from a source. In the context of examinations, the raw data might be described as a raw score. Raw data is primarily unstructured or unformatted repository data. It can be in the form of files, visual images, database records or any other digital data.[13] In a lab when performingexperiment if a provision is made to note the temperature in a reaction mixture every minute and recorded in computer system or record sheet are called raw data.

3. **Static Data:** This document comprises static data and may take the form of an electronic image or a paper record. It is data that does not alter after recording, such as a paper or pdf record that permits little or no interaction between the user and the record matter. For example, once printed or converted to fixed or constant pdfs, chromatographic data lose their ability to be recovered or allow a more complete survey of baselines. [14]

4. **Dynamic Data:** A chromatogram, for example, where the integration parameters (peaks) can be changed, is an example of a data layout that enables the recording of essential relationships between the user and the content recorded. Additionally, it enables the user to modify formulas or entries in record sheets used to compute analytical results or other data, such as computed yield. [1, 9,14]

5. **Electronic Data:** Digitalization for raw materials maintenance can take many shapes from online certificates of analysis (CoAs) to electronic batch records (EBR). Companies in the pharma industry are now approaching towards more digitalization at various stages of manufacturing. Automated systems benefit us by minimizing the double verification which is mandatory for manual papers. Electronic data includes data from ERP software for controlling lab scale data, quality systems data warehousing data and other records maintained in pharmaceutical industry.[14, 25]

6. **Audit Trail:** It a system that traces the detailed transactions relating to any item in an accounting. record. It is secure, computer-generated, time-stamped electronic recordof the changes that have been made to a database or file. An Audit trail is source of progressive, sequential, consecutive set of records that furnishes secured documentary evidences for series of activities that may have altered or modified at any time of specific procedure, test, operation or event. E.g.,For system-based record: In generation and issuance of Batch manufacturing record from software audit trail will include the record of date and time, BMR number, which regulated market it belongs to, Batch size, product name etc. For a paper record, the initials of the person making the change, the date of the change, and the reason for the change—all information necessary to authenticate and ratify the change—would be documented with a single-line cross-out, preserving the legibility of the original entry. [1, 14]

7. **Backup:** In the event of a system crash or disc corruption, backup data is required. A backup is a set of duplicate electronic files that are kept as insurance in case the original data or system is lost, corrupted, or otherwise unrecoverable. It's crucial to remember that these backup files serve their purpose in storage temporarily and shouldn't be used as an archive system. [1, 14]

8. **Computerized System:** A computerized system consists of hardware, software, operating system software, and supporting documentation. Examples include automated laboratory systems, operating instructions, manuals, control systems, clinical, manufacturing or compliance monitoring database systems. One or more automated processes and/or functions are collectively controlled by a computerised system. [14]

9. **Corrective and Preventive Action (CAPA):** The word CAPA refers to actions done to eliminate or reduce nonconformities or other objectionable, disagreeable, unacceptable, unsuitable, undesirable, or out of place situations in an organisation. CAPA is a widely used concept in GxPs (good laboratory practises, good clinical practises, and good manufacturing practises), as well as a number of ISO (International Organisation for Standardisation) business standards.It is a collection of procedures, laws, or rules that must be followed by an organisation in order to account for duplicate non-conformance in production, documentation, analytical methods, or computing systems. Nonconformity can happen when a production process causes a drop in quality and is not immediately corrected. A non-conformance could be a market or consumer complaint, a machine failure, a whole quality management system failure, or a misinterpretation of written job instructions. Corrective action refers to a measure taken to get rid of the source of a nonconformity and stop it from happening again. Preventive action is defined as action taken to eliminate the source of a possible nonconformity or other unpleasant circumstance. [14]

10. **Data Governance:** Data governance refers to a way of retaining data in the format in which it is generated, documented and handled in order to assure completeness, consistency, and accuracy throughout the data life cycle. [14]

11. **Data Life Cycle:** The data life cycle is the series of eventsthat a specific unit of data goes through, from beginning of its generation or manifestation to its ultimate recording, documentation and/or deletion, withdrawal and clearance.Data life cycle management inculcates all the aspects of data generation, collection, processing, storing, retrieval, analysing, visualizing, interpretation and its archivalThere should be a structured approach for analysing, monitoring, and managing data so that the risks associated with data may be managed in a way that corresponds to its potential impact on product quality and patient safety throughout all phases of the data life cycle. The data life cycle acts as a navigational tool to assist users in finding suggestions on how to efficiently deal with their data across all stages of the data life cycle. [14]

12. **Good Documentation Practices:** Good documentation is described as standards by which data is collected and framed into document. It can be defined as measures taken safe guard the data collected whether on paper or digitalized system continue to exist as traceable, legible, attributable, intact, permanent, original and accurate throughout the document longevity. [14]

13. **GxP:** A collection of rules and standards known by the abbreviation GxP were developed to guarantee the security of products used in the life sciences while upholding the standard of operations at every stage of manufacture, control, storage, and distribution. GxP is a shared term for guidelines and regulations under good practices which ensures the quality in many domains of globally recognised paradigm such as GMP (Good Manufacturing Practice), GCP (Good Clinical Practices), GLP (Good Laboratory Practice), GDP (Good Documentation Practice), GSP (Good Storage Practice), GDP (Good Distribution Practice) and GRP (Good Review Practice)[14]

14. **Hybrid Approach:** Hybrid approach is a mix up computerized system where combination of original paper records andelectronic records that combine to give information about the complete data set which is then reviewed and maintained. E.g., while doing analysis the analysts, with generated electronic record even takes a print of the result for summarizing the total data. When hybrid approach is used its mandatory to use appropriate controls for recording data like templates, master documents, recording sheets according to SOP.The link between the original record and paper record should be such that they be legible throughout data life cycle.[16]

