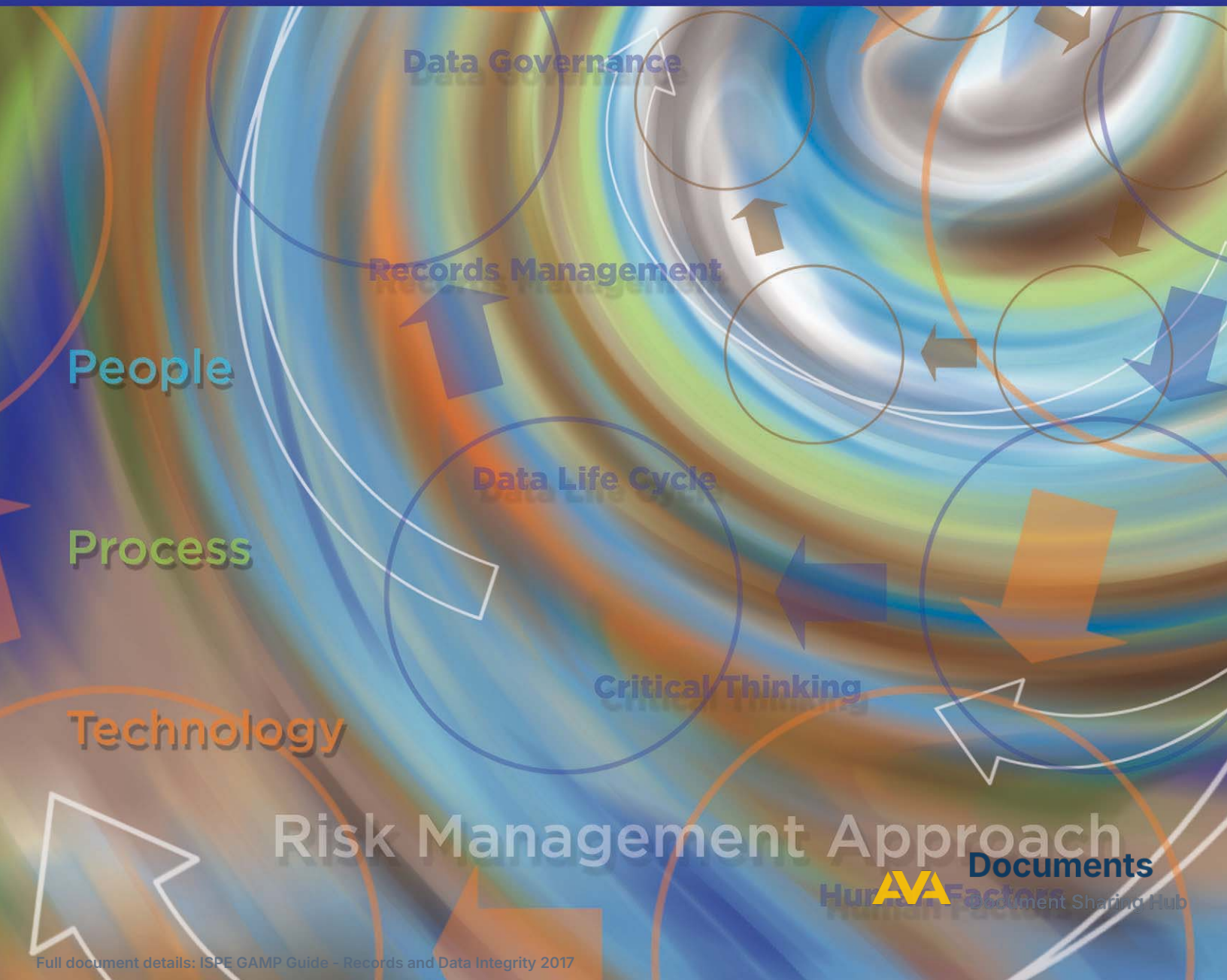


Records and Data Integrity GUIDE



Records and Data Integrity GUIDE

Disclaimer:

This Guide is intended to assist regulated companies in managing records and data throughout the data life cycle. ISPE cannot ensure and does not warrant that a system managed in accordance with this Guide will be acceptable to regulatory authorities. Further, this Guide does not replace the need for hiring professional engineers or technicians.

Limitation of Liability

In no event shall ISPE or any of its affiliates, or the officers, directors, employees, members, or agents of each of them, or the authors, be liable for any damages of any kind, including without limitation any special, incidental, indirect, or consequential damages, whether or not advised of the possibility of such damages, and on any theory of liability whatsoever, arising out of or in connection with the use of this information.

© Copyright ISPE 2017. All rights reserved.

All rights reserved. No part of this document may be reproduced or copied in any form or by any means – graphic, electronic, or mechanical, including photocopying, taping, or information storage and retrieval systems – without written permission of ISPE.

All trademarks used are acknowledged.

ISBN 978-1-936379-96-5

Preface

The importance of data integrity is reflected in recent guidance, citations, and public comments of Regulators and Health Agencies. A number of companies have suffered serious regulatory and financial consequences as a result of unacceptable data integrity practices.

Patient safety is affected by the integrity of critical records, data, and decisions, as well as those aspects concerned with physical attributes of the product. That the phrase “patient safety, product quality, and data integrity” is commonly used in regulatory and industry guidance underlines this point.

The use of information technology and computerized systems in all aspects of life sciences continues to grow and has resulted in the generation of more data to support the development and manufacture of products. Key decisions and actions are routinely being made based on this data, and the integrity of the data, whether in electronic or paper form, is of paramount importance to the industry, the regulatory agencies, and ultimately the patient.

Industry will benefit from clear guidance on ensuring that the management of records and data forms an integral part of the Quality Management System, and is compliant with GxP requirements. This Guide intends to provide such guidance and is aligned with *ISPE GAMP® 5: A Risk-Based Approach to Compliant GxP Computerized Systems*.

Acknowledgements

The *ISPE GAMP® Records and Data Integrity Guide* was produced by a Task Team led by:

Nigel Price	QCDI Ltd.	United Kingdom
Mike Rutherford	Eli Lilly and Company	USA
Sion Wyn	Conformity Limited	United Kingdom

The work was supported by the ISPE GAMP Community of Practice (CoP).

Core Team

The following individuals took lead roles in the preparation of this Guide:

Chris Clark	TenTenTen Consulting	United Kingdom
Colin Jones	Conformity Limited	United Kingdom
Tony Margetts	Factorytalk Co., Ltd.	Thailand
Mark Newton	Eli Lilly and Company	USA
Arthur “Randy” Perez	Novartis (retired)	USA
Chris Reid	Integrity Solutions Ltd.	United Kingdom
Lorrie Vuolo-Schuessler	GlaxoSmithKline	USA
Charlie Wakeham	Waters Australia Pty. Ltd.	Australia
Christopher White	Alexion Pharmaceuticals	USA
Guy Wingate	GlaxoSmithKline	United Kingdom

Regulatory Input and Review

Particular thanks go to the following for their review and comments on this Guide:

David Churchward	MHRA	United Kingdom
Stephen Grayson	MHRA	United Kingdom
Karl-Heinz Menges	Regierungspräsidium Darmstadt	Germany
Ian Thrussell	World Health Organization	United Kingdom

Subject Matter Expert Input and Review

Particular thanks go to the following for their review and comments on this Guide:

Monica Cahilly	Green Mountain Quality Assurance, LLC	USA
Robert McDowall	R.D. McDowall Ltd.	United Kingdom

The Team would like to special thanks to the Global GAMP Data Integrity Special Interest Group (SIG) for their efforts.

The Team Leads would like to express their grateful thanks to the many individuals and companies from around the world who reviewed and provided comments during the preparation of this Guide; although they are too numerous to list here, their input is greatly appreciated.

Company affiliations are as of the final draft of the Guide.



600 N. Westshore Blvd., Suite 900, Tampa, Florida 33609 USA
Tel: +1-813-960-2105, Fax: +1-813-264-2816

www.ISPE.org

AVA **Documents**
Document Sharing Hub

Table of Contents

1	Introduction	9
1.1	Background	9
1.2	Purpose	9
1.3	Scope	10
1.4	Structure of this Guide	10
1.5	Key Concepts	11
1.6	Key Terms	15
2	Regulatory Focus	17
2.1	Introduction	17
2.2	Data Integrity Requirements	17
3	Data Governance Framework	21
3.1	Introduction	21
3.2	Overview	21
3.3	Elements of the Data Governance Framework	23
3.4	Human Factors in Data Integrity	30
3.5	Data Integrity Maturity Model	31
4	Data Life Cycle.....	33
4.1	Introduction	33
4.2	Data Creation	34
4.3	Data Processing.....	35
4.4	Data Review Reporting and Use.....	36
4.5	Data Retention and Retrieval	39
4.6	Data Destruction	42
5	Quality Risk Management	43
5.1	Introduction	43
5.2	Process Risk Assessment.....	43
5.3	Quality Risk Management Approach.....	43
5.4	Product and Process Context	45
 <u>Management Appendices</u>		
6	Appendix M1 – Corporate Data Integrity Program.....	47
6.1	Introduction	47
6.2	Is a Corporate Data Integrity Program Required?.....	47
6.3	Indicators of Program Scope and Effort.....	48
6.4	Implementation Considerations.....	50
6.5	Keys to Success.....	52
7	Appendix M2 – Data Integrity Maturity Model	55
7.1	Maturity Model.....	55
7.2	Data Integrity Maturity Level Characterization	59

