Shyne Cindy Ronolo

Assuring Data Integrity towards Regulatory Compliance

A Study on Process Improvement in Data Integrity Compliance of Computerized Systems

Metropolia University of Applied Sciences

Master’s Degree

Degree Programme in Business Informatics

Master’s Thesis

23 May 2023

Abstract

Author:	Shyne Cindy Ronolo
Title:	Assuring Data Integrity towards Regulatory Compliance: A Study on Process Improvement in Data Integrity Compliance of Computerized Systems
Number of Pages:	86 pages
Date:	23 May 2023
Degree:	Master of Business Administration
Degree Programme:	Business Informatics
Instructor:	Zinaida Grabovskaia, PhL, Senior Lecturer

The objective of this thesis is to review and revise the verification process for data integrity compliance of computerized systems used in the case company to a more streamlined, efficient and risk-based approach. Data integrity compliance is a subset discipline supporting the GxP-relevant processes to assure the manufacturing of products — its quality, purity, and efficacy, and to eventually warrant the safety of the patients. Decisions are made based on relying on critical data and records therefore data, in its essence, is fundamental to be proven as credible and truthful. As digitalization increased in the pharmaceutical and life science industry, the need for data integrity controls inclusion also increased in configurations, requirements, and specifications. This study used applied action research as its research approach and relied on qualitative research methods in order to improve the verification process of the case company to a more streamlined, efficient, and risk-based approach.

Currently, the data integrity compliance process at the case company is already defined but may have the tendency to mislead the risk assessment when systems complexity is not accounted for, or difficulties in implementation due to ambiguities and missing acceptance criteria. Therefore, the current state analysis determined the focus areas for improvement based on the risks assessed at a system level down to the system functions such as its impact on product recall and lot traceability, regulatory records to be submitted, aiding the manufacturing process, and product labeling, etc. With this, an understanding of the current data integrity maturity level was a starting point for the proposed improvements.

The outcome of this thesis is a proposal on how to strengthen the verification process of data integrity compliance without compromising the quality. The data integrity compliance improvement proposal was formulated collaboratively with the participants, validated, and implemented with additional training throughout the organization to raise awareness of the significant role of each one in the culture and overall compliance to data integrity. There will always be an area for improvement particularly on this topic when more and more complex technologies will be introduced in the future.

Keywords Data Integrity Compliance, ALCOA, Systems Validation

Contents

Glossary

Archiving

PIC/S (2021) describes this as “Long term, permanent retention of completed data and relevant metadata in its final form for the purposes of reconstruction of the process or activity.” (PIC/S, 2021, p.61)

Audit Trail

PIC/S (2021) describes this as “GMP/GDP audit trails are metadata that are a record of GMP/GDP critical information (for example the creation, modification, or deletion of GMP/GDP relevant data), which permit the reconstruction of GMP/GDP activities.” (PIC/S, 2021, p.61)

Back-up

PIC/S (2021) describes this as “A copy of current (editable) data, metadata and system configuration settings (e.g. variable settings which relate to an analytical run) maintained for the purpose of disaster recovery.” (PIC/S, 2021, p.61)

Business Continuity Planning

ISPE GAMP 5 (2022) describes this as “Documented procedures that guide organizations to respond, recover, resume, and restore to a pre-defined level of operation following disruption.” (ISPE, 2022, p.393)

Computerised system

PIC/S (2021) describes this as “A system including the input of data, electronic processing and the output of information to be used either for reporting or automatic control.” (PIC/S, 2021, p.61)

Computerised System Validation

ISPE GAMP 5 (2022) describes this as “Achieving and maintaining compliance with applicable GxP regulations and fitness for intended use by:

- the adoption of principles, approaches, and life cycle activities within the framework of validation plans and reports
- the application of appropriate operational controls throughout the life of the system” (ISPE, 2022, p.393)

Data

PIC/S (2021) describes this as “Facts, figures and statistics collected together for reference or analysis.” (PIC/S, 2021, p.61)

Data Flow Map

PIC/S (2021) describes this as “A graphical representation of the "flow" of data through an information system.” (PIC/S, 2021, p.61)

Data Governance

PIC/S (2021) describes this as “The sum total of arrangements to ensure that data, irrespective of the format in which it is generated, recorded, processed, retained and used to ensure a complete, consistent and accurate record throughout the data lifecycle.” (PIC/S, 2021, p.61)

Data Integrity

PIC/S (2021) describes this as “The degree to which data are complete, consistent, accurate, trustworthy, reliable and that these characteristics of the data are maintained throughout the data life cycle. The data should be collected and maintained in a secure manner, so that they are attributable, legible, contemporaneously recorded, original (or a true copy) and accurate. Assuring data integrity requires appropriate quality and risk management systems, including adherence to sound scientific principles and good documentation practices. The data should comply with ALCOA+ principles.” (PIC/S, 2021, p.62)

Data Lifecycle

PIC/S (2021) describes this as “All phases in the life of the data (including raw data) from initial generation and recording through processing (including transformation or migration), use, data retention, archive / retrieval and destruction.” (PIC/S, 2021, p.62)

Data Quality

PIC/S (2021) describes this as “The assurance that data produced is exactly what was intended to be produced and fit for its intended purpose. This incorporates ALCOA + principles.” (PIC/S, 2021, p.62)

Data Ownership

PIC/S (2021) describes this as “The allocation of responsibilities for control of data to a specific process owner. Companies should implement systems to ensure that responsibilities for systems and their data are appropriately allocated and responsibilities undertaken.” (PIC/S, 2021, p.62)

Dynamic Record

PIC/S (2021) describes this as “Records, such as electronic records, that allow an interactive relationship between the user and the record content.” (PIC/S, 2021, p.62)

Electronic Record

ISPE GAMP 5 (2017) describes this as “Any combination of text, graphics, data, audio, pictorial, or other information representation in digital form that is created, modified, maintained, archived, retrieved or distributed by a computer system. (ISPE, 2017, p.144)

Electronic Signature

ISPE GAMP 5 (2017) describes this as “A computer data compilation of any symbol or series of symbols executed, adopted, or authorized by an individual to be the legally binding equivalent of the individual’s handwritten signature. (ISPE, 2017, p.144)

Exception Report

PIC/S (2021) describes this as “A validated search tool that identifies and documents predetermined ‘abnormal’ data or actions, which require further attention or investigation by the data reviewer.” (PIC/S, 2021, p.62)

GxP

Good “X” Practice where X is collective term for Manufacturing (M), Clinical (C), Laboratory (L), Documentation (Doc), Distribution (D), Quality (Q), Pharmacovigilance (V) etc. GxP compliance is meeting all pharmaceutical and associated life science regulatory requirements, such as those set forth in the US FD&C Act, US PHS Act, FDA regulations, EU Directives, or other applicable national legislation or regulations under which a company operates.

Good Documentation Practices (GDocP)

PIC/S (2021) describes this as “Those measures that collectively and individually ensure documentation, whether paper or electronic, meet data management and integrity principles, e.g. ALCOA+.” (PIC/S, 2021, p.62)

Hybrid Systems

PIC/S (2021) describes this as “A system for the management and control of data that typically consists of an electronic system generating electronic data, supplemented by a defined manual system that typically generate a paper-based record. The complete data set from a hybrid system therefore consists of both electronic and paper data together. Hybrid systems rely on the effective management of both sub-systems for correct operation.” (PIC/S, 2021, p.62)

Master Document

PIC/S (2021) describes this as “An original approved document from which controlled copies for distribution or use can be made.” (PIC/S, 2021, p.62)

Metadata

PIC/S (2021) describes this as “In-file data that describes the attributes of other data, and provides context and meaning. Typically, these are data that describe the structure, data elements, inter-relationships and other characteristics of data e.g. audit trails. Metadata also permit data to be attributable to an individual (or if automatically generated, to the original data source). Metadata form an integral part of the original record. Without the context provided by metadata the data has no meaning.” (PIC/S, 2021, p.63)

Primary Record

ISPE GAMP 5 (2017) describes this as “The record which takes primacy in cases where data that are collected and retained concurrently by more than one method fail to concur.” (ISPE, 2017, p.146)

Quality Unit

PIC/S (2021) describes this as “The department within the regulated entity responsible for oversight of quality including in particular the design, effective implementation, monitoring and maintenance of the Pharmaceutical Quality System.” (PIC/S, 2021, p.63)

Raw Data

PIC/S (2021) describes this as “Raw data is defined as the original record (data) which can be described as the first-capture of information, whether recorded on paper or electronically. Information that is originally captured in a dynamic state should remain available in that state.” (PIC/S, 2021, p.63)

Source Data
ISPE GAMP 5 (2017) describes this as “All information in original records and certified copies of original records of clinical findings, observations, or other activities (in a clinical investigation) used for the reconstruction and evaluation of the trail. Source data are contained in source documents (original records or certified copies).” (ISPE, 2017, p.147)

Static Record
PIC/S (2021) describes this as “A record format, such as a paper or electronic record, that is fixed and allows little or no interaction between the user and the record content.” (PIC/S, 2021, p.63)

Supply Chain
PIC/S (2021) describes this as “The sum total of arrangements between manufacturing sites, wholesale and distribution sites that ensure that the quality of medicines in ensured throughout production and distribution to the point of sale or use.” (PIC/S, 2021, p.63)

System Administrator
PIC/S (2021) describes this as “A person who manages the operation of a computerised system or particular electronic communication service.” (PIC/S, 2021, p.63)

List of Tables

List of Figures

1 Introduction

Data integrity issues have been an ongoing focal point for inspections to assess how a company puts emphasis on their processes in delivering quality products with accurate information to which patient safety heavily relies on. Consumers are expecting that approved medicines and therapeutic drugs that are available in the market can effectively treat diseases. Regulatory authorities, such as FDA, are tasked to challenge manufacturer’s capability to consistently produce safe, effective and quality drugs. Severe violations on data integrity can be grounds for rejection of Biologics License Application (BLA), warning letters that may eventually lead to complete forfeiture of the rights to manufacture and sell the product. Failure to implement sufficient controls, both system-dependent and those with human circumvention, post risks of intentional and widespread falsifications of data or inadequate access management of critical functions.

Data in itself is extensively used in several areas across different departments and managed in computerized systems, consequently, relaying accurate and complete information to one another is pivotal in order to proceed to another step or to fix an issue prior. Thus, in pharmaceutical and life science industry where the security of a human life is dependent on the efficacy of medicines and treatments to be consumed, it is crucial to assure data integrity compliance. Particularly when the pandemic highlighted the importance of data integrity where a reliable and stable supply chain is instrumental to maintain drug safety, efficacy and quality.

1.1 Business Context

FinVector is a fast-growing and internationally renowned biopharmaceutical company that develops and manufactures viral-based gene therapy products. As a pioneer in its operational field, FinVector has extensive experience of nearly 30 years in cGMP manufacturing. The company’s operations are centered in Kuopio, Finland, and it employs hundreds of professionals from nearly 40 different countries. FinVector works on cutting-edge biopharmaceuticals and has invested in building state-of-the-art development and GMP facilities for Viral Vector and Cell Therapy platforms including vaccines. FinVector is owned by Ferring Ventures, a subsidiary of Ferring Foundation B.V. FinVector is referred to as the Company throughout this document.

The Company's capabilities span from preclinical and clinical product development to process development, analytical development, GMP manufacture and aseptic fill and finish from pre-clinical through to commercial product supply. Recently, the Company acquired a huge milestone of obtaining the U.S. FDA approval for a new bladder cancer therapy.

1.2 Business Challenge, Objective and Outcome

The Company strives to minimize the regulatory risks for non-compliance, due to follow-up inspections, manufacturing expansion and shift to a more IT systems-dependent environment. For a company undergoing preparations for the transition to commercialization, the most sustainable approach is to focus on where the risks are and how these are reduced or managed. This raises a need to improve data integrity verification processes.

The Company currently follows an outdated data integrity assessment and ALCOA+ (Attributable, Legible, Contemporaneous/Complete, Original, Accurate) principle-based testing, driven with the requirement to have a documentation that may tend to obscure the goal of highlighting the risks of the computerized systems limitations and identifying the existing mitigating controls or needed future plans.

The Objective is to review and revise the verification process for data integrity compliance of computerized systems to a more streamlined, efficient and risk-based approach.

The Outcome is a proposal on how to strengthen the verification process of data integrity compliance without compromising the quality.

1.3 Thesis Outline

The scope of the thesis is the data integrity verification process improvement that must be implemented by the Quality Department with assistance of different Subject Matter Experts (SMEs) from different departments including IT. The current manufacturing facility is based in Kuopio, Finland, with the sister company, Ferring Pharmaceuticals based in Switzerland which sponsors the commercialization of the first-of-a-kind gene therapy medicine for Non-muscle-invasive Bladder Cancer (NMIBC) patients.

The study will be conducted from multiple interviews from the people involved in driving the current data integrity verification process, observations and probes of any gaps. Aligning these with regulatory guidelines and standards will be the baseline and starting point for improvements.

This Thesis is written in seven sections. Section 1 establishes the purpose of why the process improvement for data integrity compliance is the focus of study. Section 2 describes research methods and materials to be used to conduct a constructive way to analyze the problem and gather information to elaborate our focus for solutions. Section 3 reports on the results of the current state analysis. Section 4 explores literature and best practice on the topics of data integrity guidelines and standards. Section 5 presents the initial proposal on the process improvement of the data integrity compliance verification. Section 6 reports on the results of early testing and validation on the process improvement of the data integrity compliance verification. Section 7 concludes the thesis.

1.4 ALCOA+ Principle of Data Integrity

Data Integrity must be established to assure ALCOA + principles pertaining to Attributable, Legible, Contemporaneous, Original, Accurate, Complete, Consistent, Enduring and Available.

Attributable requires the user that created or performed a data-related task to be identified for reliable traceability.

Legible must build the document in a clear, human-readable and understandable form for visibility of where the data had gone through.

Contemporaneous should be established so the document is recorded at the same time as the activity has been made.

An **Original** record is needed to represent the exact data captured within the retention period free from replacement or deletion.

Accurate GxP data values must be recorded, calculated, analyzed and reported with precision and validity.

Completeness is maintaining the data features in full to contain the history of every update and amendment since creation and to prove that there’s no omission or deletion.

Consistent data handling shows that the defined changes are reflected in all other areas minimizing conflicts and confusion in the documents.

Enduring should conserve the readability and accessibility of data throughout the system lifecycle and beyond via backups and effective archival practices.

Available information must be accessible from a centralized and reliable source to promote alignment and transparency among different departments and teams.

These principles are good foundation tools to guide companies on what to follow, however, these are not separately defined and set up in computerized systems on a per principle basis, rather it is grouped per controls.

2 Method and Material

This section describes the research approach, research design, and data collection and analysis methods used in this Thesis.

2.1 Research Approach

Today, the coexistence of multiple research philosophies and methodologies abundantly provide researchers with many options to direct their academic endeavors. Applied sciences is an academic discipline aimed to achieve an application improvement in the business and management realm that rewards a pragmatic approach. Throughout this research, the data collection techniques are administered in multiple interviews with relevant individuals who are involved and have expertise on the topic. The data analysis procedures that are rooted in data categorizations, groupings and levels are the key approach for qualitative analysis using non-numerical data. The philosophical assumptions of the research is that it will seamlessly be included in the operational workplace environment such as continuous improvement initiatives, and where the active participation of the team will eventually contribute to a common and deeper understanding of the topic and to eventually land on the most realistic and practical means of the process improvements. (Saunders 2019.)

Some of the characteristics of qualitative research studies rely on how the participants interpreted the purpose, the contents, the interconnectivities of the data collected from different resources including the data analysis of the current situation and the best practice. This will theoretically create the conceptual framework for the improvements. It is important to hear each contribution as there are no metrics to quantitatively assess the data gathered, thus any ambiguities or unstructured ideas in the open discussions

Among the research strategies especially in the business field are Action research, Case Study research, Ethnography, Grounded theory and Narrative inquiry, etc. In Action research, the researcher strategize as this will be more helpful enhance collaboration with the team and come up with improvements that will make sense to almost everyone in the organization who have knowledge of the business processes and its current maturity.

In this Thesis, the research strategy is Applied action research (Kananen 2013) as the goal of the study is to solve a problem – an observed weaknesses in the process. Applied action research is more suitable for this topic in order to facilitate continuous improvements. The main concern is to provide practical results with improvement for the better so users can make the most benefit and rewarding experience out of the activities that they're complied to perform. This study should dive-deep into how the process came about, its foundation and basis, the practical experience of the relevant people, what works and what doesn't to identify where the gaps are as a starting point for the process improvements. Thus, the thesis is the field study of collecting and analyzing the current data to determine potential areas of improvement.

In this Thesis, the research methods and techniques to be used are interviews with key participants of the process, the SMEs and knowledgeable parties who are competent to contribute to the discussion of the as-is and to-be situations, and some elements of the quantitative calculations. The outcome is the development proposal for a change towards efficiency and streamlining the process so resources can allot more effort to value-adding tasks.

Applied action research is more suitable for the topic to analyze and evaluate what the existing information that the company have in order to facilitate continuous improvements. A mixed of qualitative and quantitative elements with more focus on the former will be used in this research to explore the assessment and verification process. This aims to not only refine the process but make it more understandable and useful in long-term. Decisions can tend to be misleading if the study will be solely based on quantitative research.

The research topic for data integrity compliance is heavily based on regulatory standards, guidelines, and framework. The guidelines and regulatory standard references will act as lighthouse to avoid any subjective explorations and misinterpretations. The guidelines and regulatory standard references will act as lighthouse to avoid any subjective explorations and misinterpretations, and these references are known and readily available within the company. As helpful as these references are, there are no specific methods mentioned suitable for different types of organization. The Company's maturity level should still be considered to weigh the most appropriate approach to achieve process improvements. supported with justification as to why it is more fitting to interpret the standards in such a manner.

2.2 Research Design

In order to systematically approach the study, the research design of the thesis starts with setting a thesis objective which is to review and revise the verification process for data integrity compliance.

Figure 1 below shows the research design of this study.

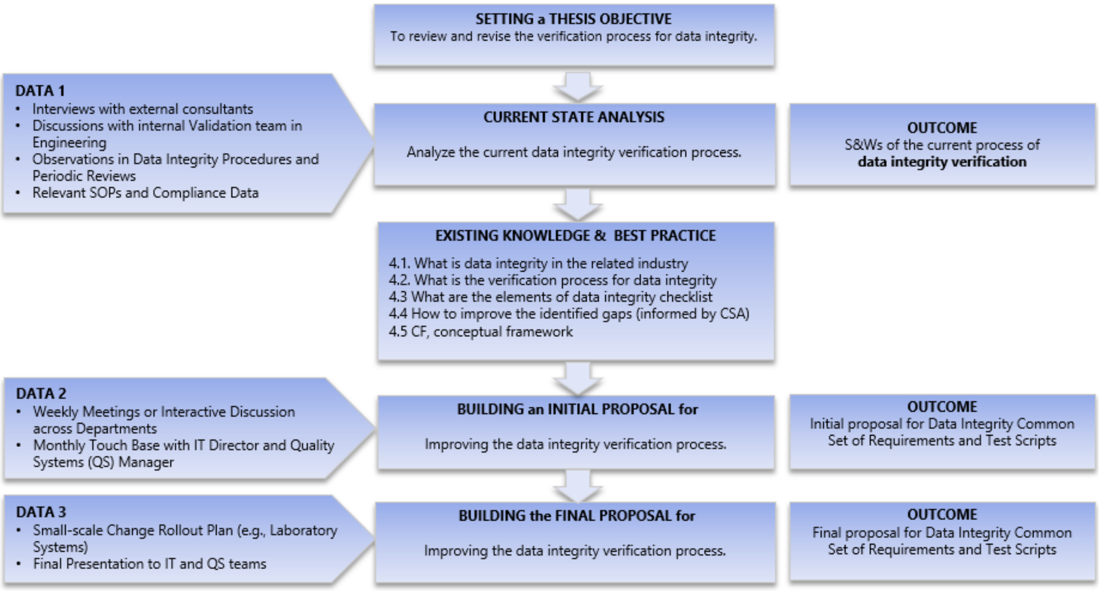

SETTING a THESIS OBJECTIVE
To review and revise the verification process for data integrity.
DATA 1
• Interviews with external consultants
• Discussions with internal Validation team in Engineering
• Observations in Data Integrity Procedures and Periodic Reviews
• Relevant SOPs and Compliance Data
CURRENT STATE ANALYSIS
Analyze the current data integrity verification process.
OUTCOME
S&Ws of the current process of data integrity verification
EXISTING KNOWLEDGE & BEST PRACTICE
4.1. What is data integrity in the related industry
4.2. What is the verification process for data integrity
4.3 What are the elements of data integrity checklist
4.4 How to improve the identified gaps (informed by CSA)
4.5 CF, conceptual framework
DATA 2
• Weekly Meetings or Interactive Discussion across Departments
• Monthly Touch Base with IT Director and Quality Systems (QS) Manager
BUILDING an INITIAL PROPOSAL for
Improving the data integrity verification process.
OUTCOME
Initial proposal for Data Integrity Common Set of Requirements and Test Scripts
DATA 3
• Small-scale Change Rollout Plan (e.g., Laboratory Systems)
• Final Presentation to IT and QS teams
BUILDING the FINAL PROPOSAL for
Improving the data integrity verification process.
OUTCOME
Final proposal for Data Integrity Common Set of Requirements and Test Scripts

Figure 1. Research design of this thesis.

As shown in Figure 1, the next part of the research design is the current state analysis of the existing data integrity verification process. The analysis is based on different data sources with the goal to review and revise the current data verification process for data integrity. The outcome of the analysis identifies the strengths and weaknesses of the existing data integrity verification process to understand the strengths that must still be kept and weaknesses that must be worked on for improvements. These areas are focused on in the next section of exploring the existing knowledge and relevant best practice in the pharmaceutical and life science industry.

Subsequently, the existing knowledge and relevant best practice are discussed on the topics of data integrity maturity model, control areas to focus for data integrity verification such as audit trail, electronic records, electronic signatures, security and access controls, backup, disaster recovery and archival, data privacy and data protection, and lastly, the approach to hybrid situations. All of these are based on guidelines and standards that are relevant and useful to the potential areas for improvement. The outcome of literature and best practice search is the conceptual framework for guiding the proposal building.

Thereafter, building an initial proposal is done with another set of data with the goal of improving the data integrity verification process via co-creations with the case company's stakeholders. Finally, the final proposal comes up in the form of the validated data integrity assessment, common set of requirements and test scripts.

2.3 Data Collection and Analysis

The data collection for this thesis will be conducted in three major rounds. Data 1 for the current state analysis is gathered by conducting interviews with external consultants, discussions with internal validation team, observations in data integrity procedures and periodic reviews, and comparing the current practices against the relevant Standard Operating Procedures (SOPs) and guidelines on data integrity compliance.

Data 2 is gathered in weekly meetings, interactive discussions across departments and monthly touch base with IT Director and QS Manager. Data 3 includes a small-scale change rollout plan (e.g. Analytical Development Systems) and the validation session in the form of a final presentation and discussion with IT and QS teams.

Table 1. Details of Data collections 1-3 used in this study.

	Participants / role	Data type	Topic, description	Date, length	Documented as
	Data 1, for the Current state analysis (Section 3)				
1	Respondent 1: External consultant (1)	Teams meeting	Interview about current process related to the respondent experiences, and point of view on the potential improvements	Feb 2022, 2hours	Field notes
2	Respondent 2: External consultant (2)	Teams meeting	Interview about best practices in the industry based on respondent experiences with other projects	Feb 2022, 1hour	Field notes
3	Respondent 3: IT Director	Face-to-face Interview	Interview about current process related to the respondent experiences, and point of view on the focus areas	March 2022, 45mins	Field notes
4	Respondent 4: IT Senior Architect	Face-to-face Interview	Interview about current process related to the respondent experiences, and point of view on the focus areas	March 2022, 30mins	Field notes
5	Respondent 5: QA/QS Manager	Face-to-face Interview	Interview about current process related to the respondent experiences, and point of view on the focus areas	June 2022, 45mins	Field notes
6	Respondent 6: Senior QA/QS Specialist	Face-to-face Interview	Interview about current process based on the respondent experiences, and lessons learned from the completion of data integrity projects	August 2022, 8hours	Field notes and approved documents
7	Respondent 7: Senior Validation Engineer	Group Meetings and Interview	Group discussions and interview about current process related to the respondent experiences	August 2022, 2hours	Field notes and approved documents
8	Respondent 8: Analytical Development Team Lead	Group Meetings and Interview	Group discussions and interview about current process related to the respondent experiences	August 2022, 2hours	Field notes and recording and approved documents
	Data 2, for Proposal building (Section 5)				
1	Participants 6-8: Senior QA/QS Specialist Senior Validation Engineer Analytical Development Team Lead	Workshop/ discussion	Proposal building of the process improvement on data integrity compliance verification	October 2022, 3hours	Field notes
2	Respondent 3: IT Director	E-mail exchanges for review	Proposal building of the process improvement on data integrity compliance verification	October 2022, 1 hour	Field notes

3	Respondent 5: QA/QS Manager	Face-to-face Interview	Proposal building of the process improvement on data integrity compliance verification	October 2022, 1 hour	Field notes
	Data 3, from Validation (Section 6)				
1	Respondent 6-8: Senior QA/QS Specialist Senior Validation Engineer Analytical Development Team Lead	Group interview/ Final presentation	Review and validation of the process improvement on data integrity compliance verification process	October 2022, 5hours	Field notes

As seen from Table 1, Data 1 includes on-site and remote meetings with the participants who worked closely on the data integrity procedures. This contributed to the current state analysis of how data integrity compliance assessment and verification processes are performed. Respondents of interviews and open discussions range from external consultants to internal respondents who are the assigned authors, reviewers and approvers in the documents (listed in Table 2 below). The documents listed in Table 2 contain supporting evidence to help understand the rationale behind why it was performed that way, any scientific foundation used as the basis of the process and scrutinize its effectivity on achieving the objectives.

Table 2. Internal documents used in the current state analysis, Data 1.

	Name of the document	**Number of pages/other content**	**Description**
A	SOP-GEN-038 Data Integrity	18 pages	Standard Operating Procedure
B	SOP-VAL-024 Data Integrity Assessment	16 pages	Standard Operating Procedure
C	IT-PLA-20-004 Initial Data Integrity and Data Governance Assessments	15 pages	Project Plan
D	IT-PLA-20-004-ADD1 Initial Data Integrity and Data Governance Assessment Addendum 1	6 pages	Project Plan
E	IT-SUM-20-004 Initial Data Integrity and Data Governance Assessment	6 pages	Project Summary Report

In the next round, Data 2 was collected to gather suggestions for developing the proposal. As it is difficult to separate the regulatory standards and best practices as reference to present the proposal to improve current state analysis, the data listed in Table 3 below are included as foundation for streamlining the process. Included in the discussion are the main changes highlighted compared with the existing one, the reason behind why it was appropriate to make the changes sustainable for the company's maturity and projected growth, the efficiency enhancements and focus readjustments to a more significant points of interest for data integrity.

Table 3. Standards and guideline documents used as reference for the Proposal building, Data 2.

	Name of the document	**Number of pages/other content**	**Description**
A	PIC/S Guidance on Good Practices For Data Management and Integrity in Regulated GMP/GDP Environments	63 pages	Regulatory Guidelines
B	PIC/S Guidance on Good Practices For Computerized Systems in Regulated "GXP" Environments	50 pages	Regulatory Guidelines
C	FDA Data Integrity and Compliance With CGMP Guidance for Industry	10 pages	Regulatory Guidelines
D	APIC Practical risk-based guide for managing data integrity	54 pages	Regulatory Guidelines
E	ISPE GAMP 5 – Records and Data Integrity Guide, 2017	152 pages	Regulatory Guidelines
F	ISPE GAMP 5 – A Risk-Based Approach to Compliant GxP Computerized Systems, Second Edition, 2022	404 pages	Regulatory Guidelines
G	Medicines & Healthcare products Regulatory Agency (MHRA) 'GXP' Data Integrity Guidance and Definitions	21 pages	Regulatory Guidelines
H	EudraLex The Rules Governing Medicinal Products in the European Union Volume 4 Good Manufacturing Practice - Guidelines on Good Manufacturing Practice specific to Advanced Therapy Medicinal Products	90 pages	Regulatory Guidelines

I	EudraLex The Rules Governing Medicinal Products in the European Union Volume 4 Good Manufacturing Practice Medicinal Products for Human and Veterinary Use – Annex 11: Computerized Systems	5 pages	Regulatory Guidelines
J	ICH Q9 Quality Risk Management	20 pages	Regulatory Guidelines
K	ICH Q10 Pharmaceutical Quality System	20 pages	Regulatory Guidelines

In the third round, Data 3 was collected when conducting validation (piloting / testing) of the initial proposal. This includes feedback from the pilot project of any additional improvements of data integrity compliance verification process — was the change easy to follow and adapt to, did it overcomplicate the course of actions, was the objective of data integrity compliance met or the simpler procedures compromise its effectivity, etc.

The findings from the current state analysis are discussed in Section 3 below.

3 Current State Analysis of Data Integrity Compliance Procedures at the Case Company

This section analyzes the current procedures of assuring data integrity compliance in the case company. This section starts with understanding the current procedures and ends with identifying the areas for improvement.

3.1 Overview of the Current State Analysis

The current state analysis focuses on the current procedures of assuring data integrity compliance starting with the 1) assessment to identify the relevant computerized systems that removes those low-risk or low GxP impact that do not need to undergo further testing and 2) the verification procedures it follows to check if the computerized system has the appropriate controls and supporting test evidence if necessary. The sequence of topics is chronologically analyzed and grouped into these two major parts of the process. The aim is to highlight the strengths and weakness of the current procedures, so improvements can be redirected towards enhancement and not replacement.

This section starts with understanding the current procedures by an active participation as a reviewer and observer in ongoing meetings to conduct data integrity assessment and verification. In addition, retrospectively studying the approved and completed documentation to aid in further understanding about its effectiveness. The process unfolds during meetings scheduled in conference rooms with the involved respondents. Among the strengths are the data governance, periodic reviews and overall culture of willingness to adhere to regulatory standards such as good documentation practices. Some of the noted weaknesses are the need for a more comprehensive risk assessment, ALCOA+ principle-based testing instead of verifying the systems controls first to identify the IT-dependent controls as recommended for hybrid situations.

