Organisation for Economic Co-operation and Development

ENV/CBC/MONO(2021)26

Unclassified

English - Or. English

20 September 2021

ENVIRONMENT DIRECTORATE
CHEMICALS AND BIOTECHNOLOGY COMMITTEE

OECD SERIES ON PRINCIPLES OF GOOD LABORATORY PRACTICE AND COMPLIANCE MONITORING

Number 22
Advisory Document of the Working Party on Good Laboratory Practice on GLP Data Integrity

JT03481133

OECD Environment, Health and Safety Publications

Series on Principles of Good Laboratory Practice (GLP) and Compliance Monitoring

No. 22

Advisory Document of the Working Party on Good Laboratory Practice on GLP Data Integrity

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate
ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT
Paris 2021

ALSO PUBLISHED IN THE SERIES ON PRINCIPLES OF GOOD LABORATORY PRACTICE AND COMPLIANCE MONITORING

- *No. 1, OECD Principles of Good Laboratory Practice (as revised in 1997)*
- *No. 2, Revised Guides for Compliance Monitoring Procedures for Good Laboratory Practice (1995)*
- *No. 3, Revised Guidance for the Conduct of Laboratory Inspections and Study Audits (1995)*
- *No. 4, Quality Assurance and GLP (as revised in 1999)*
- *No. 5, Compliance of Laboratory Suppliers with GLP Principles (as revised in 1999)*
- *No. 6, The Application of the GLP Principles to Field Studies (as revised in 1999)*
- *No. 7, The Application of the GLP Principles to Short-term Studies (as revised in 1999)*
- *No. 8, The Role and Responsibilities of the Study Director in GLP Studies (as revised in 1999)*
- *No. 9, Guidance for the Preparation of GLP Inspection Reports (1995)*
- *No. 10, The Application of the Principles of GLP to Computerised Systems (1995)*
- *No. 11, The Role and Responsibilities of the Sponsor in the Application of the principles of GLP (1998)*
- *No. 12, Requesting and Carrying Out Inspections and Study Audits in Another Country (2000)*
- *No. 13, The Application of the OECD Principles of GLP to the Organisation and Management of Multi-Site Studies (2002)*
- *No. 14, The Application of the Principles of GLP to in vitro studies (2004)*
- *No. 15, Establishment and Control of Archives that Operate in Compliance with the Principles of GLP (2007)*

- *No. 16, Guidance on the GLP Requirements for Peer Review of Histopathology (2014)*
- *No. 17, The Application of GLP Principles to Computerised Systems (2016)*
- *No. 18, OECD Position Paper Regarding the Relationship between the OECD Principles of GLP and ISO/IEC 17025 (2016)*
- *No. 19, The Management, Characterisation and Use of Test Items (2018)*
- *No. 20, Guidance Document for Receiving Authorities on the Review of the GLP Status of Non-Clinical Safety Studies (2019)*
- *No. 21, OECD Position Paper Regarding Possible Influence of Sponsors on Conclusions of GLP Studies (2020)*

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 38 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in eleven different series: **Testing and Assessment**; **Good Laboratory Practice and Compliance Monitoring**; **Pesticides**; **Biocides**; **Risk Management**; **Harmonisation of Regulatory Oversight in Biotechnology**; **Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents;** and **Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/chemicalsafety/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

Also published in the Series on Principles of Good Laboratory Practice and Compliance Monitoring: link

For this and many other Environment, Health and Safety publications, consult the OECD's World Wide Web site (www.oecd.org/chemicalsafety/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

FOREWORD

This advisory document was developed by the OECD Working Party on Good Laboratory Practice (GLP). The development of the document was initiated and led by the UK and included a drafting group under the leadership of Stephen Vinter (UK) and Thomas Lucotte (France Medical Products). The drafting group included representatives from Argentina, Austria (Medical Products), Belgium, Denmark (Medical Products), Italy, Mexico, the Netherlands, the US (EPA) and the US (FDA). The process included a public comment period and review and endorsement of the document by the Working Party on Good Laboratory Practice.

This document is published under the responsibility of the Chemicals and Biotechnology Committee which agreed to its declassification on 8 September 2021.

Table of Contents

1. Background

One of the fundamental purposes of the Principles of Good Laboratory Practice (GLP) is to ensure the quality and integrity of test data related to non-clinical safety studies**.**

The way in which study data, supporting human, animal and environmental safety assessment, is generated, handled, reported, retained and archived has continued to evolve in line with the introduction and ongoing development of supporting technologies. This includes the increasing use of electronic data capture, integration and automation of systems and other technologies. Systems can range from manual processes with paper records to the use of complex computerised systems. However, the main purpose of the requirements of the Principles of GLP remains the same in having confidence in the quality, the integrity of the data and being able to reconstruct activities performed during the conduct of non-clinical safety studies.

2. Introduction

The following overarching aspects apply to this document:

1. This document provides guidance for test facilities or test sites that conduct GLP studies or GLP study phases.

   For the purposes of this document, the term ‘test facility’ includes test sites; the term ‘study’ includes study phases; and the term ‘study director’ is extended to cover the responsibilities of principal investigator where this is appropriate.