8	Appendix M3 – Human Factors	67
8.1	Introduction	67
8.2	Corporate and Local Cultures	67
8.3	Classification of Incidents.....	68
8.4	Human Error.....	69
8.5	Data Falsification and Fraud	70
8.6	Impartiality.....	71
8.7	Behavioral Controls.....	71
9	Appendix M4 – Data Audit Trail and Audit Trail Review	75
9.1	Introduction	75
9.2	Regulatory Background.....	76
9.3	Application and Use of Audit Trails.....	77
9.4	Audit Trail Review	79
9.5	Technical Aspects and System Design	79
10	Appendix M5 – Data Auditing and Periodic Review	81
10.1	Introduction	81
10.2	Auditing for Data Integrity.....	81
10.3	Periodic Review	82
10.4	Other Reviews.....	83
10.5	Documenting Review Processes	83
11	Appendix M6 – Inspection Readiness	85
11.1	General Procedures	85
11.2	Key Information for Regulatory Inspections	86
12	Appendix M7 – Integrating Data Integrity into Existing Records Management Processes	91
12.1	Introduction	91
12.2	Record Creation.....	92
12.3	Active Records.....	92
12.4	Semi-active Records.....	92
12.5	Inactive Records	92
 Development Appendices		
13	Appendix D1 – User Requirements	93
13.1	Introduction	93
13.2	Business Process.....	93
13.3	General Data Integrity Requirements.....	94
14	Appendix D2 – Process Mapping and Interfaces	99
14.1	Introduction	99
14.2	Process Flowcharts.....	99
14.3	Data Flow Diagrams.....	102
14.4	How Much Is Needed?.....	103

15 Appendix D3 – Risk Control Measures for Records, Data, and Electronic Signatures.....	105
15.1 Introduction	105
15.2 Record and Data Controls.....	105
15.3 Electronic Signature Controls.....	105
15.4 Implementation of Record and Data Controls	107
15.5 Rigor of Controls	110
16 Appendix D4 – Data Integrity Concerns Related to System Architecture	111
16.1 Data Resides on a Local Hard Disk	111
16.2 Internally Managed Central Database.....	112
16.3 Internally Managed Distributed Data.....	112
16.4 Outsourced Managed Services.....	113
17 Appendix D5 – Data Integrity for End-User Applications.....	117
17.1 Introduction	117
17.2 Data Integrity for Spreadsheets	117
17.3 Data Integrity for PC Databases	119
17.4 Data Integrity for Statistical Tools.....	120

Operation Appendices

18 Appendix O1 – Retention, Archiving, and Migration.....	121
18.1 Introduction	121
18.2 Retention Options	121
18.3 Protection of Records.....	121
18.4 Record Aging and Risk.....	122
18.5 Archival	122
18.6 Hybrid Situations and Archives	123
18.7 Audit Trail Considerations	124
18.8 Alternative Systems	125
18.9 Converting Electronic to Alternative Format or Alternative Media Hybrids.....	126
19 Appendix O2 – Paper Records and Hybrid Situations.....	131
19.1 Paper Records	131
19.2 Hybrid Situations	133
19.3 Use of Forms to Enforce Procedures.....	135

General Appendices

20 Appendix G1 – References	137
21 Appendix G2 – Glossary	141
21.1 Acronyms and Abbreviations.....	141
21.2 Definitions	143

1 Introduction

1.1 Background

The impact of record and data integrity issues can be significant on a regulated company. It can result in recalls of products, warning or untitled letters, import alerts, injunctions, seizures, Application Integrity Policy Invocations/legal action, and ultimately the potential for patient harm. These regulatory actions can also have a significant financial impact.

There has been increased regulatory focus on all aspects of data integrity, including publication of specific regulatory guidance on the topic, and increased number of citations in the area.

For the purposes of this Guide:

- Regulated data is information used for a regulated purpose or to support a regulated process.
- *"Metadata is data that describes the attributes of other data, and provide context and meaning. Typically, these are data that describe the structure, data elements, inter-relationships, and other characteristics of data."* [1].
- A regulated record¹ is a collection of regulated data (and any metadata necessary to provide meaning and context) with a specific GxP purpose, content, and meaning, and required by GxP regulations. Records include instructions as well as data and reports.
- *"Data Integrity is defined as the extent to which all data are complete, consistent and accurate throughout the data life cycle."* [1]
- The integrity of records depends on the integrity of underlying data, and signatures executed to electronic records should be trustworthy and reliable. See Appendix D3.

This Guide addresses paper records, electronic records, and hybrid situations, while encouraging a move away from hybrid situations, wherever practical.

1.2 Purpose

This *ISPE GAMP® Guide: Records and Data Integrity* provides principles and practical guidance on meeting current expectations for the management of GxP regulated records and data, ensuring that they are complete, consistent, secure, accurate, and available throughout their life cycle. This approach is intended to encourage innovation and technological advance while avoiding unacceptable risk to product quality, patient safety, and public health.

This Guide is intended as a stand-alone Guide. It is aligned with *ISPE GAMP® 5: A Risk-Based Approach to Compliant GxP Computerized Systems* [3]. This Guide has been designed so that it can be used in parallel with guidance provided both in *ISPE GAMP® 5* [3] and other *ISPE GAMP® Good Practice Guides* [4].

Although the scope of this document is wider, it replaces the *ISPE GAMP® Good Practice Guide: A Risk-Based Approach to Compliant Electronic Records and Signatures*.

¹ Within the US regulatory framework regulated electronic records, and associated signatures, are subject to 21 CFR Part 11 [2]. For further information see Appendix D3.

1.3 Scope

This guide addresses the integrity of GxP records and data used within the regulated life science industries including pharmaceutical, biological, and medical devices. The guidance is intended for regulated companies and suppliers of systems, products, or services in this area, as well as a useful reference for regulators.

Applicable life science regulations and guidance have been taken into account, and the following publications have been specifically considered:

- US Code of Federal Regulations (CFRs) [5] covering GCP, GLP, GMP, and medical devices
- US 21 CFR Part 11 [2] and associated guidance
- Relevant sections of EU GMPs including Chapter 4 [6] and Annex 11 [7]
- MHRA GMP Data Integrity Definitions and Guidance for Industry (Revision 1.1, March 2015) [1]
- MHRA GxP Data Integrity Definitions and Guidance for Industry (Draft version for consultation July 2016) [8]
- FDA Draft Guidance for Industry: Data Integrity and Compliance with CGMP [9]
- ICH Q9 Quality Risk Management [10]
- ICH Q10 Pharmaceutical Quality System [11]
- WHO Annex 5: Guidance on Good Data and Record Management Practices [12]
- PIC/S Draft Guidance: Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments (Draft 2 published August 2016) [13]

This Guide provides a method for managing risk to record and data integrity. Regulated companies may already have established risk management activities and tools, and this Guide does not intend or imply that these existing methods should be discarded, rather that they continue to be used as appropriate within the context of the overall risk management process described. Other methods or techniques giving documented evidence of adequate control, and ensuring appropriate security and integrity, may also be acceptable.

This Guide may also be useful in other regulated areas such as cosmetics and food, or in other areas or sectors where data integrity is important.

1.4 Structure of this Guide

This Guide contains this introduction, a main body, and a set of appendices. It has been structured to meet the needs of various readers, and contains, in increasing level of detail:

1. Data integrity requirements, critical areas of regulatory focus and concern, and key concepts
2. A framework for data governance and the importance of human factors
3. A complete data life cycle approach as part of a Quality Management System (QMS), from creation to destruction
4. Further information on how to apply the Quality Risk Management (QRM) approach from *ISPE GAMP® 5* [3] to record and data integrity

5. More detailed information, including “how to” guidance for specific topics, in a series of management, development, and operation appendices

1.5 Key Concepts

This section describes key concepts that apply throughout this Guide.

1.5.1 Risk Management Approach

A holistic and flexible risk management approach should be used to ensure the integrity of records and data. This is achieved by the application of appropriate controls to manage identified risks within the context of the regulated process. The effort required to assess and manage risk should be commensurate to the level of risk. Critical thinking and analysis skills should be applied to identify and adequately control risks to patient safety, product quality, and data integrity.

The QRM approach defined in *ISPE GAMP®* 5 [3], following ICH Q9 [10], (and also detailed in Section 5) can be applied to identifying, assessing, and managing risks to data and record integrity.

A full understanding of the regulated process to be supported, including the intended use of data within the process, is fundamental. Data integrity cannot be achieved without a complete understanding of the data flow.

1.5.2 Data Governance

Data governance is the sum total of arrangements to ensure that data is recorded, processed, retained and used to ensure a complete, consistent, and accurate record throughout the data life cycle [1]. Data governance ensures formal management of records and data throughout the regulated company. Data governance encompasses the people, processes, and technology required for effective data handling. See Section 3.

1.5.3 Data Life Cycle

All phases in data life cycle from initial data creation, capture, and recording through processing (including transformation or migration), review, reporting, retention, retrieval, and destruction should be controlled and managed in order to ensure accurate, reliable, and compliant records and data. See Section 4.

Details of the life cycle will vary depending on the type of documentation. See Section 4.

Two main types of documents are defined by EU GMP Chapter 4 [6]:

1. Instructions (directions or requirements) type, e.g., specifications, manufacturing formulae, processing, packaging, and testing instructions, SOPs, protocols, and technical agreements
2. Record/report type, e.g., batch records, laboratory testing results, certificates of analysis, reports

Regulated data should be controlled and managed, and integrity of the data ensured, e.g., following the principles and requirements described in this Guide. All regulated data is subject to GxP requirements for data integrity and good documentation practices.

A primary record is the record which takes priority in cases where data is collected and retained concurrently by more than one method, and the data does not correspond. The primary record attribute should be defined and documented, and should not be changed on a case by case basis. The UK MHRA [1] defines a primary record as:

“The record which takes primacy in cases where data that are collected and retained concurrently by more than one method fail to concur.”

Risk management principles should be used to ensure that the primary record provides the greatest accuracy, completeness, content, and meaning [1].

For example, high resolution or dynamic (electronic) data should be designated as a primary record in preference to low resolution or static (printed/manual) data. All relevant data should be considered when performing activities such as a risk-based investigation into data anomalies (e.g., out of specification results) [1].

1.5.4 Key Concepts Summarized by ALCOA and ALCOA+

Both the WHO Guidance [12] and the draft PIC/S Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments [13] indicate that key concepts described by the ALCOA (and ALCOA+) acronyms can help to support record and data integrity.

The WHO Annex 5: Guidance on Good Data and Record Management Practices [12] states:

“The basic building blocks of good GXP data are to follow GDocP and then to manage risks to the accuracy, completeness, consistency and reliability of the data throughout their entire period of usefulness – that is, throughout the data life cycle.

Personnel should follow GDocP for both paper records and electronic records in order to assure data integrity.”

PIC/S [13] further states:

“The application of GDocPs may vary depending on the medium used to record the data (ie. physical vs. electronic records), but the principles are applicable to both...” and “Some key concepts of GDocPs are summarised by the acronym ALCOA...”