To benchmark how the Company conducts data integrity compliance procedures as the starting point for process improvement, it was fundamental to start the analysis with what instructions the company follows — is it grounded on the basis of regulatory standards. From the instructions, examine the completed assessment and verification documents to substantially check how it effectively meets the objectives. Lastly, engage with people involved in this process to actively probe the purpose of every action step and analyze. With that, the following steps were performed:

1. Scrutinize the current and effective Standard Operating Procedures (SOP). The Company has two (2) relevant SOPs for the chosen topic, one is SOP-GEN-038 which tackles the concept, terminologies and the overall approach for data integrity controls across all operational activities and SOP-VAL-024 which covers the assessment and verification procedures in a computerized system. The latter SOP is the focus of this study, and the former is out-of-scope and will be kept as it is. Interview the author, reviewer and approver of the document about their perspective and corroborate the contents, if possible.

2. Analyze related project plans (IT-PLA-20-004 and IT-PLA-20-004-ADD1) or summary reports (IT-SUM-20-004) that aim to complete the annual data integrity compliance procedures. Examine the completed and documented assessments and verification and raise any questions or clarifications with the assigned signatories, if necessary. This strongly demonstrates the flaws of the current process and will be used as the starting point for the proposal.

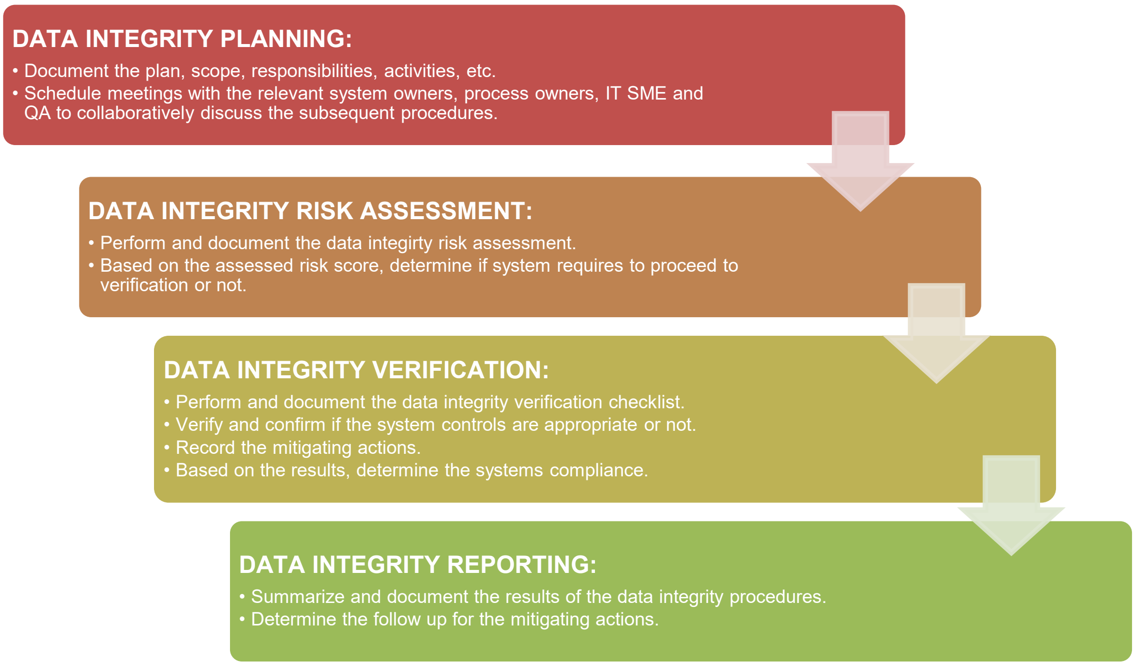
3. Observe and analyze any current or ongoing data integrity compliance procedures, and act as an active participant with the IT SME role. The goal is to understand what initially triggers the execution, who facilitates and who else takes part in it, when does it start, how long it normally gets finalized and approved, what are the setbacks and difficulties, where is it conducted and other important points that needs attention to. During these procedures, ask and confirm the intentions of why it is being done currently. Attempt to collaborate, listen to their justifications and seek their point of view.

Based on this analysis, the thesis points to the identified strengths and weaknesses in the current procedures of assuring data integrity compliance as a basis for proposing the improvement actions.

3.2 Description of the Data Integrity Compliance Procedures

In the context of the case company, Data integrity Compliance is required for each systems in use, of changed use or being brought into use. The assessment and verification will determine if the data produced by the system is managed in compliance with the relevant regulations such as EU GMP Annex 11 and US FDA 21 CFR Part 11.

Figure 2 below shows the current data integrity compliance procedures.



DATA INTEGRITY PLANNING:

- Document the plan, scope, responsibilities, activities, etc.
- Schedule meetings with the relevant system owners, process owners, IT SME and QA to collaboratively discuss the subsequent procedures.

DATA INTEGRITY RISK ASSESSMENT:

- Perform and document the data integirty risk assessment.
- Based on the assessed risk score, determine if system requires to proceed to verification or not.

DATA INTEGRITY VERIFICATION:

- Perform and document the data integrity verification checklist.
- Verify and confirm if the system controls are appropriate or not.
- Record the mitigating actions.
- Based on the results, determine the systems compliance.

DATA INTEGRITY REPORTING:

- Summarize and document the results of the data integrity procedures.
- Determine the follow up for the mitigating actions.

Figure 2. Current Data Integrity Compliance Procedures.

As shown in Figure 2, the data integrity compliance procedures are governed under a project plan and eventually wrapped up in a summary report. Scheduled meetings are initiated with the relevant system owners, process owners, IT SME and QA to collaboratively discuss, perform and document the data integrity risk assessment and verification. Any noted mitigating actions resulting from the procedures are also included and monitored periodically.

Currently, the core data integrity compliance procedures are performed into two parts. The first part is a questionnaire with an equivalent scoring method to assess if the system requires compliance with data integrity regulations. The second part is a questionnaire checklist to confirm if system controls are acceptable or not, with rationale to be documented for each. At the end of the procedure, the summary and conclusion indicate whether the system is indeed compliant with data integrity. In cases where there are risks in the system of not being compliant, further action plans should be established to assure system will be in the future.

For example, the data integrity procedures are conducted within a project plan (IT-PLA-20-004 and IT-PLA-20-004-ADD1) as a response to a finding where the process is in place, however, it was not fully performed for all the systems particularly the newly acquired ones. It was then completed in a project summary report (IT-SUM-20-004) via numerous meetings and discussions with the relevant business SMEs, system owners, process owners, QA representatives, etc.

Presently, the data integrity risk assessment is required for each system in use, of changed use or being brought into use at the Company. Defining this flexibility in the procedure must be reinforced when making a revision to keep this seamless transition. The current procedure allows two approaches when the data integrity assessment is initially performed and if there will be subsequent updates. The assessment will determine if data produced by the system is required to be created and managed in compliance with the relevant regulations. Aside from the data criticality itself, data storage (e.g. permanent or temporary), data type (e.g. electronic or paper), data functions (e.g. creation, modification, or deletion), system-level criticality, risks and complexity, and the built-in controls within the system should also be taken accounted for in the data integrity risk assessment which seemed to be missing in the current process.

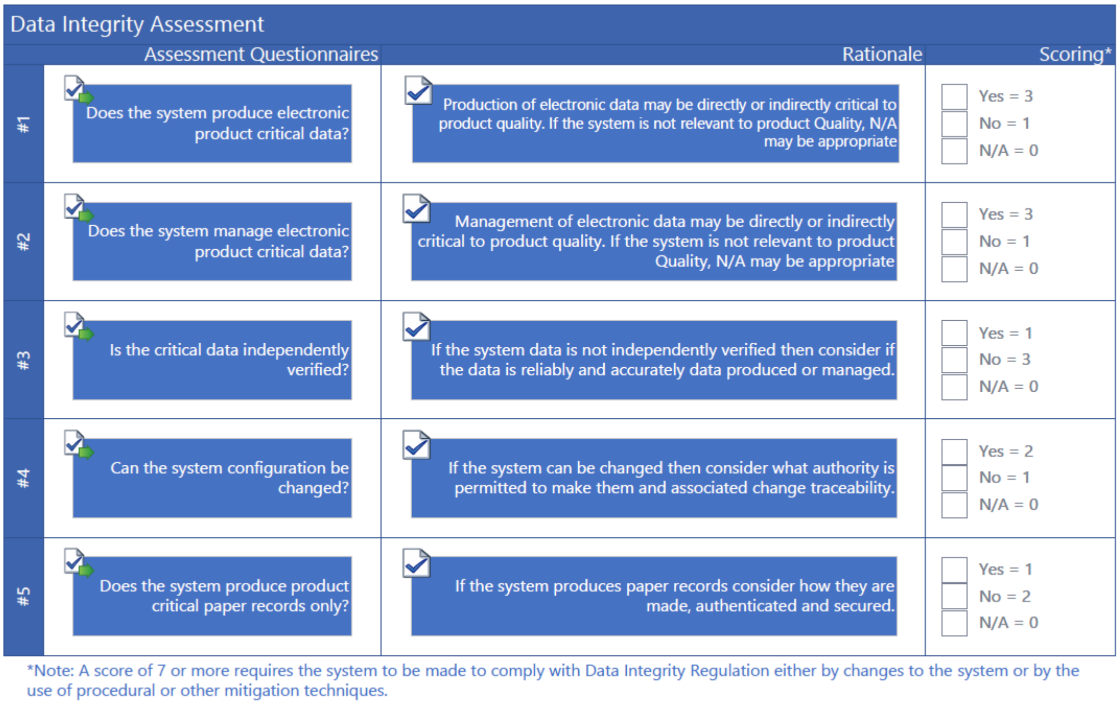
Data integrity Risk Assessment (SOP-VAL-024-A01) is a template used to record the risk assessment, whereas Data Integrity Verification (SOP-VAL-024-A02) is template used to record the actual verification results, any action plans resulting from the verification that needs to be developed and overall, the conclusion of the data integrity compliance of the system.

3.2.1 Data Integrity Assessment and Verification

Presently, the Data integrity assessment only contains five questionnaires (refer to Figure 2) looking into any critical data that was produced or managed in the system, any independent verifications manually performed outside the system, option of system configuration changes and whether it is fully handling paper records instead of a hybrid with electronic records. Scoring that ranges from zero (0) to three (3) having the highest as more of a risky element is not based on any scientific sources, but rather a simplistic approach to sort the systems on the data criticality and risks associated to data retention. Therefore, selecting the systems that scored seven (7) and above in total is required to go through verification, must be proven compliant to data integrity regulations or by use of procedural or other mitigating controls.

However, another emerging issue about inconsistent asset management resulted in this data integrity assessment as a potential misestimation of risks. There are basic tools such as weighing scale that were incorrectly identified as a computerized system but are forced to conduct data integrity verification procedures since it produces critical data. Though it does not have any system controls to be verified in the first place, but still scored high enough to proceed with the next stage for verification.

Figure 3 shows the mixed scoring method where each answer from the assessment questionnaires has an equivalent quantifiable metric. The higher the score, whether the response is a yes or a no, meant that it posts more risks to data integrity.



Data Integrity Assessment

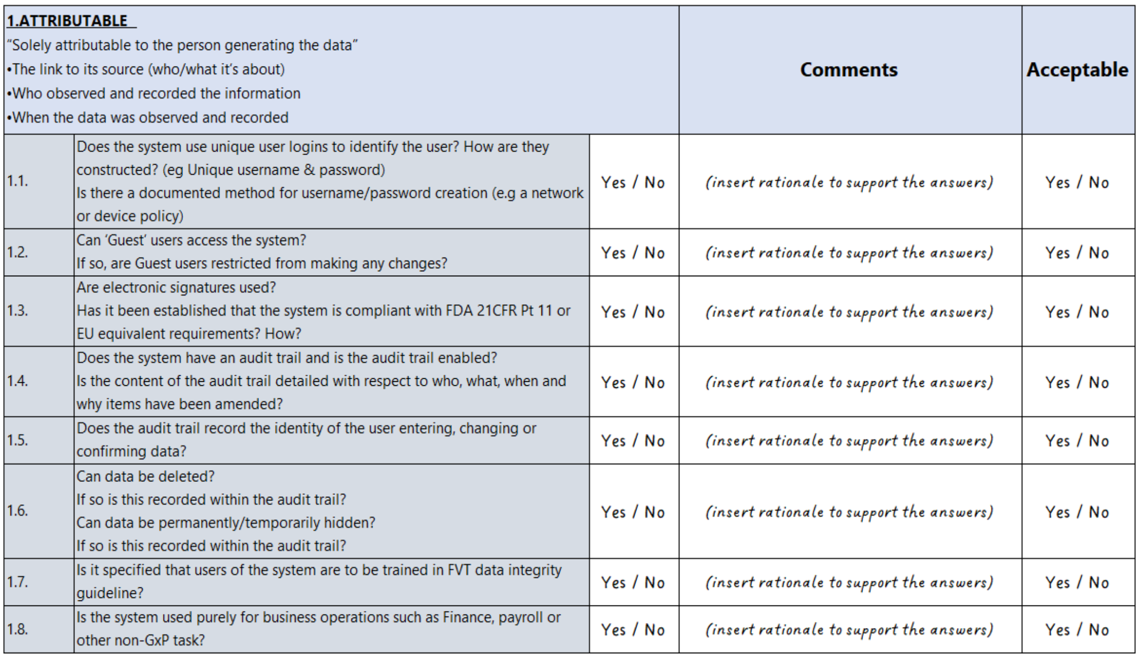
	Assessment Questionnaires	Rationale	Scoring*
#1	Does the system produce electronic product critical data?	Production of electronic data may be directly or indirectly critical to product quality. If the system is not relevant to product Quality, N/A may be appropriate	□ Yes = 3 □ No = 1 □ N/A = 0
#2	Does the system manage electronic product critical data?	Management of electronic data may be directly or indirectly critical to product quality. If the system is not relevant to product Quality, N/A may be appropriate	□ Yes = 3 □ No = 1 □ N/A = 0
#3	Is the critical data independently verified?	If the system data is not independently verified then consider if the data is reliably and accurately data produced or managed.	□ Yes = 1 □ No = 3 □ N/A = 0
#4	Can the system configuration be changed?	If the system can be changed then consider what authority is permitted to make them and associated change traceability.	□ Yes = 2 □ No = 1 □ N/A = 0
#5	Does the system produce product critical paper records only?	If the system produces paper records consider how they are made, authenticated and secured.	□ Yes = 1 □ No = 2 □ N/A = 0

*Note: A score of 7 or more requires the system to be made to comply with Data Integrity Regulation either by changes to the system or by the use of procedural or other mitigation techniques.

Figure 3. Current Data Integrity Risk Assessment.

As seen from Figure 3, presently, for any risks associated with production of electronic data, management of data critical to product quality, independent verification of data, capability of systems configurations to be changed and paper records management. Continuously, as presented in Figure 3, a metric total of seven (7) or more sets the threshold to sort out the systems that goes through the next phase. Systems that are assessed to have risks require data integrity procedures and mitigation techniques.

Figure 4 shows the data integrity verification procedures that are grouped into ALCOA+ principle beginning with the **Attributable** concept.

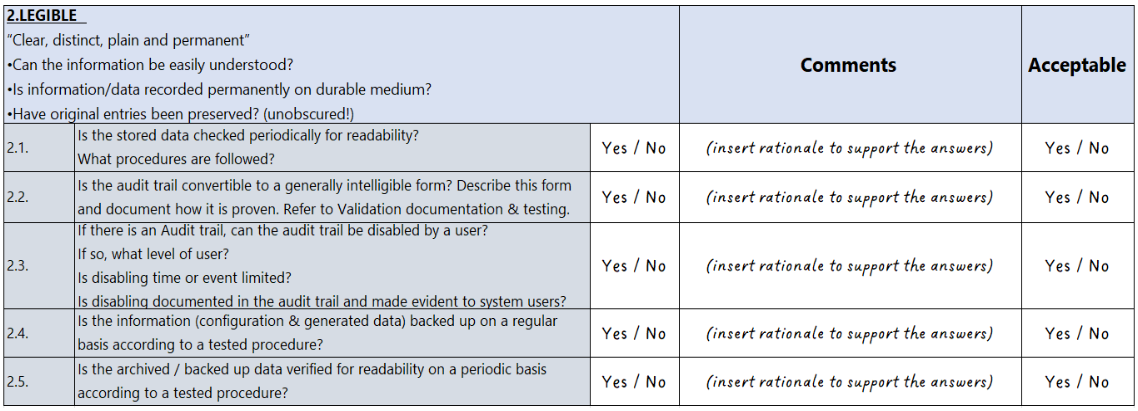


1.ATTRIBUTABLE “Solely attributable to the person generating the data” •The link to its source (who/what it's about) •Who observed and recorded the information •When the data was observed and recorded			**Comments**	**Acceptable**
1.1.	Does the system use unique user logins to identify the user? How are they constructed? (eg Unique username & password) Is there a documented method for username/password creation (e.g a network or device policy)	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.2.	Can 'Guest' users access the system? If so, are Guest users restricted from making any changes?	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.3.	Are electronic signatures used? Has it been established that the system is compliant with FDA 21CFR Pt 11 or EU equivalent requirements? How?	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.4.	Does the system have an audit trail and is the audit trail enabled? Is the content of the audit trail detailed with respect to who, what, when and why items have been amended?	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.5.	Does the audit trail record the identity of the user entering, changing or confirming data?	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.6.	Can data be deleted? If so is this recorded within the audit trail? Can data be permanently/temporarily hidden? If so is this recorded within the audit trail?	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.7.	Is it specified that users of the system are to be trained in FVT data integrity guideline?	Yes / No	*(insert rationale to support the answers)*	Yes / No
1.8.	Is the system used purely for business operations such as Finance, payroll or other non-GxP task?	Yes / No	*(insert rationale to support the answers)*	Yes / No

Figure 4. Current Data Integrity Verification – ALCOA+ (Attributable).

As seen from Figure 4, presently, questions that will confirm and verify whether the system has functions that can attributably trace transactions to the specific user who generated it, the user who may reviewed or observed it and the date and time stamp for the full trail of recording events, to name a few.

Figure 5 continuously shows the data integrity verification procedures that are grouped into ALCOA+ principle with the next concept, **Legible**.

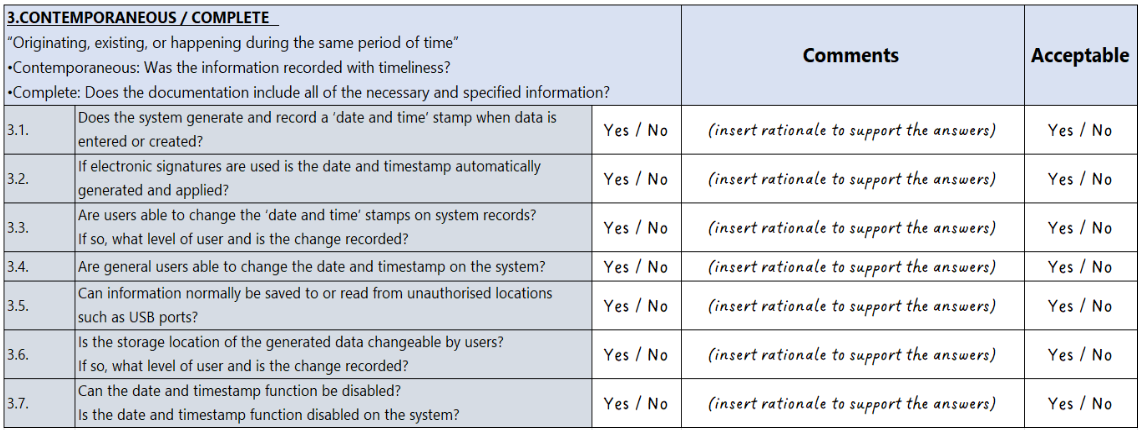


2.LEGIBLE “Clear, distinct, plain and permanent” •Can the information be easily understood? •Is information/data recorded permanently on durable medium? •Have original entries been preserved? (unobscured!)			**Comments**	**Acceptable**
2.1.	Is the stored data checked periodically for readability? What procedures are followed?	Yes / No	*(insert rationale to support the answers)*	Yes / No
2.2.	Is the audit trail convertible to a generally intelligible form? Describe this form and document how it is proven. Refer to Validation documentation & testing.	Yes / No	*(insert rationale to support the answers)*	Yes / No
2.3.	If there is an Audit trail, can the audit trail be disabled by a user? If so, what level of user? Is disabling time or event limited? Is disabling documented in the audit trail and made evident to system users?	Yes / No	*(insert rationale to support the answers)*	Yes / No
2.4.	Is the information (configuration & generated data) backed up on a regular basis according to a tested procedure?	Yes / No	*(insert rationale to support the answers)*	Yes / No
2.5.	Is the archived / backed up data verified for readability on a periodic basis according to a tested procedure?	Yes / No	*(insert rationale to support the answers)*	Yes / No

Figure 5. Current Data Integrity Verification – ALCOA+ (Legible).

As seen from Figure 5, presently, questions that will confirm and verify whether the data being managed in the system are readable, definite, and in a generally understandable format. Legible section is closely related to Original section and may be difficult to separate, thus, verifying if the data is permanent, unobscured and preserved in its original from were also included. In addition, periodic reviews, backup and restore, and archival are also mentioned to be verified in this section.

Figure 6 continuously shows the data integrity verification procedures that are grouped into ALCOA+ principle with the following combined concepts, **Contemporaneous** and **Complete**.

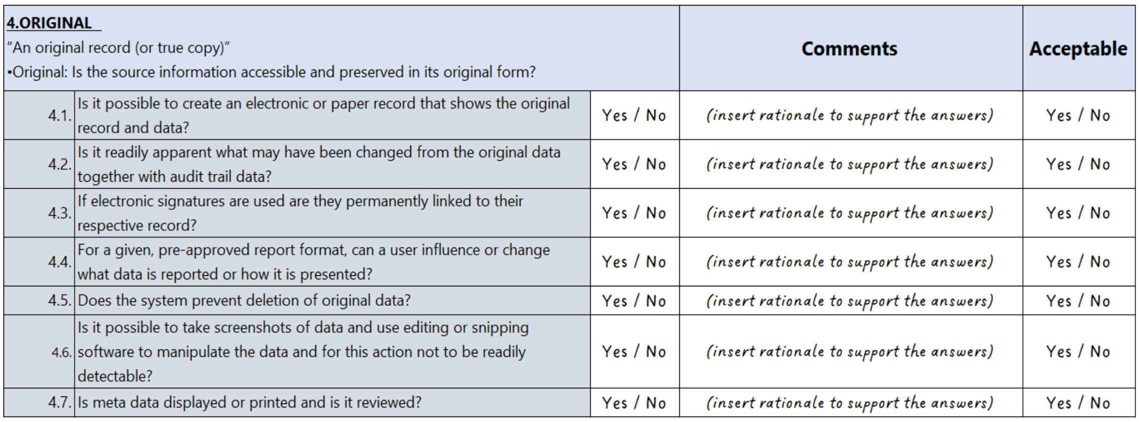


3.CONTEMPORANEOUS / COMPLETE "Originating, existing, or happening during the same period of time" •Contemporaneous: Was the information recorded with timeliness? •Complete: Does the documentation include all of the necessary and specified information?			**Comments**	**Acceptable**
3.1.	Does the system generate and record a 'date and time' stamp when data is entered or created?	Yes / No	*(insert rationale to support the answers)*	Yes / No
3.2.	If electronic signatures are used is the date and timestamp automatically generated and applied?	Yes / No	*(insert rationale to support the answers)*	Yes / No
3.3.	Are users able to change the 'date and time' stamps on system records? If so, what level of user and is the change recorded?	Yes / No	*(insert rationale to support the answers)*	Yes / No
3.4.	Are general users able to change the date and timestamp on the system?	Yes / No	*(insert rationale to support the answers)*	Yes / No
3.5.	Can information normally be saved to or read from unauthorised locations such as USB ports?	Yes / No	*(insert rationale to support the answers)*	Yes / No
3.6.	Is the storage location of the generated data changeable by users? If so, what level of user and is the change recorded?	Yes / No	*(insert rationale to support the answers)*	Yes / No
3.7.	Can the date and timestamp function be disabled? Is the date and timestamp function disabled on the system?	Yes / No	*(insert rationale to support the answers)*	Yes / No

Figure 6. Current Data Integrity Verification – ALCOA+ (Contemporaneous / Complete).

As seen from Figure 6, presently, questions that will confirm and verify whether the system can support functions to ensure timeliness and completeness of the data being recorded. Thus, date and time stamp including the settings are the key focus in this section to check if this can be manipulated. In addition, the storage location and authorized user levels are also verified to check if data are saved in a centralized manner within their allowed permissions.

Figure 7 continuously shows the data integrity verification procedures that are grouped into ALCOA+ principle with the **Original** concept.

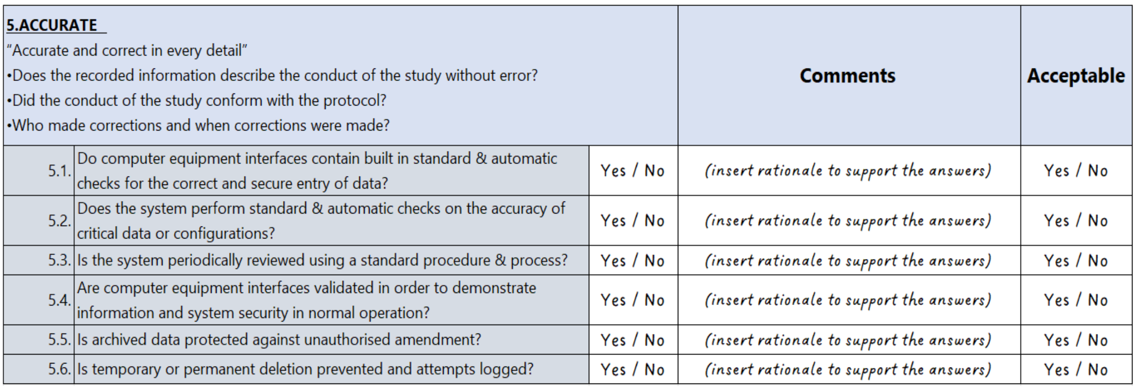


4.ORIGINAL "An original record (or true copy)" •Original: Is the source information accessible and preserved in its original form?			**Comments**	**Acceptable**
4.1.	Is it possible to create an electronic or paper record that shows the original record and data?	Yes / No	*(insert rationale to support the answers)*	Yes / No
4.2.	Is it readily apparent what may have been changed from the original data together with audit trail data?	Yes / No	*(insert rationale to support the answers)*	Yes / No
4.3.	If electronic signatures are used are they permanently linked to their respective record?	Yes / No	*(insert rationale to support the answers)*	Yes / No
4.4.	For a given, pre-approved report format, can a user influence or change what data is reported or how it is presented?	Yes / No	*(insert rationale to support the answers)*	Yes / No
4.5.	Does the system prevent deletion of original data?	Yes / No	*(insert rationale to support the answers)*	Yes / No
4.6.	Is it possible to take screenshots of data and use editing or snipping software to manipulate the data and for this action not to be readily detectable?	Yes / No	*(insert rationale to support the answers)*	Yes / No
4.7.	Is meta data displayed or printed and is it reviewed?	Yes / No	*(insert rationale to support the answers)*	Yes / No

Figure 7. Current Data Integrity Verification Template – ALCOA+ (Original).

As seen from Figure 7, presently, questions that will confirm and verify whether the system can assure the data to be in its original state, free from unauthorized alterations or unnecessary ambiguities. True copies are also verified if the same information is held compared to the original record lending its credibility.

Figure 8 continuously shows the data integrity verification procedures that are grouped into ALCOA+ principle with the **Accurate** concept.



5.ACCURATE "Accurate and correct in every detail" •Does the recorded information describe the conduct of the study without error? •Did the conduct of the study conform with the protocol? •Who made corrections and when corrections were made?			**Comments**	**Acceptable**
5.1.	Do computer equipment interfaces contain built in standard & automatic checks for the correct and secure entry of data?	Yes / No	*(insert rationale to support the answers)*	Yes / No
5.2.	Does the system perform standard & automatic checks on the accuracy of critical data or configurations?	Yes / No	*(insert rationale to support the answers)*	Yes / No
5.3.	Is the system periodically reviewed using a standard procedure & process?	Yes / No	*(insert rationale to support the answers)*	Yes / No
5.4.	Are computer equipment interfaces validated in order to demonstrate information and system security in normal operation?	Yes / No	*(insert rationale to support the answers)*	Yes / No
5.5.	Is archived data protected against unauthorised amendment?	Yes / No	*(insert rationale to support the answers)*	Yes / No
5.6.	Is temporary or permanent deletion prevented and attempts logged?	Yes / No	*(insert rationale to support the answers)*	Yes / No

Figure 8. Current Data Integrity Verification Template – ALCOA+ (Accurate).

As seen from Figure 8, presently, questions that will confirm and verify whether the system can assure its records are reflective of the authentic data which is free from errors. Any changes or corrections made are properly accounted for. This section also checks any potential errors in transferring or migrating data for interfacing systems. Accurate concept also resembles the same in Original concept with the requirement that records should be protected against unauthorized amendments.

As seen from Figures 4 to 8, presently, the data integrity verification procedures are grouped in the ALCOA+ principle and listed in a confirmation questionnaire format. There is a column to record a yes or no answer with another column for the comments to elaborate any supporting details and another column to determine if the combination of existing controls and mitigating controls are acceptable to substantiate that the system meets the regulatory requirements and remain suitable for its continued use:

Table 4. Summary Objectives of Data Integrity Verification Sections based on ALCOA+.

ALCOA+ Objectives for Data Integrity Verification Process	
1	**Attributable** section pertains to verifying that the system is capable of solely attributing the data generation to the person by linking it to its source of what exactly happened, who observed and recorded the information and when the data occurred.
2	**Legible** section establishes if the system data is clear, easily understood, distinct, plain and permanent in a way that the information is recorded on a durable medium with the original entries having been preserved and remain unobscured.
3	**Contemporaneous** section verifies that the originating and existing event is happening during the same period of time it was recorded. Complete section checks that the documentation includes all of the necessary and specified information.
4	**Original** section corroborates the data is indeed the original record or true copy in the system or on-hand. This also includes checking that the source information is accessible and preserved in its original form.
5	**Accurate** section authenticates the correct data in every detail verifying that the recorded information describes the conduct of the study without error, conforms

	with the protocol and includes who made any corrections and when these corrections were made.