15. **Master Controlled Documents:** Mater controlled documents are approved documents which are in custody of QA department and can be issued for operational work after doing requisite issuance formalities. Examples of such documents are master SOP, Validation Protocols, Manufacturing\Production Records, master packing recordQC specifications etc.[16]

16. **Original:** Original records are the preferable versions of records and refer to the first generation of data. Records should be archived to the greatest extent practicable for an original record.[16]

17. **Post-Dating:** Post-dating means assign a date later than the actual one to (Enter a data before activity is performed).[16]

18. **Back- Dating:** It means is executing a document and then dating it with an earlier date than the actual date of execution.[16]

19. **True Copy:** A true copy is a copy made from the original record. It is a certified document showing all the details of the original documents but is not original document. It confirms that the it contains the same data or exact data mirror imaging to original data.[16]

Pharmaceutical industry is the vital segment of health care system as it deals with manufacturing of medication for patients intended for safe and therapeutically active with good consistent quality. The creation, certification, registration, commercialization, and life-cycle governance of pharmaceutical products all depend on the integrity of data management, which is a critical component of good manufacturing practises (GMP). The GDPs help us avoid collecting inaccurate data when making and analysing pharmaceutical products. Product quality and patient safety would be affected, either directly or indirectly, by this. GDP compliance is required by both the US and European regulatory agencies, the EMA (European Medicines Agency) and the USFDA (United States Food and Drug Administration). A number of international organisations, including the World Health Organisation (WHO), Health Canada, and EudraLex (European union collection of standards for fundamental legislation governing medical products), together with the United States Pharmacopoeia (USP), have released certain guidelines corresponding to GDPs.Not only the regulatory requirements but maintenance of authentic records is also important as documentations of activities in the pharmaceutical industry allows critical evaluation of internal procedures and continuous improvement of process and minimizes the time for repetitive detailed study from start to end.[17]

The WHO states that the goal of good documentation practises is to:

- Set a predefined specifications and operational strategies for all materials and processes of manufacturing and analysis,
- Set job responsibilities of employees based on their expertise field
- Ensures release of a product to be done after all the necessary information to authorized persons
- Maintain documented evidence for future investigations and audits in a legible manner
- Maintain constant availability of data for statistical and validation analysis.

For an organisation to be compliant with regulatory authorities and to enhance market value it is important to:
- Provide requisite resources for fulfilling complete documentation,
- Ensure that the documents are prepared with concept of ALCOA+. (Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Enduring and Available)

- Fill the gap of any missing data for existing documents providing basis of forecasting what is to be done in future
- Continuous training should be provided to employees on existing and current GDPs to be followed in manufacturing ang quality control departments.

Quality Assurance department needs to review all the documents before release of product to market [17].

III. TYPES OF DOCUMENTATION IN PHARMACEUTICAL INDUSTRY[18]

Both printed electronic forms and electronic systems are considered documentation..

1. Standard Operating Procedures(SOP)
2. Records/Worksheets
3. Annexures
4. Quality management system documents
5. Production and packaging instructions
6. Standard operating procedures
7. Records.
8. Licencing documents
9. Master documents
10. Technical agreements
11. Confidentiality agreements
12. Training records
13. Qualification Documents(URS, DQ, FAT, OQ, PQ, IQ)
14. Quality Manual
15. Issuance Documents
16. Annual product quality review (APQR)documents
17. BMR/BPR
18. Validation protocols and reports
19. Deviation reports
20. Audit plans
21. Validation Master Plans and validation documents
22. Test material related documents including test material receipt, product specification, and reports
23. Documents pertaining to personnel, such as training records
24. Documents related to facilitywhich includes floor plans, HVAC plans, and environmental specifications
25. Planned Deviation, unplanned deviations and system failure investigation
26. Change control
27. Logbooks, worksheets, and notebooks

In pharmaceutical industry it is said that "**If It Is Documented It Is Done**". So, recording and maintenance of data plays a very important part in pharmaceutical industry.

IV. CONCEPT OF ALCOA TO ALCOA+

For prevention of data integrity issues concept of ALCOA was utilized by pharmaceutical companies. ALCOA is concept to implement for the data guarding and sturdiness in pharmaceutical industries. In 1990's Stan Woollen used the acronym ALCOA when he worked for the agency to help him remember compliance terms relevant to data quality which has been widely associated with data integrity by FDA. [19]ALCOA defines a framework to achieve and maintain data integrity, especially important for ensuring good manufacturing practices in regulated industries. The theory of ALCOA can aid to furnish an audit trail data that encapsulates details such as additions, deletions, or alterations of data in an electronic record without obscuring the original record. [20]As an improvement over this, additionally, new characteristics were identified which are critical to data integrity apart from regular good documentation practices core characteristics. This launched the improved ALCOA now known as ALCOA (+).[21]The term ALCOA which is acronym stands for Attributable, Legible, Contemporaneous, Original and Accurate. Later ALCOA was outstretched to ALCOA-C or ALCOA+ (Fig 1). Complete, Consistent, Enduring and Available (CCEA) was added to ALCOA in 2010. They are considered as the heart of Good Documentation Practices because of the specific significance they hold. [22]
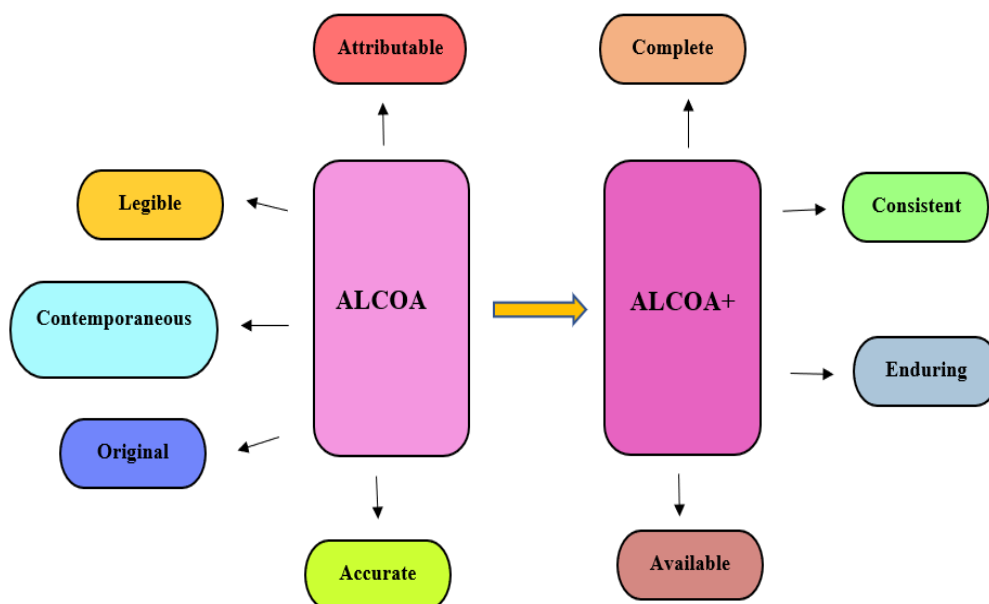