Along with the additional key concepts described by ALCOA+, PIC/S [13] goes on to state:

“Together, these expectations ensure that events are properly documented and the data can be used to support informed decisions.”

Tables 1.1 and Table 1.2 provide information of how key concepts described by ALCOA and ALCOA+ should be applied throughout the data life cycle.

Table 1.1: ALCOA

Principle	Data Expectation
Attributable	• Attributable to the person or system generating the data • Identify the person or system performing an activity that creates or modifies data • Linked to the source of the data
Legible	• Readable and permanent • Accessible throughout the data life cycle • Original data and any subsequent modifications are not obscured
Contemporaneous	• Recorded or observed at the time the activity is performed
Original	• Original data is the first recording of data, or a “true copy” which preserves content or meaning
Accurate	• Free from error • No editing performed without documented amendments • Conforming to truth or standard

Table 1.2: ALCOA+

Principle	Data Expectation
Complete	• All data, and relevant metadata, including any repeat or re-analysis performed
Consistent	• Application of good documentation practices throughout any process • The application of date and time stamps in the expected sequence
Enduring	• Recorded in a permanent, maintainable form for the retention period
Available	• Available and accessible for review, audit, or inspection throughout the retention period

1.5.5 Critical Thinking

Critical thinking is a systematic, rational, and disciplined process of evaluating information from a variety of perspectives to yield a balanced and well-reasoned answer. Critical thinking allows the effective interpretation of data and situations while avoiding personal biases, assumptions, and other factors [14].

The application of critical thinking skills allows the identification of gaps in data governance and processes, and assists in challenging the effectiveness of behavioral, procedural, and technical controls in achieving data integrity.

Critical thinking is an important component of data integrity, and many regulators are trained in critical thinking to help them more quickly to identify and assess risk to product quality and patient safety. PIC/S Guidance [13] states:

“Critical thinking skills should be used by inspectors to determine whether control and review procedures effectively achieve their desired outcomes.”

Elements of critical thinking include:

- Analyzing situations through gathering relevant details, and reviewing them carefully and objectively through applying knowledge and experience
- Gathering and evaluating information from different sources, understanding links between concepts and ideas, and identifying inconsistencies and errors in reasoning
- Analyzing situations and solving problems consistently, systematically, and logically
- Evaluating information in an open-minded manner to better interpret and understand all available data and signals
- Challenging and questioning ideas and assumptions in a rational and balanced manner
- Comparing, contrasting, and testing alternatives based on ambiguous, incomplete, or partial information
- Creating, developing, and applying new models from experience
- Designing new processes to meet changing process, technical, and regulatory needs

Critical thinking encourages a product quality driven and patient focused approach, rather than a document driven and purely compliance focused approach.

This Guide encourages regulated companies to apply critical thinking and promote its application through leadership, awareness, and training. This Guide encourages the application of critical thinking as part of a holistic top-down risk-based approach.

The identification of appropriate and effective controls within a specific process context, in accordance with an understanding of risks to patient and product, is encouraged.

1.5.6 GxP Computerized System Life Cycle

Data integrity is underpinned by well-documented, validated GxP computerized systems, and the application of appropriate controls throughout both the system and data life cycles.

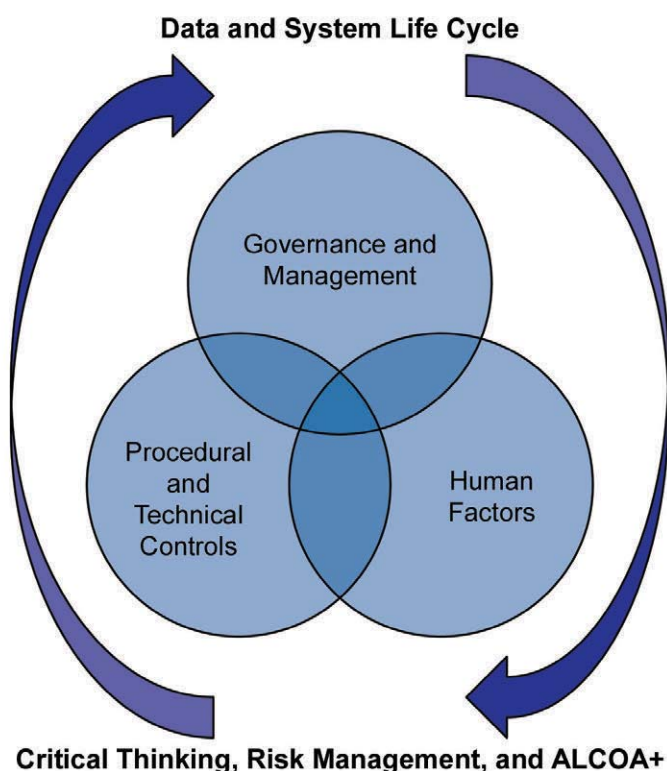
Multiple GxP computerized systems may be involved in supporting a data life cycle, as the data may be passed from system to system. To ensure data integrity all GxP computerized systems should be trustworthy and validated for intended use.

A system life cycle approach, such as described in *ISPE GAMP® 5* [3], should be applied to each GxP computerized system. Record and data integrity should be built-in and maintained throughout the GxP computerized system life cycle phases, from concept through project and operations, to retirement. The GxP computerized system life cycle activities should be scaled based on the complexity and novelty of the system, and potential impact on product quality, patient safety, and data integrity.

The inherent risk of standalone GxP computerized systems may be greater than for enterprise GxP computerized systems. Standalone GxP computerized systems may require different approaches. There should be sufficient effort placed on identifying and managing standalone systems, based on level of impact and vulnerability, when designing controls, defining the data life cycle, and applying the Data Governance Framework.

1.5.7 Summary of the Key Concepts

Figure 1.1: Key Concepts



Data is managed through a controlled data management life cycle within a Data Governance Framework. A holistic risk management approach is applied to manage risks to data integrity and to ensure that the principles of ALCOA+ are met. The data life cycle may be supported by one or more computerized systems that should be trustworthy and compliant.

1.6 Key Terms

Regulated Data

Information used for a regulated purpose or to support a regulated process.

Metadata (MHRA, 2015 [1], MHRA, 2016 [8])

"Metadata is data that describes the attributes of other data, and provide context and meaning. Typically, these are data that describe the structure, data elements, inter-relationships and other characteristics of data." [1] "It also permits data to be attributable to an individual (or if automatically generated, to the original data source)." [8]

Regulated Record

A collection of regulated data (and any metadata necessary to provide meaning and context) with a specific GxP purpose, content, and meaning, and required by GxP regulations. Records include instructions as well as data and reports.

Atypical / Aberrant / Anomalous Result (MHRA Out Of Specification Investigations Guidance [15])

"Results that are still within specification but are unexpected, questionable, irregular, deviant or abnormal. Examples would be chromatograms that show unexpected peaks, unexpected results for stability test point, etc."

2 Regulatory Focus

2.1 Introduction

It is a regulatory expectation that GxP data and records are complete, consistent, reliable, accurate, that their content and meaning are preserved, and that they are available and usable for the required retention period.

Specific areas impacting data integrity, which have been of particular regulatory focus and concern include:

- Lack of basic access control and security measures allowing unauthorized changes
- Shared user logins
- Missing or disabled audit trails
- Lack of contemporaneous recording of activities
- Failure to investigate data discrepancies
- Testing into compliance
- Incomplete collection, retention, and review of data for quality decisions
- Overwriting or deletion of original data
- Data falsification

The FDA [9] states:

“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.”

2.2 Data Integrity Requirements

This section provides an overview of the data integrity expectations on regulated companies based on published regulatory guidance documents. It is an overview and should not be considered as all inclusive.

Regulated companies should have confidence in the quality and the integrity of the data used to make decisions impacting product quality and patient safety.

Data integrity controls for records should ensure that the accuracy, completeness, content, and meaning of data is maintained throughout the data life cycle. The principles of ALCOA+ should be applied.

MHRA states:

“The effort and resource applied to assure the validity and integrity of the data should be commensurate with the risk and impact of a data integrity failure to the patient or environment.” [8]

“The data governance system should be integral to the pharmaceutical quality system...” [1]

Regulated companies should implement meaningful and effective holistic strategies to manage risks to data integrity, based upon their process understanding and knowledge of technologies. Critical analysis skills should be applied to identify and adequately control risks to data integrity, and to investigate and address root causes if failures occur.

The impact on quality may be determined by considering the type of decisions or activities influenced by the data. Risk to data reflects its vulnerability to unauthorized or inadvertent deletion or amendment, and the opportunity for detection during routine review. Risk to data is typically increased by complex, inconsistent processes, with open ended outcomes that are open to subjective interpretation, compared to simple tasks that are consistent, well defined, and objective [8].

Regulated companies should maintain an inventory of systems generating and maintaining data and the capability of each system. An inventory of documents should be maintained within the Quality Management System (QMS) [6].

Data governance activities should be integral to the QMS [1]. Data governance activities should be supported by an organizational culture that enforces data integrity, led by senior management leadership and behavior. Senior management should:

- Take primary responsibility for data integrity by initially understanding the capabilities and limitations of existing processes, methods, environment, personnel, and technologies
- Subsequently ensure the allocation of necessary resources to address any identified limitations and maintain data integrity

As part of data governance, regulated companies should take the following three types of steps to achieve an acceptable level of data integrity:

- Behavioral
- Procedural
- Technical

2.2.1 Behavioral Steps

Senior management should take responsibility for establishing and maintaining a culture which supports data integrity. Data governance should address the:

- Ownership of data throughout the data life cycle
- Training of personnel in the importance of data integrity principles [8]
- Creation of a working environment that:
 - Enables visibility of errors, omissions, and atypical (aberrant) results [8]
 - Encourages transparent investigation and analysis

Personnel should be trained in the importance of data integrity, and in the methods of detecting data integrity issues, as part of routine GxP training programs. See Section 3.3.7.

Senior management should also establish arrangements to ensure personnel are not subject to commercial, political, financial, and other pressures or conflicts of interest that may adversely affect the quality of their work and integrity of their data [12].