As seen from Table 4, this summarizes the objectives of data integrity verification grouped into the ALCOA+ principles. On the surface, some of these principles tend to be interrelated and seem to mention the same system controls to verify. The objective of each principle is clear, whether Attributable, Legible, Contemporaneous/Complete, Original or Accurate. Each covered its respective control areas to magnify and assess whether the system indeed meet the concepts.

At the end of the template is the conclusion for the overall results of data integrity verification with another field to include any required further actions and the justification. These further actions must be revisited to check if it is already executed such as an SOP revision request to incorporate additional controls to mitigate a system limitation.

3.2.2 Scope

Presently, the scope of the data integrity compliance are all computerized systems that handles and manages data that has a GxP impact to patient safety and product quality. The GxP impact is separately assessed at the planning stage of computerized systems validations and will not be a relevant topic for this study and will not be subject to any revisions.

The improvement aims to be carried out prospectively upon the revised effective date of the SOP-VAL-024 Data Integrity Assessment, and not respectively provided that it’s irrelevant to revise the approved and completed data integrity verification documents for systems in validated state and in production use. However, if remaining tasks to assure data integrity compliance are still open with any existing Corrective Actions and Preventive Actions (CAPA), the proposed change can also influence the said tasks.

Defining this scope is essential to minimize abruptions and the risks of overdoing document revisions that may exceed the costs of rolling the change against its benefits. The transition of the usage from the current to a revised process must promote continuity of the operations.

3.2.3 Roles and Responsibilities

As an observation, the data integrity procedures are currently highly driven by Quality Assurance (QA) Specialist who is assigned and tasked for this initiative. The individual facilitates the meetings, most of the times even schedules them, instructs the system owners and process owners on what to do, and guides the people to understand the rationale based on the regulatory standards.

A separate validation team under the Engineering department is mostly handling the computer systems validation procedures including the data integrity compliance tasks. Process owners have direct involvement during testing alongside the validation SME and will either have administrator roles in the system and thus be knowledgeable of its behavior and controls. System owners who are normally the Department Heads are involved in monitoring, reporting and approval of the deliverables. End-users who will most likely be using the system after it is completely validated are usually involved and part of the performance qualification but may or may not participate in the data integrity compliance procedures. An IT SME is also encouraged to be included in the working sessions to review if the appropriate controls are available in the system, otherwise, a mitigating control is highly recommended to have as a workaround.

Best practice is to collaboratively have a strong stakeholder engagement involving business process owners, quality assurance and every SMEs in key supporting technical groups. They will most likely have the data integrity understanding and awareness to maintain the system compliant during the operational use of the system.

3.2.4 Document Management Tools

The currently used document management tool is the QMS and EDMS application that holds the approved data integrity assessment and verification documents. The working files are managed within the organizational communication application to be collaboratively shared with the relevant participants for any input or review comments. The approvals were previously in wet signatures and the transition to electronic signatures made the documentation more expeditious delivery and completion.

Summing up, the data integrity assessment and verification are jointly performed and documented. The first part, which is the assessment identifies the system’s risks with electronic data, its criticality to product quality, independent verification, systems configurations to and paper records. What can also be considered in the assessment are system's complexity and whether the data storage is permanent or temporary. Subsequently, the second part is the verification which confirms the system's functions and controls to support and mitigate these identified data integrity risks. The data integrity verification process is grouped into the ALCOA+ principles which tend to overlap and repeat the same concept or system controls. These gaps are the focus of the analysis in the next section of this thesis.

3.3 Analysis & Key Findings from the Current State Analysis

The analysis focused on finding the efficiency opportunities by enhancing the data integrity assessment and verification procedures. To support the analysis, responses were collected from participants and key responsible members during group discussion, interviews and online meetings. The responses were collected in minutes of meeting (no recordings) and the exact citations cannot be shared due to confidentiality restrictions and the aim to maintain the anonymity of the personnel.

The following key findings were highlighted as topic areas in the current state: a) for the strengths of the process, these must be kept and still observed, whereas b) for the noted weaknesses, these must be highlighted, further analyzed and improved. However, identified topics that may need further in-depth investigation so that to look for potential process improvements, especially the misdirected risk assessment and ALCOA+ principle-based verification process.

3.3.1 Areas of Strengths

Among the strengths is the current **two-stage process**, i.e. the risk assessment followed by the verification, which is a good start in data integrity procedure. This sorts the low risk from high and critical risk systems to channel the time and resources appropriately.

One of the strengths is the **data governance** measures in place that are conducive for managing the data contained within the computerized systems. This includes electronic records and other relevant GxP data classified according to its data types, criticality, and confidentiality. Data types are itemized into its source and origins, whether primary data, true copy, raw data, processed data, metadata, audit trail data, backup data, or archived data. Data are also categorized into its confidentiality whether data is public, internal, or confidential. Data criticality is assessed alongside its GxP impact whether it is GxP critical, business critical or non-critical.

One of the strengths is the stable and existing process to maintain the system compliant throughout its operational phase by means of **periodic review tasks** conducted to verify its continued validated state. These may include system audit trail review, user access review, backup and restore tests, data readability checks, performance monitoring, status checks of incidents, deviation, CAPAs and Change Control, etc.

Another strength is the **Good Documentation Practices** (GDocP) which aims to produce a records that clearly and consistently documents all activities performed to ensure compliance with GMP and internal procedures. It also ensures the traceability of all components related to the manufacturing process and the data integrity of handwritten information. Continuously, the paper records are kept within the required **data retention** periods and managed accordingly based on the standards and regulations.

The purpose of highlighting the strengths is not to start from scratch, but to make use of what the Company already have in place and utilize these. Some of the underlying controls mentioned in the verification stays as a reference to check if they are covered, and if omitted should be justified. The omission must not compromise the data integrity compliance, but also must not enforce unnecessary verification outside the system's capability scope.

3.3.2 Misdirected Risk Assessment

The data integrity risk assessment included the data criticality but neglected to consider the system's complexity and the storage capability of maintaining data temporarily or permanently. When these elements are ignored, the risk assessment for data integrity compliance may be misdirected.

One of the respondents cited that the current risk assessment tends to overestimate the criticality just because the system has an indirect impact to product quality. The respondent continued that the generated information from the system is included as attachments to release assays which are considered critical data, despite having the system as a very simple measurement tool. Another respondent supported that there are no issues on having the critical data as an element in the risk assessment; however, if the complexity of the system is ignored including the capability of permanent storage of electronic data, the subsequent verification process will not be applicable and descriptive justification will be needed if the requirement is not met.

One of the external consultants confirmed that **the risk assessment is not scientifically based** on any scoring system defined to segregate systems applicable for data integrity verification. The consultant continued that it is generally based on the risk exposure and data criticality alone, **with minimal considerations on the system complexity and controls**. He explained further that even the set threshold of scoring of seven (7) or higher is not based on any regulatory resource, but a mere professional estimate.

The Senior QA/QS Specialist mentioned that one of the line items in the risk assessment which confirms “Does the system data meet relevant cGMP requirements?” will be more appropriate in the conclusion or by the end of the verification process. Most of the respondents agreed with this observation. During the conducted interviews, the QA/QS Manager informed that the SOP for data integrity assessment and verification is scheduled for periodic revision, hence a window is available to implement the necessary improvements and fixes.

3.3.3 Inconsistencies in Asset Management

The asset management for computerized systems is inconsistently recorded with a prominent observation of having individual components considered as one computerized systems each. As more and more systems are integrated, the important data points throughout the data lifecycle and the associated controls where the data is managed must be deliberated holistically to understand the focus areas of data integrity compliance. A reliable inventory of computerized systems is the foundation of assessing the risks in data integrity, data security and data flow of inputs and outputs.

Looking at the list of GxP systems that are relevant for data integrity compliance procedures, it was prominent that there’s **an inconsistent identification** of high-level computerized system against the components such as software, equipment, hardware and other assets. According to ISPE GAMP 5 – A Risk-Based Approach to Compliant

GxP Computerized Systems, where a computer system is regarded as one component of a wider manufacturing process or system, particularly in an integrated QbD environment, specific and separate computerized system validation may not be necessary (ISPE, 2022).

The Senior Validation Engineer explained for instance that a simple equipment with an embedded firmware **does not have a separate database** to store any GxP data permanently, rather it's just used as a tool to measure particle contaminants and prints this out in a report. He elaborated further that this **equipment is recorded as one asset** in the GxP computerized systems list and there are similar other equipment with the same model having a separate record, separate qualifications, thus requiring separate data integrity assessment and verification testing. As one can argue that these equipment require different calibration, the system controls for electronic records may just be managed in one centralized software. The IT Infrastructure Architect stated that there is a separate software list but these are used primarily to monitor the required licenses and version, but these are not linked to a specific equipment or hardware yet. Furthermore, he communicated the need for a visual IT Infrastructure mapping of all the business applications to have a comprehensive understanding of the network layouts and their interconnections to other systems. He added that in order to outline this, a reliable list of assets should be defined first including the production equipment, software, network components and other assets.

Another respondent agreed and corroborated that it is definitely an improvement that needs to be worked on. Admittedly, she confirmed that the **GxP computerized systems are not listed on the same level** having some system components listed each as a single system. She further substantiated in the discussion that the information on the GxP system metadata cards tend to be outdated and should be regularly reviewed. This was then included as part of the tasks in the periodic reviews that are being conducted. Due to the nature of this improvement, another project was initiated to secure the availability and schedule of the resources.

3.3.4 ALCOA+ Principles-Based vs System Controls-Based Verification

The ALCOA+ principle is a good guideline to follow but does not align with the required system controls to be verified for data integrity compliance. In order to establish a more cohesive and efficient approach of data integrity verification process, grouping the system controls together is more advisable and comprehensible for a system's language instead of the ALCOA+ principle.

One of the respondents pointed out that the data integrity verification procedures are grouped in the ALCOA+ principle which was difficult to follow since it is **not aligned with the computerized system controls.** The respondent reinforced this argument with how the data integrity verification procedures are documented. She continued on the discussion that the audit trail functionality, an example of a system control to which has its own well-defined section in the regulatory standards and guidelines was **scattered into multiple sections** as shown in Current Data Integrity Verification – ALCOA+ (Attributable) (refer to Figure 4) section, where audit trail controls such as enabled settings and its minimum required metadata are captured. However, she added that in Current Data Integrity Verification – ALCOA+ (Legible) (refer to Figure 5) section, audit trail controls are brought up again and this time it covers the capability to extract in readable reports and its restrictions for disabling the settings. She concluded that this meant having all the participants revisit the same system control that was already addressed in the beginning principle just because it will be tackled again in the subsequent section of a different principle.

This was a common observation from the respondents that it seemed **repetitive, incohesive and difficult to follow**. Despite having the control to be verified in the Attributable section different from the one in the Legible section, grouping the same controls based on the standards are a lot easier to follow and comprehend. The test objectives to accomplish can also be more cohesive during the data integrity verification process. A respondent reasoned out that computerized systems are not built to honor and setup the controls according to ALCOA+ principle, rather into system functionalities and security settings. Not all systems are built for data integrity compliance because there are other clients outside the pharmaceutical and life science industry that are not necessarily required to be compliant with FDA and other regulations. The external consultant suggested that the changes to group the verification procedures into system controls can facilitate the learning curve for the users who will repeatedly perform it.

3.3.5 Repetitive Efforts in Data Integrity as Part of CSV

Repetitive efforts and duplicate work tend to occur since data integrity assessment and verification are separately managed as a process from CSV, despite being tested and reported within the CSV deliverables. The siloed processes may create inefficiencies and distract the resources from achieving the valuable goal of data integrity compliance.

The Senior Validation Engineer indicated that data integrity-related requirements are part of the User Requirement Specifications which is one of the systems validation deliverables. He confirmed that this needs to be sent to vendor beforehand, so they will be aware of what parameters setup and other configurations to be defined in the system. One of the respondents exclaimed that data integrity verification procedures must be aligned with a specific requirement to meet, may it be audit trail, electronic signatures and user access controls. She continued on that if these are intentionally repeated across multiple systems, it will be advisable to have the data integrity-related requirements as predefined with associated risk assessment to filter its applicability. With these arguments and discussion, most of the respondents agreed to avoid duplicate efforts in the computerized systems validation.

Another frequency to assure data integrity is done periodically according to system's risks and criticality. These periodic review procedures have been successfully established in the Company including system audit trail review, user access review, backup and disaster recovery restore test, data readability test, etc.

One of the respondents convinced others to complete the data integrity compliance procedures alongside the computerized systems validation as an efficient move to not repeat the testing, unless system changes necessitate a revision. She continued on that if it is already done during one of the actions in computerized systems validation, the best option is to not do it again and just include a document reference.

The MHRA GMP Data Integrity Definition and Guidance for Industry states the principle that "The effort and resource applied to assure the integrity of the data should be commensurate with the risk and impact of a data integrity failure to the patient or environment." (MHRA, 2018, p.4). Estimating the level of effort and required resources to implement data integrity compliance begins at the planning stage of computerized systems validation as well. A respondent contended that in order to save the efforts during operational or performance qualification, the same regulatory requirements as the ones for data integrity compliance must be fulfilled once.

3.3.6 Lack of Direction and Guidance

The verification process is at risk of being unclear and insufficient for an individual without any background on data integrity compliance to comprehend. The requirements, action steps and acceptance criteria or expected results are not separately identified and documented which can potentially be used across different systems in a routinary manner. With this defined, uncertainties in the process can be minimized.

During the group discussions, one of the respondents commented that the questionnaires seemed to lack guidance since these are written with an underlying expectation that the process owners and other participants in the data integrity procedures are aware of the regulatory standards. Another respondent exclaimed that all the employees are required to be trained for the general concepts of data integrity as she remembered herself during onboarding. However, she added that these alone are still insufficient and difficult to incorporate if an individual has to be involved in testing the data integrity compliance of a computerized system. As a baseline, there should be user requirements stating what is expected of the system and align them with a step-by-step procedures or instructions that a person who has no knowledge can follow.

The Analytical Development Team Lead criticized that the questionnaires in the data integrity verification process do not exactly inform the user of what is acceptable or not. She continued that it may be difficult for a first-time reader who does not have knowledge or basic understanding of data integrity’s purpose. She explained that meetings are always required to be scheduled so these gray areas that still need to be confirmed with Quality and other SMEs can be raised to assist in identifying what controls must exactly be in place within the system. The Senior QA/QS Specialist and QA/QS Manager proposed that negative and challenge testing should be included in validating computerized systems that can be included in the data integrity compliance procedures.

3.4 Summary of the Current State Analysis Results

FDA states that “CGMP regulations and guidance allow for flexible and risk-based strategies to prevent and detect data integrity issues. Firms should implement meaningful and effective strategies to manage their data integrity risks based upon their process understanding and knowledge management of technologies and business models” (FDA, 2016, p.1). Based on this guidance, the Company still has that freedom to select the approach that makes sense in its maturity. The current data integrity compliance procedures in place comprise of both positive and negative elements in terms of efficiency and effectivity. What makes sense to facilitate the growth of the Company is not a process that may compensate controls but that which eventually produce documentation just for the sake of it. The focus should be accentuating high risks relating to data integrity and find solutions to mitigate it within or outside the systems.

3.4.1 Strengths and Weaknesses of Data Integrity Compliance Procedures

Among the strengths of the current data integrity compliance procedures that the Company should sustain is the two-stage of having a risk assessment first before proceeding to the verification proper. The QA/QS manager and the external consultants are convinced that this is still a good control to separate only the relevant computerized systems to manage. The Analytical Development Team Lead and QA/QS Senior Specialist both agreed during the discussions that most of the verification procedures have controls that will most like be kept, while others with inherent risks and those outside the system's control can be removed with justification.

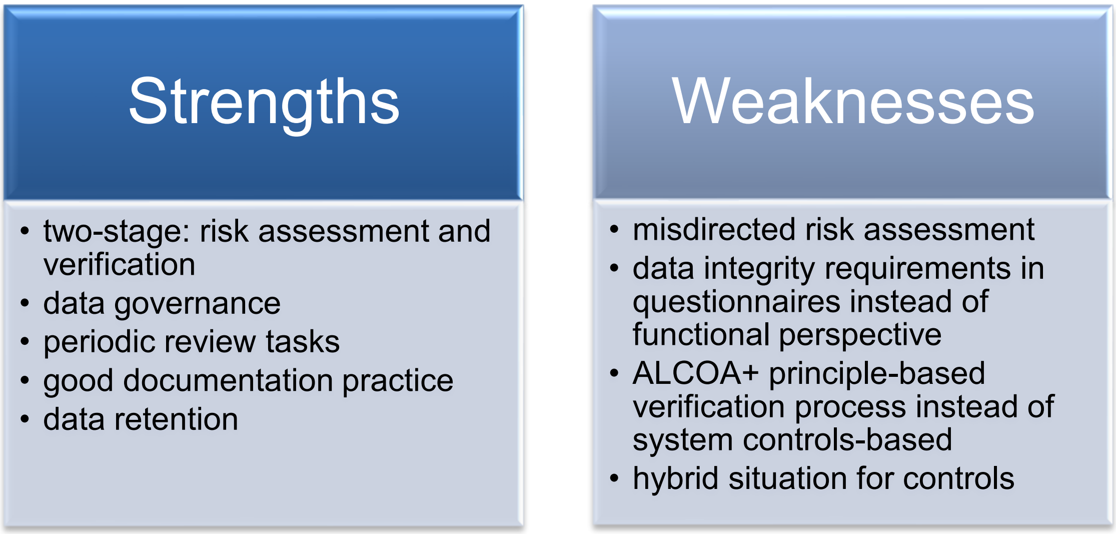
According to ISPE GAMP Guide for Records and Data Integrity, "Data governance encompasses the people, processes, and technology required to achieve consistent, accurate, and effective data handling. Data governance provides the structure within which appropriate decisions regarding data related matters may be made according to agreed models, principles, processes, and defined authority." (ISPE, 2017, p.21) The Company has strong procedures in place to manage and monitor data governance to assure data integrity. Also relevant is the subject for data life cycle which includes all phases from creation, processing, review, approval, usage, retention, retrieval and destruction.

The IT Director is assured that periodic tasks including systems periodic review, user access review, audit trail review, data readability review, backup and disaster recovery restore tests, are effectively carried out with collaboration from different departments. This is expected to be performed annually, every two (2) or three (3) years based on the risks in the system. It's one of the conducive activities that remind the Company of the importance of data integrity compliance during the operational phase of the system.

The QA/QS Senior Specialist discussed that there’s a strength in the Company’s good documentation practice (GDocP) including data retention and archival of paper records. He continued that these are all kept in a controlled and safe environment limited only to authorized personnel. One of the respondents exclaimed that based on the data readability tests performed recently during the periodic review, these successfully proven that the GDocP is applied and followed. The QA/QS Senior Specialist further added to the discussion that of course, there are still minor slip ups during the documentation, but these are carefully addressed with proper corrective actions such as refresher training.

A respondent during the interviews scrutinized the current process in detail saying that the data integrity risk assessment needs revision since it not based on any scientific scoring. The external consultants agreed that the current data integrity assessment questionnaires can be improved. Currently it only considers the following: 1) the data criticality, 2) any independent verification conducted outside the system, 3) capability of system configurations to be changed and 4) whether system fully manage data as paper records. The former respondent acknowledged this weakness that the assessment questionnaires did not take into account the complexity of the systems, and whether the data is permanent, temporary or transferred to another location or system. The discussion yielded towards an overhaul of this assessment and the respondents opted to look for more extensive risk-based approach.

The Analytical Development Team Lead expressed her uncertainty on why the data integrity requirements are stated in questionnaires instead of a user or functional perspective. She further added that this opens the procedure to multiple interpretations and inconsistencies. Another respondent concurred to her sentiment that the verification process must be grouped per system controls, not based on principles though it is a very useful guideline to follow. She explained that another weakness she had observed is the hybrid situation — both system controls and manual controls are managing data integrity compliance, however the verification process is not chronologically tested in the system level first. She argued that this makes it difficult to pull out the relevant risks and system limitations that are vital to be mitigated.



Strengths

- two-stage: risk assessment and verification
- data governance
- periodic review tasks
- good documentation practice
- data retention

Weaknesses

- misdirected risk assessment
- data integrity requirements in questionnaires instead of functional perspective
- ALCOA+ principle-based verification process instead of system controls-based
- hybrid situation for controls

Figure 9. Strengths and Weaknesses of the current data integrity compliance procedures.

3.4.2 Selected Focus Areas

Upon probing the current state analysis, it was evident that there is clear opportunities for improvement in the risk assessment down to the verification process. The core areas of this study are the weaknesses identified such as 1) misdirected risk assessment as this is crucial in identifying the applicable controls and level of efforts for each system, 2) replacing the ALCOA+ principle-based with the system controls-based verification for efficiency purposes and 3) enabling users with concrete requirements and instructional procedures to combat lack of direction and guidance. The other topics, though most of the respondents agreed, should also be addressed will be kept as out of scope to avoid drifting off from the core focus of the study.

With these as the focal points, the next section is helpful to align the process improvements on data integrity compliance.

4 Existing Knowledge & Best Practice on Data Integrity Compliance

This section discusses the best practice on data integrity compliance derived from standards, regulations, guidelines and existing knowledge from performing data integrity procedures for multiple companies in the industry. Since there are rules to follow in this industry, we already have the principle as a framework. However, the principle itself cannot simply be followed as computerized systems are not all built to accommodate the regulatory requirements and address them properly. Thus, understanding and professional judgment are still significantly essential to determine where the system has its limitations and where a workaround and human-dependent controls come in.

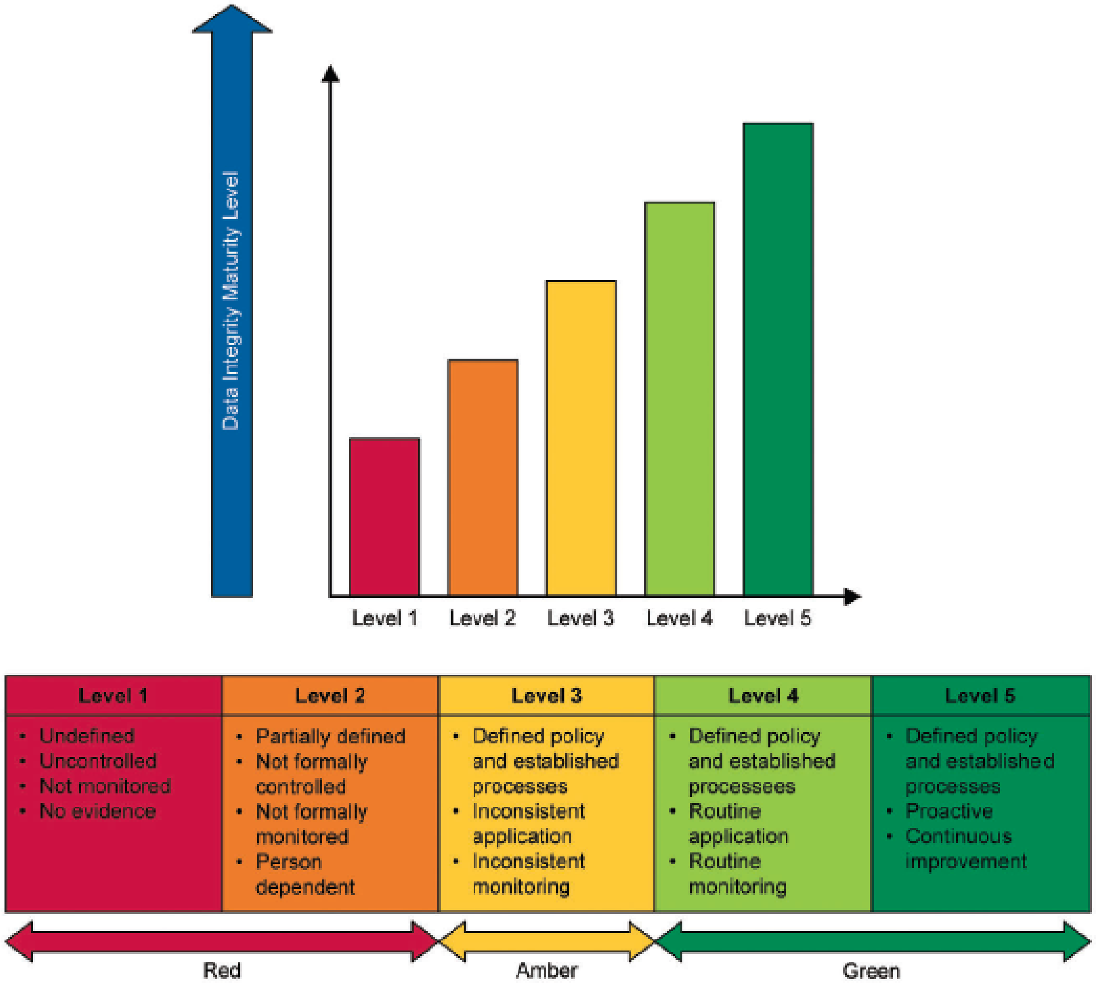
4.1 Data Integrity Maturity Model

Based on ISPE Guideline (2017), "Regulated companies should consider implementing a corporate data integrity program to identify, remediate, and manage potential risks to data integrity". (ISPE, 2017 p.47). Data integrity compliance has an increasing regulatory focus and importance due to its impact and risks to product recalls, warning letters, legal actions resulting from potential harm to consumers or patients, etc.

Continuously according to ISPE Guideline (2017) when discussing Indicators of Program Scope and Effort for a Corporate Data Integrity Program, "In order to design and implement an appropriate corporate data integrity program, regulated companies should first understand their current state and acceptability of control based on risk to data integrity." (ISPE, 2017 p.48) Since reviewing the requirements and procedural controls are a non-measurable concept, the data integrity maturity model is the best approach for evaluating an organization's current state.

To implement continuous improvement, it is crucial to know the starting point of the Company based on the Data Integrity Maturity model. This is an invaluable indicator of where to better allocate resources and distribute efforts. Currently, the data integrity maturity level is at Level 3 where there is a defined policy and established practices. However, due to misleading system assessment, vague requirements and incoherent instructions, there's a possibility of inconsistent application depending on the voices and ideas of the present SMEs in the room. As a result of this, inconsistent monitoring is inevitable too. The achievable goal is to move to Level 4 where there is a routine application and routine monitoring. This can be executed with a baseline of methodical risk assessment, objective data integrity requirements rooted from regulatory standards, and clear verification instructions that can be used across systems to test the controls and emphasize the risks to be mitigated.

Figure 10 shows an example of a data integrity maturity model (ISPE, 2017). It encapsulates the entire spectrum of the key elements in data integrity as it describes the process areas that should be assessed and the maturity factors for each of these areas. According to ISPE GAMP Guide for Records and Data Integrity, “The maturity model may also be used as a rapid and efficient, but relatively detailed management indicator, enabling regulated companies to focus resources and effort effectively. This general approach is flexible and may be structured several ways, e.g., by geographical area, site, or department.” (ISPE, 2017, p.55)



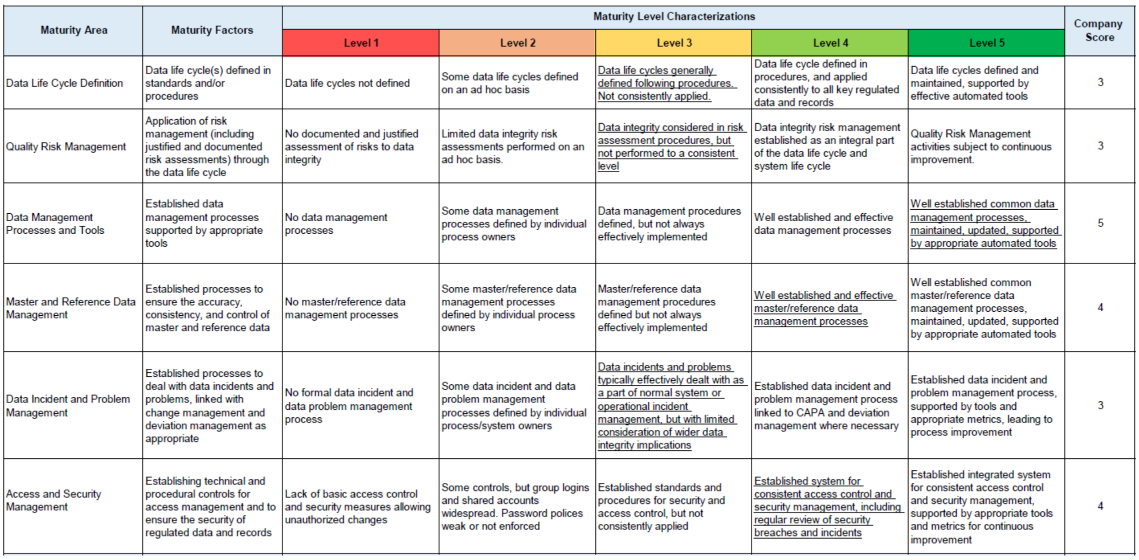
Level 1	Level 2	Level 3	Level 4	Level 5
• Undefined • Uncontrolled • Not monitored • No evidence	• Partially defined • Not formally controlled • Not formally monitored • Person dependent	• Defined policy and established processes • Inconsistent application • Inconsistent monitoring	• Defined policy and established processees • Routine application • Routine monitoring	• Defined policy and established processes • Proactive • Continuous improvement

Figure 10. Data Integrity Maturity Model (ISPE, 2017).

As seen from Figure 10 based on ISPE GAMP 5 Guide on Records and Data Integrity, the determination of the Company on its data integrity maturity level is a necessary starting point to assess for improvements. The levels can range from the lowest Level 1 (Red) where there is no defined policies or appropriate procedures to manage, control and monitor data integrity compliance, up to the highest Level 5 (Green) where there is a defined policy and established processes, proactive and continuous improvements. Currently, the Company being at Level 3 (Amber) where there is defined policy and established processes, however, there is also an inconsistent application and inconsistent monitoring to be both considered as the key focus of the research for improvements.

Based on ISPE GAMP 5 Data Integrity, the Process Areas to gauge the Maturity Factors are as follows: Culture, Governance and Organization, Strategic Planning and Data Integrity Program, Regulatory, Data Life Cycle and its Supporting Processes. Among these Maturity Factors, the study will concentrate on the last two as these relates more to data integrity compliance procedures as part of computerized systems validation.