Figure 1: Outstretching to ALCOA to ALCOA+

Following are the features for Quality records considering Good Documentation Practices in place.

1. **Attributable:** When and Who performed an activity? It indicates that the any data should be traceable and determinable regard of when it is performed and who performed it. It should be capable to discover and find out the emergence and the past of data. The record should also be able to identify the person who initiated, collected, preserved, or managed the information. This produces trace-ability in urgent situations. Consequently, accountable is sometimes regarded as traceable. [[22, 23,24]

Example: System user ID sharing and password sharing, Analysis done by two person and one person signing

2. **Legible:** Data should be readable throughout its entire life cycle? Data should be readable throughout its entire cycle. The document should be simple to read, accurate, and intelligible, in other words, relatable and readable. This is necessary because data must be readable and understandable for years or even decades after it has been captured. History for the particular data should be relevant and understandable even after a gap of long time period.[19, 25]
   Example: Hand writing should be readable by others, Write over's- usage of pencils or erasers/ correction fluids

3. **Contemporaneous:** Is the documentedat the time of the activity? Contemporaneous means Parallel, Simultaneous, or Concurrent. The data should be recorded at the time when activity is performed. Data should be recorded at the same time they are generated, collected and observed(include time and date stamps for electronic records). Data should never be backdated, or forms completed with expected results prior to execution.[26, 30]
   Example: Data should be recorded at the time of activity, Back dating or forward dating (Non-compliance with GDP)

4. **Original:** Is the information gathered original or certified copy? The original word refers to the availability or existence of raw data, whether on paper or electronic data. Sometimes the prints which are generated electronically or digitally fade away over a period of time, in those cases the document should be scanned or photocopied and saved for future references.Original data refers to the paper or electronic media on which the data point was first recorded, such as logbooks, protocol, registers,notepad, worksheets, data archives, or software tools. It is best to get in the habit of entering the data directly into the main register or original notebooks rather than writing it down on a piece of paper and finishing it later. This could lead to more errors and impair the ability to authenticate original data. Original data is defined as data or information that is documented at the time of data generation and includes all subsequent data required to assure the quality and proper conduct of GxP activities. A second person verifies a certified copy by comparing it to the original, ensuring that it is exact, correct, and complete and maintaining the original content and importance in its original form. [27, 30]
   Example: Altering/ modifying and deleting original data, Analytical results written on new worksheet as original data got impaired (smudged or torn off or discarded)

5. **Accurate:** Are there any errors or editing without documented amendments? Consistent, factual, error free and recorded as it is. For validity of data, it should be error free. In case of amendments there should be evidence for accompanying support for changes made. The quality of data should be maintained such that any changes made during any time of data life cycle has enough proof for supporting the change made. If any editing is done it should be done considering GDP.[26, 27]
   Example: System adjusted to get passing results to avoid OOS (out of specification), Data from passing run analysis is used for another sample to expect a result within specification

6. **Complete:** Is all data documented viz. any test, repeat or re-analysis performed? Data for any repeat or reanalysis performed on the sample should be noted correctly.Evidences of the data for any reanalysis performed should not be deleted(Evidence: Audit Trail). Evidence should be such that, when reconstructing or reforming the events, data must be sufficient and complete for required information. This means that there should not be any breach when remodelling of requisite information. All paper and electronic data, including all tests (original and retest), must be properly recorded, clearly identifying when and who performed the test, and ensuring that nothing is missing.[25, 30]
    Example: Information worksheets without name of person, instrument ID, date and time etc, meta data support is not mentioned for the original raw data collected

7. **Consistent:** Whether all the components of the investigations are always carried out similarly? Consistent with reference to the events in chronological order.Records should be maintained in such a way that data should reproduce correct information consistently at any time. Any information that will be entered into a database must be valid and adhere to all the established guidelines for that system. Data consistency does not guarantee that the transaction is accurate in every way. The main purpose of it is to find any programme faults that caused the violation. [21, 23]Data should be collected with the use of a system that enforces the use of approved data gathering and analysis methodologies, reporting templates, and laboratory workflows.
    Example: Batch records not filled on time by the operators, Flash results which tend to disappear before analysts can jot down the result.

8. **Enduring:** Is all the datarecorded systematically in laboratory notebooks or in validated systems? The data or information must be maintained in such a way that it should remain unviolated, pristine, approachable and accessible throughout their specified retention period. Paper or electronic data should be transcribed in validated software systems and defined laboratory notebooks, registers, log books, spreadsheets, worksheets, annexures etc. Data should not be recorded in any scrap papers.[19]
    Example: Tampered batch records, deletion or modification of existing system data due to upgradation of system.