The quality manual, or equivalent document, should include a quality policy statement of management's commitment to an effective QMS. The policy statement should include a code of ethics and a code of proper conduct which are intended to assure the reliability and completeness of data, including mechanisms for personnel to report any quality and compliance questions or concerns to management [12].

Implementation of an effective quality culture and data governance may be different in different locations. A single approach to quality management or data governance may not be effective in all situations, e.g., if cultural differences and dynamics challenge the societal acceptability of open reporting of problems, and challenging of hierarchy.

2.2.2 Procedural Steps

Regulated companies should implement systems and procedures to:

- To minimize the potential risk to data integrity [1]
- Identify the residual risk, using risk management techniques following the principles of ICH Q9 [10, 1]
- Assess risks associated with any third parties. This should include:
 - Consideration of the control of associated risks
 - How the risks may be addressed in contracts and quality/technical agreements

Data which is manually recorded requires a high level of supervision. Supervisory measures or technical controls should be considered to reduce risk. Examples include second person verification at the same time as the data entry is made or cross checks of related information sources [8].

Records should be accessible at locations where regulated activities take place. Ad hoc data recording and later transcription to official records should be discouraged and should not be necessary [1].

Access to blank paper templates for original data recording should be controlled, where appropriate. Reconciliation of blank paper templates and documentation may be necessary to prevent recreation of a record without control and traceability.

Blank forms (including worksheets, laboratory notebooks, and master production control records) should be controlled by the Quality Unit or by another document control method. For example, numbered sets of blank forms may be issued, as appropriate. The numbering of completed forms should be compared to the numbering of all issued blank forms to ensure that they match. Incomplete or erroneous forms should be kept as part of the permanent record along with written justification for their replacement [9].

Regulated companies should allow and encourage correct performance of tasks and accurate recording of data, as required; regulated companies should control physical aspects such as space, equipment, and the timing of events to support these activities. Regulated records and data (including hybrid situations) should be held in secured areas. Appropriate access should be provided to all relevant data for personnel performing data checking activities.

2.2.3 Technical Steps

GxP computerized systems should be validated for intended use, supporting infrastructure qualified, and equipment qualified and calibrated as necessary.

Regulated companies should apply access controls to ensure that people have access only to functionality that is appropriate for their job role, and that actions are attributable to a specific individual [1]. Controls should prevent the unauthorized deletion or modification of regulated data and records, inside or outside the software application (e.g., limiting user rights). Regulated companies should be able to demonstrate the access levels granted to individual staff members and ensure that historical information regarding user access level is available [1]. User access arrangements should prevent (or if necessary log) unauthorized data amendments.

Basic access control measures should be established (e.g., unique usernames and private passwords, and policies and processes for their management). Shared logins or generic user access should not be used as they do not allow for traceability to an individual. Where the computerized system design supports individual user access, this function should be used [1].

Appropriate segregation of duties should be applied. Account privileges should be limited to those required for individuals to perform their duties, e.g., users, supervisors, Quality Unit and administrators.

Elevated privileges permitting activities such as data deletion, database amendment, or system configuration changes (e.g., system administrator rights), should not be assigned to individuals with a direct interest in the data (e.g., data generation, or data review or approval) [1].

Where possible automated data capture techniques should be applied to minimize the risk of data transcription error. Appropriately controlled and synchronized clocks should be available for recording timed events.

Computerized systems should be designed to ensure that the execution of critical steps is recorded at the same time as they are performed [8], and are individually traceable. The EU GMP Annex 11 [7] states that:

“Consideration should be given, based on a risk assessment, to building into the system the creation of a record of all GMP-relevant changes and deletions (a system generated “audit trail”).”

Audit trails are required when users create, modify, or delete regulated records during normal operation [16]. The reason for change or deletion of regulated data should be documented and be consistent with regulatory expectations. The EU GMP Chapter 4 [6] states that:

“Any alteration made to the entry on a document should be signed and dated; the alteration should permit the reading of the original information. Where appropriate, the reason for the alteration should be recorded.”

Audit trail functionality should be available, enabled, and verified. Routine data review should include audit trail review where relevant.

Data transfer or migration should be designed and validated to ensure that data integrity principles are maintained. Copies of records should preserve the integrity (accuracy, completeness, content, and meaning) of the original record. Backup and recovery processes should be validated and periodically tested. Security controls should be in place to ensure the data integrity of the record throughout the retention period. Security controls should be validated, where appropriate.

Archival arrangements should be established for the long-term retention of regulated data (this should be accurate and complete) in compliance with legislation. The procedures for destruction of data should consider data criticality and any applicable regulatory or legal requirements [1].

Where outsourced or cloud services are used, attention should be paid to understanding the service provided, ownership, retrieval, retention and security of data [8], and the role quality agreements can play in such understanding.

The physical location where the data is held, including the impact of any laws applicable to that geographic location, should be considered. The responsibilities of the service provider (particularly quality related aspects and requirements) should be defined in a technical/quality agreement or contract [8].

3 Data Governance Framework

3.1 Introduction

This section describes a framework for data governance, covering:

- Definition and overview of data governance
- Elements of data governance
- Importance of human factors in data integrity
- Maturity levels for data governance

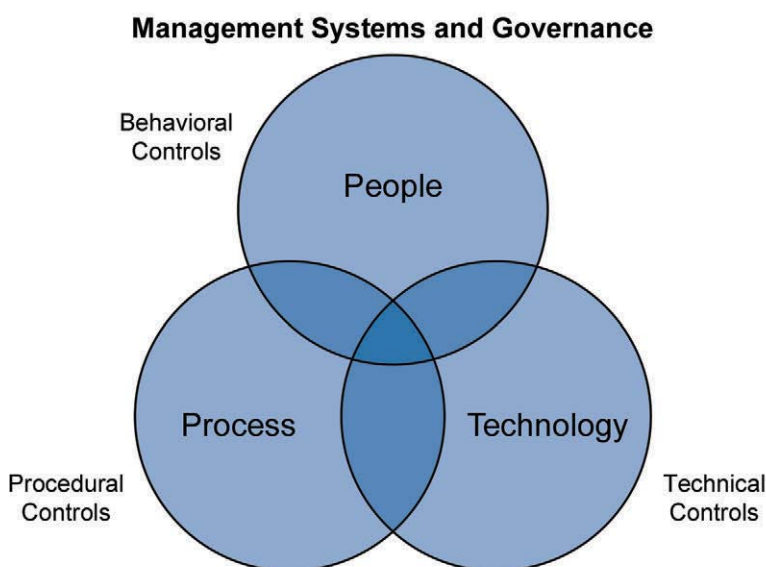
3.2 Overview

Data governance may be defined as (MHRA, 2015 [1]):

"The sum total of arrangements to ensure that data, irrespective of the format in which it is generated, is recorded, processed, retained and used to ensure a complete, consistent and accurate record throughout the data life cycle."

Data governance ensures formal management of records and data throughout the regulated company. 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. It may also be considered as a quality assurance and control approach for applying rigor and discipline to the process of managing, using, protecting, and improving organizational information.

Figure 3.1: Elements of the Data Governance Framework



Elements of data governance are closely related to regulatory requirements. Data governance activities should be integral to the QMS [1]. A specific Corporate Data Integrity Program may be necessary to address any gaps in a regulated company's existing Quality Management System.

Senior management commitment and leadership is considered essential to the effectiveness of the data governance framework. Lack of explicit and demonstrable senior management commitment risks ineffective data governance.

Training should be provided for personnel on the importance of data integrity principles and policies [8]. All personnel should be encouraged to report instances of data integrity failures, bad practice, or falsification, without fear of penalty. Such reports should be fully and transparently investigated by senior management, including root cause analysis and the establishment of prevention or detection measures.

Data governance should also address data ownership and responsibilities throughout the data life cycle [1]. Individuals accountable and responsible for specific data, and its integrity and compliance, at various stages of the data life cycle should be defined and documented.

Regulated companies may recognize the need to manage data as a corporate asset. An executive level role, such as a Chief Data Officer (CDO) may be appointed to oversee this area.

The specification, design, validation, and operation of processes and systems should meet the defined requirements for regulated data integrity. This should include ensuring appropriate control over intentional, unintentional, authorized, and unauthorized changes to regulated data.

The risk to data integrity, especially as it may be related to risk to product quality and product safety should be managed by an established Quality Risk Management process, and defined as part of the QMS. This should include consideration of the risk to data integrity associated with any outsourcing of activities or use of service providers, which should be assessed and managed through appropriate formal agreements. Risk associated with use of summary reports or data should be considered, and this may include a review of the mechanisms used to generate and distribute summary data and reports, based on risk.

The data governance approach should be holistic, proportionate, and integrated (MHRA, 2015 [1]):

"The data governance system should be integral to the pharmaceutical quality system described in EU GMP chapter 1. The effort and resource assigned to data governance should be commensurate with the risk to product quality, and should also be balanced with other quality assurance resource demands. As such, manufacturers and analytical laboratories are not expected to implement a forensic approach to data checking on a routine basis, but instead design and operate a system which provides an acceptable state of control based on the data integrity risk, and which is fully documented with supporting rationale."

Manual systems and paper based records may be a key area of data integrity failure. Risks associated with manual systems, including risks at the interface between manual and computerized systems, uncontrolled copies, and multiple inconsistent copies, should also be considered. Computerized systems related activities are only one part of the broader governance framework, and equivalent considerations are required for paper based systems and processes. See Appendix O2.

Human factors are a critical aspect of an effective Data Governance Framework, including the topics of cultural differences, human error, understanding and awareness, and motivation and behavior. See Section 3.4 and Appendix M3.