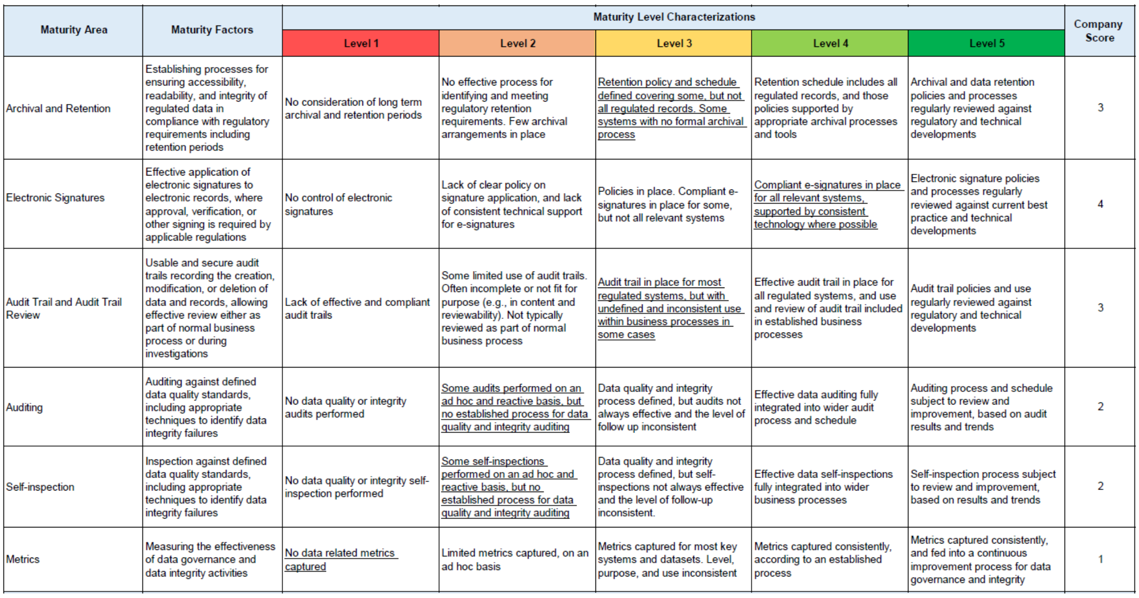
Figure 11 shows the Company Score based on the Maturity Factors within the Maturity Areas of Data Life Cycle Definition, Quality Risk Management, Data Management Processes and Tools, Master and Reference Data Management, Data Incident and Problem Management, and Access and Security Management. Scoring the highest is the established data management processes supported by appropriate tools. On the other hand, the areas that can improve on a more consistent application are the data lifecycle, risk assessment procedures, and data incidents and problem management.



Maturity Area	Maturity Factors	Maturity Level Characterizations					Company Score
		Level 1	Level 2	Level 3	Level 4	Level 5	
Data Life Cycle Definition	Data life cycle(s) defined in standards and/or procedures	Data life cycles not defined	Some data life cycles defined on an ad hoc basis	Data life cycles generally defined following procedures. Not consistently applied.	Data life cycle defined in procedures, and applied consistently to all key regulated data and records	Data life cycles defined and maintained, supported by effective automated tools	3
Quality Risk Management	Application of risk management (including justified and documented risk assessments) through the data life cycle	No documented and justified assessment of risks to data integrity	Limited data integrity risk assessments performed on an ad hoc basis.	Data integrity considered in risk assessment procedures, but not performed to a consistent level	Data integrity risk management established as an integral part of the data life cycle and system life cycle	Quality Risk Management activities subject to continuous improvement.	3
Data Management Processes and Tools	Established data management processes supported by appropriate tools	No data management processes	Some data management processes defined by individual process owners	Data management procedures defined, but not always effectively implemented	Well established and effective data management processes	Well established common data management processes, maintained, updated, supported by appropriate automated tools	5
Master and Reference Data Management	Established processes to ensure the accuracy, consistency, and control of master and reference data	No master/reference data management processes	Some master/reference data management processes defined by individual process owners	Master/reference data management procedures defined but not always effectively implemented	Well established and effective master/reference data management processes	Well established common master/reference data management processes, maintained, updated, supported by appropriate automated tools	4
Data Incident and Problem Management	Established processes to deal with data incidents and problems, linked with change management and deviation management as appropriate	No formal data incident and data problem management process	Some data incident and data problem management processes defined by individual process/system owners	Data incidents and problems typically effectively dealt with as a part of normal system or operational incident management, but with limited consideration of wider data integrity implications	Established data incident and problem management process linked to CAPA and deviation management where necessary	Established data incident and problem management process, supported by tools and appropriate metrics, leading to process improvement	3
Access and Security Management	Establishing technical and procedural controls for access management and to ensure the security of regulated data and records	Lack of basic access control and security measures allowing unauthorized changes	Some controls, but group logins and shared accounts widespread. Password polices weak or not enforced	Established standards and procedures for security and access control, but not consistently applied	Established system for consistent access control and security management, including regular review of security breaches and incidents	Established integrated system for consistent access control and security management, supported by appropriate tools and metrics for continuous improvement	4

Figure 11. Data Integrity Maturity Level Score on Data Life Cycle and Data Life Cycle Supporting Processes (Part 1) (ISPE, 2017).

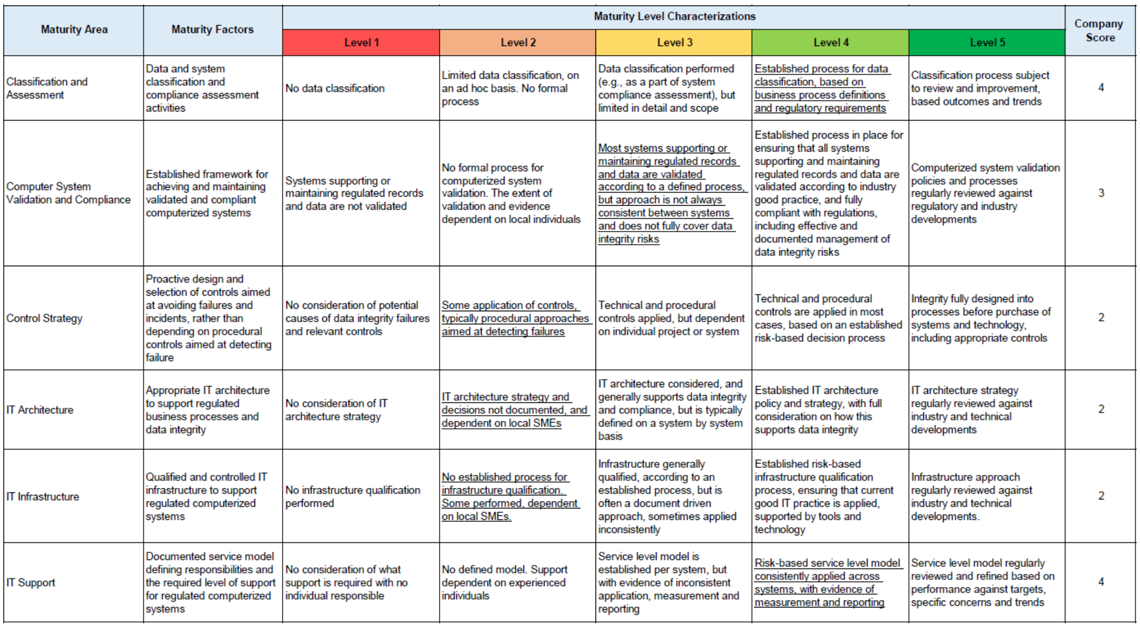
Continuously, Figure 12 shows the Company Score based on the Maturity Factors within the Maturity Areas of Archival and Retention, Electronic Signatures, Audit Trail and Audit Trail Review, Auditing, Self-inspection and Metrics. As some are assessed as sufficient but can still improve, the weakest areas brought up are the Metric, Auditing and Self-inspection specific for any data integrity failures. However, these can only be established once the routine procedures are consistently implemented.



Maturity Area	Maturity Factors	Maturity Level Characterizations					Company Score
		Level 1	Level 2	Level 3	Level 4	Level 5	
Archival and Retention	Establishing processes for ensuring accessibility, readability, and integrity of regulated data in compliance with regulatory requirements including retention periods	No consideration of long term archival and retention periods	No effective process for identifying and meeting regulatory retention requirements. Few archival arrangements in place	Retention policy and schedule defined covering some, but not all regulated records. Some systems with no formal archival process	Retention schedule includes all regulated records, and those policies supported by appropriate archival processes and tools	Archival and data retention policies and processes regularly reviewed against regulatory and technical developments	3
Electronic Signatures	Effective application of electronic signatures to electronic records, where approval, verification, or other signing is required by applicable regulations	No control of electronic signatures	Lack of clear policy on signature application, and lack of consistent technical support for e-signatures	Policies in place. Compliant e-signatures in place for some, but not all relevant systems	Compliant e-signatures in place for all relevant systems, supported by consistent technology where possible	Electronic signature policies and processes regularly reviewed against current best practice and technical developments	4
Audit Trail and Audit Trail Review	Usable and secure audit trails recording the creation, modification, or deletion of data and records, allowing effective review either as part of normal business process or during investigations	Lack of effective and compliant audit trails	Some limited use of audit trails. Often incomplete or not fit for purpose (e.g., in content and reviewability). Not typically reviewed as part of normal business process	Audit trail in place for most regulated systems, but with undefined and inconsistent use within business processes in some cases	Effective audit trail in place for all regulated systems, and use and review of audit trail included in established business processes	Audit trail policies and use regularly reviewed against regulatory and technical developments	3
Auditing	Auditing against defined data quality standards, including appropriate techniques to identify data integrity failures	No data quality or integrity audits performed	Some audits performed on an ad hoc and reactive basis, but no established process for data quality and integrity auditing	Data quality and integrity process defined, but audits not always effective and the level of follow up inconsistent	Effective data auditing fully integrated into wider audit process and schedule	Auditing process and schedule subject to review and improvement, based on audit results and trends	2
Self-inspection	Inspection against defined data quality standards, including appropriate techniques to identify data integrity failures	No data quality or integrity self-inspection performed	Some self-inspections performed on an ad hoc and reactive basis, but no established process for data quality and integrity auditing	Data quality and integrity process defined, but self-inspections not always effective and the level of follow-up inconsistent.	Effective data self-inspections fully integrated into wider business processes	Self-inspection process subject to review and improvement, based on results and trends	2
Metrics	Measuring the effectiveness of data governance and data integrity activities	No data related metrics captured	Limited metrics captured, on an ad hoc basis	Metrics captured for most key systems and datasets. Level, purpose, and use inconsistent	Metrics captured consistently, according to an established process	Metrics captured consistently, and fed into a continuous improvement process for data governance and integrity	1

Figure 12. Data Integrity Maturity Level Score on Data Life Cycle and Data Life Cycle Supporting Processes (Part 2) (ISPE, 2017).

Continuously, Figure 13 shows the Company Score based on the Maturity Factors within the Maturity Areas of Classification and Assessment, Computer System Validation and Compliance, Control Strategy, IT Architecture, IT Infrastructure, and IT Support. Few highlighted as strong maturities are the Classification and Assessment and IT Support. On the other hand, the weakest areas are the Control Strategy, IT Architecture, IT Infrastructure. However, these again can only be established once the routine procedures are consistently implemented.



Maturity Area	Maturity Factors	Maturity Level Characterizations					Company Score
		Level 1	Level 2	Level 3	Level 4	Level 5	
Classification and Assessment	Data and system classification and compliance assessment activities	No data classification	Limited data classification, on an ad hoc basis. No formal process	Data classification performed (e.g., as a part of system compliance assessment), but limited in detail and scope	Established process for data classification, based on business process definitions and regulatory requirements	Classification process subject to review and improvement, based outcomes and trends	4
Computer System Validation and Compliance	Established framework for achieving and maintaining validated and compliant computerized systems	Systems supporting or maintaining regulated records and data are not validated	No formal process for computerized system validation. The extent of validation and evidence dependent on local individuals	Most systems supporting or maintaining regulated records and data are validated according to a defined process, but approach is not always consistent between systems and does not fully cover data integrity risks	Established process in place for ensuring that all systems supporting and maintaining regulated records and data are validated according to industry good practice, and fully compliant with regulations, including effective and documented management of data integrity risks	Computerized system validation policies and processes regularly reviewed against regulatory and industry developments	3
Control Strategy	Proactive design and selection of controls aimed at avoiding failures and incidents, rather than depending on procedural controls aimed at detecting failure	No consideration of potential causes of data integrity failures and relevant controls	Some application of controls, typically procedural approaches aimed at detecting failures	Technical and procedural controls applied, but dependent on individual project or system	Technical and procedural controls are applied in most cases, based on an established risk-based decision process	Integrity fully designed into processes before purchase of systems and technology, including appropriate controls	2
IT Architecture	Appropriate IT architecture to support regulated business processes and data integrity	No consideration of IT architecture strategy	IT architecture strategy and decisions not documented, and dependent on local SMEs	IT architecture considered, and generally supports data integrity and compliance, but is typically defined on a system by system basis	Established IT architecture policy and strategy, with full consideration on how this supports data integrity	IT architecture strategy regularly reviewed against industry and technical developments	2
IT Infrastructure	Qualified and controlled IT infrastructure to support regulated computerized systems	No infrastructure qualification performed	No established process for infrastructure qualification. Some performed, dependent on local SMEs.	Infrastructure generally qualified, according to an established process, but is often a document driven approach, sometimes applied inconsistently	Established risk-based infrastructure qualification process, ensuring that current good IT practice is applied, supported by tools and technology	Infrastructure approach regularly reviewed against industry and technical developments.	2
IT Support	Documented service model defining responsibilities and the required level of support for regulated computerized systems	No consideration of what support is required with no individual responsible	No defined model. Support dependent on experienced individuals	Service level model is established per system, but with evidence of inconsistent application, measurement and reporting	Risk-based service level model consistently applied across systems, with evidence of measurement and reporting	Service level model regularly reviewed and refined based on performance against targets, specific concerns and trends	4

Figure 13. Data Integrity Maturity Level Score on Data Life Cycle and Data Life Cycle Supporting Processes (Part 3) (ISPE, 2017).

As a summary from the assessment of the current situation of the Company against the Data Integrity Maturity model, it is crucial to also understand the root cause of why these areas are not yet in the desired maturity level. Having the root cause that are similarly observed among the maturity areas, the common prerequisite of these weaker areas for improvement are the data integrity procedures not being consistently applied.

4.2 Control Areas to Focus for Data Integrity Verification

The following control areas are identified as generally mentioned across different regulatory standards and guidelines such as PIC/S Guidance, EudraLex Volume 4 Annex 11, UK MHRA, US FDA 21 CFR: Part 11, Part 211 and Part 820, APIC Guidance, ISPE GAMP 5, and GDPR. The key data integrity elements are grouped according to its control areas and how it is commonly managed within the system in the industry:

4.2.1 Audit Trail

A system-generated, secure, date and time-stamped electronic record that enables the reconstruction of the sequence of events connected to the creation, processing, alteration, or deletion of an electronic record is referred to as an audit trail, according to the FDA. (FDA, 2016, p.3) The audit trail is a type of metadata that, according to the UK MHRA, comprises details about actions related to the creation, modification, or deletion of GXP records. Without obscuring or overwriting the original record, an audit trail enables secure recording of life-cycle characteristics like creation, additions, deletions, or revisions of information in a record, whether it be electronic or paper-based. Regardless of the media used to store the record, an audit trail makes it plausible to reconstruct the historical context of such events, capturing the "who, what, when, and why" of the action. (MHRA, 2018, p.13)

PIC/S (2021) states that the capability function of audit trail must be demonstrated and substantiated during the computerized systems validation to verify that transactions for deletions and changes in the GxP critical data are corresponding to the manual activities recorded and also manages to establish ALCOA+ principles. (PIC/S, 2021, p.45)

Continuously according to the PIC/S Guideline (2021), the functionalities of audit trail should be enabled and restricted for other unauthorized users to delete, modify, or deactivate. In cases when administrative users have system permissions to have audit trail settings removed, changed or disabled, the audit trail should reflect an additional automated log entry stating that this event indeed happened. (PIC/S, 2021, p.45)

According to the UK MHRA, audit trails including the archived records maintained within the retention period must aid reconstruction of every data handling and data processing activities, irrespective of whether the results are eventually used as a supporting document for regulatory compliance or business requirements. To make sure that the processing parameters are not being changed in order to obtain a more desirable outcomes, it should be apparent if the data processing has been repeated with deliberate alteration of the processing settings. (MHRA, 2018, p.11).

Additionally, according to UK MHRA, the access of system administrator must be limited to the least reasonable group of individuals, taking into consideration the complexity, nature, size and structure of the organization. It shouldn't be possible to utilize the default system administrator account on a regular basis. Individuals provided with access to system administrator role should log in using an identifiable credentials that permits the actions in the audit trail(s) to be traceable towards a particular person. The purpose is to avoid granting access to users who have a conflict of interest and allowing them to make changes outside their authority that would be difficult to trace back to them. (MHRA, 2018, p.17)

FDA 21 CFR Part 11 states under the controls for closed systems that individuals who use it for creation, modification, management, or transfers of electronic records must implement controls and procedures intended to preserve the truthfulness, data integrity, and necessarily, the confidentiality of electronic records, to ensure that the signer cannot effortlessly denounce that the authorized record is not authentic. Similarly the same controls are required for open systems with the exception of the confidentiality of electronic records, which must be kept from the point of their creation to the point of their receipt. (FDA, 2023).

In continuation according to FDA 21 CFR Part 11, these controls within closed systems must, at the very least, be able to produce complete and accurate copies of records that are understandable and intelligible, and are in an electronic format appropriate for reviews and inspection by regulatory authorities. (FDA, 2023)

FDA 21 CFR Part 11 further states that there should be a control for a system-generated and secure audit trails to objectively capture the date and time of the user's actions and data inputs when creating, modifying, processing or deleting electronic records. Data previously recorded shall not be obscured or overwritten by the subsequent record revisions. These records of audit trail must be kept at a minimum required period until the subject electronic records must be accessible to the agency for examination, investigations and replication. (FDA, 2023)

FDA 21 CFR Part 11 also states that appropriate controls for the use of systems documentation consist of (1) sufficient controls over the dissemination of, permissions to, and usage of system operation and maintenance documentation (2) change control processes and any revisions to keep chronological records of system documentation in its development and modification phase. (FDA, 2023)

EudraLex states that there must be consideration provided on a risk-based assessment to integrating the system-generated audit trail function for the transactional records of deletions and changes that may have a GxP impact. The reason for any deletion or modification of data with GxP relevance needs to be recorded. The ability to transform audit trails into a widely understandable format and regular review are both necessary. (EudraLex, 2011, p.4) EudraLex further promotes the idea of an audit trail by indicating in its guideline that the traceability data must be maintained as documents available for audit. It is appropriate for them to be stored separately from the batch processing record as long as they are easily accessible and clearly associated to the relevant medical product. The storage system must make sure that access to traceability data is available in any adverse reaction that a patient may encounter. (EudraLex, 2017, p.37)

As stated in FDA guidelines, FDA suggests that the audit trails covering the changes to GxP critical data should be reviewed with every record prior to final approval of the record. Audit trails that are regularly examined should document the change history of the final product test results and findings, sample identification, sample run sequences, and significant process parameter changes, among other things. Based on the complexity of the system and its intended use, the FDA encourages periodically scheduled audit trail reviews to be performed based on GxP risks. (FDA, 2016, p.6)

4.2.2 Electronic Records

The regulated user must conduct a risk assessment for the purposes of identifying all the electronic data that have relevance to GMP/GDP including its criticality, which are created and managed by the computerized systems, with respect to the PIC/S Guideline (2021). When determined, the GxP critical data must be audited and verified by the regulated user to ensure that operations were carried out correctly and to check for any alterations such as modification, deletion, or overwriting to the original data in electronic records or the generation of any pertinent unreported data. Modifications to GxP data should be appropriately authorized. (PIC/S, 2021, p.49)

As described by the MHRA, the initial or origin source recording of information or data, such as the electronic raw data file from a computerized system or the original paper record of a manual observation, as well as all the follow-up information necessary to completely recreate or reconstruct the conduct of the GXP activity. Static or dynamic original records are both acceptable. A static record format is one that is stationary, fixed and permits little to no interaction between the user and the record content, such as a paper or electronic record. For instance, once chromatographic records are printed or converted to static electronic formats, they can no longer be reprocessed or used to view baselines at a deeper depth. Electronic records, for example, provide for a dynamic connection between the user and the record's content. Consider electronic

For instance, electronic records maintained as database formats allow the user or reviewer with the necessary access permissions to track, trend, reprocess, query and expand the baseline of the data to view the integration with more clarity such as chromatography data maintained as electronic records. (MHRA, 2018, p.12)

The MHRA states that the risks and their mitigation should be documented where it is not feasibly reasonable to retain the original copy of source data, (e.g. MRI scans, where the source machine is beyond the study sponsor's control and the operator has limitations to only provide summary statistic reports). When recording data that necessitates manual observation (such as the results of a manual titration or the visual interpretation of environmental monitoring plates), the process should be risk assessed based on its criticality. Depending on how crucial the result is, this will determine whether a second contemporaneous verification check is necessary or whether it can be recorded using another method. (MHRA, 2018, p.12)

4.2.3 Electronic Signatures

Electronic signatures are regarded as the legally binding equal of handwritten signatures, according to FDA's answer regarding their use in place of handwritten signatures. Companies that utilize electronic signatures are urged to monitor the security measures employed to guarantee that they can pinpoint a particular individual who signed the documents electronically. (FDA, 2016, p.8)

The FDA claims that electronic signatures with proper controls can be utilized as an alternative to initials or handwritten signatures in every CGMP regulated record. (FDA, 2016, p.8). The FDA added that an electronic signature that complies with this criteria has the mandatory procedures for securely linking the signature with the related electronic record. This is consistent with part 11, which specifies the conditions under which electronic signatures are comparable to handwritten signatures in terms of legal force. Companies that utilize electronic signatures should keep track of the security measures taken to guarantee that they can pinpoint the particular individual who signed the documents or other GxP records electronically. (FDA, 2016, p.8)

Under the section on linking signatures to records, US FDA 21 CFR Part 11 states that electronic signatures and handwritten signatures carried out for electronic records must be associated to their corresponding electronic records in order to protect signatures from being copied, excised, or otherwise reused and transmitted to falsify an electronic record by any means possible. (FDA, 2023)

The electronically signed records must clearly indicate the name of the user who signed, the date and time when the electronic signature was executed, and the purpose or reason behind the electronic signature, namely authoring, reviewing, or approving the document. This requirement is based on US FDA 21 CFR Part 11 under the Signature manifestations section. These formerly indicated specifications for signature manifestations are also needed to comply with the rules and regulations for electronic records, including readability in the display of user interface along with any reports and generated printouts. (FDA, 2023)

The general standards for electronic signatures in US FDA 21 CFR Part 11 state that every electronic signature must be unique to an individual and must never be reused by, or reassigned to, anyone else. (FDA, 2023)

The US FDA's 21 CFR Part 11 also states that prior to an organization establishing, assigning, certifying, or otherwise permitting a person's electronic signature, or any components of such electronic signature, the organization must verify the identity of that specific individual. (FDA, 2023)

In addition, the US FDA 21 CFR Part 11 states that individuals who are using electronic signatures must, beforehand or at the point of time of use, certify to the agency that the electronic signatures performed within their system on or after August 20, 1997, have been intended to as legally binding equivalent of the traditional handwritten signatures. A conventional handwritten signature must be used to sign certification, which may either be submitted electronically or on paper. On the FDA website's page on Letters of Non-Repudiation Agreement, information regarding where to send the certification is indicated. At the request of agency, people who use electronic signatures must produce further certification or evidence, establishing that a particular electronic signature is equal and legally binding of the signer's handwriting signature. (FDA, 2023)

Electronic signatures that are not based upon biometrics must use at least have two distinct identification elements, such as user identification code and password according to US FDA 21 CFR Part 11 under the section on electronic signature components and controls. It should only be used exclusively by their actual owners. In addition, it should only be administered and carried out so that any effort to use an individual's electronic signature by anybody other than its authentic owner requires a minimum of two individuals to collaborate. (FDA, 2023)

More specifically when using a minimum of two distinct identification components in accordance with US FDA 21 CFR Part 11, for cases when an individual executes a series of signings in one, continuous period of controlled system access, the initial signing must be carried out using all electronic signature components; the following signatures must be executed using at least one electronic signature component and can only be executed by, and configured to only be used by, the assigned individual to electronically sign. (FDA, 2023)

Furthermore, in accordance with US FDA 21 CFR Part 11, in cases when electronic signatures are not performed in one, continuous period of controlled system access, each signing must be performed using all the electronic signature components by the designated individual. (FDA, 2023)

Additionally, in accordance with US FDA 21 CFR Part 11, electronic signatures with the usage of biometrics must have security mechanisms to ensure that this is not allowed be used by anyone other than their rightful owners. (FDA, 2023)

It is recommended by EudraLex that document and other data management systems must be configured to record the identity of the users who are tasked to enter, change, review, confirm, approve or delete data, including the date and timestamp. EudraLex continues by stating that electronic records may be electronically signed within the confines of the company, electronic signatures are supposed to have the same force as handwritten signatures and be inextricably connected to the relevant record with the inclusion of the date and time that they were executed. (EudraLex, 2011, p.4)

A digital signature whether it is biometric or non-biometric reflects the signatory, as per the MHRA. In legal terms, this ought to be comparable to the signer's handwritten signature. (MHRA, 2018, p.14)

In addition, according to MHRA (2018), appropriate controls on the use of electronic signatures should have considerations including how signature is attributed to a specific individual, how the signing procedure is recorded in the computerized system to ensure that it cannot be manipulated or modified without invalidating the status or signature of the entry, how the signatures are linked to the relevant entry records created including its ways of verification checks, and lastly the security measures on ensuring that electronic signatures can only be applicable to the actual user account registered on it. (MHRA, 2018, p.14-15)

It is required, according to the UK MHRA, that appropriate verification of the signatures processing managed in a GxP computerized system shall be performed to demonstrate the suitability and oversight of signed records is maintained. The metadata linked to an electronic signature should be preserved with the accompanying document or records whenever a paper or PDF copy of an electronically signed document is made. (MHRA, 2018, p.15)

The usage of electronic signatures must be compliant with the international standards requirements according to UK MHRA. Where the risk assessment specifies this technique of authentication, the usage of a more advanced electronic signatures must also be taken into account. The term "signature manifestations" refers to a display within the viewable record that indicates who signed it, their title, the date (and time, if appropriate), and the purpose of the signature (e.g., validated, qualified, reviewed or approved). E-signature systems or electronic signatures must support these required signature manifestations. (MHRA, 2018, p.15)

Also according to UK MHRA, an uploaded image of a footnote or signature which denotes that the document has already been electronically signed, by a means other than the validated electronic signature procedures, is not acceptable and will require additional controls. If a document is electronically signed, the signature's related metadata must be stored within the required retention period. (MHRA, 2018, p.15)

4.2.4 Security and User Access Controls

The Security and User Access Controls should generally be established as a best practice to help manage overall data integrity compliance. FDA guidelines (2016) suggest that that your restrictions must be in place to control the ability to modify specifications, process parameters, manufacturing or test methods and techniques if the system settings and functions allows it, for instance, by granting permissions to change GxP data to only authorized and trained users. FDA advises designating a separate system administrator job, along with any permissions to change files and settings, apart from those employees in charge of record contents. For each CGMP computer systems used in operation, the FDA suggests keeping a list of approved users and their access credentials to aid with the security access controls. (FDA, 2016, p.5)

FDA (2016) guidelines also instructed that alternative strategies for security controls must be implemented if a separate and independent security role assignments are not feasible for small-scale enterprises with operations and limited staff, such as medical gas facilities. For instance, FDA advises having a second person evaluate settings and content in the exceptional circumstances where the same individual is required to hold the system administrator role and be accountable for the records' contents. The Agency advises the person to double-check settings and their own work if a second-person assessment is not available (FDA, 2016, p. 5).

Procedures and controls related to security of electronic records are listed in FDA 21 CFR Part 11 under the controls for closed systems. Examples are including limited system access to authorized individuals, utilizing operational system checks to enforce permitted sequencing of events and actions, as per suitable situations. Furthermore, employing the authority checks to verify that only authorized individuals can access the computerized system input or output device, use the systems functionalities, sign records electronically, modify a record, or perform the operational tasks and activities. (FDA, 2023)

People who use electronic signatures via user credentials and passwords must make use of controls to assure information security and data integrity, according to US FDA 21

CFR Part 11 under Controls for identification codes/passwords. According to US FDA 21 CFR Part 11, this involves maintaining the distinctive features of the coupled user identification code and password, such that no two persons have the same combination of user identification code and password. (FDA, 2023)

US FDA 21 CFR Part 11 states that securing that identification code and password issuances are routinely checked, revised or recalled, for situations to cover such events as password aging, is a continuous requirement for the aforementioned controls. (FDA, 2023)

Another control also mentioned in US FDA 21 CFR Part 11 is defined as adhering to loss management procedures to electronically deactivate and deauthorize any missing, misplaced, stolen or otherwise potentially compromised key cards, security tokens, and other devices that contain or produce information for user verification code or passwords, and to issue temporary or permanent replacements using robust and adequate controls. (FDA, 2023)

In addition to this, US FDA 21 CFR Part 11 listed controls including the use of transactional protection measures to avoid unauthorized use of passwords and user identification codes, and to monitor, detect, notify and report in an urgent manner and immediate timing for any attempts at accessing, or unauthorized systems to security unit and to organizational management as appropriate." (FDA, 2023)

And lastly for relevant control listed in US FDA 21 CFR Part 11, there should be an initially and periodically performed testing of devices, such as security tokens or key cards, which contain or produce user identification code or password information in order to verify that they are still operating and functioning as intended and have not been manipulated in any unauthorized manner." (FDA, 2023)

Physical and/or logical restrictions should be in place, according to EudraLex, to restrict access to computerized systems to registered users. The use of keys, access cards, uniquely personal codes with passwords, biometrics, or access restrictions to data storage locations and computer equipment, hardware, and other assets are some of the suitable methods for preventing unauthorized entry to the systems. The criticality of the computerized system determines the scope of security controls. (EudraLex, 2017, p.30) EudraLex emphasizes that there must be controlled and documented records for creation, modification, termination or cancellation of access authorizations. Data and document management systems should be configured to monitor and identify the individual users or anyone who enter, modify, confirm, or delete data, as well as the date and time. (EudraLex, 2011, p.4)

Furthermore, EudraLex states that formal agreements must be in place between the manufacturer and any third parties, and such arrangements must contain explicit definitions of the third party's responsibilities, when such third parties (for example, suppliers or service providers) are used to perform tasks like providing, installing, configuring, integrating, validating, maintaining (e.g. via remote access), modifying, or retaining a computerized system or related service, or for data processing. (EudraLex, 2011, p.2)

Pursuant to UK MHRA, the entirety of access controls utilization to ensure that individuals have access only to functionalities that are appropriate for their job position, and that activities are attributed to a particular person. Companies must be able to show the access levels given to specific employees and guarantee that historical information about the granted user access levels remains accessible. If the system is unable to record this information, an external record must be maintained. Both the operating system and application levels should have access controls. If adequate safeguards are in place to preserve data integrity (e.g., no modification, deletion, or creation of data outside the application is allowed), then individual login at the operating system level may not be necessary. (MHRA, 2018, p.16)

Apart from that, according to UK MHRA, generic user access or shared logins must not be utilized for shared logins computerized systems in generating, modifying, or storing GXP data. This function must be used when the system configuration allows for individually accounted user access. It might be necessary to purchase more licenses for this. Systems (like MRP systems) that are partially used for GXP purposes yet contain GXP-applicable features like approved vendors, inventory status, location, and historical transactions must undergo the proper assessment and control. In case of single or a limited number of user logins supported by some computerized systems, this is still acknowledged. Whenever there are no adequate alternative computerized system available, a third-party software or a paper-based method may provide equal controls such as historical versioning, traceability and audit trails.