9. **Available:** Can the data be accessed for review, audit or inspection over the lifetime of the record? Availability of the data severs the basic purpose of saving it so that it can be referred when needed. Data should be readily available when required for inspection and audit purposes. During regular inspection of authorities, the requested data should be produced instantaneously in legible format. Electronic and Paper data should be neatly arranged in a manner so that retrieval and recovery is easy. [19, 30]
    Example: Back up for the files are not available and deleted in case of OOS results, Archival of data is not done until data retention period

V. SIGNIFICANT ISSUES OF DATA INTEGRITY

Majority of the data integrity issues and Form 483 are due failure in quality control instruments.

1. **There are Numerous Causes of Data Integrity Problems. Some of them Include the Following. [28]:**

   - No support record for raw data or loss of data during changes made in digitalized system
   - Inaccurate and incomplete records or discarding the data of repeated tests, trial runs, sample runs
   - Using CoA(Certificate of Analysis) of one batch to release another
   - Backdating
   - Manipulating integration parameters of chromatogram to get pass results
   - Deletion/manipulation of electronic records
   - Shutting down audit trail
   - Signing instead of concerned person authorized for the work
   - Deficient/In complete computer validation.
   - Activities not recorded contemporaneously
   - Sharing Login IDs by other analysts. By doing this work done by individual analyst cannot be identified.
   - Ditto markings are used
   - Use of a signature stamp
   - Not using ink in the manner prescribed by protocol
   - Incorrect ink was used for entries during the spillage, resulting in unintelligible data.
   - Failure to identify person who made corrections of changes in logbooks
   - Arcane original data
   - Usage of pencil
   - Erroneous/Fallacious records
   - Correct dating of handwritten changes not done
   - Multiple line-throughs, write-overs, and the use of "white-out" or another masking technique

   The most common GDP violations occur when corrections of the errors are done during recording of information.

2. **Following Methods should be Carried Out for Correction of Documentation (Figure 2):**

   - A single line should be drawn across the mistake
   - Mention the correction next to the error
   - Explain the error,
   - Date and sign of the individual carrying out the correction
   - Mention time in 24 hr format

Recording Time	Recording Date	Recorded by	Verified by
~~16:30~~ * 16:32	09th June 2020	NSK 09th June 2020	NSS 09th June 2020
16:35	09th June 2020	NSK 09th June 2020	NSS 09th June 2020
~~16:41~~ * 16:43	09th June 2020	KPS 09th June 2020	NSS 09th June 2020

* ® NSK 09the June 2020	# ® KPS 09the June 2020

Figure 2: Method for Correction of Document

These common errors should be highlighted in training of good documentation practices. [9, 28] As it is said **"Prevention is better than cure"** industries must follow the Good Documentation Practices actively.

VI. HOW TO DECREASE THE RISK TO DATA INTEGRITY?

1. **All Computer Systems should Comply to 21 CFR Part 11 Guidelines:** The FDA regulation 21 CFR Part 11 is concerned with electronic records. Make sure the electronic data is accurate, comparable, and consistent with paper records. Data must adhere to quality requirements in order to be kept in digital systems.

2. **Software Development Lifecycle Must be Followed:** To supervise the quality task performance, a software development lifecycle methodology can be followed. This methodology aids to investigate different functions of software lifecycle phases which embodies software testing, software development, installation and amalgamation of ongoing system maintenance. All digitalised systems should be suitably qualified, developed, quantified, analysed, investigated, and on routine basis they should be validated.

3. **Validation of Computer Systems:** Software system validation is essential because it gives proof that a certain manufacturing procedure regularly produces products that meet pre-established requirements and quality standards.

4. **Implement Audit Trails:** The date and time of data entries, changes, individuality and deletions made should be recorded by a secure, digitally-generated, time-coined audit trail. Audit trails provides evidence that data entered in electronic records are true, ethical, meets necessary data conversation, and promises that information recorded have not been edited, modified or deleted.

5. **Software for Detecting Errors should be Used:** There must be an incorporation of automated inspection software that will authenticate critical documents which will assure systems meticulousness. Manual examining or inspections are not sufficing to furnish that data are error-free.

6. **User System Access should be Maintained to Secure Your Records:** All systems containing at least two different pieces of information should have a secured login that only allows concerned parties access in order to ensure data integrity.

7. **Maintenance of Backup and Recovery Procedures are Important:** In case of unexpected data loss or system errors backup and recovery strategy are necessary. This practice guarantees the physical and critical recovery of database ensuring veracious restoration of data.

8. **SOPs and Logical Controls should be Followed to Design a Quality Management System (QMS):** Standard operating procedures and guidelines collectively builds a Quality Management System which lays foundation for ensuring quality into process by systematically organising the process. For ensuring clear liability it is mandatory to follow efficient procedures.

9. **Establishment of a Vendor Management Qualification Program is Necessary:** For ensuring that all the products manufactured are genuine and authentic it is mandatory to examine all the vendors supplying raw materials. Initial validation followed by continuous verification must be carried for ascertaining quality production. Organizations must interact with vendors for new or updated data integrity procedures they follow.

10. **Training of Users and Maintenance of Records Must be Done Appropriately:** For a job to be performed in predetermined manner precise and appropriate training must be given with the help of expertise's. Proof of the training delivered must be documented.

