DATA INTEGRITY: AN OVERVIEW

Nikita R. Nikam1*, Priyanka R. Patil2, Mr. Rohan R.Vakhariya3, Dr. S. K. Mohite4 and Dr. C. S. Magdum5

1,2 Department of Pharmaceutics, Rajarambapu College of Pharmacy, Kasegaon, Tal. Walwa, Dist. Sangli, Maharashtra, India. 415404
3 Department of Quality Assurance, Rajarambapu College of Pharmacy, Kasegaon, Tal. Walwa, Dist. Sangli, Maharashtra, India. 415404
4,5 Department of Pharmaceutical Chemistry, Rajarambapu College of Pharmacy, Kasegaon, Tal. Walwa, Dist. Sangli, Maharashtra, India. 415404

DOI: http://dx.doi.org/10.24327/ijrsr.2020.1106.5372

ABSTRACT

Data integrity is essential to regulatory enforcement and the fundamental justification for the US-published 21 CFR Part 11. Food and Drug Administration (FDA) Regulators wanted to make sure that industry collects accurate data during the drug production lifecycle and through marketing consider the number of warning letters recently issued on data integrity by inspectors around the world. FDA uses the acronym ALCOA to describe their electronic data standards. ALCOA stands for Original, Attributable, Legible, Contemporary and Correct. This article addresses the concepts, relevance, benefits, drawbacks, regulatory guidelines and data integrity forms. It also provides details about the risk of data privacy and different approaches can mitigate risk. The article is going to be very much useful for industry to maintain all data accurate and safe.

INTRODUCTION

Data can be presented by manually recording an observation, results or other data and information on paper or by using electronic records using equipment and equipment connected to a computerized system. A combination of manual and electronic systems can also be used.1

The same as DI copies other data sets (such as photos, videos, 400 DVDs, images and chromatography plates), with all the additional controls necessary for formats such as photography or scanning 402. There should be a rationale for documenting the choice of such a method.1

Data integrity is a guarantee of the accuracy and consistency of data throughout its life 2 and is an important aspect of the design, implementation and use of any system that stores, processes, or retrieves data. The term has a broad scope and can have different meanings under the same umbrella of a computer, depending on the specific context. It is used as a proxy term for data quality 3, while data validation is a prerequisite for data integrity.4 Data integrity is against data corruption.5

Principle of data integrity

Systems will be set up and enforced to ensure that all data collected, processed and recorded are in compliance with guideline principles. Data should be provided: 1

A = assignable to the person who generates the data
L = legible and persistent
C = contemporary
O = original record (or true copy certified)
A = accurate

ALCOA

Attributable

The identity of the eligible record holder must be documented. This is usually done by personal signature for paper records and dating records with their signatures. Since the file you sign may be a legal document, the meaning of your signature must be clearly understood. The signature must be independent of a
particular person and the tendency to sign the name or initials of another is fraudulent and must be taken very seriously.  

**Legible**

Recordings that cannot be read or understood have no value and may not exist. All records must be created to be consistent with the grammatical assemblies which must be perfectly consistent. It is good to avoid buzzwords, groups and abusers, as these are likely to change over time and are often not understood in a particular locality. It is always a good idea to look at anyone else's case as it can highlight most of the ambiguities.  

**Contemporaneous**

All records must be created at the time an activity occurs. For example, delaying writing to the end will inevitably affect the accuracy of this recording, as the details may not be forgotten or remembered.  

**Original**

All records must be original; the information must be recorded directly on the document. This avoids the possibility of detecting errors when transcribing information into documents. If information is printed by the instrument, the print is the original recording and must be signed, dated and attached to the recording.  

**Accurate**

The record should reflect what actually happened. Any modification must be carried out without masking or destroying the original information, the use of blanks or modifications of liquids is prohibited. Any modification made to the file must be signed by the person making the modification and dated to indicate when it occurred and a written explanation must also be provided. Be aware that records may be required after you leave the company and may not be contacted for clarification.  

**Importance of Data Integrity**

The importance of accurate and reliable data while ensuring the safety and quality of drugs. As the regulator has emphasized the integrity of the data for many years, the FDA and other global regulators emphasize.  

**Advantages**

**Control of data redundancy**

The database system tries to eliminate the extremes by integrating files. Although redundancy is not completely eliminated by the database system, it does control the amount of redundancy in the database.  

**Data consistency**

The Database approach reduces the risk of inconsistency by removing or managing redundancy. It ensures consistent maintaining of all copies of the data.  

**More information from the same amount of data**

Thanks to the integration of the data used in the database system, it is possible to obtain additional information for the same data.  

**Sharing of data**

The database is a part of the enterprise as a whole and can be accessed by all registered users.  

**Improved data integrity**

The integrity of the database ensures the validity and consistency of the data stored. Integrity is often expressed in terms of restrictions, which is the consistency rule prohibiting the violation of a database.  