Alternative systems' suitability needs to be demonstrated and supported by evidence. Given that hybrid systems are susceptible to non-attributable data changes, increased data review is essential. Companies are expected to be installing or using systems that meets or complies current regulatory requirements. (MHRA, 2018, p.16-17)

The UK MHRA consistently advises that system administrator access must be restricted to the most limited number possible taking into account the nature, size and structure of the organization. It shouldn't be possible to utilize the default system administrator account on a regular and operational basis. Individuals having system administrator access should log in using unique credentials that enable actions in the audit trail(s) to be associated with a particular person. The purpose of this is to avoid granting access to users who might have a conflict of interest so they cannot make unauthorized changes that might not separately identify them. Individuals with a direct interest in the data such as generation, reviews, confirmation and approval, should not be given System Administrator permissions, which allow for actions like data deletion, database revision, or system configuration changes. Depending on the state of clinical study data, some people may need their access rights updated. For instance, after data management procedures are finished, the records and data are locked by removal of modification access permissions. Within the system, this ought to be verified and demonstrated. (MHRA, 2018, p.17)

4.2.5 Backup, Disaster Recovery and Archival

According to Practical risk-based guide for managing data integrity released by APIC, all GxP-related data must have formal data backup procedures and methods that must be established, documented, validated, and routinely tested based on the requirements of the business, to which backup storage frequency shall be determined. (APIC, 2019, p.21)

In continuation according to APIC, system must have a written archival strategy in place. If system modifications have an impact towards the capability to read or process existing files, GxP data and related meta data must be archived. At system decommissioning, GxP data must be archived for compliance. (APIC, 2019, p.23)

According to FDA (2016) guidance, the backup file should be in a format that is consistent with the original structure and compatibility with the source format, and should contain the data including any associated metadata. Contrast this with backup copies, which might be made during routine computer usage and temporarily stored for disaster recovery purpose (for instance, in case of a computer breakdown or other disruption readiness). (FDA, 2016, p.4)

In accordance with US FDA 21 CFR Part 11, individuals who use closed systems to create, modify, maintain, or transmit electronic records must use protocols and safeguards intended to ensure the authenticity, integrity, and, when necessary, the confidentiality of those records, as well as to make sure the signer cannot easily retract the signed document as being a counterfeit. (FDA, 2023) The US FDA's 21 CFR Part 11 has consistently incorporated protocols and verification, such as protecting documents to allow for their prompt and accurate retrieval of records throughout the data retention period. (FDA, 2023)

Backup is the process of replicating records, data, configurations, and software in order to protect the systems against a possible unavailability of compromise of integrity of the original records, according to ISPE GAMP 5 (2022). When the need arises, restore is initiated to the process of recovering the documents, data, configurations, or software. Archiving and retrieval processes must not be confused with backup and restore. DR is supported by backup, but access to records in a readable and understandable format within the records retention term is supported by archives. (ISPE, 2022, p.319)

According to ISPE GAMP 5 (2022), procedures should specify the backup and restoration strategy, as well as how backup failures will be handled. Technology for backups should be based on company requirements. Scheduling of backup must align with disaster Recovery Point Objective (RPO). To ensure proper operation, backup technologies and procedures should be periodically tested. Storage locations for backups should be kept independently of the main location. Geographical separation must be risk-based and take into account environmental dangers like storms, hurricanes, earthquakes and other natural disasters. (ISPE, 2022, p.319)

Furthermore, according to ISPE GAMP 5 (2022), archiving is the process of relocating documents or data from a computerized system to another secured site or system, frequently safeguarding them from future alterations. Records that have been archived should still be retrievable for legal, compliance or business needs. It is acceptable to use cloud storage options for archived records. (ISPE, 2022, p.341)

Moreover, according to ISPE GAMP 5 (2022), GxP documents and data must be protected against unauthorized or unintentional access, alteration, or deletion for the duration of the mandated retention period using physical and/or logical security measures. The archiving procedures should make sure that the record's content and meaning remain intact, alongside the electronic signature, audit trail, and other metadata necessary for comprehending the record. (ISPE, 2022, p.341)

The maintained records and data must be initially and re-checked for accessibility, endurance, readability, and completeness, according to ISPE GAMP 5 (2022). During an inspection, regulators should have appropriate access to archived GxP documents in a timely manner. (ISPE, 2022, p.341)

Data must be protected against loss, theft, and damage through both physical and electronic measures, according to EudraLex. It is important to verify the readability, accuracy and accessibility of stored data. Data should be available to access at all times during the retention period. All relevant data should be periodically backed up. The integrity, accuracy, and ability for restoration of backup data must be examined during validation and routinely monitored. (EudraLex, 2011, p.3)

Additionally, EudraLex establishes that in order to maintain the availability of computerized systems supporting crucial operations, arrangements should be made through manual or alternative methods to guarantee the continuity of support for those processes in unlikely event of a system breakdown. The amount of time needed to implement the alternative arrangements should be based on risk and suitable for the system in consideration and the business process it supports. These arrangements need to be properly evaluated and documented to verify the data integrity, accessibility, and readability. The capability of retrieving the data should be assured and evaluated if pertinent changes are going to be made to the system (such as to equipment, hardware or software). (EudraLex, 2017, p.5)

Data retention may be used for archiving secured data for long-term storage and compliance or backup data for disaster recovery and restoration, according to UK MHRA. Arrangements for document retention should protect data from intentional or unintentional loss or manipulation. Security policies and procedures must be in place and, validated when appropriate during data transfers/migration in order to protect the integrity of the records during the retention period. Provided that there is an established procedure in place to ensure that the result is a true copy generated in paper format may be maintained by employing a validated scanning technique. Procedures for data destruction should take into account the criticality of the data and any applicable statutory retention requirements. (MHRA, 2018, p.17)

The UK MHRA consistently asserts that a designated secure space or facility such as a cabinet, room, building, or computerized system is conducive for the long term, retention of data and metadata for the intention of verifying operational activities or procedures. (MHRA, 2018, p.18)

Furthermore, according to UK MHRA, archived documents may be the original record or a "true copy" and should be safeguarded to prevent unauthorized eradication or modification in addition to unintentional harm from pests and fire. Archive arrangements must be created so that data and metadata allow retrieval and readability throughout the needed retention time. To confirm the ongoing maintenance of legacy computerized systems, the process of archiving electronic data should be certified, and for historical systems, the capacity to periodically examine data should also be verified. References between physical and electronic records must be kept where hybrid records are preserved so that complete event verification is feasible for the duration of the retention term. (MHRA, 2018, p.18)

According to UK MHRA, when legacy systems can no longer be supported, maintenance of the software must be updated for data accessibility needs (for as long as is practical based on the particular retention requirements); a virtual environment could assist with this. With the legacy data is getting more obsolete, migration to an alternate file format that preserving the 'true copy' features of the data may be necessary. Options should be evaluated based on risk and the long-term significance of the data where migration with full original data functionality is not physically feasible. File format should be chosen taking into account during migration when the risks of balancing the lowered dynamic data capability (such as data interrogation, trending, re-processing, etc.) against the long-term accessibility. It is acknowledged that switching to a file format that loses some features and/or the capability of dynamic data may be necessary to retain accessibility. (MHRA, 2018, p.18)

A copy of the most recent (editable) data, metadata, and system configuration settings should also be retained for disaster recovery, according to UK MHRA. Processes for backup and recovery should be verified and tested on a regular basis. Each backup should be checked to make sure it is working as intended, for example, by matching the data size transmitted with the original record. Backups for disaster recovery do not eliminate the need for a long-term, permanent storage of data and metadata in its final state for the purposes of process or activity verification. (MHRA, 2018, p.18)

All necessary records must be maintained at the manufacturing establishment or other location reasonably accessible to authorized officials of the manufacturer and to FDA employees designated to conduct inspections, according to US FDA CFR Part 820 under the General Requirements for Records. These documents, including those not kept at the inspected location, must be easily accessible for FDA employees to examine and copy. These documents must be legible and stored to reduce deterioration and prevent loss. It is still essential to back up any records kept in automated data processing systems. (FDA, 2023)

Continuously based on US FDA CFR Part 820, the manufacturer may identify which of the data may be considered as public information to support FDA in determining the limitations in disclosures. All records required by this part must be kept for the time specified in the device's design and estimated lifespan, but in no less than two years after the manufacturer's release of the product for commercial distribution. (FDA, 2023)

For exceptional cases according to US FDA CFR Part 820, an employee with executive responsibility must certify in writing that the management reviews, quality audits, and supplier audits required by this part have been carried out, documented, on the dates that they were performed, and that any necessary corrective action has been taken, upon request from a designated employee of the FDA. (FDA,2023)

In terms of IT Suppliers and Service Providers, MK UHRA recommends that care should be taken to understanding the service provided, ownership, retrieval, retention, and security of data where "cloud" or "virtual" services are employed. It is important to account for the actual data location, including any regulations that may have an impact. A technical agreement or contract should specify the obligations of the contract giver and acceptor. As a result, the data owner and national competent authorities shall always have immediate access to data (including audit trails and metadata) upon request. Contracts with service providers should specify who is responsible for data archiving and readability during the retention period. (MHRA, 2018 p.19)

Additionally, according to UK MHRA, suitable measures must be in place, including validation and change control information, to enable the restoration of the program or system to its initial validated condition. Contracts should have provisions for business continuity that have been tested. Risk should be used to determine whether the service provider needs to be audited. (MHRA, 2018 p.19)

According to US FDA CFR Part 211 for automatic, mechanical, and electronic equipment, suitable controls must be used over computer or associated systems to ensure that only authorized employees implement changes to master production and control records or other records. The accuracy of all data input into and output from the systems, any formulas connected to it, and any records or data must be verified. The level and frequency of input/output verification must take the computer system's reliability and complexity into account. Except in cases where specific data, such as calculations made in connection with laboratory analysis, are eliminated by computerization or other automated procedures, a backup file of data entered into the computer or associated system must be kept. In such cases, a documented record of the software must be maintained along with the necessary validation information to ensure that backup data are accurate and complete and that they are safe against modifications, unintentional erasures, or loss, hard copy or alterations. (FDA,2023)

4.2.6 Data Protection and Data Privacy

Generally as best practice, a data protection impact assessment is mandatory whenever the processing involves a high risk in terms of a person's rights and freedoms, e.g. when new technologies are used. Good Documentation Practices, Data Governance, Data Lifecyle Management and Periodic Review are already managed in separate procedures in the Company and despite having an effect on data integrity compliance, these are deemed not required to be topics for revision or improvements, and will remain as is.

According to European Union (EU) for General Data Protection Regulation (GDPR) specifically under Article 6 that processing of personal data can be only be lawful if “the data subject has given consent to the processing of his or her personal data for one or more specific purposes”. (EU GDPR, 2023)

Additionally, lawful processing based on EU Guideline for GDPR is applicable when “processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract", "processing is necessary for compliance with a legal obligation to which the controller is subject", "processing is necessary in order to protect the vital interests of the data subject or of another natural person" and "processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller". (EU GDPR, 2023)

In continuation according to EU Guideline for GDPR, "processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child." (GDPR, 2023) This is provided that this is not applicable to processing conducted by public authorities as part of their operational tasks. (GDPR, 2023)

4.3 Approach to Hybrid Situations

PIC/S Guideline (2021) states that hybrid systems need particular and supplementary controls because of their complexity and increased susceptibility to data manipulation. Due to this, hybrid systems should be avoided wherever possible, and instead should be replaced when the company maturity allows it. (PIC/S, 2021, p.52) Some old computerized systems tend to have multiple limitations and the only option for controls are paper-based records and management, for example due to incomplete or lack of audit trail functionality, changes can only be monitored through physical log book that are required to be periodically reviewed and checked for good documentation practices and inconsistencies.

In case of hybrid situations, the manual processes are more prone to risks of falsification and unauthorizes changes since system controls are established as preventive controls. Attention should be given to the interaction between the manual and computerized system, according to PIC/S Guideline (2021). Due to the challenges in applying a manual process consistently, careful tracking must be placed on verifying: 1) The extent of qualification and/or validation of the computerized system; and, 2) The robustness of controls applied to the management of the manual element of the hybrid system. (PIC/S, 2021, p.53).

4.4 Conceptual Framework

To introduce the Conceptual framework, Figure 14 is significant to support the context of the key areas of improvements for data integrity compliance procedures where most of this are to be performed prior the release of the system in the Project phase.

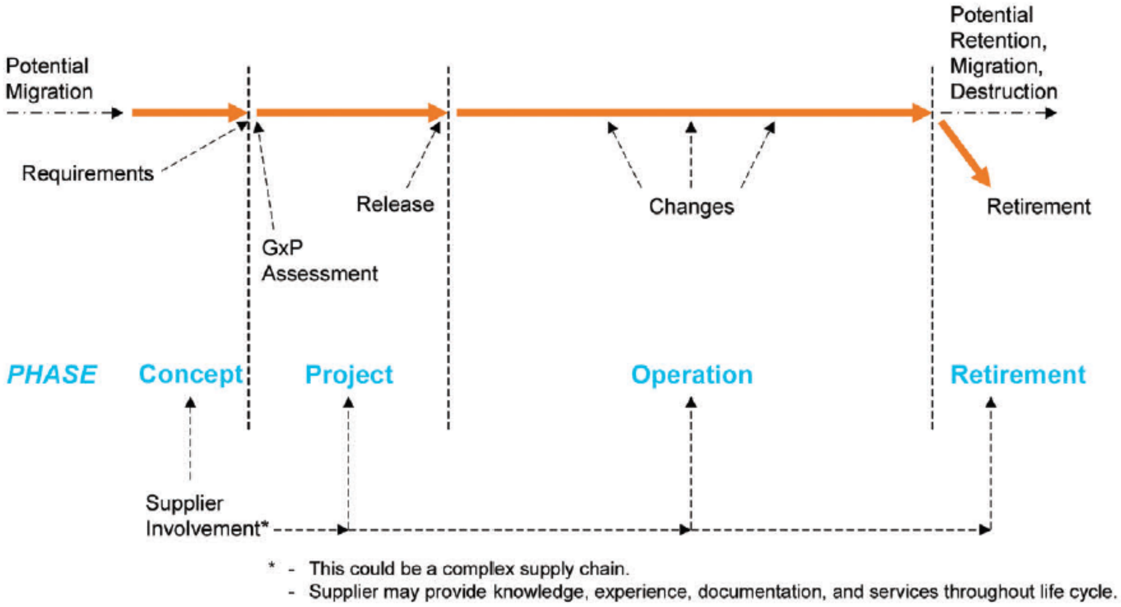

Potential Migration
Requirements
GxP Assessment
Release
Changes
Potential Retention, Migration, Destruction
Retirement
PHASE
Concept
Project
Operation
Retirement
Supplier Involvement*
* - This could be a complex supply chain.
- Supplier may provide knowledge, experience, documentation, and services throughout life cycle.

Figure 14. Computerized System Life Cycle Phases (ISPE, 2017).

As shown in Figure 14, the different life cycle phases of a computerized systems give readers a bird’s eye view of where the data integrity compliance procedure lies. The data integrity assessment is mainly performed during the Concept phase to which requirements are built. Subsequently, the data integrity verification is heavily performed during the Project phase where computerized systems are validated prior releasing to operational use. The changes within the Operational phase still have some data integrity elements even towards until the retirement such as readability checks. However, these are removed from the scope for the purposes of defining the focal point of this study.

Figure 14 shows the Data Integrity Compliance Framework.

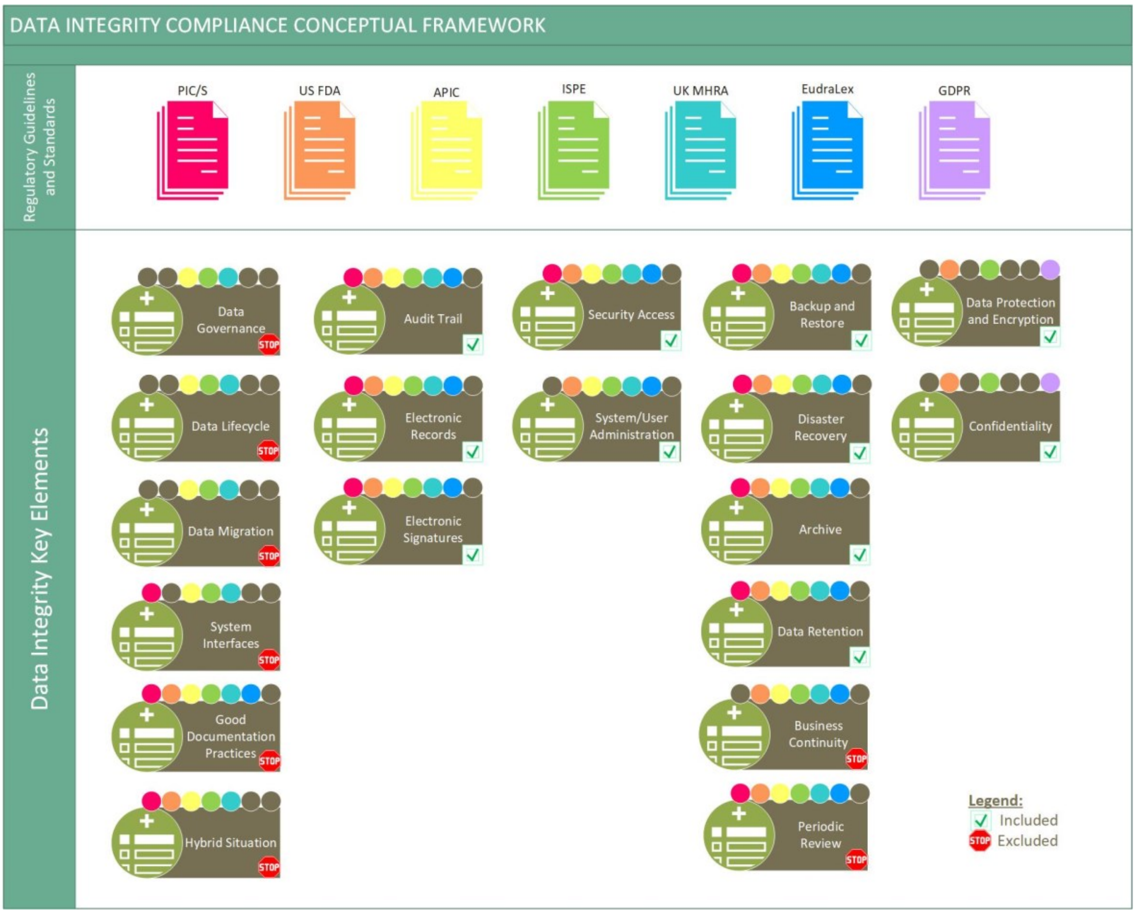


Figure 15. Data Integrity Compliance, Conceptual Framework.

As shown in Figure 15, the colors of the source on each element meant that these are acknowledged as regulatory requirements commonalities from the regulatory guidelines and standards such as PIC/S, US FDA, APIC, ISPE, UK MHRA, EudraLex and GDPR. These all pointed into several data integrity key elements, thus it only showed that the more it was mentioned, the more grounded it can be as requirements thus can also be routinely formulated. Some of these elements were removed from the thesis research due to 1) out of scope areas that may expound the study in a larger scale, or 2) identified as strengths and are part of separate processes. Therefore, the ones that are labelled as included will be the CF that summarizes the directions for improvement efforts into the following areas: Audit Trail, Electronic Records, Electronic Signatures, Security Access, System or User Administration, Backup and Restore, Disaster Recovery, Archival, Data Retention, Data Protection and Encryption and Confidentiality.

5 Building Proposal for Data Integrity Compliance Procedures for the Company

The current state analysis brought up the areas for improvement in the data integrity verification process — to keep the strengths unchanged and to replace the weaknesses into a more defined and understandable approach. Guided with the best practice from regulatory standards and the existing knowledge from multiple companies in the same industry, the underlying requirements for data integrity controls had been repeatedly mentioned and worthy of establishing in the process. The key points are narrowed down and explained in the following sections.

5.1 Overview of the Proposal Building Stage

This section presents the steps in the Proposal building for this study which focuses on the improvements to the data integrity compliance procedures when validating computerized systems. The focus for the improvement is to redefine the risk assessment by the inclusion of system's complexity to remove the non-applicable controls instead of exhausting the functionality out of the system, to streamline the process by removing repetitive efforts in the verification steps and documentation, and to restructure the process into a straightforward and easy to follow instructions.

The IT department led by the IT Compliance team brainstorm the proposal together, collectively gathered the data and entrusted the thesis researcher to develop the Initial proposal-1 based on regulatory guidance and best practice (can be found in Section 5.2). The proposed draft was discussed with the other stakeholders and the Initial proposal-2 was formed (described in Section 5.4, based on stakeholder suggestions).

A starting point for the Company on its data integrity maturity level based on ISPE GAMP 5 Guide on Records and Data Integrity is at Level 3 Amber where there is defined policy and established processes, inconsistent application and inconsistent monitoring. Improvements must be applicable to the Company's current maturity to move it to at least Level 4 Green where inconsistencies will be eliminated and there will be routine application and routine monitoring. In order to achieve this improvement, the weakest maturity areas must be considered as an emphasis for changes and maintain

Data integrity risk assessment has to be defined as this is crucial in identifying the applicable controls and level of efforts for each system. System categorization and its alignment to minimum required controls can direct what should really just be the focus of verification testing. It will remove the unnecessary time and efforts to look for controls that the system is not even capable of. Thus, the time and efforts are redirected to identification of risks, formulating the mitigation, managing and monitoring these risks until there's a need to suspend, retire and replace the system overall.

Replacing the ALCOA+ principle-based with the system controls-based verification for efficiency purposes is another opportunity for improvement. Instead of the ALCOA+ principles, the grouping for verification procedures will be more understood if this is based on system controls which is aligned with the concepts mentioned in the regulatory standards. Also in the same terminologies that are more in line with the language that systems can understand, may it be in functionalities or settings.

Another improvement is enabling the users with concrete requirements and instructional procedures to combat lack of direction and guidance. The proposal aims to remove the ambiguities when verification results are acceptable or not. As per observations, there are instances where multiple meetings are scheduled and rescheduled just to go through and answer the clarifications from each user. So instead of a yes or no confirmation checklist for data integrity verification, it should be in an instructional format where there are action steps, expected results as the acceptance criteria, a blank section for documenting the actual result during the test execution and the result column to determine if it meets the test objective as pass and fail if it did not. The procedure must be as lean as possible without any repetitions and the requirement must be clear with its purpose.

Gathered from the discussions with the respondents, the mixture of simple to complex computerized systems is a key information that is missed out in the data integrity risk assessment. APIC has a flowchart framework that is helpful and beneficial to categorize the systems into relevant and applicable areas for data integrity verification. This covers the missing risk assessment of systems' complexity and whether data storage is temporarily or permanently kept.

Table 5 below lists the step-by-step action plan on how to implement and roll out the proposal for the data integrity compliance procedures in the Company:

Table 5. Data Integrity Compliance Procedures – Proposed Rollout Plan.

Data Integrity Compliance Procedures Proposal Rollout Plan	
First	Risk assessment will be changed according to the system categorization provided by APIC to identify the required data integrity controls. In addition, the applicability of whether there is electronic signatures or not, and whether the system handles personal data for privacy and data protection regulations. To make this easier, the results of assessment will be recorded via checkbox option.
Second	From the selected system category based on complexity, the GxP data criticality should be included within a range of risks of High Medium and Low. This is already assessed via an existing process (SOP-VAL-011) that determines if the system has a direct, indirect or no impact to product quality, thus also to patient safety. Classifying it further into these segmented levels also adapt to APIC system categorization.
Third	All the applicable system controls that are deemed applicable based on the assessment is recorded in a summary table. In cases that there are previously completed data integrity compliance procedures using the current effective version of the templates that will eventually be overwritten by the proposal, there are option for references in the previous assessment to only account and test for the major changes.
Fourth	The control area grouping for data integrity verification is conducive to remove those controls that are not relevant or applicable, and also to assist in selectively testing what needs to be repeated from the previous assessment in case of major upgrades. It is important to build ease in routine procedures and repetition, thus an instructional method is laid out with data integrity requirements as the basis for test objective. In addition, for ease in recording the actual results, recommendation to include references in the operational qualification or performance qualification with checkbox for pass and fail results are included.

Fifth	Room for risks monitoring and IT-dependent controls are encouraged to include further actions when one or more of the tests failed. This can be summarized in the Further Actions section alongside the Summary and Conclusion section, and will be monitored during the completion or during the periodic reviews, whichever comes first.
Sixth	As per advise by the QA manager, conducting an organization-wide training of the new process should be scheduled with the integration of other procedures related to data integrity procedures during the operational lifecycle of the system.
Seventh	The change to operational from project management basis should be implemented. Previously this has been managed as a project as a response to an audit finding which makes this procedure a corrective measure. The goal must be shifted to operational where ownership of the systems and its data integrity controls are practiced, any remaining risks are documented and monitored until the data integrity maturity level allows it. Thus, assistance is offered to the individual(s) after transferring and clarifying the ownership of the responsibilities.

5.2 Proposal Draft-1 based on Regulatory Guidance & Best Practice

This Proposal focuses on the improvements on data integrity compliance procedures for the risk assessment and verification with the supporting Figures below:

- Figures 16 – 24 details the improvements on Data Integrity Risk Assessment

- Figures 25 – 41 details the improvements on Data Integrity Verification

First, Figure 16 shows the system categorization as a flowchart with an aligned guidance on the minimum system requirements per category. Based on the Practical risk-based guide for managing data integrity issued by Active Pharmaceutical Ingredients Committee (APIC) that once the system is identified, it can be further categorized based upon the GxP data that is generated in and by the system (APIC, 2019). The Company has a criticality and risk assessment in place which also analyze the data severity and can be set into High, Medium and Low with High having the most impact to product quality and patient safety.