11. **Conduct Internal Audits to Evaluate Procedures:** For a company to thrive in competitive business environment, routine internal audits aids to understand the flaws in regime followed by the firm which empower firms to continuously prosper. Self-inspection by the company under an internal auditor who is trusted consultant assigned for advising upper management on how to best manage the company's quality management system (QMS). [29, 30]

VII. CODE OF FEDERAL REGULATIONS (CFR) AND DATA INTEGRITY [31-36]

With respect to 21 CFR Guidelines includes:

- Part 210 – Current Good Manufacturing Practice in Manufacturing, Processing, Packing, or Holding of Drugs; General.
- Part 211 – Current Good Manufacturing Practice for Finished Pharmaceuticals.
- Guidance for Industry Part 11, Electronic Records; Electronic Signatures — Scope and Application. Table 1 represents data integrity issues related to 21 CFR guidelines in pharmaceutical industry. [23]

Table 1: 21 CFR Guidelines Associated Data Integrity Issues

Sl. No.	21 CFR Guideline No.	Guidelines	Data Integrity Issues in Pharmaceutical Industry
1	211.192	Failure to adequately investigate any inexplicable disparity or a batch's failure to meet any of its standards, whether or not the batch has already been delivered, or any component of a batch failing to meet any of its specifications	OOS and partial batch rejection were the results of the investigations for a tablet with a specified unit dose (mg) being insufficient. Investigation revealed an OOS for a certain lot's tablet thickness. An evaluation of the production processes and the partially released completed lot that wasn't subjected to stability was left out of the assessment.
2	211.188(b)(11)	Failure to keep track of batch production and control records that contain proof of completion for each important step in the production, processing, packing, or storage of each batch of pharmaceutical product.	Batch Manufacturing Record of the firm lacks to include appropriately statement and control of the processing steps at each stage of manufacturing/ production of products.
3	211.194(a)(8)	Failure to confirm that laboratory records contained complete data generated from all tests necessary to assure adherence to established specifications and standards	In a laboratory for Viscosity testing the quantity samples is critical. Results will be low if sample is too less and results will be high if sample is more. But the firm failed to mention the amount of the sample used for the test.
4	211.68	Failure to implement sufficient controls over computer or connected systems to ensure that only authorised people make modifications to control records and master production, or other records.	For computer or digitalized systems there is not much appropriate controls on its utilization by individuals. The user access controls or password protection with relation to the state of the goods in inventory were not effectively validated by the software meant for drug material and inventory control, drug production scheduling, and control over the distribution of finished products.

5	211.100(b)	Failure to adhere to written manufacturing and process control methods intended to ensure that the pharmaceuticals you create have the identity, strength, quality, and purity they claim or are represented to have, and to record the same at the time of performance.	Firm fails to follow manufacturing process and online entry control parameters that will define the quality and purity of the product to conform with predetermined specifications.
6	211.160	Failure to follow and record the necessary laboratory control procedures at the time of performance	Any extra tests or procedures that the firm does that are not covered by a client-specific process must be listed in the Supplemental Method Information Sheet (SMIS). A SMIS then becomes a component of the company's official testing procedure. The material should be introduced to the viscometer in little amounts dispersed over the cone and plate rather than all at once, according to laboratory literature on the bulk release testing procedure. In addition, lab SMIS states that the quantity on the sample cup is important for units of measurement.
7	211.165(a)).	Your business does not have a suitable laboratory evaluation of satisfactory compliance to the drug product's final requirements, including the identity and concentration of each active ingredient, prior to release, for each batch of a drug product.	The average of three samples tested for each batch is used to determine potency results in accordance with the sampling technique. The company did not establish an acceptance standard for the standard deviation of test results or the outcomes of individual tests. Therefore, current procedure allowed the release of sterile medicinal products even while individual potency test results were sub- or super-potent in comparison to the potency release criterion of the finished drug product.

VIII. STATISTICAL TRENDS OF DATA INTEGRITY VIOLATIONS IN PHARMACEUTICAL INDUSTRIES

1. **Statistical Trends of Data Integrity Issues:** The Food and Drug Administration's Bioresearch Monitoring (BIMO) Program's warning and closeout letters were issued between the US fiscal years 2007 and 2018. The letters were analysed by grouping regulatory violations into topics. For the purposes of this study, 300 warning letters in total were examined. All warning letter categories contained instances of failure to follow and maintain processes as well as poor documentation practises. The premarket side of the pharmaceutical sector continues to experience severe data integrity and other regulatory compliance challenges, despite the fact that the number of warning letters has reduced over the past decade and inspection outcomes have been improving. [37].

Table 2and Figure3unveils Data Integrity Associated Warning Letters by different countries for the calendar year 2008 to 2018, along with a cumulative total[38].The pie graph Figure 3 shows that since 2008, the number of nations receiving warning letters about data integrity has increased, with over 80% of those warning letters occurring in the last four calendar years. The warning letter-related sites in 2018 were located in 11 different nations. The number reached its pinnacle in the calendar year 2017, and it will be remarkable to observe if it declines once more in 2019 like it did in 2018. [38].

Table 2: Data Integrity Related Warning Letters by Different Countries

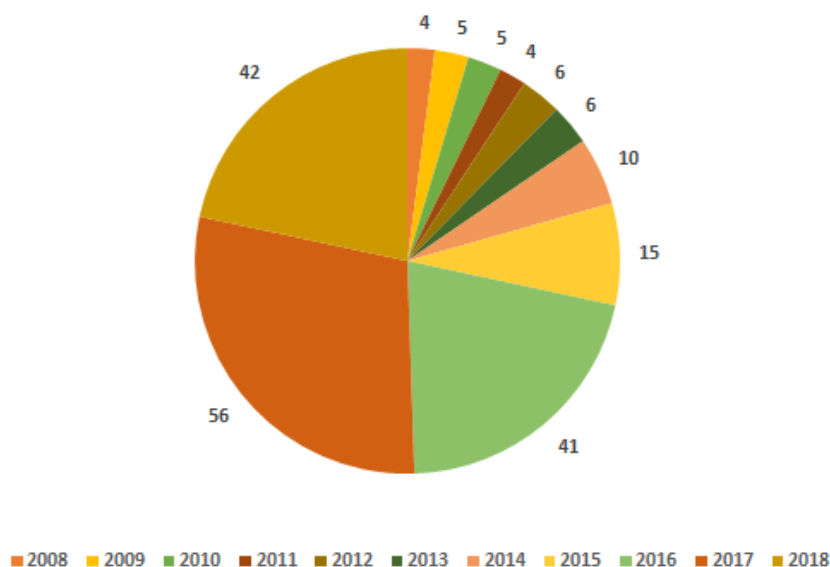|	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total
China	1	1	3	1			2	2	14	19	15	58
India	1	1		2		6	7	10	9	12	6	54
US	1	2	1	1	1				7	15	8	36
Europe		1					1	2	6	3	1	14
Japan	1								2	1	3	7
South Korea										2	4	6
Canada			1	1						2	1	5
Mexico					2					1	1	4
Brazil									3			3
Thailand							1					1
UAE					1							1
Jamaica					1							1
Singapore										1		1
Australia											1	1
Taiwan											1	1
Dominican Republic											1	1
Total	4	5	5	4	6	6	10	15	41	56	42	194