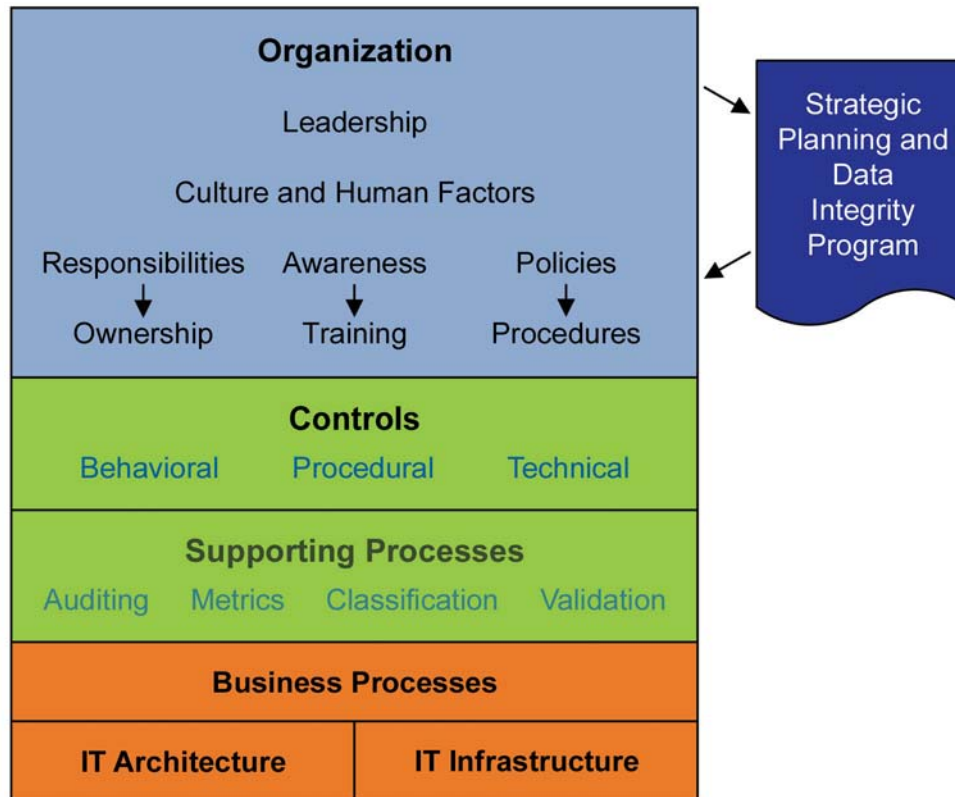
3.3 Elements of the Data Governance Framework

The overall Data Governance Framework consists of the following elements:

- Goals and objectives
- Organization and data ownership
 - Leadership and management responsibility
 - Roles and responsibilities
 - Policies and standards
 - Awareness and training
- Strategic planning and data integrity program
- Data life cycle
 - Quality risk management
 - Data management
 - Data incident and problem management
 - System access and security management
- Supporting processes
 - Auditing
 - Metrics
 - Classification
 - Validation
- IT architecture and infrastructure
- Maturity level model

The regulated company should define clear data governance goals and objectives, supported by a specific data integrity program, if necessary. Senior management should establish appropriate organizational structures to achieve these objectives, with clear policies, standards, and procedures, supported by appropriate training. Appropriate controls should be applied throughout the data life cycle, based on Quality Risk Management. Suitable supporting process (including system validation) and suitable IT architecture and infrastructure should be established. See Figure 3.2.

Figure 3.2: Data Governance Framework



3.3.1 Scope and Objectives

Effective data governance requires regulated companies to be clear on the scope and objectives. General goals and objectives for data governance may include:

- Increasing consistency and confidence in decision making
- Decreasing compliance risk
- Improving data security and privacy
- Maximizing the potential business value of data
- Clarifying information ownership and accountability for data quality
- Minimizing or eliminating re-work
- Optimizing process effectiveness

Data governance activities may have a different scope and objective depending on the nature, situation, and the business context of the regulated company, and it is likely that some focus and prioritization will be required. Although the overall structures and concepts in this guide would be appropriate for many regulated companies, the discussion will concentrate on focus areas for a regulated company, specifically around compliance, regulated data quality, and managing risks to data integrity, and therefore, product quality and patient safety.

Specific data governance goals and objectives for a regulated company may include:

- Demonstrating fitness for intended use through computerized system validation
- Assessing and controlling regulated data integrity risk
- Effective compliance with GxP regulations
- Minimizing inspection risk
- Compliance with various data privacy laws and regulations
- Ensuring adequate data security and access control
- Achieving these objectives effectively throughout a wide range of sites encompassing many local cultures and circumstances

For a regulated company, the key objectives should be product quality and patient safety, for which appropriate data governance delivering acceptable data integrity is considered a prerequisite.

Data governance goals, objectives, and scope should be defined and communicated by senior management, based on significant input from Business Process Owners, and Quality Unit and Information Technology functions.

3.3.2 Leadership and Management Responsibility

Senior management with executive authority has a responsibility across all levels of the regulated company to:

- Promote the requirements for data integrity
- Provide appropriate resources
- Resolve issues
- Define priorities
- Ensure that data integrity expectations are achieved

Senior management should also make personnel aware of the relevance of data integrity and the importance of their role in protecting the safety of patients and the reputation of the organization [12].

Senior management should lead by example and reinforce the messages by positive action, rewarding appropriate behavior, and taking the necessary management action when data integrity expectations and policies are not met.

It may be helpful to create a data governance council, or equivalent, ensuring adequate input from Business Process Owners, QU, and IT.

A data governance council, or equivalent, could play a key role in:

- Defining policies
- Taking decisions on roles and accountabilities
- Leading initiatives aimed at raising awareness

- Dealing with serious data related problems or incidents

Such a body would typically be led by a member of executive management, e.g., the chief data officer.

3.3.3 Organization and Data Ownership

Data ownership and responsibilities should be defined in the data governance framework and wider QMS. Individuals accountable and responsible for specific data, and its integrity and compliance, at various stages of the data life cycle should be defined and documented.

Data governance should not be regarded as primarily an IT issue, and IT is primarily a supporting role. While it may be an important role for computerized systems, IT has typically no involvement for non-electronic records and data.

Effective data governance in regulated companies requires communication and co-operation between Business Process Owners, Quality Assurance, IT department, and other technical support departments such as engineering, with sufficient support and leadership from senior management.

3.3.4 Key Performance Indicators

Leadership teams should monitor the progress of Data Integrity Programs, covering assessment and remediation activities, as well as monitoring the changing risk profile as work proceeds.

Typically, a small number of Key Performance Indicators (KPIs) will be identified for tracking in a summary dashboard. The nature and content of KPI Dashboards will vary depending on the specific needs, points of focus, and the priority of the data integrity work within the regulated company. Data for the selected KPIs should be readily available to avoid creating excessive work to collect supporting information. KPIs should be a predictor of the challenges ahead and should not just be to show what has been achieved.

KPIs should be clearly defined to promote consistency when the supporting data is collected. Ambiguous definitions can lead to misunderstandings when deciding exactly what to count and not count within the KPI being measured. KPIs should not give an over optimistic or pessimistic view that could impede effective management. Although leadership teams may like simple presentation of data (e.g., Red, Amber, Green status) over simplification should be avoided. Reporting should be accurate and based on metrics. Leadership should also be aware of different levels of risk tolerance across the regulated company.

3.3.5 Roles and Responsibilities

Two key roles associated with regulated Computerized Systems are defined in EU GMP Annex 11 [7] and *ISPE GAMP® 5* [3]:

- Process Owner
- System Owner

The terms and definitions used in specific organizations and the boundaries between such roles may vary. The use of the terms in this guide and the role descriptions below are aligned with the definitions in *ISPE GAMP® 5* [3] and EU GMP Annex 11 [7].

Process Owner

This is the owner of the business process or processes being managed. The process owner is ultimately responsible for ensuring that the computerized system and its operation is in compliance and fit for intended use in accordance with applicable SOPs. The process owner also may be the system owner. The process owner may be the de facto owner of the data residing on the system (data owner) and therefore, ultimately responsible for the integrity of the data. Process owners are typically the head of the functional unit using the system.

Specific activities may include:

- Approval of key documentation as defined by plans and SOPs
- Providing adequate resources (personnel including SMEs, and financial resources) to support development and operation of the system
- Ensuring adequate training for end users
- Ensuring that SOPs required for operation of the system exist, are followed, and are reviewed periodically
- Ensuring changes (including business process changes) are approved and managed
- Reviewing assessment/audit reports, responding to findings, and taking appropriate actions to ensure GxP compliance
- Ensuring that processes/systems are fit for the intended business use, and support data integrity
- Ensuring that data integrity risks are identified and controlled to acceptable levels

System Owner

The system owner is responsible for the availability, and support and maintenance, of a system and for the security of the data residing on that system. The system owner is responsible for ensuring that the computerized system is supported and maintained in accordance with applicable SOPs. The system owner also may be the process owner (e.g., for IT infrastructure systems or systems not directly supporting GxP).

For systems supporting regulated processes and maintaining regulated data and records the ownership of the data resides with the GxP process owner, not the system owner.

The system owner acts on behalf of the users. The system owner for larger systems will typically be from IT or Engineering functions. Global IT systems may have a global system owner and a local system owner to manage local implementation.

Specific activities may include:

- Approval of key documentation as defined by plans and SOPs
- Ensuring that SOPs required for the maintenance of the system exist and are followed
- Ensuring adequate training for maintenance and support staff
- Ensuring changes (including technical changes) are managed
- System life cycle management, including system upgrade and replacement planning

- Ensuring the availability of information for the system inventory and configuration management
- Providing adequate resources (personnel including SMEs, and financial resources) to support the system
- Reviewing audit reports, responding to findings, and taking appropriate actions to ensure GxP compliance, in conjunction with the process owner
- Ensuring that systems are supported and maintained such that they are fit for the intended business use, and support data integrity
- Ensuring that data integrity risks are identified and controlled to acceptable levels

Data Steward

The term data steward is usually used in the context of data governance. The term may be used differently and may apply to different roles in the data governance framework. It may be a functional role, or included as part of a wider job description. Data stewardship activities may be embedded in the responsibilities of other roles, rather than being a new and specific individual role.

In this Guide a data steward is defined as a person with specific tactical coordination and implementation responsibilities for data integrity. A data steward is responsible for performing data usage, management, and security policies as determined by wider data governance initiatives, such as acting as a liaison between the IT department and the business.

Data stewards are typically members of the operational unit or department creating, maintaining, or using the data, e.g., personnel in the laboratories who generate, manage, and handle the data. Segregation of duties should seek to minimize conflict of interest in the data steward role, e.g. avoiding the granting of unnecessary administrator privileges to individuals responsible for functional review and approval of GxP data.