**Improved maintenance**

The database system provides data freedom. Changes to the data structure in the database will affect the application program, which will facilitate maintenance of the database application.  

**Increased concurrency**

The database can effectively manage interactive access to data. This guarantees no user interference so that no information is lost or loss of integrity.  

**Disadvantages**

**Complexity**

Database management is very complex system software. All parties must be aware of its effectiveness and take full advantage of it. Therefore, training is essential for administrators, designers and users.  

**Size**

Database management systems consume large amounts of main memory as well as large amounts of disk space to operate efficiently.  

**Performance**

The database method is used to provide many applications rather than specific applications, some applications may not run as quickly as before.  

**Higher impact of a failure**

Centralizing the database approach increases the vulnerability of the system. When you respond to all users and the availability of the application database, a component failure can hamper operations and seriously affect customer service.  

**Cost of conversion**

When switching from a file system to a database system, the enterprise must incur additional costs for the acquisition of equipment and training costs.  

**World Regulatory Guidance on Data Integrity**

**USFDA-21-CFR**

21-CFR (Code of Federal Regulation) is a coding of the common and permanent rules published in the Federal Register by the federal government's executive department and agency. The title of CFR 21 is reserved for Food and Drug Administration regulations. Each CFR title / volume is updated once per calendar year on April 1 of each year.
MHRA

The guidance on the topic of data integrity expectations for the pharmaceutical industry is complementary to existing EU GMPs for active substances and dosage forms. In the pharmaceutical quality system, data integrity is central to ensuring that medicines are of the required quality.11,13

Difference between data security and data integrity:

<table>
<thead>
<tr>
<th>Table 1 Difference between data security and data integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Security</strong></td>
</tr>
<tr>
<td>Data security defines the prevention of data corruption</td>
</tr>
<tr>
<td>through the use of controlled access mechanisms.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Data security deals with the protection of data.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Data security is making sure only the people who have</td>
</tr>
<tr>
<td>should have access to the data are the only ones who can</td>
</tr>
<tr>
<td>access the data.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Authentication/authorization, encryption and masking are</td>
</tr>
<tr>
<td>some of the popular means of data security.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

For example, if you have an account in the “yahoo.com”, then you have to give your correct username and password to access your account or email. Similarly, when you insert your ATM card into the Automated Teller Machine (ATM), the machine reads your ID number printed on the card and then asks you to enter your pin code (or password). In this way, you can access your account.

For example, a balance for any account must not be less than zero. Such constraints are enforced in the system by adding appropriate code in application programs. But, when new constraints are added, such as balance should not be less than Rs. 5000, application programs need to be changed. But, it is not an easy task to change programs whenever required.

TGA

The Australian Regulatory Body Therapeutic Goods Administration (TGA) does not require data integrity in the form of scarcity. Lack of practice or process that produced could result in a significant risk of delivering a product harmful to the user. In addition, when it appears that the manufacturer misleads, distorts or improperly engages in products or data.1,14

cGMP

Reflecting on the importance of this issue, the FDA has issued guidelines on data integrity and GMP compliance itself. The FDA reports trends that increase data integrity violations. Archiving practices in accordance with the cGMP prevent the loss or obfuscation of data. The FDA's authority for cGMP stems from section 1 of the FD&C Act, "If the methods used, or the facilities or controls used for this, its production, processing, packaging, or possession or operation or its administration are considered falsified". According to practice, it is guaranteed that such a drug will require security Is-lakhali and their identity and strength and meets the characteristics of quality and purity.1,14

WHO

Essential drugs and health products, WHO is launching data integrity guidelines to protect patients around the world. It involves a number of artists and activities, a fundamental step added to the persistence and accuracy of the data submitted by manufacturers to the National Regulatory Authority. These data must be complete, complete and accurate and truthful in order to determine the quality of the studies that support drug applications in the market. In addition, many standards must be followed, namely: good manufacturing practices (GMP), good clinical practices (GCP) and good laboratory practices (GLP).11,16

EMA

The European Medicines Agency (EMA) has announced new Good Product Practices (GMP) guidelines to ensure the integrity of data produced during testing, production, packaging, distribution and medication maintenance. Good control of data records is precise and consistent with the data generated and will help make good decisions by pharmaceutical manufacturers and regulatory authorities.11,17

Types of data integrity

There are two types of integrity of the data: physical and logical integrity. Both are a collection of processes and methods that implement data integrity in hierarchical and relational databases.

Physical integrity

Physical integration is the preservation of the completeness and accuracy of data during archiving and retrieval. Physical integrity is compromised when natural disasters end, power outages or hackers disrupt database operations. Due to human errors, storage erosion and many other issues, it is impossible for data controllers, system programmers, seas application programmers and internal auditors to obtain accurate data.18,19

Logical integrity

The data does not change due to the logical integrity used differently in the relational database. Logical integrity data protects against human error and hackers, but in a different way from physical integrity. There are four types of logical integrity.