Figure 3: Pie Chart for Data Integrity Related Warning Letters of Different Countries

Table 3displays a clear picture of major of countries involved in data integrity warning letters viz. CHINA, INDIA and UNITED STATES. FDA did not classify numerous violations as "conclusions" or "data integrity remediation" that required a response from the pharmaceutical industry. Warning letters sent to API manufacturers do not mention 21 CFR 211. The FDA agency stated objective of concentrating on the assessment of predicate rule requirements is still followed by the citation of regulations. [38]

Table 3: Data Integrity Related Warning Letters with Reference to Geographical Locations

Country	Total Number, 2008-2018	% of Total, 2008-2018
China	58	30%
India	54	28%
United States	36	19%
Europe	14	7%
Rest of the World	32	16%

Poor documentation practises and failure to follow and maintain processes were the most prevalent violations identified across all warning letter categories Table 4. Approximately 50% of all global drugs 483s that were issued between 2014 and 2018 were reported due to data integrity issues. The Big Data and AI Analytics company Govzilla discovered this information. Data integrity issues were identified in 79% of API warning letters sent out globally during this time. Defects in data integrity have been seen to significantly increase in recent years. [39]

Table 4: Most Recurrent Data Integrity Related Warning Letters

21 CFR Reference	Frequency of Citations	Caption of CFR Section
211.194	10	Review of entire data, Lab Records
211.188	6	Batch Manufacturing and control records
211.165(a) and (b)	5	Testing and release for distribution
211.192	5	Deviations, Risk managements, Investigations, Production records review
211.68	2	Automatic, Mechanical and Electronic Equipment

For violations of 21 CFR Part 11, very few form 483s and warning letters were issued. The majority of the flaws were attributed to failure to follow 21 CFR Parts 210, 211, and 212's cGMP regulations.

2. **Two Major Violations were Cited for CFR Code 211.68 and 211.194.**

- Part 211.68 outlines the specifications for "Automatic, Mechanical and Electronic Equipment." Frequently used references in this field include:

  ➢ Mismatch of the printed data with actual audit trail.
  ➢ Data was not legible and correctly backed up for future reference.

- Part 211.194 is referenced when industries failed to study and consider all relevant facts when making lot release decisions. Frequently used citations in this field include:

  ➢ For meeting the predefined acceptance criteria laboratory analysts delete or manipulate data.
  ➢ Identification for out of specification results becomes difficult for investigations since companies fail to review critical data and/or metadata
  ➢ The test results of analysis are either deleted or corrupted by company which do not have sufficient supportive data.

3. **Other Most Recent Inadequacies Include:**

- Making use of 'pre-injections' of product samples outside of whole sample sets to check whether findings meet acceptance requirements and, if not, deleting/ignore results.
- turning off audit trails to hide results
- Deletions or alterations of results
- Data reduction through the use of integration suppression settings that would probably result in an OOS
- No proper justification for Aborting test runs

The FDA has promoted the use of "independent" data integrity evaluations as part of the process for resolving detected concerns. Several instances include:

- On August 10th, 2018, Kyowa Hakko Bio Co., Ltd.'s manufacturing facility in Japan received a warning notice that stated the following: "We recommend that a qualified third party with specific expertise in the area where potential breaches were identified should evaluate all data integrity lapses."
- On March 26, 2019, a warning letter was issued to the Georgia-based manufacturing facility of Winder Laboratories, LLC. Stating the following: "Your quality system does not adequately ensure the accuracy and integrity of data to support the safety, effectiveness, and quality of the drugs you manufacture…. We strongly recommend that you retain a qualified consultant to assist in your remediation."
- On June 13th, 2019, Akorn Inc.'s Illinois manufacturing site received a warning notice that stated: "Your quality system does not adequately ensure the accuracy and integrity of data to support the safety, effectiveness, and quality of the drugs you manufacture…. We acknowledge that you are using a consultant to audit your operation and assist in meeting FDA requirements."

4. Benefits Data Integrity Assessment:

- Builds up focus of Organization on Data Integrity issues– Braces the fact that all employees are focused on compliance of data integrity concerns.
- Satisfaction of identifying data integrity concerns – Once the data integrity issues have been identified it becomes easy to work on compliance.
- Expenditure reductions and Time saving – Data integrity issues identified and rectified internally are much more significant than identified by regulatory authorities and are also less time consuming for redrafting.
- Can concentrate on enhancing Business – Can stay focused on well-being of business in effective way rather than spending time and expenditure on data integrity issues if solved proactively.[39]

Regulatory organisations like the WHO and FDA have published guidance on appropriate documentation practises and how data should be collected and handled as a result of an increase in data integrity breaches during 2019. Draft document of guideline on data integritywas published by WHO in October 2019.