3.3.6 Policies and Standards

Data governance policies and standards should be established and communicated to all relevant staff.

Data policies define the expectations and rules covering the integrity, security, quality, and use of data during its life cycle. Data standards provide more detail on structure, format, definition, and use of data.

It should be clear who has responsibility for defining, reviewing, approving, and monitoring compliance with policies and standards. Such policies and standards may be developed by a data governance council, or similar.

Based on the policies and standards, practical procedures, typically in the form of SOPs should be established, defining key activities and processes related to data integrity and providing details on how to achieve the defined policies and standards. Examples include procedures for handling adverse event and complaint data and evidence, manual chromatography integration practices, and batch record assembly and review.

These policies, standards, and procedures as described above should be incorporated as an integral part of the overall QMS, and unnecessary duplication should be avoided.

3.3.7 Awareness and Training

Regulated companies should ensure sufficient training in the importance of data integrity principles and data governance activities, and awareness and training on regulatory requirements and organizational policies and standards.

The aim is to achieve a state where all staff routinely follow accepted data integrity principles and practices, from a position of awareness and understanding, rather than depending on policing and technical controls to prevent users from doing the wrong thing.

For further information on training see Appendix M3.

3.3.8 Technology and Tools

Data governance technology and tools may be used to automate the definition, management, and enforcement of business rules at the data level.

Technology may assist in improving data quality and fitness for intended use by providing tools for data standardization and cleansing. Systems should be designed and configured to enforce integrity and consistency rules, ensuring conformance to defined policies and standards of the organization, and applying technical controls to minimize risks to data integrity. Other tools may include data reporting and visualization tools.

3.3.9 Strategic Planning and Data Integrity Program

Data governance initiatives and programs should be strategic and high level to provide a clear vision and direction to the regulated company, while also ensuring that critical immediate actions are prioritized, facilitated, and delivered. Effective data integrity initiatives and programs need senior management sponsorship at an appropriate level.

Short term needs should be addressed (particularly critical compliance requirements) while the wider aspects of data governance in the organization are being developed and overall maturity level increases.

Data governance programs should be scaled based on the size and complexity of the business unit, level of compliance risk, and potential impact on product quality and patient safety (MHRA, 2015 [1]).

"The effort and resource assigned to data governance should be commensurate with the risk to product quality, and should also be balanced with other quality assurance resource demands."

Communication should clearly link the Data Integrity Program with immediate business objectives or regulatory compliance challenges and requirements, so that the value of the program is obvious to all stakeholders. A communication plan, and if necessary, a change management plan should be established to ensure ongoing stakeholder engagement and understanding, and a smooth transition to new ways of working.

A process should be established for systematically incorporating learning points, and building them into the program and sharing with stakeholders. A repository of items such as templates, checklists, example citations, and FAQs on an internal information sharing and collaboration site, or similar, should be considered.

For further information on corporate data integrity programs, see Appendix M1. For further information on knowledge and information sharing, see Appendix M3.

3.3.10 Data Life Cycle and Data Management

The data life cycle, should be defined in standards and procedures. See Section 3.3.6.

Appropriate data management functions should implement the policies and standards established by the data governance framework, including:

- Data quality management
- Master and reference data management

- Data incident management
- Data inventory management

These activities and responsibilities should be integrated into existing roles and functions, where possible.

Data architecture models (which may be conceptual, logical, or physical) may be defined in the form of data flow diagrams, entity relationship diagrams, or system architecture diagrams. These together with data standards and procedures define how the data life cycle should be implemented within the regulated company. Higher level business process mapping and business process definition are prerequisites for the successful development of detailed data related flows and diagrams. See Appendix D2.

Data quality relates to the data's fitness to serve its intended purpose in a given context within a specified business or regulatory process. Data quality management activities address aspects including accuracy, completeness, relevance, consistency, reliability, and accessibility. See Section 1.5.4.

Data quality management enforces the established standards to ensure that data meets the relevant business definitions and rules of the data governance framework. The data incident management process should ensure that data problems and errors are identified, evaluated, resolved, and closed in a timely manner. Data should be safeguarded throughout its life cycle by appropriate system access and security management procedures.

Systems and procedures should be established to minimize the potential risk to data integrity, identify the residual risks, and apply the principles of ICH Q9 [10].

3.4 Human Factors in Data Integrity

Consideration of various human factors is considered critical for effective data integrity.

Cultural considerations can refer to a corporate culture (the model within which an organization operates) or to a local geographic culture (the moral and behavioral norm within a country or region).

Openness and a willingness to discuss difficult situations can support an environment where failing results are seen as a group problem needing to be resolved.

Management should help employees to achieve the openness around data integrity that is needed for compliance.

Data integrity issues often arise from genuine human error; however, regulators do not distinguish between human error and data falsification when assessing the impact of a data integrity failure.

Personal gain or self-interest has been the motivator in several high profile fraud cases. The extent and impact of falsification can be magnified if collusion is involved, but geographic and corporate cultures can influence the degree to which collusion may be prevented.

Robust technical controls within all of the data generation, collection, processing or storage systems, coupled with effective data review processes, can reduce opportunities for fraud.

The main foundation for a high level of data integrity is the knowledge and understanding of what data integrity is, the importance it has for an organization, and the personal role each employee has in protecting it.

For further information see Appendix M3.

3.5 Data Integrity Maturity Model

Regulated companies should focus on modifying their processes and systems to use appropriate available technical controls, and evaluate systems for gaps prior to use.

Where feasible, regulated companies should design record and data integrity into their processes before purchasing systems and technology. Purchased systems should be able to be configured to provide adequate data integrity.

The data integrity maturity model described in Appendix M2 is a simple representation of the regulated company. It is based on the status of essential elements of effective processes for data integrity.

4 Data Life Cycle

4.1 Introduction

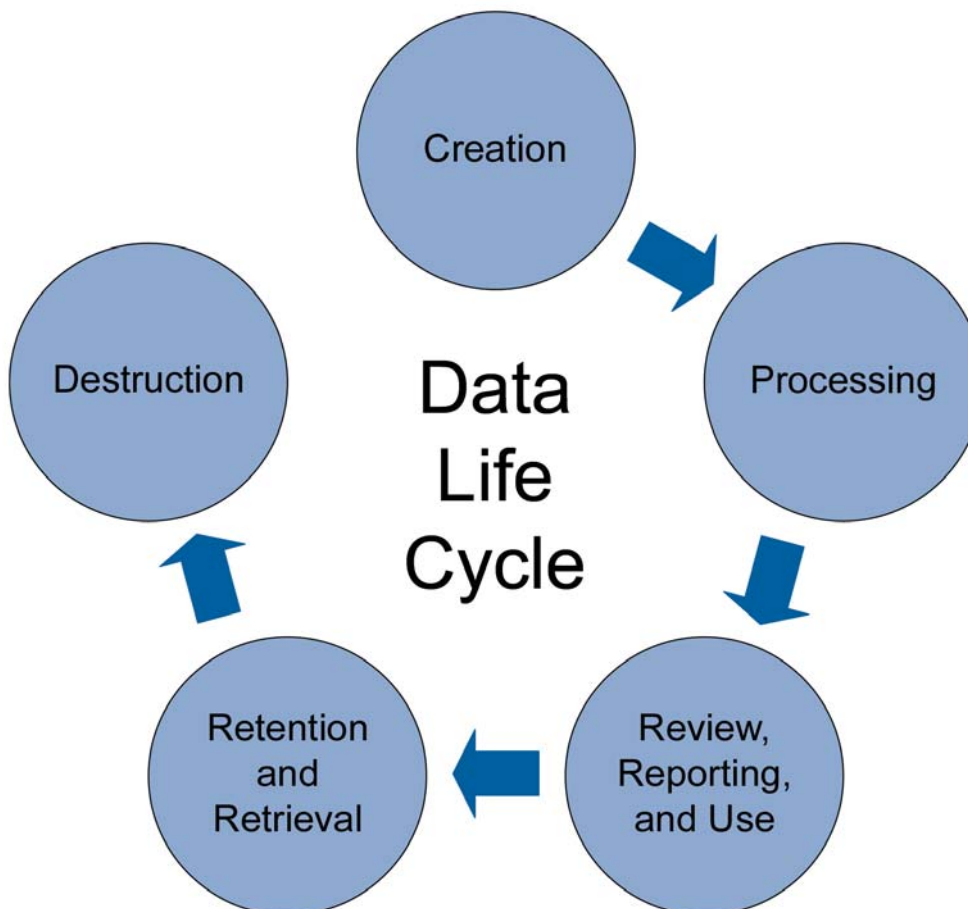
This section describes a generic data life cycle model suitable for all types of data and records. It also describes the activities and requirements for each phase in the data life cycle. Variations within the life cycle phases for different record types are also addressed.

Data integrity should be ensured throughout the data life cycle. Data integrity controls for data and records should ensure that they remain attributable, legible, contemporaneous, original, and accurate (ALCOA) throughout the data life cycle. In addition, data and records should be complete, consistent, enduring and available (ALCOA+). See Section 1.5.4.

The data life cycle includes all phases from initial creation of the data, through processing, use, retention and retrieval to eventual destruction, as shown in Figure 4.1. The data life cycle also supports instructions (e.g., specifications, procedures and templates) and records/reports (including data and results).

Different life cycle phases may need different controls, e.g., second person verification of data during creation or audit trail review during processing and/or use.

Figure 4.1: Data Life Cycle



Each life cycle phase can have an impact on data integrity. Robust risk-based business processes should be defined and implemented and data flows should be understood. This understanding, and these risk-based business processes, should be used to identify, assess, mitigate, and communicate potential data integrity issues throughout the data life cycle.

A specific data life cycle should be defined, based on a thorough understanding of the supported business process. Quality risk management should be applied throughout the data life cycle.

GxP computerized systems supporting the data life cycle and business process should be validated for their intended use, including verification of compliance with regulatory requirements and expectations for data integrity. Regulated companies should ensure that their compliance activities include appropriate specification and verification of data integrity requirements and controls. Regulated companies should not rely only on supplier qualification packages [13].

For further details of verification activities see the *ISPE GAMP® Good Practice Guide: A Risk-Based Approach to Testing of GxP Systems* [17]. For further details on all aspects of system operation see *ISPE GAMP® Good Practice Guide: A Risk-Based Approach to Operation of GxP Computerized Systems* [18].