Entity integrity

Entity integrity depends on creating unique values that identify the primary key or data elements to ensure that the data entry is not listed more than once and that the fields in a table are not empty. It is a characteristic of relational systems that stores data in arrays that can be linked and used in different ways.

Referential integrity

Referential integrity refers to a series of processes that ensure that data is stored and used consistently. The rules built into the structure of the foreign key usage database ensure that only appropriate changes, additions, or deletions are found. Rules may include the deletion of duplicate data entries, ensuring the accuracy of the data, and the termination of access to unapproved data.

Domain integrity

Domain integrity is a set of processes that ensure the accuracy of each piece of data in the domain. In this context, a domain is
a set of acceptable values which can be included in a column. These may include constraints and other measures that limit the nature, type and extent of the data entered.

**User-defined integrity**

User-defined integrity includes rules and restrictions that the user has created to meet their specific needs. Sometimes the integrity of the entity, context and domain is not enough to protect the data. Specific business principles are also to be taken into account and used in data integrity initiatives.18, 19

**Data integrity risks**

There is an assortment of factors that can affect the integrity of the data stored in a database. A few examples include:

- **Human error**: Data integrity is compromised when a person enters incorrect information, duplicates or deletes data, does not follow the correct protocol, or makes mistakes while executing the process for information security purposes.

- **Transfer errors**: A transfer error has occurred when data cannot successfully transfer from one location in a database to another. Transfer errors occur when a piece of data is present in the destination table but not in a relational database's source table.

- **Bugs and viruses**: Spyware, malware, and viruses are software pieces that could invade a computer and change, erase, or steal data.

- **Compromised hardware**: Sudden computer or server crashes, and problems with how a computer or other system operates, are examples of significant failures and can suggest that the hardware is being compromised. Compromised hardware may incorrectly or incompletely render data, restrict or disable access to data, or make information difficult to use.18

**Risks to data integrity can easily be minimized or eliminated by doing the following**

**Protect Your Documents**

By applying constraints to the files you will secure the records. Vendors like Adobe and Microsoft have the ability to limit behavior that can be taken on documents-from read-only to making only certain changes. Also, you can encrypt your documents and PDFs by password.

**Introduce Access Controls**

Without official access and with malicious intent, personal can cause serious harm to data. Implementing a minimum privilege model, where access is accessible only to users who need access to the data, is a very popular form of access control.

**Implement Backup and Recovery Procedures**

Regular backups of data are critical to preserving data integrity and avoiding valuable loss of data. Data which is periodically backed up can be restored in its original form in case of a breach.

**Leverage Audit Trails**

A time-stamped, computer-generated audit trail tracks the name, date and time of data entries, modifications and deletions. In case data is compromised, audit trails provide the breadcrumbs which lead to the problem source.

**Passwords**

Use passwords to prevent unauthorized access to all storage devices where data is stored. Do not place the passwords on your computer or post-it notes.

**Ensure Security**

Cyber-security measures like file encryption can also help to reduce the loss of data integrity. Anti-malware software is also recommended to block malicious attacks and files from accessing your computer and manipulating your data.20

**Data integrity as applied to various industries**

The U.S. Food and Drug Administration have created draft guidelines for the pharmaceutical manufacturers required to adhere to U.S. data integrity. Code of Federal Regulations 21 CFR Parts 210–212.21 Outside the United States, the United Kingdom (2015), Switzerland (2016) and Australia (2017) have provided similar guidelines on data integrity.22

Various standards for medical device manufacturing address data integrity either directly or indirectly, including ISO 13485, ISO 14155 and ISO 5840.23

In early 2017, the Financial Industry Regulatory Authority (FINRA), noting data integrity problems with electronic trading and money movement surveillance systems, announced that priority should be given to 'developing a data integrity program to track the quality of the data submitted'.24 In early 2018, FINRA said it would expand its data integrity policy to the "information change management policies and procedures" and the assessments of Treasury securities.25, 26, 27

Cloud storage companies have long faced significant challenges in ensuring accuracy or provenance of customer data and in monitoring violations.28

**Impact of data integrity on pharmaceutical industry**

Data integrity in the pharmaceutical industry includes correct data generation and recording, protection of data against accidental or deliberate changes, deletion or loss of data by falsification. The degree of impact from faulty data may vary based on the discipline, and therefore the nature of the investigation, ensuring accurate and truthful data collection is essential to maintain research integrity. Improperly collected or incorrect data can affect the ability of the researcher to answer questions correctly and repeat or confirm studies, confuse other researchers and influence decision-making and lead to wasted resources etc. Data integrity is essential in all areas of academic and industrial science research; however, it is even more critical in the healthcare sector.29