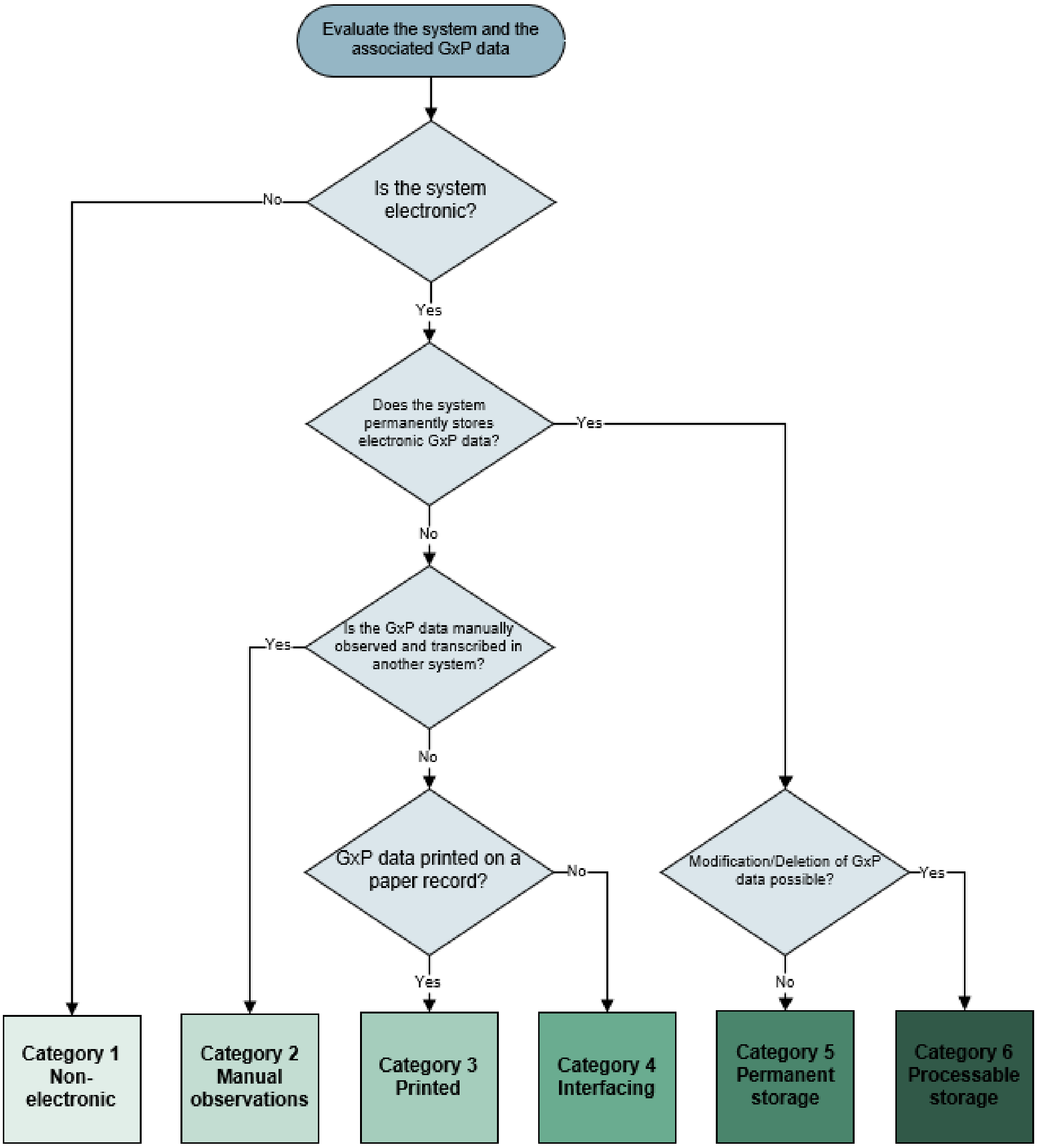
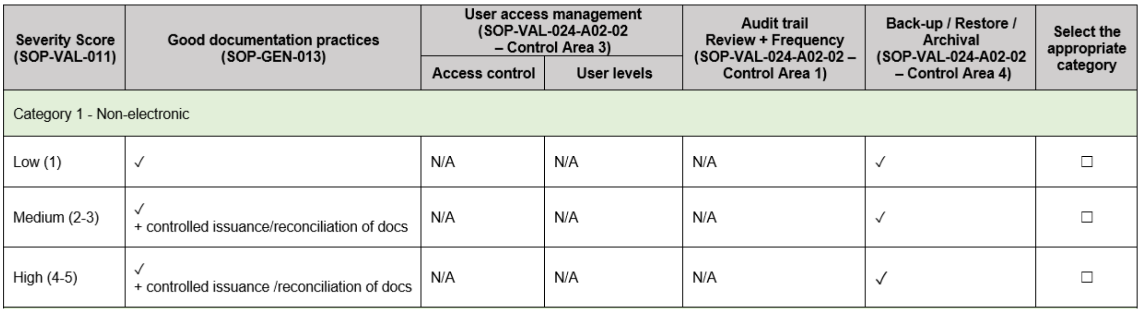

Evaluate the system and the associated GxP data
Is the system electronic?
No
Yes
Does the system permanently stores electronic GxP data?
Yes
No
Is the GxP data manually observed and transcribed in another system?
Yes
No
GxP data printed on a paper record?
No
Yes
Modification/Deletion of GxP data possible?
Yes
No
Category 1 Non-electronic
Category 2 Manual observations
Category 3 Printed
Category 4 Interfacing
Category 5 Permanent storage
Category 6 Processable storage

Figure 16. Data Integrity Risk Assessment: Flowchart for System Categorization.

As shown in Figure 16, on each of the categories that the system will be classified, there are applicable control areas that will be relevant for the data integrity verification procedures such as good documentation practices, user access management both including the access controls and minimum user levels, audit trail review and frequency, backup and restore, and archival.

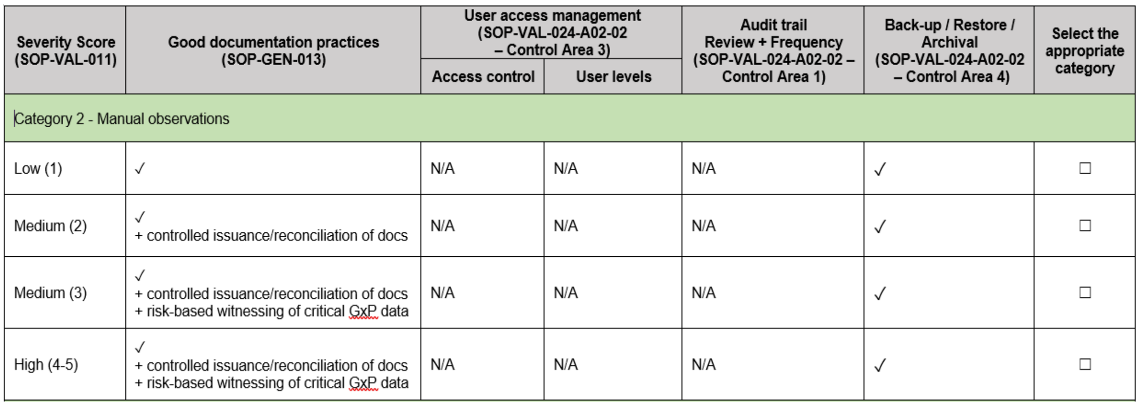
Second, Figure 17 shows the associated controls when the system is assessed as Category 1: “**Non-electronic**” system. In this category, data is purely managed in paper records thus there’s no electronic GxP data stored. The relevant and applicable controls are good documentation practices (GDocP) and periodic review within the retention period and security for archived records. Typical examples are bag sealers, pH paper, density meters, CAPA logbook (APIC, 2019, p.12). Audit trail, user levels and access controls are not applicable.



Severity Score (SOP-VAL-011)	Good documentation practices (SOP-GEN-013)	User access management (SOP-VAL-024-A02-02 – Control Area 3)		Audit trail Review + Frequency (SOP-VAL-024-A02-02 – Control Area 1)	Back-up / Restore / Archival (SOP-VAL-024-A02-02 – Control Area 4)	Select the appropriate category
		Access control	User levels			
Category 1 - Non-electronic						
Low (1)	✓	N/A	N/A	N/A	✓	□
Medium (2-3)	✓ + controlled issuance/reconciliation of docs	N/A	N/A	N/A	✓	□
High (4-5)	✓ + controlled issuance /reconciliation of docs	N/A	N/A	N/A	✓	□

Figure 17. Data Integrity Risk Assessment: Minimum Requirements for System Category 1.

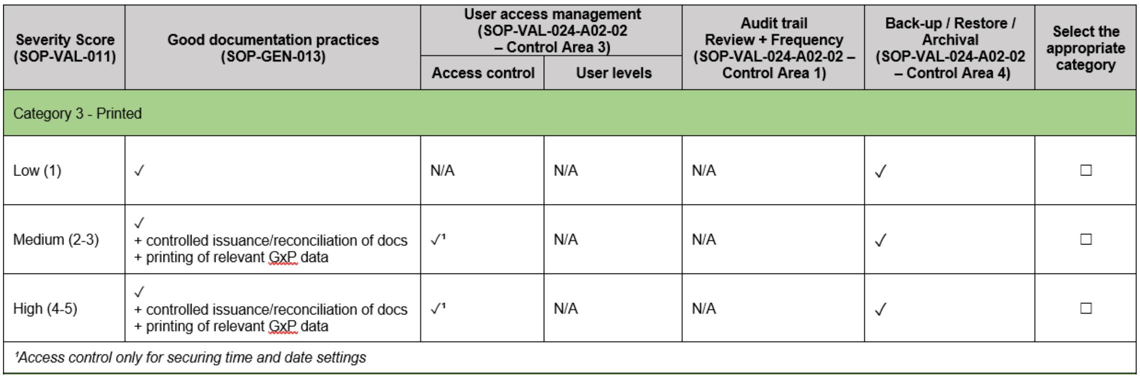
Third, Figure 18 shows the associated controls when the system is assessed as Category 2: “M**anual observations**” system. In this category, data can be in electronic form but it does not stay in the system due to a one-time recording for manual observations or manual transfers on to papers records. Thus the generated GxP data in the system is not stored and goes back to purely manual controls similar to Category 1 and should observe GDocP and periodic review within the retention period and security for archived records. Typical examples could include pH meters, balances, polarimeters with manual adjustable a wavelength, pressure gauge with display (APIC, 2019, p.12). Audit trail, user levels and access controls are not applicable.



Severity Score (SOP-VAL-011)	Good documentation practices (SOP-GEN-013)	User access management (SOP-VAL-024-A02-02 – Control Area 3)		Audit trail Review + Frequency (SOP-VAL-024-A02-02 – Control Area 1)	Back-up / Restore / Archival (SOP-VAL-024-A02-02 – Control Area 4)	Select the appropriate category
		Access control	User levels			
Category 2 - Manual observations						
Low (1)	✓	N/A	N/A	N/A	✓	□
Medium (2)	✓ + controlled issuance/reconciliation of docs	N/A	N/A	N/A	✓	□
Medium (3)	✓ + controlled issuance/reconciliation of docs + risk-based witnessing of critical GxP data	N/A	N/A	N/A	✓	□
High (4-5)	✓ + controlled issuance/reconciliation of docs + risk-based witnessing of critical GxP data	N/A	N/A	N/A	✓	□

Figure 18. Data Integrity Risk Assessment: Minimum Requirements for System Category 2.

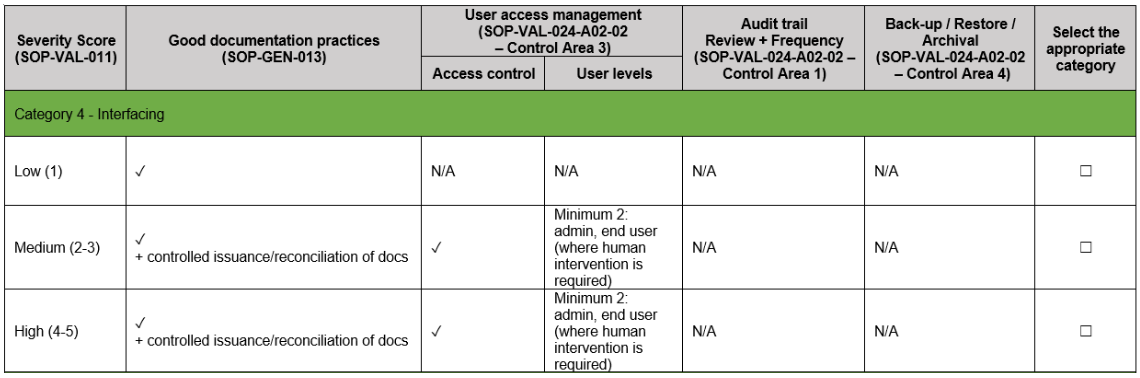
Fourth, Figure 19 shows the associated controls when the system is assessed as Category 3: “**Printed**” system. In this category, the generated GxP data is not intended to be kept and stored within the system but printed out. There will be some limited adjustable input data to be performed manually but still no permanent data to be managed. Typical examples could be potentiometric titrators not connected to a PC, balances with printer (APIC, 2019, p.12). Access control should be managed for securing time and date settings to avoid risks of manipulation or unauthorized modifications. Audit trail and user levels are not applicable.



Severity Score (SOP-VAL-011)	Good documentation practices (SOP-GEN-013)	User access management (SOP-VAL-024-A02-02 – Control Area 3)		Audit trail Review + Frequency (SOP-VAL-024-A02-02 – Control Area 1)	Back-up / Restore / Archival (SOP-VAL-024-A02-02 – Control Area 4)	Select the appropriate category
		Access control	User levels			
Category 3 - Printed						
Low (1)	✓	N/A	N/A	N/A	✓	□
Medium (2-3)	✓ + controlled issuance/reconciliation of docs + printing of relevant GxP data	✓[1]	N/A	N/A	✓	□
High (4-5)	✓ + controlled issuance/reconciliation of docs + printing of relevant GxP data	✓[1]	N/A	N/A	✓	□
[1]*Access control only for securing time and date settings*						

Figure 19. Data Integrity Risk Assessment: Minimum Requirements for System Category 3.

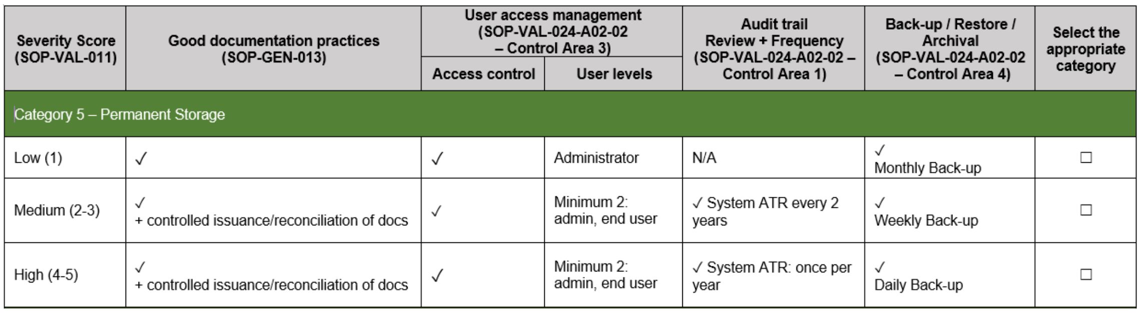
Fifth, Figure 20 shows the associated controls when the system is assessed as Category 4: “**Interfacing**” system. In this category, the generated GxP data is not intended to be kept and stored within the system but interfaced to another system such as the ones in Category 5 or 6. Therefore, controls such as audit trail and backup are not necessary since the system just acts as a bridge for transferring GxP data. There will be some limited adjustable input data to be performed manually but still no permanent data to be managed. Typical examples could be temperature sensors (APIC, 2019, p.12).



Severity Score (SOP-VAL-011)	Good documentation practices (SOP-GEN-013)	User access management (SOP-VAL-024-A02-02 – Control Area 3)		Audit trail Review + Frequency (SOP-VAL-024-A02-02 – Control Area 1)	Back-up / Restore / Archival (SOP-VAL-024-A02-02 – Control Area 4)	Select the appropriate category
		Access control	User levels			
Category 4 - Interfacing						
Low (1)	✓	N/A	N/A	N/A	N/A	□
Medium (2-3)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user (where human intervention is required)	N/A	N/A	□
High (4-5)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user (where human intervention is required)	N/A	N/A	□

Figure 20. Data Integrity Risk Assessment: Minimum Requirements for System Category 4.

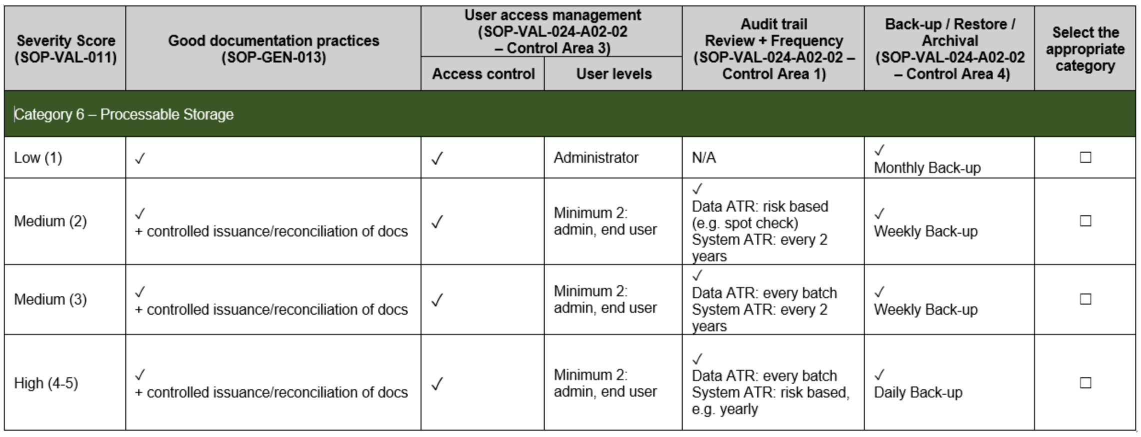
Sixth, Figure 21 shows the associated controls when the system is assessed as Category 5: "**Permanent storage**" system. In this category, the GxP data are permanently stored and not intended to be modified by the user to generate results such as static GxP data. If modification/deletion transactions are not available in the system, data integrity controls are in the maintenance and data inputs. Thus, audit trail, user levels, access controls are now applicable with the rigorous frequency depending on the criticality, risks and suitability. Examples could include UV instruments or IR instruments used for identification testing, in line particle size and TOC testing (APIC, 2019, p.12).



Severity Score (SOP-VAL-011)	Good documentation practices (SOP-GEN-013)	User access management (SOP-VAL-024-A02-02 – Control Area 3)		Audit trail Review + Frequency (SOP-VAL-024-A02-02 – Control Area 1)	Back-up / Restore / Archival (SOP-VAL-024-A02-02 – Control Area 4)	Select the appropriate category
		Access control	User levels			
Category 5 – Permanent Storage						
Low (1)	✓	✓	Administrator	N/A	✓ Monthly Back-up	□
Medium (2-3)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user	✓ System ATR every 2 years	✓ Weekly Back-up	□
High (4-5)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user	✓ System ATR: once per year	✓ Daily Back-up	□

Figure 21. Data Integrity Risk Assessment: Minimum Requirements for System Category 5.

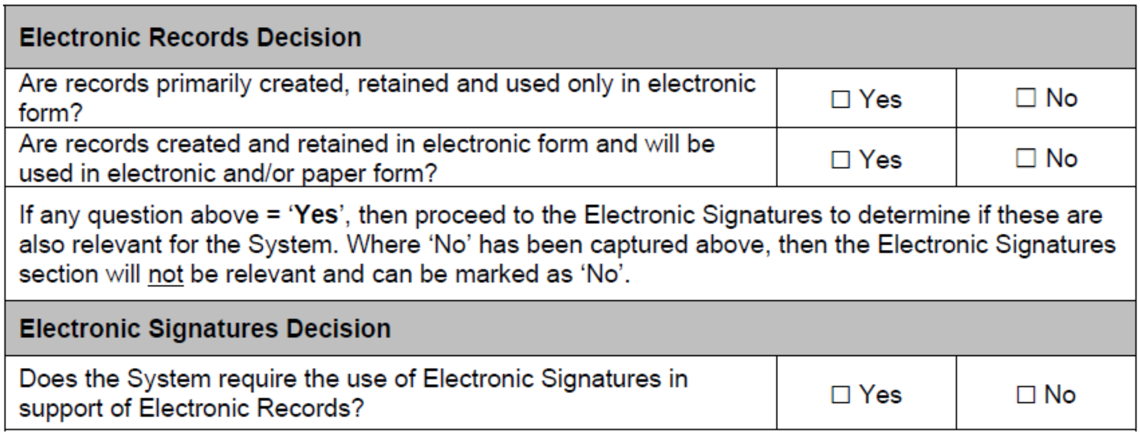
Seventh, Figure 22 shows the associated controls when the system is assessed as Category 6: "**Processable Storage**" system, this holds the highest level of controls and required monitoring since GxP data are permanently stored and can be processed by the user to generate results. Thus, audit trail, user levels, access controls are also applicable as Category 5 with the most rigorous frequency but still dependent on its criticality, risks and suitability. Examples could be MES systems, ERP systems, chromatographic data systems, electronic deviations management system (APIC, 2019, p.12).



Severity Score (SOP-VAL-011)	Good documentation practices (SOP-GEN-013)	User access management (SOP-VAL-024-A02-02 – Control Area 3)		Audit trail Review + Frequency (SOP-VAL-024-A02-02 – Control Area 1)	Back-up / Restore / Archival (SOP-VAL-024-A02-02 – Control Area 4)	Select the appropriate category
		Access control	User levels			
Category 6 – Processable Storage						
Low (1)	✓	✓	Administrator	N/A	✓ Monthly Back-up	□
Medium (2)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user	✓ Data ATR: risk based (e.g. spot check) System ATR: every 2 years	✓ Weekly Back-up	□
Medium (3)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user	✓ Data ATR: every batch System ATR: every 2 years	✓ Weekly Back-up	□
High (4-5)	✓ + controlled issuance/reconciliation of docs	✓	Minimum 2: admin, end user	✓ Data ATR: every batch System ATR: risk based, e.g. yearly	✓ Daily Back-up	□

Figure 22. Data Integrity Risk Assessment: Minimum Requirements for System Category 6.

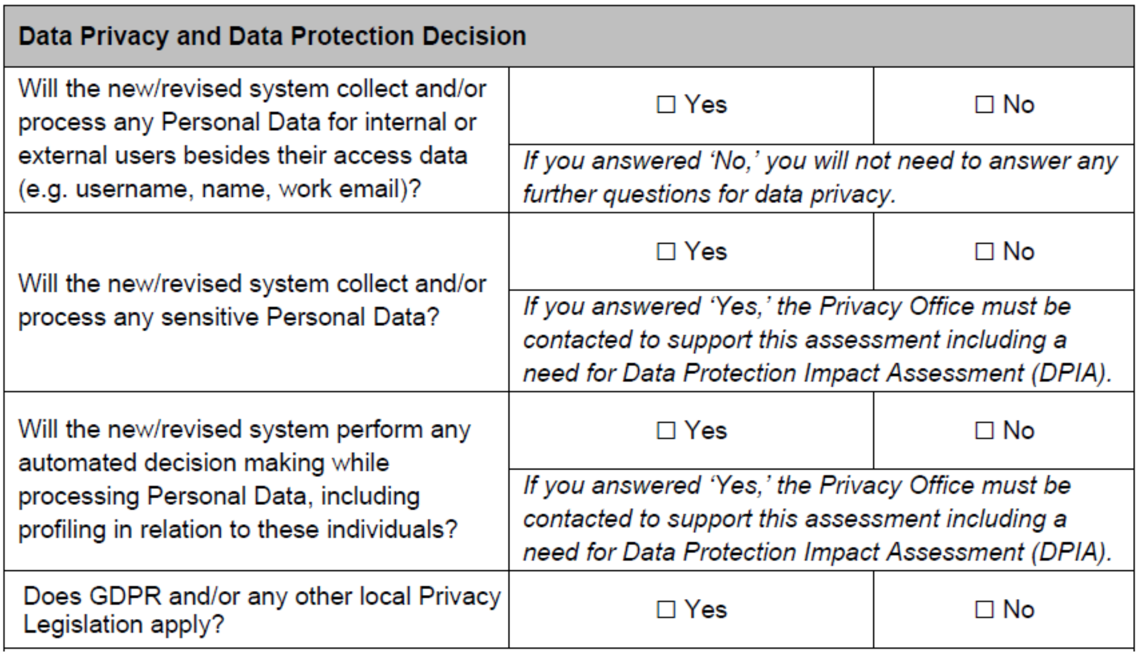
Eights, Figure 23 shows the applicability when the system is assessed as to have controls for Electronic Records and also the Electronic Signatures to which signatures executed to electronic records are equivalent to handwritten signatures executed on paper. This must distinctively be identified to determine if it the data integrity controls are expected in the system features or outside the system controls.



Electronic Records Decision		
Are records primarily created, retained and used only in electronic form?	□ Yes	□ No
Are records created and retained in electronic form and will be used in electronic and/or paper form?	□ Yes	□ No
If any question above = ‘**Yes**’, then proceed to the Electronic Signatures to determine if these are also relevant for the System. Where ‘No’ has been captured above, then the Electronic Signatures section will <u>not</u> be relevant and can be marked as ‘No’.		
Electronic Signatures Decision		
Does the System require the use of Electronic Signatures in support of Electronic Records?	□ Yes	□ No

Figure 23. Data Integrity Risk Assessment: FDA 21 CFR Part 11 Compliance Assessment

Nineth, Figure 24 shows the applicability when the system is assessed to be compliant with the General Data Protection Regulation, another regulatory authority that requires proper control on the collection, storage and management of personal data, commonly applicable to those systems handling patient data during clinical trials.



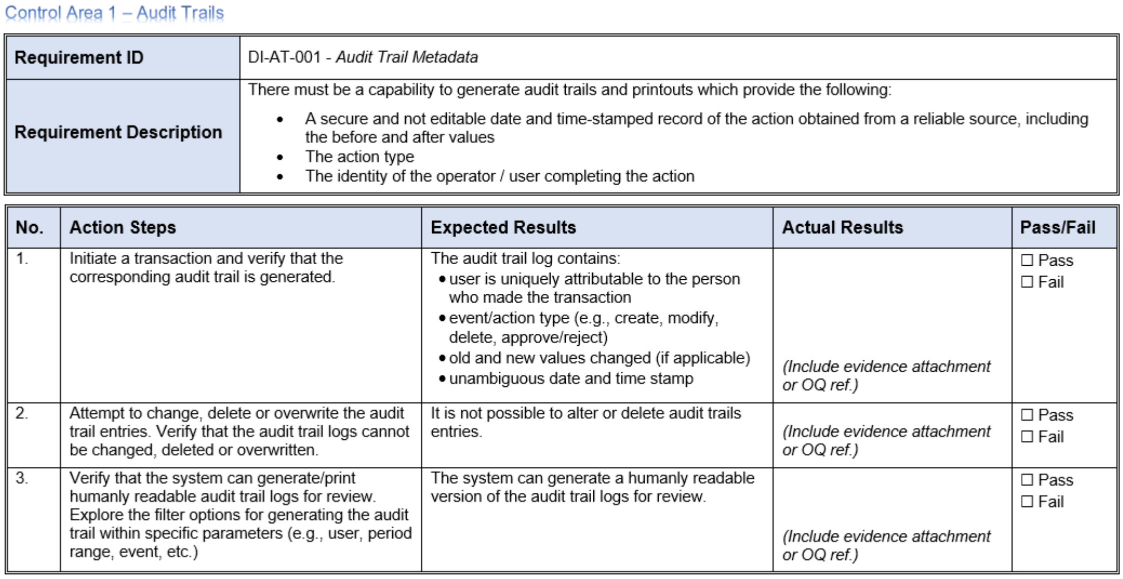
Data Privacy and Data Protection Decision		
Will the new/revised system collect and/or process any Personal Data for internal or external users besides their access data (e.g. username, name, work email)?	☐ Yes	☐ No
	If you answered 'No,' you will not need to answer any further questions for data privacy.	
Will the new/revised system collect and/or process any sensitive Personal Data?	☐ Yes	☐ No
	If you answered 'Yes,' the Privacy Office must be contacted to support this assessment including a need for Data Protection Impact Assessment (DPIA).	
Will the new/revised system perform any automated decision making while processing Personal Data, including profiling in relation to these individuals?	☐ Yes	☐ No
	If you answered 'Yes,' the Privacy Office must be contacted to support this assessment including a need for Data Protection Impact Assessment (DPIA).	
Does GDPR and/or any other local Privacy Legislation apply?	☐ Yes	☐ No

Figure 24. Data Integrity Risk Assessment: GDPR Assessment.

As a summary, using a more comprehensive risk assessment that expound further to expected and applicable system controls, the subsequent verification procedures can be carried out more efficiently instead of forcing this controls to systems that do not even store or manage data indefinitely.

After the data integrity risk assessment, the applicable data integrity verification procedures ensue and grouped into five control areas: 1) Audit Trails, 2) Electronic Records and Electronic Signatures, 3) User Access Management, 4) Backup Restore and Archival, and 5) Data Protection and Data Privacy.

Figure 25 shows the data integrity verification procedures for Audit Trail Metadata under Control Area 1 – Audit Trails that demonstrate the systems' capability of generating audit trails and printout containing the required metadata based on regulatory guidelines and standards.



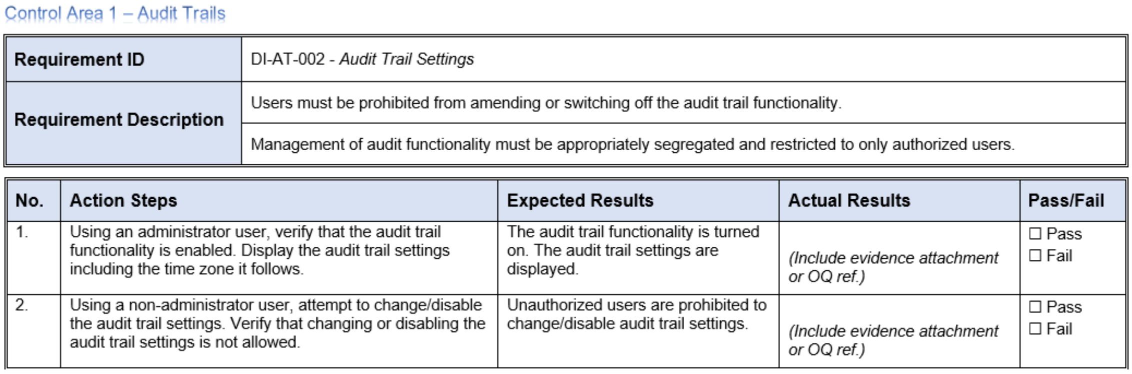
Control Area 1 – Audit Trails

Requirement ID	DI-AT-001 - *Audit Trail Metadata*
Requirement Description	There must be a capability to generate audit trails and printouts which provide the following: • A secure and not editable date and time-stamped record of the action obtained from a reliable source, including the before and after values • The action type • The identity of the operator / user completing the action

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Initiate a transaction and verify that the corresponding audit trail is generated.	The audit trail log contains: • user is uniquely attributable to the person who made the transaction • event/action type (e.g., create, modify, delete, approve/reject) • old and new values changed (if applicable) • unambiguous date and time stamp	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Attempt to change, delete or overwrite the audit trail entries. Verify that the audit trail logs cannot be changed, deleted or overwritten.	It is not possible to alter or delete audit trails entries.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
3.	Verify that the system can generate/print humanly readable audit trail logs for review. Explore the filter options for generating the audit trail within specific parameters (e.g., user, period range, event, etc.)	The system can generate a humanly readable version of the audit trail logs for review.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 25. Data Integrity Verification: Control Area 1 – Audit Trail Metadata.

Figure 26 shows the data integrity verification procedures for Audit Trail Settings under Control Area 1 – Audit Trails that demonstrate the systems' capability of prohibiting users from amending or disabling audit trail functions, or that these modifications are restricted only to authorized system users.