Data associated with a product or process may cross various boundaries and interfaces throughout the data life cycle [13]. These boundaries may include transfer:

- Of data between systems
- Between a manual process and a computerized system
- To cloud-based applications and storage
- Across organizational boundaries, e.g., between production, QC and QA (internal boundaries)
- Between regulated companies and third parties, e.g., service providers (external boundaries)

Risks associated with such transfers should be considered and appropriate controls established to prevent loss or modification.

4.2 Data Creation

Data capture or recording should ensure that data of appropriate accuracy, completeness, content, and meaning is collected and retained for its intended use [8]. Data integrity may be compromised at the point of creation. If the original data is not reliable then its integrity cannot be ensured. Data creation should provide accurate data which is needed throughout the business process for decisions based on that data.

Data can be created either by entry of new data, captured by the system from an instrument, device, another system, or captured manually. Where data is created by an instrument or measuring system, the instrument should be adequately maintained and calibrated, and should have the range, resolution, linearity, and sensitivity to accurately measure and/or detect the sample attributes under evaluation.

Data should be captured and saved at the time of the activity (contemporaneously), and prior to proceeding to the next activity in the process. Where possible, automated data capture techniques should be applied to minimize the risk of data transcription error. Appropriately controlled and synchronized clocks should be available for recording timed events [8]. Time and date stamps used should be explicit within the context of their use and should be protected from unauthorized change.

All data (even if it has been excluded) should be retained, and be available for review in a format that allows the validity of the decision to exclude the data to be confirmed [8]. If any data is excluded from GxP decision making processes or investigations (e.g., the product release process), there should be a valid, documented, scientific justification for its exclusion [9]. Data may be excluded only where it can be demonstrated through robust science that the data is anomalous or non-representative. This justification should be documented and considered during data review and reporting [8].

Where the capability of the electronic system permits dynamic storage, high resolution or dynamic (electronic) data should be collected in preference to low resolution or static (printed/manual) data [8]. Data should not be stored electronically in temporary memory in a manner that allows for manipulation, before being stored safely and securely.

Data should be attributable to a specific individual, where relevant. The need for verification of data entry, e.g., via second person verification or through technical means such as data validation or barcoding, should be considered based on specific requirements, intended use, and potential impact of data errors within the process.

Risks to data integrity will be influenced by the degree to which data can potentially be manipulated. Individuals with a direct interest in the data should not be granted system level privileges (e.g., permitting data deletion, database modification).

Data should be stored in the predefined location and format. Where the same information is recorded concurrently in more than one location or format, the process/data owner should define where the primary record is retained [1]. Data should be secured from modification by unauthorized persons and should be changed or deleted only in accordance with regulatory expectations and requirements. Any changes should be recorded in the audit trail.

4.3 Data Processing

During this phase, data is processed to obtain and present information in the required format. Processing should occur in accordance with defined and verified processes (e.g., specified and tested calculations and algorithms), and approved procedures.

Record/report type documents are typically processed prior to review and reporting. Process data should not be manipulated to achieve a more desirable end point. If data processing has been repeated, with iterative modification of processing parameters, this should be made apparent [8]. All data relating to iterative processing runs should be reviewed during routine data verification and stored for the duration of the regulatory retention period.

If data is reprocessed, written procedures should be established and followed. Each result should be retained for review, including all data relating to iterative processing runs. For most laboratory analyses, reprocessing data should not be needed on a regular basis. Sampling, testing, or processing should not be performed with the goal of achieving a specific result or to overcome an unacceptable result (e.g., testing different samples until the desired passing result is obtained; this practice is sometimes referred to as testing into compliance) [9].

The impact of data processing on product quality and patient safety will vary by product and business process. The rigor of the controls and verification required for the data processing step should be determined by a documented and justified risk assessment. Topics to consider include, where appropriate:

- Impact of data on product quality and decision making
- Requirements for independent second person verification by a qualified individual
- Opportunity for data modification or deletion
- Preventing original data from being overwritten or deleted

- Preventing data in displays and printouts from being obscured (e.g., by inappropriate annotation)
- Limiting processing to authorized personnel
- Securing processing parameters from unauthorized change
- Logging of parameter changes
- Reporting of data processing errors (either by the system or manually)
- Reconstruction of all data processing activities
- Ensuring traceability of user defined parameters
- Ability of user to influence what data is reported (e.g., user can select what data to print)
- Ability of user to determine presentation of data (e.g., adjusting scale/resolution of graphical reports)
- Use of calculations as validity checks
- Use of calculations for transformation of data
- Exclusion of data during processing should be justified and documented. Excluded data should be retained with the original data set and be available for subsequent review [8].

4.4 Data Review Reporting and Use

During this phase data is used for informed decision making. Data review, reporting, and use should be performed in accordance with defined and verified processes and approved procedures. Data review and reporting is typically concerned with record/report type documents.

4.4.1 Data Review

Data review (including second person review as required by regulation) should determine whether predefined specifications, targets, limits, or criteria have been met. The review should be based on a thorough process understanding (and where applicable system understanding) and impact on product quality and/or decision making, and outcomes and conclusion documented.

The process for the review and approval of data should be described in a procedure.

Original records subject to review should be defined. Reviews should be based upon original data or a true copy (preserving content and meaning). Data review should include a review of relevant metadata and GxP data audit trails, where appropriate [8]. Audit trails are considered part of the history of associated records. Personnel responsible for the review of regulated records should review the audit trails that capture changes to critical data associated with the record. Audit trails that capture changes to critical data should be reviewed with each record and before final approval of the record [9].

All data should be considered, recognizing that data may be stored in different locations. This includes atypical, suspect, rejected, invalid, or deleted data, along with any justifications. Excluded data should be supported by a documented justification.

Routine data review should evaluate [8]:

- The integrity of an individual data set
- Compliance with established organizational and technical measures
- Any data risk indicators (e.g., data amendment, or orphan data)

Data review procedures should cover:

- Method for review
- Method for approval, e.g., by use of an electronic signature
- The meaning of review and approval signatures to ensure persons understand their responsibilities
- Requirements for review by quality assurance (e.g., under US GMPs, any data created as part of a CGMP record must be evaluated by the quality unit as part of release criteria) [9]
- Handling errors or omissions identified during data review
- Managing data corrections or clarifications
- Managing atypical, erroneous, or invalid results

Second person reviews should focus on the overall process from data creation to calculation of reportable results. Such reviews may cross system boundaries as well as the associated external records and may include verification of any calculations used.

Review by exception should be based on validated data processing routines that cannot be influenced by the user.

Periodic review or audit of data governance measures should assess effectiveness of established organizational and technical measures, and should also consider the possibility of unauthorized activity [8].

Where data reviews are conducted by a third party the regulated company should ensure respective roles and responsibilities are documented and agreed by both parties.

4.4.2 Audit Trail Review

Regulated companies should establish a documented process for review of audit trails, including as a part of second person review. These reviews should form part of the routine data review/approval process and are usually performed by the operational area which has generated the data (e.g., clinical, laboratory, manufacturing).

The requirement for audit trail review, including the frequency and rigor and roles and responsibilities, should be based on a documented risk assessment taking into account the business process and criticality of the data, the complexity of the system and its intended use, and the potential impact on product quality and patient safety.

Audit trail reviews should be performed by an individual who has an understanding of the business process and the impact of the actions recorded. They are an effective means of verifying that changes are made by authorized users and for detecting potential data integrity issues.

For more information on review of audit trails, see Appendix M4.

4.4.3 Data Reporting

Data reporting procedures should ensure the consistency and integrity of results. The procedures should:

- Define what data is to be included in the data set used for reporting the results, i.e., the complete data set
- Address report layout and formatting requirements
- Describe use of reports for GxP decisions
- Address report, review, and approval requirements
- Ensure all data is included in the dataset, unless there is documented justification for excluding it
- Address manual data entry
- Address handling and investigation of atypical results
- Address suspect, rejected, invalid, or deleted data
- Describe how corrective and preventative actions are established for invalid runs, failures, repeats, and other atypical data
- Consider controls to prevent and detect data manipulation
- Encourage problem identification and solving

Particular attention is required when users are able to influence the reporting of data, e.g., testing into compliance should be avoided.

Poorly designed or defined processes, test methods, and reports can create opportunities for data reporting errors, and the potential for erroneous decisions.

Summary reports are limited as they may not contain all data and there is a risk that data issues may not be included, and therefore, not reviewed. Where data summaries are used for reporting, there should be documented verification of these summaries in accordance with original data.

Trending and appropriate metrics can be used to identify and investigate potential data integrity issues.

4.4.4 Data Distribution

Data should be accessed by, and distributed to, authorized individuals and other systems supporting the business process. Interfaces between business process systems should be designed and verified to report failures, prevent data loss, and enable recovery.

Version control should be applied to ensure that personnel have access to the appropriate and up to date versions required in order to perform regulated activities. Instruction type documents should have clear effective dates.

The need for formal confirmation of data receipt should be considered.

4.5 Data Retention and Retrieval

4.5.1 General Requirements

During this phase, data should be retained securely. Data should be readily available through the defined retention period in accordance with defined and verified processes and approved procedures.

Applicable regulatory requirements, other laws and legislation, and internal regulated company policies should be met. This includes retention and privacy requirements. Regulated companies should be aware that legislation such as local laws may have specific requirements to be met (e.g., blood products have different retention periods in the US, Europe, and Japan). Legal admissibility of information stored electronically should also be considered.

Data, including all the original data and associated metadata required to maintain GxP content and meaning, should be defined and stored in a secure location that has adequate physical and electronic protection against deliberate or inadvertent alteration or loss throughout the retention period. Access should be limited to authorized persons and adequate environmental protection from damage should be provided (e.g., from water and fire).

Data, and relevant associated metadata, should be readily traceable and accessible throughout the retention period. The relationships between data and associated metadata should be preserved securely to support future queries or investigations, including reconstruction of GxP activities. Data and associated metadata may reside in separate locations and possibly on different media.

Retrieval of records and copies of records made (including those made for archival purposes), should preserve the content and meaning of the record, and continue to meet relevant GxP regulatory requirements.