When these data are used to support research, development, registration and advertising of pharmaceutical products for human and animal health use, there is potential to cause harm to patients. This is due to the fact that the compromised data will influence the quality of the final product, i.e. medicine, because the quality of the raw materials in process materials and the finished goods is assured by laboratory data. Later, products of low quality can be unsafe for patients.30
The Quality Assurance (QA) and Quality Control (QC) methods will assure data integrity. These approaches are applied at different studies levels. For example, prior to data collection, the QA activities take place by establishing quality systems, optimizing protocols, Formal Operating Procedures (SOPs), and training personnel involved in data collection, which avoid any data integrity issues prior to data collection. On the other hand, after data collection, the QC activities, monitoring, analysis and corrective action relating to the data take place. Depending on appropriate internal protocols and SOPs the data will be checked for unintentional or deliberate data problems in a timely manner by the correct subordinates, line managers or research directors. A clear communication procedure between the workers and the supervisor or the research director would reduce the risk of incorrect data generation.31

Data integrity issues

Warning Letters, Statement of Non-Compliance and Consent Decrees

The regulatory authorities have issued several warning letters, notification of non-compliance and consent decrees to pharmaceutical manufacturing facilities after identifying data integrity issues. If the regulatory authorities take these types of action, it will affect the company's ability to get approval of new drug product for sale, loss of trust by regulatory authorities. Additionally, a condition may occur in which the company needs to minimize the production or keep the goods on site. This will lead to a lack of pharmaceutical products and a loss of consumer trust.32

Import Alert, Product Recalls and Seizure of Products

Drug products which have issues with data integrity is called adulterated drug products. US FDA will restrict them from being allowed on the US market for these adulterated drug products. Import alerts inform FDA field staff and the public that the agency has ample evidence to require Products that appear to be in violation of FDA laws and regulations to be detained without Physical Examination. These violations may concern the product, distributor, shipper and/or other details.33

Once a drug is approved, the approval committees within the agency will review the data, and may require additional studies, audits, or inspections to ensure the integrity of the data.34

When an FDA-regulated drug product is either defective or potentially harmful, the most successful way of protecting the consumer is to note that withdrawing it from the market or fixing the issue.

Loss of Regulatory Trust

When issues surrounding data integrity arise, they are likely to result in loss of regulatory trust. This will result in further frequent inspections of the plant, hoping to see more evidence to support claims, and making it unlikely that a company can get approval for the average issues they may wish to conduct.34

Need to Appoint Third Party Consultants for Data Integrity

If a warning letter has been given to the pharmaceutical facility by the US FDA, FDA suggests hiring a third-party specialist who is skilled in identifying data integrity problems to assist the company with this assessment and to assist with the overall compliance of the company with cGMP. Usually the process of identifying data integrity issues and meeting the regulatory requirement through a consultant is time-consuming and expensive too.

Debarment and Imprisonment (for Individuals involved in data integrity issue)

A case study understands well how data integrity issues can affect individuals involved in data integrity issues. A former Quality Control Department vice president and three supervisory chemists at the now-defunct generic drug manufacturer in New Jersey, Able Laboratories, pleaded guilty to a scheme involving systematic falsification and manipulation of their drug test results. Because of this all the 500 Able laboratories employees, Inc. had lost their job.

Reasons of Data Integrity Issues to Occur

There is a common misconception that data integrity problems arise only from intentional fraud actions. Deliberate fraud act is a single factor that may contribute to data integrity issues but there are several other elements that are equally responsible for data integrity issues. A few elements are mentioned below,35

- Time / Work Pressure
- Insufficient education and understanding
- Fear for Mistakes
- Performance pressure
- Am told by Leader / Manager to do the activity which is against cGMP procedures
- Reputation Company Culture or Accepted behavior
- Money
- And others.

CONCLUSION

The integrity of data ensures that electronic data is kept intact. Reports are just as good as the data on which they are based. The integrity of data can also apply to information outside of the computer world. Whether digital or written, maintaining data integrity forms the basis for successful business decision taking. As data becomes a commodity it needs to be at the forefront of the mind ensuring its consistency with minimal effort. After all, the more data you have at your disposal, the more that your business can grow. The quality of data produced by any controlled lab is a prime factor in determining the lab's reputation. Recall that inspections and audits can only review; finding one instance of falsification poses the question of how many more cases of non-compliance exist. Therefore ensuring data integrity is of great importance to every organization's analytical scientists, managers and quality assurance, because the implications of getting it wrong are very costly and it will take a long time to restore regulatory confidence.

References

1. World Health Organization, working document QAS/19.819, Guideline on Data Integrity, October 2019, 14-16.
34. https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm049070.htm.

How to cite this article:
DOI: http://dx.doi.org/10.24327/ijrsrc.2020.1106.5372

*******