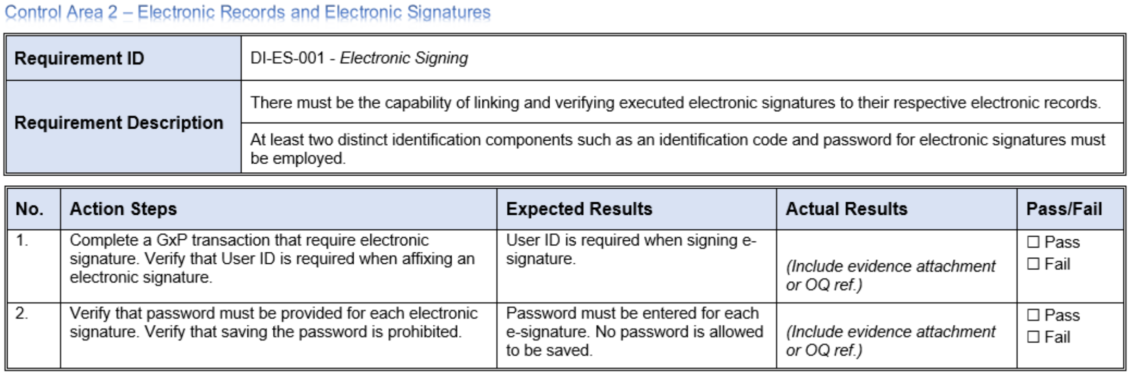
Control Area 1 – Audit Trails

Requirement ID	DI-AT-002 - *Audit Trail Settings*
Requirement Description	Users must be prohibited from amending or switching off the audit trail functionality.
	Management of audit functionality must be appropriately segregated and restricted to only authorized users.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Using an administrator user, verify that the audit trail functionality is enabled. Display the audit trail settings including the time zone it follows.	The audit trail functionality is turned on. The audit trail settings are displayed.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Using a non-administrator user, attempt to change/disable the audit trail settings. Verify that changing or disabling the audit trail settings is not allowed.	Unauthorized users are prohibited to change/disable audit trail settings.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 26. Data Integrity Verification: Control Area 1 – Audit Trail Settings.

Figure 27 shows the data integrity verification procedures for Electronic Signing under Control Area 2 – Electronic Records and Electronic Signatures that demonstrate the linkage of the two and the systems' capability to require the user a minimum of two distinct identification components.



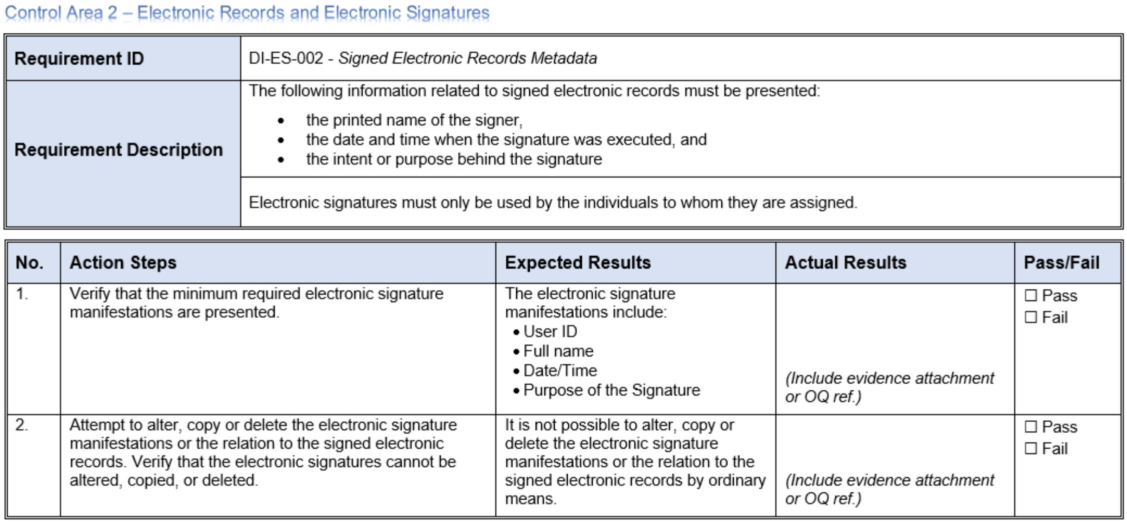
Control Area 2 – Electronic Records and Electronic Signatures

Requirement ID	DI-ES-001 - *Electronic Signing*
Requirement Description	There must be the capability of linking and verifying executed electronic signatures to their respective electronic records.
	At least two distinct identification components such as an identification code and password for electronic signatures must be employed.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Complete a GxP transaction that require electronic signature. Verify that User ID is required when affixing an electronic signature.	User ID is required when signing e-signature.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that password must be provided for each electronic signature. Verify that saving the password is prohibited.	Password must be entered for each e-signature. No password is allowed to be saved.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 27. Data Integrity Verification: Control Area 2 – Electronic Signing.

Figure 28 shows the data integrity verification procedures for Electronic Records Metadata under Control Area 2 – Electronic Records and Electronic Signatures that demonstrate the systems' capability to provide the minimum metadata of electronic signatures and its restrictions of the assigned signatory.



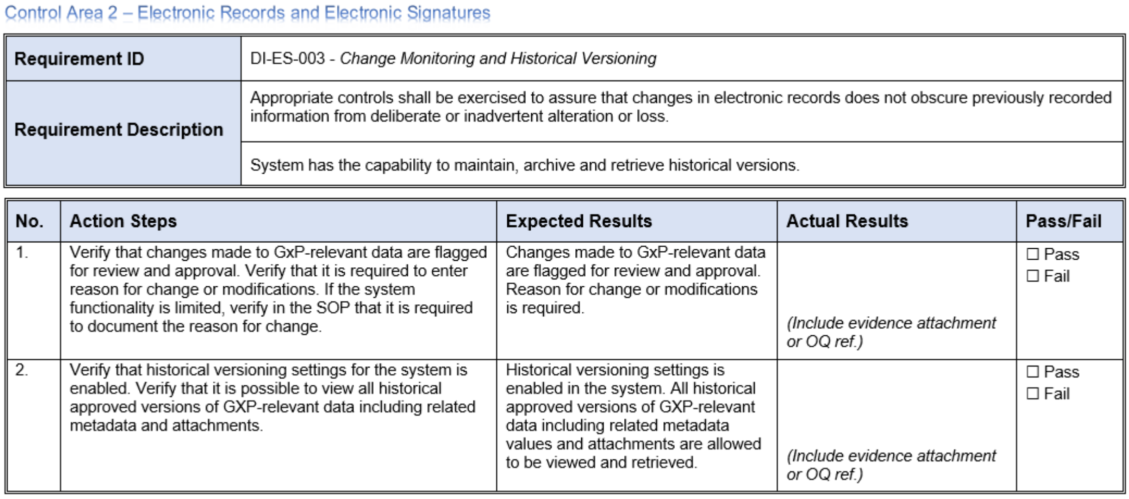
Control Area 2 – Electronic Records and Electronic Signatures

Requirement ID	DI-ES-002 - *Signed Electronic Records Metadata*
Requirement Description	The following information related to signed electronic records must be presented: • the printed name of the signer, • the date and time when the signature was executed, and • the intent or purpose behind the signature
	Electronic signatures must only be used by the individuals to whom they are assigned.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that the minimum required electronic signature manifestations are presented.	The electronic signature manifestations include: • User ID • Full name • Date/Time • Purpose of the Signature	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Attempt to alter, copy or delete the electronic signature manifestations or the relation to the signed electronic records. Verify that the electronic signatures cannot be altered, copied, or deleted.	It is not possible to alter, copy or delete the electronic signature manifestations or the relation to the signed electronic records by ordinary means.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 28. Data Integrity Verification: Control Area 2 – Signed Electronic Records Metadata.

Figure 27 shows the data integrity verification procedures for Change Monitoring and Historical Versioning under Control Area 2 – Electronic Records and Electronic Signatures that demonstrate the systems' capability to prohibit users to obscure data from deliberate or inadvertent alteration or loss, and its capability to archive and retrieve the version history of a record.



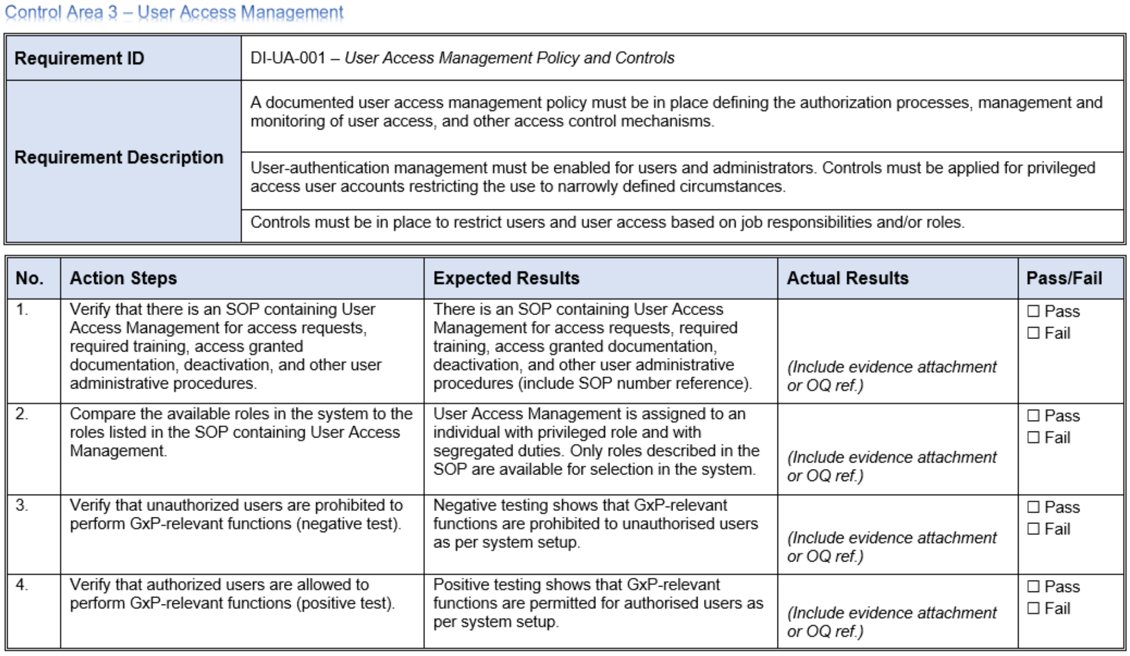
Control Area 2 – Electronic Records and Electronic Signatures

Requirement ID	DI-ES-003 - *Change Monitoring and Historical Versioning*
Requirement Description	Appropriate controls shall be exercised to assure that changes in electronic records does not obscure previously recorded information from deliberate or inadvertent alteration or loss.
	System has the capability to maintain, archive and retrieve historical versions.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that changes made to GxP-relevant data are flagged for review and approval. Verify that it is required to enter reason for change or modifications. If the system functionality is limited, verify in the SOP that it is required to document the reason for change.	Changes made to GxP-relevant data are flagged for review and approval. Reason for change or modifications is required.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that historical versioning settings for the system is enabled. Verify that it is possible to view all historical approved versions of GXP-relevant data including related metadata and attachments.	Historical versioning settings is enabled in the system. All historical approved versions of GXP-relevant data including related metadata values and attachments are allowed to be viewed and retrieved.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 29. Data Integrity Verification: Control Area 2 – Change Monitoring and Historical Versioning.

Figure 30 shows the data integrity verification procedures for User Access Management Policy and Controls under Control Area 3 – User Access Management to check the documented process in place to support the user access controls, and to demonstrate the systems' capability in restricting users based on its assigned permissions and roles.



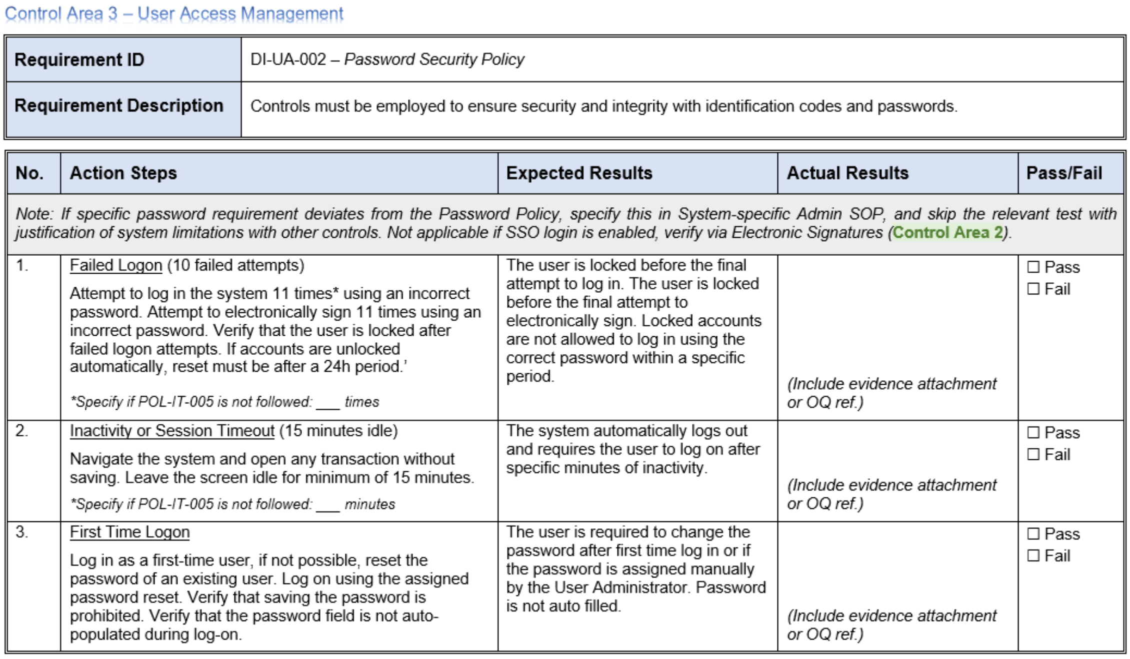
Control Area 3 – User Access Management

Requirement ID	DI-UA-001 – *User Access Management Policy and Controls*
Requirement Description	A documented user access management policy must be in place defining the authorization processes, management and monitoring of user access, and other access control mechanisms.
	User-authentication management must be enabled for users and administrators. Controls must be applied for privileged access user accounts restricting the use to narrowly defined circumstances.
	Controls must be in place to restrict users and user access based on job responsibilities and/or roles.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that there is an SOP containing User Access Management for access requests, required training, access granted documentation, deactivation, and other user administrative procedures.	There is an SOP containing User Access Management for access requests, required training, access granted documentation, deactivation, and other user administrative procedures (include SOP number reference).	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Compare the available roles in the system to the roles listed in the SOP containing User Access Management.	User Access Management is assigned to an individual with privileged role and with segregated duties. Only roles described in the SOP are available for selection in the system.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
3.	Verify that unauthorized users are prohibited to perform GxP-relevant functions (negative test).	Negative testing shows that GxP-relevant functions are prohibited to unauthorised users as per system setup.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
4.	Verify that authorized users are allowed to perform GxP-relevant functions (positive test).	Positive testing shows that GxP-relevant functions are permitted for authorised users as per system setup.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 30. Data Integrity Verification: Control Area 3 – User Access Management Policy and Controls.

Figures 31 and 32 shows the data integrity verification procedures for Password Security Policy under Control Area 3 – User Access Management that demonstrate the systems' capability to implement password security based on Company's security policies.

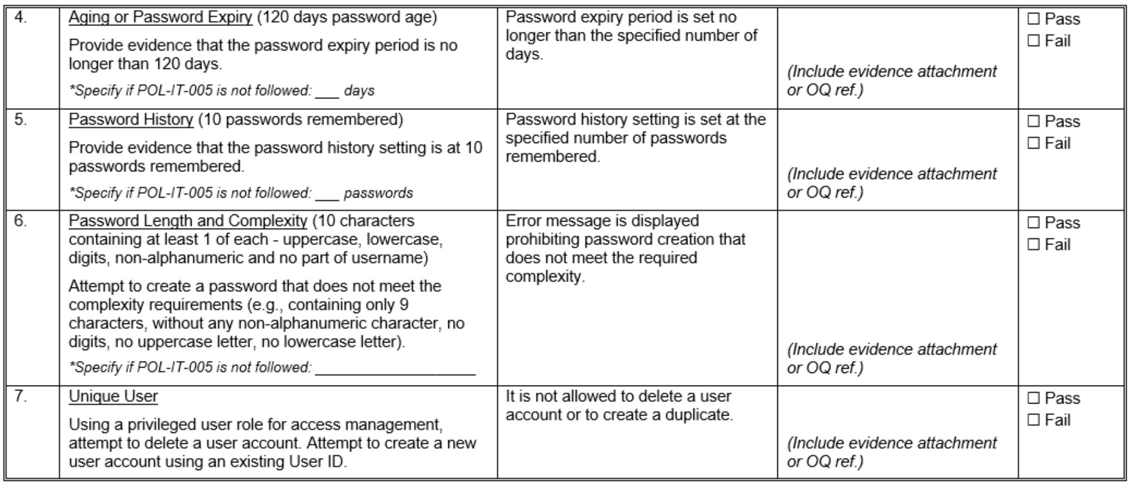


Control Area 3 – User Access Management

Requirement ID	DI-UA-002 – *Password Security Policy*
Requirement Description	Controls must be employed to ensure security and integrity with identification codes and passwords.

No.	**Action Steps**	**Expected Results**	**Actual Results**	**Pass/Fail**
*Note: If specific password requirement deviates from the Password Policy, specify this in System-specific Admin SOP, and skip the relevant test with justification of system limitations with other controls. Not applicable if SSO login is enabled, verify via Electronic Signatures (***Control Area 2***).*				
1.	Failed Logon (10 failed attempts) Attempt to log in the system 11 times* using an incorrect password. Attempt to electronically sign 11 times using an incorrect password. Verify that the user is locked after failed logon attempts. If accounts are unlocked automatically, reset must be after a 24h period.' **Specify if POL-IT-005 is not followed: ___ times*	The user is locked before the final attempt to log in. The user is locked before the final attempt to electronically sign. Locked accounts are not allowed to log in using the correct password within a specific period.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Inactivity or Session Timeout (15 minutes idle) Navigate the system and open any transaction without saving. Leave the screen idle for minimum of 15 minutes. **Specify if POL-IT-005 is not followed: ___ minutes*	The system automatically logs out and requires the user to log on after specific minutes of inactivity.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
3.	First Time Logon Log in as a first-time user, if not possible, reset the password of an existing user. Log on using the assigned password reset. Verify that saving the password is prohibited. Verify that the password field is not auto-populated during log-on.	The user is required to change the password after first time log in or if the password is assigned manually by the User Administrator. Password is not auto filled.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

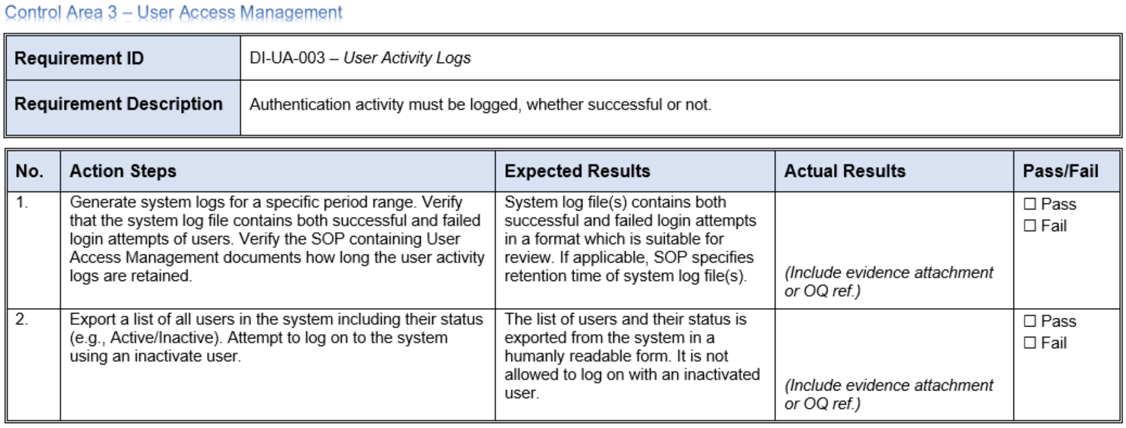
Figure 31. Data Integrity Verification: Control Area 3 – Password Security Policy (Part 1).



4.	Aging or Password Expiry (120 days password age) Provide evidence that the password expiry period is no longer than 120 days. **Specify if POL-IT-005 is not followed: ___ days*	Password expiry period is set no longer than the specified number of days.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
5.	Password History (10 passwords remembered) Provide evidence that the password history setting is at 10 passwords remembered. **Specify if POL-IT-005 is not followed: ___ passwords*	Password history setting is set at the specified number of passwords remembered.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
6.	Password Length and Complexity (10 characters containing at least 1 of each - uppercase, lowercase, digits, non-alphanumeric and no part of username) Attempt to create a password that does not meet the complexity requirements (e.g., containing only 9 characters, without any non-alphanumeric character, no digits, no uppercase letter, no lowercase letter). **Specify if POL-IT-005 is not followed: ___________________*	Error message is displayed prohibiting password creation that does not meet the required complexity.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
7.	Unique User Using a privileged user role for access management, attempt to delete a user account. Attempt to create a new user account using an existing User ID.	It is not allowed to delete a user account or to create a duplicate.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 32. Data Integrity Verification: Control Area 3 – Password Security Policy (Part 2).

Figures 33 shows the data integrity verification procedures for User Activity Logs under Control Area 3 – User Access Management that demonstrate the systems' capability to generate user activity logs for traceability and aid in periodic reviews or data integrity compliance monitoring.



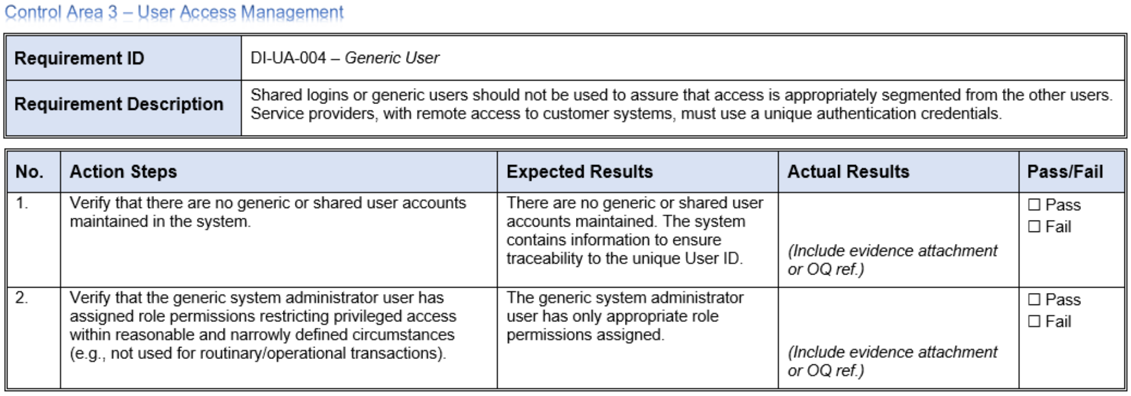
Control Area 3 – User Access Management

Requirement ID	DI-UA-003 – *User Activity Logs*
Requirement Description	Authentication activity must be logged, whether successful or not.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Generate system logs for a specific period range. Verify that the system log file contains both successful and failed login attempts of users. Verify the SOP containing User Access Management documents how long the user activity logs are retained.	System log file(s) contains both successful and failed login attempts in a format which is suitable for review. If applicable, SOP specifies retention time of system log file(s).	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Export a list of all users in the system including their status (e.g., Active/Inactive). Attempt to log on to the system using an inactivate user.	The list of users and their status is exported from the system in a humanly readable form. It is not allowed to log on with an inactivated user.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 33. Data Integrity Verification: Control Area 3 – User Activity Logs.

Figures 34 shows the data integrity verification procedures for Generic User under Control Area 3 – User Access Management that demonstrate the systems' capability to appropriately segment user accounts that are traceable and attributable to a specific user, thus discouraging or prohibiting shared logins or generic users.



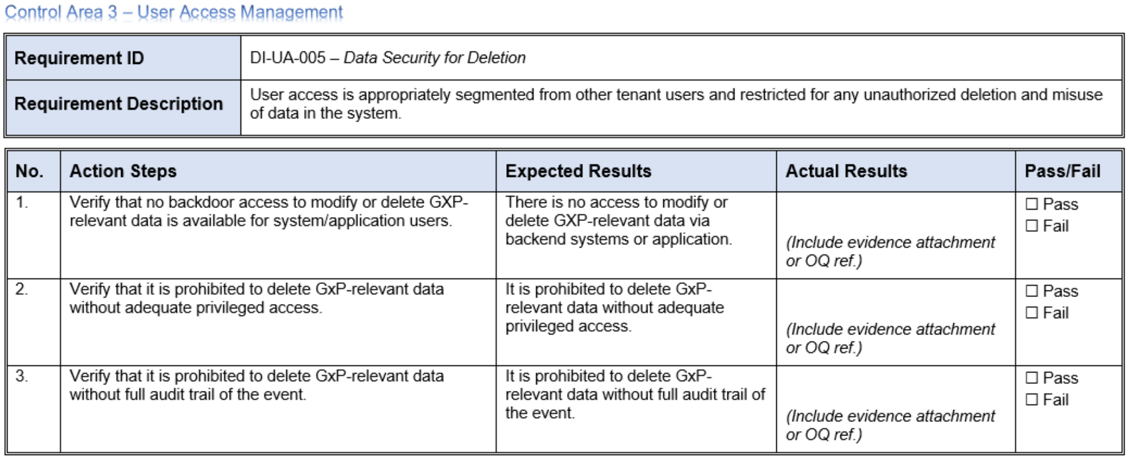
Control Area 3 – User Access Management

Requirement ID	DI-UA-004 – *Generic User*
Requirement Description	Shared logins or generic users should not be used to assure that access is appropriately segmented from the other users. Service providers, with remote access to customer systems, must use a unique authentication credentials.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that there are no generic or shared user accounts maintained in the system.	There are no generic or shared user accounts maintained. The system contains information to ensure traceability to the unique User ID.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that the generic system administrator user has assigned role permissions restricting privileged access within reasonable and narrowly defined circumstances (e.g., not used for routinary/operational transactions).	The generic system administrator user has only appropriate role permissions assigned.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 34. Data Integrity Verification: Control Area 3 – Generic User.

Figures 35 shows the data integrity verification procedures for Data Security for Deletion under Control Area 3 – User Access Management that demonstrate the systems' capability to manage external users outside an organization domain to only limit to appropriate tasks and to restrict from any unauthorized transactions such as deletion, abuse and misuse of privileged access.



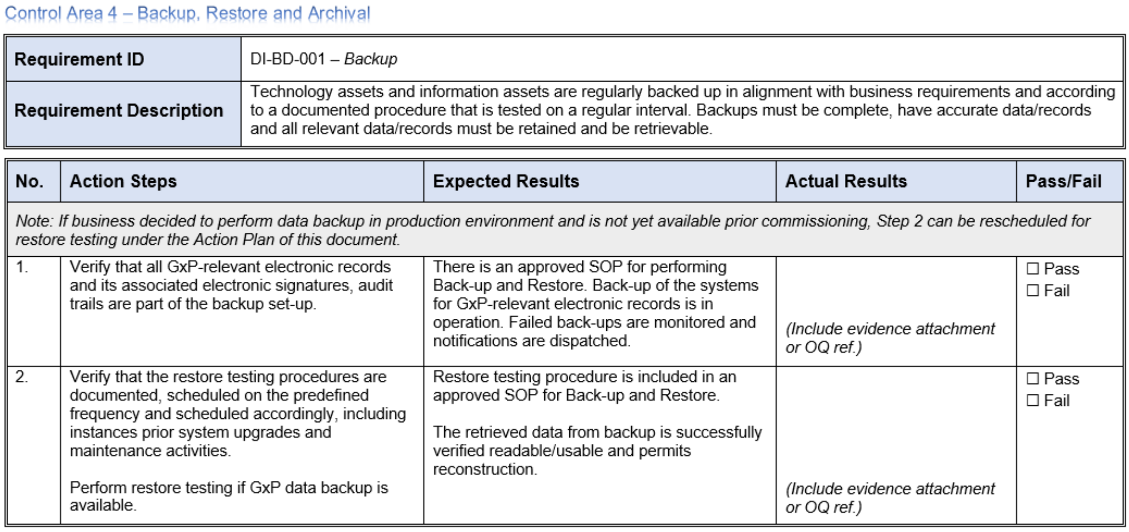
Control Area 3 – User Access Management

Requirement ID	DI-UA-005 – *Data Security for Deletion*
Requirement Description	User access is appropriately segmented from other tenant users and restricted for any unauthorized deletion and misuse of data in the system.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that no backdoor access to modify or delete GXP-relevant data is available for system/application users.	There is no access to modify or delete GXP-relevant data via backend systems or application.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that it is prohibited to delete GxP-relevant data without adequate privileged access.	It is prohibited to delete GxP-relevant data without adequate privileged access.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
3.	Verify that it is prohibited to delete GxP-relevant data without full audit trail of the event.	It is prohibited to delete GxP-relevant data without full audit trail of the event.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 35. Data Integrity Verification: Control Area 3 – Data Security for Deletion.

Figures 36 shows the data integrity verification procedures for Backup under Control Area 4 – Backup, Restore and Archival that demonstrate the systems' capability to create a reliable, complete and accurate backup that are retained and retrievable.



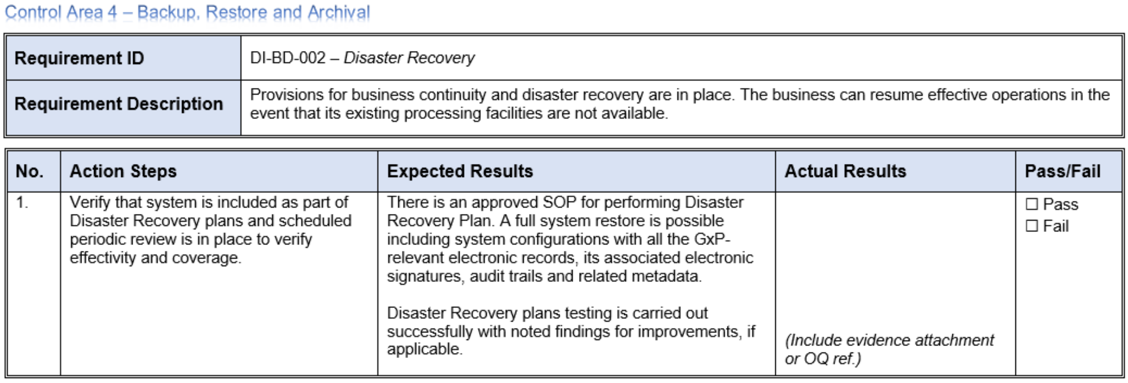
Control Area 4 – Backup, Restore and Archival

Requirement ID	DI-BD-001 – *Backup*
Requirement Description	Technology assets and information assets are regularly backed up in alignment with business requirements and according to a documented procedure that is tested on a regular interval. Backups must be complete, have accurate data/records and all relevant data/records must be retained and be retrievable.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
Note: If business decided to perform data backup in production environment and is not yet available prior commissioning, Step 2 can be rescheduled for restore testing under the Action Plan of this document.				
1.	Verify that all GxP-relevant electronic records and its associated electronic signatures, audit trails are part of the backup set-up.	There is an approved SOP for performing Back-up and Restore. Back-up of the systems for GxP-relevant electronic records is in operation. Failed back-ups are monitored and notifications are dispatched.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that the restore testing procedures are documented, scheduled on the predefined frequency and scheduled accordingly, including instances prior system upgrades and maintenance activities. Perform restore testing if GxP data backup is available.	Restore testing procedure is included in an approved SOP for Back-up and Restore. The retrieved data from backup is successfully verified readable/usable and permits reconstruction.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 36. Data Integrity Verification: Control Area 4 – Backup.