Paper records may be retained electronically, e.g., by scanning, provided the copy is verified as a true copy, following an established validated process. Regulated companies may discard the original record once a verified true copy (preserving content and meaning) has been made, e.g., where the original record is not permanent.

Following changes to the system (hardware or software) the regulated company should verify that the retained data can still be accurately retrieved.

Changes to data should be amended only by authorized persons in accordance with an approved process. A record of all changes should be retained indicating when and how the original data was changed. There should be controls in place to ensure that previous versions of data are not inadvertently restored or otherwise made available.

Data retention and retrieval procedures should consider the following, where relevant:

- Address data synchronization where the system architecture involves storing data on multiple servers
- Disaster recovery, including backup and restore
- Durability of media used for storage
- Use of encryption to enhance security (dependent on criticality of data and risks to that data)
- Environmental controls

For further details on operational processes and controls see *ISPE GAMP® Good Practice Guide: A Risk-Based Approach to Operation of GxP Computerized Systems* [18].

Where a third party is involved in regulated data retention and retrieval activities, the regulated company should ensure that a documented assessment of the third party and the risks is undertaken. This should demonstrate that adequate controls are in place and residual risks are understood and accepted.

A formal agreement should be established defining the respective roles and responsibilities and covering:

- Procedures to follow
- Physical location
- Applicable regulations and laws and their impact
- Monitoring
- Incident management
- Business continuity
- Backup
- Data destruction

The same retention and retrieval requirements apply irrespective of whether the data is electronic or paper based.

When a system is retired, or can no longer be supported, regulated companies should consider how regulated data will continue to be retained and retrieved throughout the remainder of the retention period, considering:

- Criticality of data
- Practicality of maintaining existing software (e.g., in a virtual environment)
- Options for migrating the data to another system or to an alternative file format, while retaining the original content and meaning
- Risk of retaining the information without metadata, e.g., retaining only reports written to PDF or paper

Regulated companies may choose to retain records in formats other than the original, provided content and meaning is preserved, and GxP regulations are met. The ability to retain records in a format that ensures the dynamic nature of the data throughout the retention period is not always possible or cost effective, due to the difficulty of migrating the data over time. Decisions should consider the balance between the need for long term accessibility and the level of ongoing retrieval functionality required (e.g., need for dynamic searching, trending, reprocessing). Original records can only be deleted if GxP regulations are fully satisfied and the content and meaning of the records are preserved, following a defined procedure including management authorization.

Migration should be based on a defined process, including a documented risk assessment, and managed within the framework of a data migration plan and report. See *ISPE GAMP® 5* [3].

When migrating data, regulated companies should make an informed risk-based decision regarding the migration of metadata, including the audit trail along with the data. This decision should be based upon business requirements and regulatory expectations. If the audit trail is integral to understanding the data, it should be maintained as part of the migrated data. A decision not to migrate an audit trail should be justified based on risk, and documented.

Where legacy systems are retained for data retention and retrieval, the regulated company should periodically verify that data can still be retrieved.

4.5.2 Backup and Restore

Regular backups of all relevant data, including associated metadata required to maintain GxP content and meaning, should be performed in accordance with a documented process in order to allow recovery in the case of system failure, data corruption, or loss. The regulated company should ensure that the backup process is designed such that regulated data is not lost or corrupted.

Data backup should have controls commensurate with those controls for the original data to prevent unauthorized access, modification, or deletion. Backups should be held in a physically separate and secure location. The process and technology for restoring data should be based on the criticality of the data and the required restoration time.

The process for backup and the ability to restore data, should be verified before use and monitored periodically for accessibility, readability, and accuracy through the retention period.

Where relevant metadata is stored separately to the regulated data (e.g., in audit trails or separate files), it should be ensured that all required data is available and can be successfully restored.

4.5.3 Archiving

Archiving involves the long term, permanent retention of data and associated metadata for the purposes of review or investigation throughout the retention period, which may include the reconstruction of a process or activity [1].

Archiving should be performed in accordance with defined and verified processes and approved procedures, that meet the general requirements for retention and retrieval. See Section 4.5.1.

These procedures should consider where appropriate:

- Determination of storage media life expectancy
- Multiple copies
- Management of stored media
- Indexing of stored records
- Impact of system upgrade on stored records
- Impact of changing technology on stored records
- Ability to reprocess data where required

Archived data should be periodically checked for accessibility and readability.

For more information on archiving of records, see Appendix O1.

4.6 Data Destruction

The data destruction phase involves ensuring that the correct original data is disposed of after the required retention period in accordance with a defined process and approved procedures. The procedure should consider:

- Retention requirements
- Method of disposal, ensuring deletion from all systems and physical locations
- Conditions of disposal
- Measures to prevent inadvertent disposal of data that is still required
- Documentation required to demonstrate timely disposal (for business reasons)
- Limiting disposal functionality to a restricted number of responsible individuals

Data disposal should account for all local legislative retention requirements and for all locations that reference the data.

Retention requirements may differ by jurisdiction, or there may be a litigation hold on some data in some countries/regions. There may also be conflict between applicable laws.

Data disposal should not occur without verification that the regulated record is not in a hold status to support litigation. Distributed copies should also be destroyed of when a record is disposed.

Data should not normally be retained beyond the defined retention period.

5 Quality Risk Management

5.1 Introduction

Quality risk management is a systematic process for the assessment, control, communication, and review of risks. It is an iterative process used throughout the entire computerized system life cycle from concept to retirement, and throughout the data life cycle from creation to destruction. Risks to data and record integrity should be identified and managed along with other quality and safety risks by adopting a risk management approach based on an understanding of the process.

The evaluation of the risk to quality should be based on scientific knowledge and ultimately linked to the protection of the patient. The level of effort, formality, and documentation of the quality risk management process should be commensurate with the level of risk.

Some records and data may reside on more than one system during their life cycle, and quality risk management activities should start at the business process level, at a level higher than individual systems.

5.2 Process Risk Assessment

A process risk assessment (also known as business process risk assessment) is a non-system specific high level assessment at the business process or data flow, which may occur before systemspecific quality risk management activities. An equivalent risk assessment from a data flow (rather than business process flow) perspective may be performed, using the same approaches and techniques, and with the same benefits.

The process risk assessment is aimed at identifying key high level risks to patient safety, product quality and data integrity, and identifying the required controls to manage those risks. Typically, at this stage no assumptions are made about the nature or exact functionality and design of the computerized system(s) that will support the process.

The process risk assessment provides valuable input to subsequent quality risk management activities. Typical inputs to the process risk assessment include:

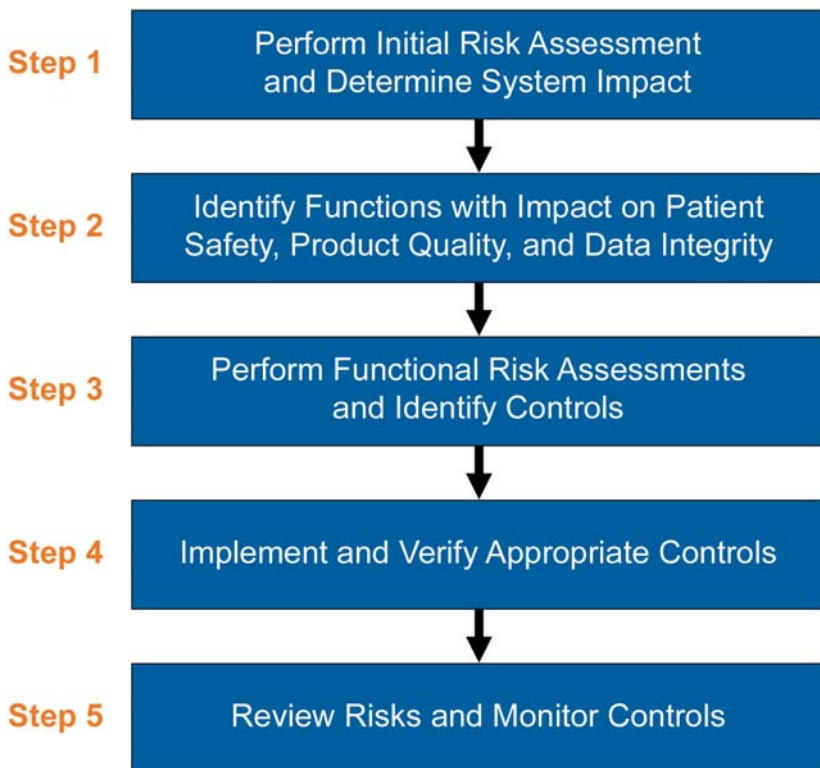
- Defined business process scope
- Process descriptions and/or diagrams
- Identified regulatory requirements for the proposed process scope
- Identified company quality requirements

5.3 Quality Risk Management Approach

The activities required to manage the specific risks associated with records and data should form part of the normal individual GxP computerized system life cycle. Decisions on the extent of validation and data integrity controls should be based on a justified and documented risk assessment [7].

ISPE GAMP® 5 [3] describes a five-step approach for quality risk management, based on ICH Q9 [10].

Figure 5.1: Quality Risk Management Approach



Some records may be stand alone, maintained on a file share, and may not be associated with any one specific computerized system. The risks associated with such records should be considered and addressed, typically as part of process (or data flow) risk assessment.

Step 1 – Perform Initial Risk Assessment and Determine System Impact

An initial risk assessment should be performed based on an understanding of business processes and process risk assessments, user requirements, regulatory requirements, and known functional areas. This should include initial identification of important data, records, and signatures.

Step 2 – Identify Functions with Impact on Patient Safety, Product Quality, and Data Integrity

Functions which have an impact on patient safety, product quality, and data integrity should be identified by building on information gathered previously, referring to relevant specifications, and taking into account project approach, system architecture, and categorization of system components.

Based on defined business processes, during development of user requirements and during subsequent functional, configuration, or design specifications, the associated data, records, and signatures should be defined and documented.

This activity typically begins in Step 1, and continues iteratively through the data life cycle. A data flow analysis is useful in supporting this activity and in determining the role of each item in regulated processes. The primary record should be identified. The emphasis should be on identifying data and records at the regulated process level, rather than physical database records, or table fields.