IX. HIGHLIGHTS OF WHO GUIDELINE ON DATA INTEGRITY[40]

1. Guideline covers basic concepts, information and recommendations for promoting GxP in documentation and record keeping which will facilitate compliance with Data Integrity.
2. This guideline does not cover Data integrity on medical devices, however designated as 'GxP'.
3. this guideline should be studied in conjunction with WHO GxP guidelines and publication, given that it has been harmonised with other published materials.
4. In order to identify and evaluate risk areas, data integrity risk assessment should be carried out in accordance with current GMP practises that provide a risk-based approach over the life cycle of data.

5. Both contract givers and contract acceptors must abide by this rule. Since contract acceptors are the ones who give them the data, it is the contract giver's responsibility to guarantee that the contract acceptor conforms with the principles emphasised in this guidance.
6. The effectiveness of controls can be verified by identification and review of data with risk-based controls.

X. CONCLUSION

In pharmaceutical industry data integrity plays an important role in regard of maintaining quality product because improper practice allows inferior quality of product to reach patients, so it becomes mandatory responsibility of pharmaceutical industry to ensure efficacy, safety and quality of products. Data Integrity is the most common issue with majority of pharmaceutical industries. Concept of ALCOA+ helps maintain data integrity necessary for following Good Manufacturing Practices. Now a day pharmaceutical industryhas started relying on digital and automated systems, hence it becomes necessary to validate them for minimizing errors related to addition or deletion of data. It eliminates the need to check each and every step of the manufacturing and distribution of pharmaceuticals..Following the ALCOA+ principles is the best way to achieve this goal. Numerous regulatory bodies suggest of using ALCOA as tool to reduce errors and ensure GDP. Warning letters and closeout letters issued by the Food and Drug Administration's Bioresearch Monitoring (BIMO) Program, Big Data and AI Analytics firm Govzilla and article by Barbara Unger suggested most of the form 483 citations were due to data integrity issue. An inflation in multiple numbers of data integrity related violations during recent years roused a vital requisite to publish guidelines on good documentation practices by regulatory bodies.

XI. ACKNOWLEDGEMENT

XII. ABBREVIATIONS

1 FDA- Food and Drug Administration
2 cGMP- Current Good Manufacturing Practices
3 MHRA- Medicines and Healthcare Products Regulatory Agency
4 WHO- World Health Organisation
5 PIC/S- Pharmaceutical Inspection Convention and Pharmaceutical Inspection Co-operation Scheme
6 GDP- Good Document Practices
7 CFR- Code of Federal Regulations
8 KB- Kilobyte
9 ID- Identity
10 ERP- Enterprise Resource Planning
11 BMR- Batch Manufacturing Record

12 GxP- Good Practices. (x-Variables)
13 SOP- Standard Operating Procedures
14 OOS- Out of Specification
15 URS-User Requirement Specification
16 DQ- Design Qualification
17 OQ- Operational Qualification
18 PQ- Performance Qualification
19 IQ- Installation Qualification
20 FAT-Factory Acceptance Testing

REFERENCES

[1] Data Integrity and Compliance with Drug CGMP Questions and Answers Guidance for Industry, December 2018. Available at: https://www.fda.gov/media/119267/download

[2] S. Schmitt, Assuring Data Integrity for Life Sciences., River Grove, IL: DHI Publishing, LLC. 2016. Available at: https://store.pda.org/TableOfContents/18016_TOC.pdf

[3] GAMP Guide Records and Data integrity. 2017, Tampa, FL: International Society for Pharmaceutical Engineering. Available at:https://ispe.org/product-types/gamp-good-practice-guides

[4] GAMP Good Practice Guide: Data Integrity - Key Concepts. 2018, International Society for Pharmaceutical Engineering: Tampa, FL. Available at: https://ispe.org/product-types/gamp-good-practice-guides

[5] Technical Report 80: Data Integrity Management System for Pharmaceutical Laboratories. 2018, Parenteral Drug Association (PDA): Bethesda, MD20814 USA. Available at: http://15476078.s21d-15.faiusrd.com/61/1/ABUIABA9GAAgtars8wUo_Ivc5gE.pdf

[6] R. D. McDowall. Data Integrity and Data Governance: Practical Implementation in Regulated Laboratories. 2019, Cambridge: Royal Society of Chemistry. November 2018

[7] Saket Yeotikar. Important Aspects of Good Documentation Practices (GDP) and 9 Principles of ALCOA+, Quality Management. November 28, 2020. Available at: https://pharmagxp.com/quality-management/good-documentation-practices-gdp/

[8] R. D. McDowall. Data Integrity: What Are Some of the Key Issues? Jan 17, 2019 Available at: https://www.technologynetworks.com/informatics/articles/data-integrity-what-are-some-of-the-key-issues 314165#:~:text=Data%20Integrity%3A%20What%20Are%20Some%20of%20the%20Key,without%20contributing%20to%20product%20quality%20and%2For%20data%20quality

[9] CAPT Sharon K. Pederson. Data Integrity Issues & Concerns PDA Meeting National Expert of Pharmaceutical Inspections Food and Drug Administration Medical Products and Tobacco Program Operations Branch. St. Louis, MO February 06, 2017 Available at: https://www.pda.org/docs/default-source/website-document-library/chapters/presentations/missouri-valley/data-integrity-issues-concerns.pdf?sfvrsn=406c808e_4

[10] Barbara Killinger. Integrity: Doing the Right Thing for the Right Reason. Mcgill-Queen's University Press (2007) Available at: Barbara Killinger, Integrity: Doing the Right Thing for the Right Reason - PhilPapers

[11] The True Meaning of Integrity 08/01/2016, ET Updated Dec 06, 2017 Available at: https://www.huffpost.com/entry/the-true-meaning-of-integ_b_11273420

[12] Available at: https://wikimili.com/en/Metadata#On_the_Internet

[13] Available at: https://en.wikipedia.org/wiki/Raw_data

[14] Annex 5 Guidance on good data and record management practices WHO Technical Report Series No.996, 2016 Available at: https://www.who.int/medicines/publications/pharmprep/WHO_TRS_996_annex05.pdf