Figures 37 shows the data integrity verification procedures for Disaster Recovery under Control Area 4 – Backup, Restore and Archival that demonstrate the systems' capability to be responsive in cases of disaster based on documented plans, policies and scheduled periodic restore testing.



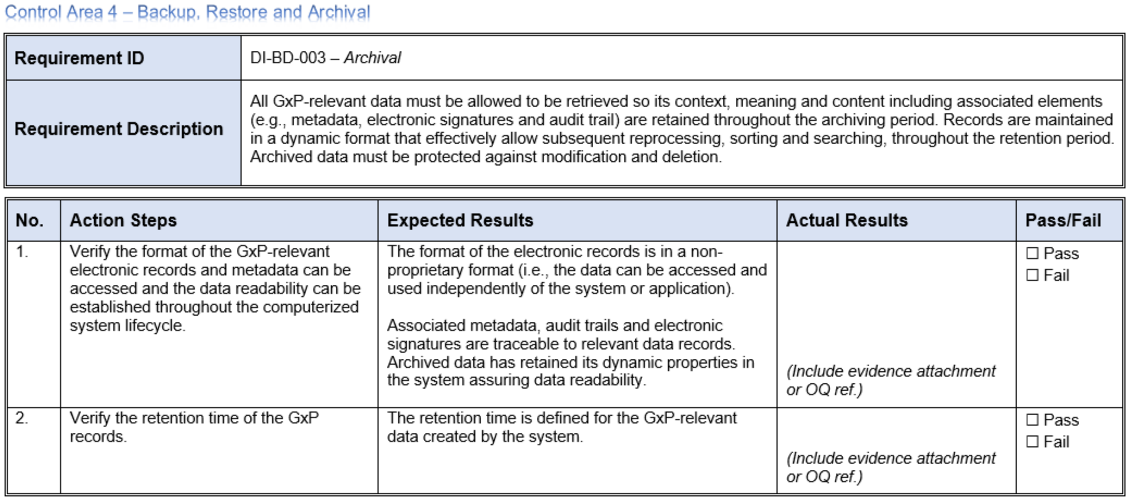
Control Area 4 – Backup, Restore and Archival

Requirement ID	DI-BD-002 – *Disaster Recovery*
Requirement Description	Provisions for business continuity and disaster recovery are in place. The business can resume effective operations in the event that its existing processing facilities are not available.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that system is included as part of Disaster Recovery plans and scheduled periodic review is in place to verify effectivity and coverage.	There is an approved SOP for performing Disaster Recovery Plan. A full system restore is possible including system configurations with all the GxP-relevant electronic records, its associated electronic signatures, audit trails and related metadata. Disaster Recovery plans testing is carried out successfully with noted findings for improvements, if applicable.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 37. Data Integrity Verification: Control Area 4 – Disaster Recovery.

Figures 38 shows the data integrity verification procedures for Archival under Control Area 4 – Backup, Restore and Archival that demonstrate the systems' capability to archive data to be kept at its original content and meaning, including its protection against modification and deletion. It must also contain functions to be searchable throughout its retention period, in cases of electronic records.



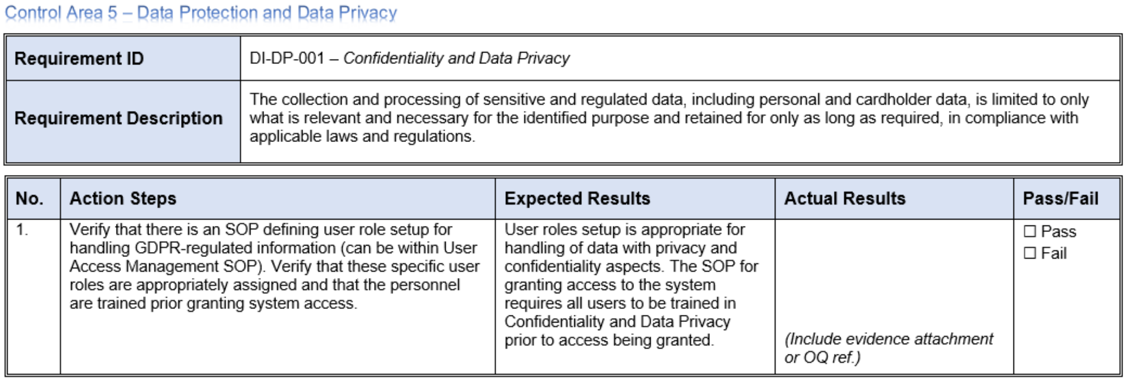
Control Area 4 – Backup, Restore and Archival

Requirement ID	DI-BD-003 – *Archival*
Requirement Description	All GxP-relevant data must be allowed to be retrieved so its context, meaning and content including associated elements (e.g., metadata, electronic signatures and audit trail) are retained throughout the archiving period. Records are maintained in a dynamic format that effectively allow subsequent reprocessing, sorting and searching, throughout the retention period. Archived data must be protected against modification and deletion.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify the format of the GxP-relevant electronic records and metadata can be accessed and the data readability can be established throughout the computerized system lifecycle.	The format of the electronic records is in a non-proprietary format (i.e., the data can be accessed and used independently of the system or application). Associated metadata, audit trails and electronic signatures are traceable to relevant data records. Archived data has retained its dynamic properties in the system assuring data readability.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify the retention time of the GxP records.	The retention time is defined for the GxP-relevant data created by the system.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 38. Data Integrity Verification: Control Area 4 – Archival.

Figures 39 shows the data integrity verification procedures for Confidentiality and Data Privacy under Control Area 5 – Data Protection and Data Privacy that demonstrate the systems' capability to have appropriate documented policies, procedures and trainings for users handling and managing sensitive and regulated data.



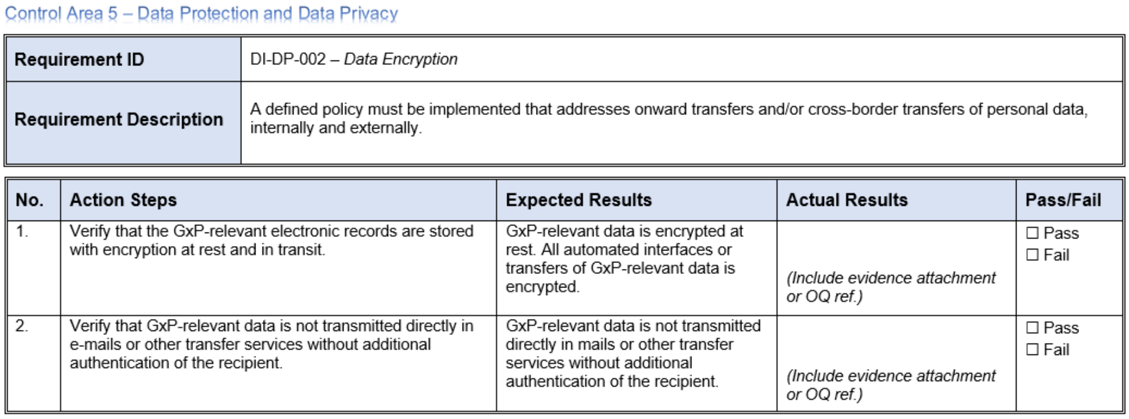
Control Area 5 – Data Protection and Data Privacy

Requirement ID	DI-DP-001 – *Confidentiality and Data Privacy*
Requirement Description	The collection and processing of sensitive and regulated data, including personal and cardholder data, is limited to only what is relevant and necessary for the identified purpose and retained for only as long as required, in compliance with applicable laws and regulations.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that there is an SOP defining user role setup for handling GDPR-regulated information (can be within User Access Management SOP). Verify that these specific user roles are appropriately assigned and that the personnel are trained prior granting system access.	User roles setup is appropriate for handling of data with privacy and confidentiality aspects. The SOP for granting access to the system requires all users to be trained in Confidentiality and Data Privacy prior to access being granted.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 39. Data Integrity Verification: Control Area 5 – Confidentiality and Data Privacy.

Figures 40 shows the data integrity verification procedures for Data Encryption under Control Area 5 – Data Protection and Data Privacy that demonstrate the systems' capability to implement data encryption for personal data tagged with confidentiality both in transit and at rest.



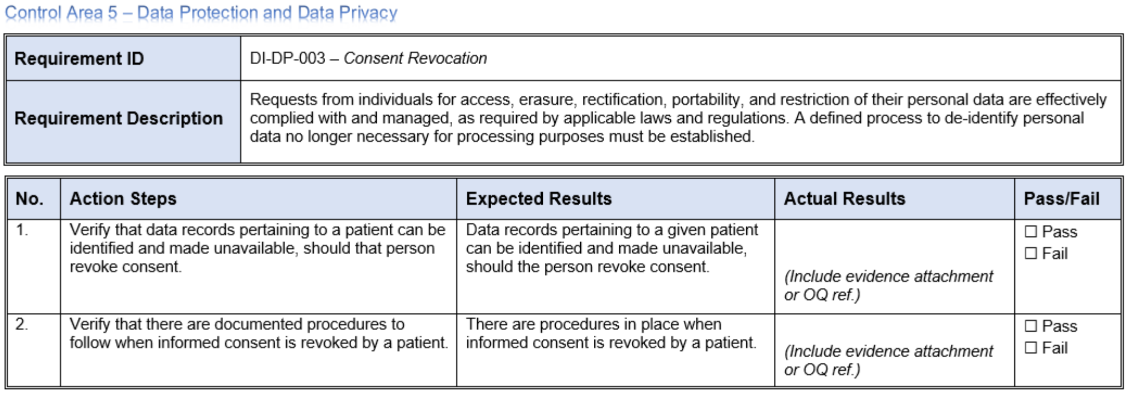
Control Area 5 – Data Protection and Data Privacy

Requirement ID	DI-DP-002 – *Data Encryption*
Requirement Description	A defined policy must be implemented that addresses onward transfers and/or cross-border transfers of personal data, internally and externally.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that the GxP-relevant electronic records are stored with encryption at rest and in transit.	GxP-relevant data is encrypted at rest. All automated interfaces or transfers of GxP-relevant data is encrypted.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that GxP-relevant data is not transmitted directly in e-mails or other transfer services without additional authentication of the recipient.	GxP-relevant data is not transmitted directly in mails or other transfer services without additional authentication of the recipient.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 40. Data Integrity Verification: Control Area 5 – Data Encryption.

Figures 41 shows the data integrity verification procedures for Consent Revocation under Control Area 5 – Data Protection and Data Privacy that demonstrate the systems' capability to effectively isolate and pull out personal data that are no longer necessary for processing specifically in cases of revocation or rectification of consent.



Control Area 5 – Data Protection and Data Privacy

Requirement ID	DI-DP-003 – *Consent Revocation*
Requirement Description	Requests from individuals for access, erasure, rectification, portability, and restriction of their personal data are effectively complied with and managed, as required by applicable laws and regulations. A defined process to de-identify personal data no longer necessary for processing purposes must be established.

No.	Action Steps	Expected Results	Actual Results	Pass/Fail
1.	Verify that data records pertaining to a patient can be identified and made unavailable, should that person revoke consent.	Data records pertaining to a given patient can be identified and made unavailable, should the person revoke consent.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail
2.	Verify that there are documented procedures to follow when informed consent is revoked by a patient.	There are procedures in place when informed consent is revoked by a patient.	*(Include evidence attachment or OQ ref.)*	☐ Pass ☐ Fail

Figure 41. Data Integrity Verification: Control Area 5 – Consent Revocation.

Next, the Initial proposal-1 was discussed with the stakeholders.

5.3 Findings from Data 2 (Suggestions from Stakeholders for the Initial Proposal – 1)

The main inputs for the Initial proposal-1 building originated from (1) Data 1 (findings from CSA, earlier), and (2) CF (input from best practices, regulatory guidelines and standards). This will then be aligned with (3) Data 2 (from this co-creation round). Inspired by these inputs, the Initial proposal-2 will be built.

5.3.1 Reinforce Data Integrity Compliance Throughout the Systems Lifecycle

After the review of external consultants, opinions were expressed as to why the process seemed to be limiting with only covering these highlighted system controls. Lack of inclusion of the periodic review activities such as system audit trail reviews, user access reviews, data readability and traceability checks during the predetermined data retention, backup and disaster recovery restore tests, business continuity plans, good documentation practices for paper records, maintaining systems inventory, and any data transfers programmed as background job runs. As these are all agreed to be important and crucial for data integrity compliance throughout the system’s lifecycle, these are also covered in other processes managed in separate SOPs and acknowledged at the beginning to be as out of scope in this study.

The most that can be done with this development proposal is to reinforce the data integrity culture by elaborating that it cannot be assured within the specific activities in the SOP improvements alone. In order to combat this, training specific for the improvements and revisions for data integrity compliance procedures will be scheduled with required participants of system users, process owners, QA, business SMEs, end-users and other relevant people that will benefit from the learnings. This will shed light on the other SOPs that covers the data integrity-related tasks to be performed periodically within the Operational phase of the system's lifecycle.

Few comments from the stakeholders similarly mentioned that the procedure looked more comprehensive, but a drastic shift from the current and soon-to-be obsolete version of the process. However, from the team who are involved on the trials and discussion, the actual execution of the procedures seemed to be easier though it seemed heavy at the first impression. Thus, the training will again be a planned resolution for this to elaborate and walkthrough the process to envision how it will be performed.

5.3.2 Ease in Transition for the Proposed Process Change

Majority of the stakeholders mentioned that the changes in the format of data integrity compliance procedures seemed to be very substantial. For those systems that currently have these activities as ongoing and following the intended previous version of the SOP, it could be difficult to execute unless a grace period for the transition is provided. There should be a more efficient option for the rollout of minor changes in the computerized systems that may affect data integrity functions. As an action plan, for instance, it can selectively use the templates in the new SOP with references to the completed procedures using the previous version of SOP particularly if it was performed accordingly with sufficient reviews and approvals. The differences between the old/current and new/proposed procedures must clearly be defined in the document and explained in the planned training for the transition. Effective dates for the revised process must also be communicated and assigned to all users in the Learning Management System to be read and understood.

Few comments are related to having more flexibility in the password security policy testing for those systems that are incapable of integrating Active Directory accounts, and limited password security restriction capabilities. The changes that were made is to remove the specific parameters in the acceptance criteria and provide users more freedom to define the password security restrictions specific to the system as long as this will be documented in the user administration procedures as well.

5.3.3 Clarity in the Interlapping Deliverables for CSV and Task Ownership

Currently, the data integrity related requirements are included within the validation deliverables as Quality requirements under the User Requirements Specifications (URS). Most of the systems has to collaboratively report these requirements to the Supplier early on during planning phase to build the configurations as preparation of Computerized Systems Validation (CSV). Some stakeholders oppose to removing the data integrity related requirements from the CSV process since it is helpful to communicate the URS to the Suppliers with its inclusion. To meet halfway, the Requirements in the Data Integrity Verification procedures will still remain as a guide and benchmark for testing objectives. It can also be an aid for formulating the URS, but will not be a replacement.

The timeline for Data Integrity Compliance has to be clarified since the resourcing are mostly allotted in the CSV processes. It was previously performed within the project plan, thus it will not be routinely achieved if nobody takes ownership of the project. To resolve this, the procedures themselves can be followed without extensive meetings to make it fully operational. Also as part of the planned training, the relevant accountable individuals have to be trained with a discussion of what phase of the CSV will the Data Integrity Risk Assessment and Verification interlaps. This should be given with enough freedom for self-learning to understand that currently, potential duplicate efforts can still be further minimized.

Data Collection 2 concentrates on identifying the suggestions from the key stakeholders. Since most of the comments were originating from the participants who are involved in the brainstorming of the proposal, these were updated and changed instantaneously such as the acceptance criteria or expected results in password security policy verification.

In summary, the key stakeholders propose to schedule training for multiple purposes such as 1) reinforcing data integrity compliance throughout the systems lifecycle, 2) facilitating users for ease in transition for proposed changes, 3) clarifying the link of data integrity procedures with the CSV deliverables and 4) promoting ownership of the responsibilities in data integrity procedures.

Table 6 below shows the inputs, comments and suggestions for the proposal.

Table 6. Key stakeholder suggestions (findings of Data 2) for Proposal building in relation to findings from the CSA (Data 1) and the Conceptual framework.

	Key focus areas from **CSA** (from **Data 1**)	Inputs from literature (**CF**)	Suggestions from stakeholders for the Proposal, summary (from **Data 2**)	Descriptions of their suggestions (in detail)
1	Misdirected data integrity risk assessment	Inclusion of system's complexity and capability of permanent or temporary storage. Applicable controls should be relevant on the system categories or the intended features and data to be managed.	Training and walkthrough to discuss the proposed changes in the data integrity compliance procedures.	Most of the participants including the QA/QS Manager suggested to conduct a training that will elaborate and clarify the changes.
2	ALCOA+ principle-based with the system controls-based verification	Control areas such as audit trail, electronic records and electronic signatures, user access management, backup, restore and archival and data protection and data privacy are among the repeated topics in data integrity compliance related for computerized systems validation.	Password security policy may be difficult to implement for some systems. Training and walkthrough to discuss the proposed changes in the data integrity compliance procedures.	The Senior QA/QS Specialist suggested removing the specific parameters in the acceptance criteria or expected results, to provide more freedom but still require controlled documentation. Most of the participants including the QA/QS Manager suggested to conduct a training that will elaborate and clarify the changes.

3	Lack of direction and guidance	Ownership and responsibilities of data integrity compliance procedures are on the system owners, process owners, business SMEs (e.g. IT) and QA. The culture of the regulated company can have an impact on the data integrity effectiveness.	Training and walkthrough to discuss the proposed changes in the data integrity compliance procedures.	Most of the participants including the QA/QS Manager suggested to conduct a training that will elaborate and clarify the changes.

As seen from Table 6,.the identified focus areas are misdirected data integrity risk assessment, ALCOA+ principle-based with the system controls-based verification, and lack of direction and guidance. The proposal for data integrity risk assessment was built to include the system's complexity and capability of permanent or temporary storage, to help filter the applicable controls that should only be relevant based on the system categories or the intended features and managed data. The proposal for data integrity verification was built to change ALCOA+ into audit trail, electronic records and electronic signatures, user access management, backup, restore and archival and data protection and data privacy that mirrors the control areas based on regulatory guidelines and standards. Last but not the least, the proposal for data integrity training is suggested to address the need for direction, guidance and lack of ownership, including the core discussion for the proposed revisions and clarity of the responsibilities.

5.4 Initial Proposal-2 after Stakeholder Suggestions

The suggestions were addressed and the data integrity compliance procedures training is scheduled after the effective date of the process. The training will discuss the improvements on the risk assessment and verification, the change of the approach from project plan to more operative tasks, the transition approach and grace period for the current/old procedures to the new/proposed procedures, the link between the CSV deliverables and data integrity procedures, the responsibilities and ownership, and the other processes that have data integrity elements but should be covered in the operational phase of the system.

6 Validation of the Proposal

This section reports on the results of the validation stage and points to further developments to the initial Proposal for data integrity compliance assessment and verification process.

6.1 Overview of the Validation Stage

Since the thesis researcher involved the IT department led by the IT Compliance team to finalize the proposal together, majority of the comments are already developed in the Initial proposal-1 based on regulatory guidance and best practices. After the remaining comments are addressed in the Initial proposal-2, these will all be addressed in the planned training. The remaining action plan is to conduct the training, respond to some further questions or clarifications brought up during the training, and include the rest of the clarifications, support and assistance in transitioning to the new improvements.

Pilot testing was conducted for the Data Integrity Compliance of a new Spectrophotometer (Spectramax Plus 384). Assistance with the team was heavily allotted during the risk assessment and verification process where most of the changes are planted. The areas that tend to be unclear to users provided several insights that helped construct the contents of the training. Having one IT Compliance team to spearhead the improvements, there are tendencies that some terminologies and processes are simple in IT- perspective but may require additional explanation for people without extensive IT background.

6.2 Developments to the Proposal (based on Data Collection 3)

The feedback from the stakeholders is that the newly proposed process is more comprehensive but something that is much needed by the people in the organization. Comparing with the previous process, the proposal may minimize the confusion as to why the procedures are being done with the inclusion of data integrity requirements, what is it for by indicating the acceptance criteria for each step and how can it be proven by listing the instructions that any reader can follow and verify within the systems.

Understandably, there are areas that may also be improved and streamlined however these are related to some other existing processes that are outside of the scope of this study. There are minor updates to define simpler and clearer terms within the document and remove those definitive remarks and references that may pose a risk of confusion for the readers. The stakeholders emphasized the significance of explaining the improvements in training and its assignment to individuals who will be involved in the data integrity compliance procedures.

6.3 Final Proposal

As a result of the minor changes and the fundamental training to be conducted, the stakeholders expressed their approval on the proposal. The SOP will be subject to revisions and will have a scheduled effectivity date with an allotment for a grace period. This will be meant for those computerized systems undergoing data integrity compliance procedures in the transitional period. The training has to be sent to the whole organization and a recording of one session for other people who will fail to attend. Necessary follow-up and periodic training must also be considered when the effectiveness of the procedures falters.

No matter how much the procedures differ from one company to another based on their needs and level of maturity, the goal of data integrity compliance remains the same — to assure that the data is truthful and credible to manufacture consistent quality product regardless of the scale, and to eventually support the decisions for the safety of consumers who entrusted regulatory bodies and companies to collectively be transparent on the quality of the drugs and therapeutic products offered to the public.

7 Conclusion

This section concludes the thesis with the executive summary, thesis evaluation and managerial implications.

7.1 Executive Summary

The objective of the thesis was to review and revise the verification process for data integrity compliance of computerized systems of the case company to a more streamlined, efficient and risk-based approach. Data integrity compliance is a subset discipline supporting the GxP-relevant processes to assure the manufacturing of products — its quality, purity, and efficacy, and to eventually warrant the safety of the patients. In companies, decisions are made based on relying on critical data and records therefore data, in its essence, is fundamental to be proven as credible and truthful. As digitalization increased in the pharmaceutical and life science industry, the need for data integrity controls inclusion also increased in configurations, requirements, and specifications. This study used applied action research as its research approach and relied on qualitative research methods in order to improve the verification process of the case company to a more streamlined, efficient, and risk-based approach.

Currently, the data integrity compliance process at the case company is already defined but may have the tendency to mislead the risk assessment when systems complexity is not accounted for, or difficulties in implementation due to ambiguities and missing acceptance criteria. Therefore, the current state analysis in this study determined the focus areas for improvement based on the risks assessed at a system level down to the system functions such as its impact on product recall and lot traceability, regulatory records to be submitted, aiding the manufacturing process, and product labeling, etc. With this, an understanding of the current data integrity maturity level was a starting point for the proposed improvements.

The outcome of this thesis is a proposal on how to strengthen the verification process of data integrity compliance without compromising the quality. The data integrity compliance improvement proposal was formulated collaboratively with the participants, validated, and implemented with additional training throughout the organization to raise awareness of the significant role of each one in the culture and overall compliance to data integrity.

The results of this thesis have redirected the focus areas of data integrity compliance at the case company not only with the criticality of the data, but also including the complexity of the systems. The data integrity risks should also be accounted for whether there is a permanent storage of the GxP records within the system to determine the expected level of electronic and/or manual controls. In addition, electronic signatures and other system features that were under the discretion of the company to be utilized or not must also be considered in the data integrity assessment. Other data integrity regulations such as the data protection to manage personal data within the system must also be included in the risk assessment.

The thesis provided a proposal to revise and improve the data integrity verification process. This proposal is aligned with the current situation of the company by maintaining the process that works, changing or enhancing the process that needs improvement, and performing the procedures in a consistent manner. Following this study, there will always be an area for improvement particularly on this topic when more and more complex technologies will be introduced in the future. Currently, the company goal sets a direction that shifts to an upscale commercialization therefore the data integrity compliance procedures as part of computerized systems validation must accommodate the maturity level of the organization that grows into a more predefined and routinary application of these procedures.

7.2 Managerial Implications (Next Steps and Recommendations toward Implementation)

The next steps would be considered after the learning curve is generally achieved among the responsible team, where data integrity culture is ingrained throughout the organization and the maturity level is at its peak. As more people become knowledgeable and aware of the systems' capabilities and limitations, systems that need replacement or more careful human intervention controls are highlighted. As a result, more rigorous and defined risk assessment will be for computerized systems.

However, the improvements will be nothing if there will not be any system ownership and IT involvement. In a hybrid situation where controls in the system must first be utilized before verifying the other manual controls or periodic tasks, it is crucial to involve the users who are responsible of the oversight, maintenance and recordkeeping of the data residing in the system. Other nice-to-have improvements such as minimizing and enhancing effective closure of deviations, incidents, and CAPAs that may have implications on data integrity. This is currently a controlled and regularly monitored process that can also be a good indicator or measure that the proposed improvements related to this study have created more appreciation and understanding of its purpose and relevance. Also included as the next steps is the removal of duplicate efforts in the CSV process, that may entail the improvement of the CSV process itself.

7.3 Thesis Evaluation

Looking back at the initial objective of the thesis, the improvement on data integrity compliance is substantially met with expected limitations experienced throughout. One of the areas for improvement is a better focal point on a specific topic rather than fixing all the problems. Several problems created a looping situation where the solution is dependent on having another separate problem fixed. The thesis happens in stages and it’s important to take a step back and process what has already been completed in the current phase before moving to the next phase. Collectively, allotting time to process what happened and repeatedly reminding oneself why the thesis topic exists in the first place are motivating for both the thesis researcher and participants to proceed to completion.

The substance of the results are rooted from the repeated revisions that may took a long process to define what will be helpful in long term but also serve as a bridge towards digitalization. When this process improvements are designed, the goal is that when we opt to a more automated tools, these procedures are easily transported and can set as a baseline for multiple projects. The most influential drive is to have an understanding for everyone who will perform it that after the execution, they will understand the risks involved and the value of why they did it.

One of the remarkable holdback is this huge hesitation at the beginning for those who are already accustomed to the current process. Fortunately for most who aimed for the improvements and are also involved mostly in the pilot project, it was altogether rewarding to receive gratitude for the guidance, to witness the growth of the people in data integrity maturity, and to hear their responses in reviewing the proposal.

7.4 Closing Words

Improvements on data integrity compliance will always be present as long as there are opportunities for more complex and newer technologies. The thesis research does not basically end here and the learning path for this topic will grow. Expected amendments on data integrity will still ensue, nevertheless, the bottom line of it all is that whichever process one follows, whether its efficient or inefficient, one should never compromise the health and well-being of a person. As it may not matter that other process will still require more time and resources, the importance of having the efficiency considerations is an extension of another life to be saved.

References

‘GXP’ Data Integrity Guidance and Definitions 2018, Medicines & Healthcare products Regulatory Agency (MHRA), viewed 12 March 2022

Data Integrity and Compliance With CGMP Guidance for Industry 2016, U.S. Department of Health and Human Services Food and Drug Administration (FDA), viewed 18 January 2022

GAMP 5 – A Risk-Based Approach to Compliant GxP Computerized Systems, Second Edition 2022, International Society for Pharmaceutical Engineering (ISPE), viewed 11 November 2022

GAMP 5 – Records and Data Integrity Guide 2017, International Society for Pharmaceutical Engineering (ISPE), viewed 05 June 2022

Good Practices for Computerized Systems in Regulated “GXP” Environments 2007, Pharmaceutical Inspection Convention, Pharmaceutical Inspection Co-operation Scheme (PIC/S) Guidance, viewed 25 August 2022

Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments 2021, Pharmaceutical Inspection Convention, Pharmaceutical Inspection Co-operation Scheme (PIC/S) Guidance, viewed 22 August 2022

Kananen, J. 2017. Design Research (Applied Action Research) as Thesis Research. JAMK University of Applied Sciences, 2013.

Practical risk-based guide for managing data integrity 2019, Active Pharmaceutical Ingredients Committee (APIC), viewed 31 January 2022

Saunders, M. N. K. 2019. Research Methods for Business Students. Pearson, Ltd.

The Rules Governing Medicinal Products in the European Union Volume 4 Good Manufacturing Practice, Guidelines on Good Manufacturing Practice

Medicinal Products for Human and Veterinary Use specific to Annex 11: Computerized Systems specific to Advanced Therapy Medicinal Products 2011, European Commission (EudraLex), viewed 21 February 2022

The Rules Governing Medicinal Products in the European Union Volume 4 Good Manufacturing Practice, Guidelines on Good Manufacturing Practice specific to Advanced Therapy Medicinal Products 2017, European Commission (EudraLex), viewed 24 February 2022

US Food and Drug Administration (FDA) 2023. 21 Code of Federal Regulations (CFR) Part 11 – Electronic Records; Electronic Signatures. Accessed 05 May 2023. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-11

US Food and Drug Administration (FDA) 2023. 21 Code of Federal Regulations (CFR) Part 211 – Current Good Manufacturing Practice For Finished Pharmaceuticals. Accessed 05 May 2023. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-C/part-211

US Food and Drug Administration (FDA) 2023. 21 Code of Federal Regulations (CFR) Part 820 – Quality System Regulation. Accessed 05 May 2023. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-820

European Union (EU) 2023. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Chapter II – Principles, Article 6 – Lawfulness of processing, Accessed 10 May 2023.https://eur-lex.europa.eu/legalcontent/EN/TXT/?uri=celex%3A32016R0679