[15] Available at: https://en.wikipedia.org/wiki/Audit_tra

[16] Good Documentation Practice (GDP) Guideline. Indian Pharmaceutical Alliance. February 2018. Available at: https://tabraizullah.files.wordpress.com/2019/03/good-documentation-practices.pdf

[17] Kumar K. Good Documentation Practices (GDPs) in Pharmaceutical Industry. J Anal Pharm, March 2017, 4(2): 00100. Available at: http://medcraveonline.com/JAPLR/JAPLR-04-00100.pdf

[18] Tim Sandle.Good Documentation Practices. Journal of Validation Technology. 20(3) Available at: https://www.ivtnetwork.com/sites/default/files/2_Good_Documentation_Practices_0.pdf

[19] What is ALCOA +(Plus)/ ALCOA+ Details / ALCOA+ Principle, August 10, 2020 Available at: https://www.pharmaguidesop.com/2020/08/what-is-alcoa-plus.html

[20] Ankur Choudhary. Principals of Data Integrity. Pharmaceutical Guidelines. 2018Available at:https://www.pharmaguideline.com/2018/04/principles-of-data-integrity.html

[21] S.W. Woollen. "Data Quality and the Origin of ALCOA". Newsletter of the Southern Regional Chapter Society or Quality Assurance, Summer 2010, Apr. 9, 2014. Available at: http://www.southernsqa.org/newsletters/Summer10.DataQuality.pdf

[22] Mark Durivage. ALCOA Principles: Data Integrity for the FDA Regulated Industry. Managing Principal Consultant and Member. Quality System Compliance, LLC. Rev 1.0 1/12/2019. Available at: https://qscompliance.com/wp-content/uploads/2019/01/ALCOA-Principles.pdf .

[23] ALCOA Principles. Quality System Compliance, LLC. January 12, 2019 Available at: https://qscompliance.com/alcoa-principles/

[24] Abhishek Ghosh. Data Integrity vs Data Consistency. January 9, 2020, Updated on January 9, 2020. Available at: https://thecustomizewindows.com/2020/01/data-integrity-vs-data-consistency/

[25] Slcontrols. What Is ALCOA+ And Why Is It Important To Validation And Data Integrity. February 27, 2019. Available at: https://slcontrols.com/what-is-alcoa-and-why-is-it-important-to-validation-and-data-integrity/

[26] Aukur Choudhary. ALCOA to ALCOA Plus for data integrity. Pharmaceutical Guidelines. Available at: https://www.pharmaguideline.com/2018/12/alcoa-to-alcoa-plus-for-data-integrity.html

[27] Anil K. Rattan. Data Integrity: History, Issues, and Remediation of Issues. PDA J Pharm Sci and Tech 2018, 72 105-116. Available at: https://journal.pda.org/content/72/2/105

[28] Shmmon Ahmed, Ashok Kumar, Dr. Abdul Hafeez. Importance of data integrity & its regulation in pharmaceutical industry. The Pharma Innovation Journal 2019; 8(1): 306-313. Available at: https://www.researchgate.net/publication/334131228_Importance_of_data_integrity_its_regulation_in_pharmaceutical_industry

[29] Ankur Choudhary, FDA's Data Integrity Issues in Indian Facilities. Pharmaceutical Guidelines. Available at: https://www.pharmaguideline.com/2015/02/FDA-data-integrity-issues-during-inspecti

[30] Mike Malz.What Is Data Integrity? Importance and Best Practices to follow for Data Integrity.Global Vision Blog. December 1, 2016. Available at: https://blog.globalvision.co/quality/12-ways-to-reduce-data-integrity-risk/

[31] Part 211 - Current Good Manufacturing Practice for Finished Pharmaceuticals. Federal Register, 1978. 43(190): p. 45014 - 45089.

[32] Medicines & Healthcare products Regulatory Agency (MHRA) 'GXP' Data Integrity Guidance and Definitions; Revision 1: March 2018 Page 1 of 21

[33] 21 CFR 11: Electronic Records; Electronic Signatures

[34] 21 CFR 211: Current Good Manufacturing Practice for Finished Pharmaceuticals

[35] 21 CFR 820: Quality System Regulation

[36] Sharon K. P. (August 2017). Data Integrity Issues & Concerns. Parenteral Drug Association (PDA). Available at: https://www.pda.org/docs/default-source/website-document-library/chapters/presentations/missouri-valley/data-integrity-issues-concernse703e9d7a8c4657391feff0000cd242a.pdf?sfvrsn=4

[37] Clinton A. R., Jennifer A. D., and Michael G. B. (2020). Data Integrity in the Pharmaceutical Industry: Analysis of Inspections and Warning Letters Issued by the Bioresearch Monitoring Program Between Fiscal. (2018). Therapeutic Innovation & Regulatory Science.54, 1123–1133; Available at: https://doi.org/10.1007/s43441-020-00129-z

[38] Barbara Unger, Unger Consulting Inc. An Analysis Of 2018 FDA Warning Letters Citing Data Integrity Failures. Guest Column. Pharmaceutical online. June 12; 2019.Available at: https://www.pharmaceuticalonline.com/doc/an-analysis-of-fda-warning-letters-citing-data-integrity-failures-0001

[39] Dale Curtis, Trends in FDA Data Integrity 483s and Warning Letters for Pharmaceutical Companies, Posted on Lab Compliance. Astrix Blog 16 February, 2020. Available at: https://astrixinc.com/trends-in-fda-data-integrity-483s-and-warning-letters-for-pharmaceutical-companies/

[40] World Health Organisation. (2019). Guideline on Data Integrity. Available at: https://www.who.int/medicines/areas/quality_safety/quality_assurance/QAS19_819_data_integrity.pdf?u