2. The guidance aims to promote a risk-based approach to the management of data that includes data risk, criticality and life cycle. Users of this document need to understand the data flows they are responsible for or involved in (as a life cycle) in order to identify data that are likely to have impact on GLP compliance. In turn, this will support the identification and the implementation of the most effective and efficient risk-based controls.

3. Data integrity is the degree to which data are complete, consistent, accurate, trustworthy and that these characteristics of the data are maintained throughout the data life cycle. Data should be collected and maintained in a secure manner, such that they are attributable, legible, contemporaneously recorded and accurate, whether raw data or a verified copy.

4. The guidance refers to the acronym ALCOA being Attributable, Legible, Contemporaneous, Original and Accurate. ALCOA has historically been regarded as the attributes of data that are suitable for regulatory purposes. ALCOA+ has been referred to in more recent times to emphasise the additional attributes Complete, Consistent, Enduring and Available. There is no difference between the expectations related to data integrity for both these terms since data governance measures should ensure that data are complete, consistent, enduring and available throughout the data life cycle.

5. The guidance addresses data integrity and not data quality since the controls required for integrity do not guarantee the quality of the data (see also definitions in section 3.4 and 3.5). Data integrity provides control over the data (i.e. whether it can be trusted), whereas data quality refers to the data characteristics that assure

that data produced are generated in compliance with applicable standards and can be used for its intended purpose.

6. This guidance should be equally applied to the control of all data types and formats. Some points are nevertheless focused, and specifically applicable, to electronic data and electronic systems.

7. This guidance should be read in conjunction with OECD Documents No 1 (*OECD Principles on Good Laboratory Practice*) (OECD, 1997[1]), No 15 (*Establishment and Control of Archives that Operate in Compliance with the Principles of GLP*) (OECD, 2007[2]), No 16 (*Guidance on the GLP Requirements for Peer Review of Histopathology*) (OECD, 2014[3]) and No 17 (*Application of GLP Principles to Computerised Systems*) (OECD, 2016[4]) and applicable national regulations. The GLP Principles that reference data integrity can be found in Section II, 1.1.2.b to e, 1.1.2.l, 1.1.2.q, 1.2.2.f, 1.2.2.g, 1.2.2.i, 1.4.3, 2.1.1.c, 3.4, 7.1, 7.4.3, 8.2.6, 8.3.3, 8.3.4, 8.3.5, 10.1 of OECD Document No 1. Where relevant complementary information is contained in this document and other documents, reference is made within the text.

3. Definitions and terms

3.1. Data

Data are quantitative or qualitative facts, figures and statistics collected for reference or analysis. These include all original records and verified copies of original records, including raw data and metadata and all subsequent transformations that are generated or recorded at the time of the GLP activity, and allow complete reconstruction and evaluation of the GLP activity.

Data can have different formats (e.g. analogue, digital) and structure, layouts (e.g. on paper or on screen), sources (e.g. chromatography charts, text, image, video, etc.), and media used to store or present (paper, DVD, photo film, tape, electronic files, etc.).

Data may be captured or recorded:

i. by manual recording, on paper or in an electronic system, of an observation or of an activity;

ii. by automatic recording, on paper (by automatic printing) or in an electronic system, using equipment that range from simple instruments through to complex highly configurable computerised systems;

iii. using a hybrid system where combinations of paper (or other non-electronic media) and electronic records constitute the raw data;

iv. on other means of media such as photography, imaging methodologies and technologies, chromatography plates, etc. that could be generated manually, or automatically or using a hybrid system.

Raw data

The Principles of GLP define raw data as all original test facility records and documentation, or verified copies thereof, which are the result of the original observations and activities in a study and allow complete reconstruction and evaluation of the GLP activities. Raw data also may include, for example, photographs, microfilm or microfiche

copies, computer readable media, dictated observations, recorded data from automated instruments, or any other data storage medium that has been recognised as capable of providing secure storage of information for a time period.

Record

A record is a piece of information (e.g. data). The term original record is used to describe the first source of information or data capture. Original records are generally raw data. If an original record meets the definition of raw data, but is not considered as such, this must be justified.

Verified copy

A verified copy is a faithful representation of the original at the time the copy is generated. A verified copy may be stored in a different format or document type to the original.

Verified copies can be generated to:

- make a duplicate of the originals to include them in different files (for example, experimental raw data common to several studies);
- extend the retention period of some data whose format does not allow preservation (e.g. thermal printouts);
- allow the retention of the data if the original cannot be kept without causing a risk to other archived materials (for example, paper raw data stained with animal fluids, chemicals etc.);
- facilitate the exchange of data;
- support archiving activities.

The most common processes to generate verified copies from static records are:

- photocopy of a paper record (paper to paper);
- scan of a paper record (paper to electronic);
- picture of a paper record (paper to picture);
- screen shot and printout of an electronic record (electronic to paper).

Derived data

Derived data are obtained and reconstructed from raw data (e.g. final concentrations as calculated by a spreadsheet relying on raw data obtained from an instrument; result tables as summarised by a Laboratory Information Management System (LIMS), etc.). Derived data are obtained by data processing.

Metadata

Metadata are data providing information used for the identification, description, and relationships of data. Metadata give data meaning, provide context, define structure, and enable retrievability across systems, and usability, authenticity, and auditability across time. For electronic data, parts of the metadata can be generated in audit trails.

Metadata form an integral part of the data. Without the context provided by metadata, the data have no or limited meaning. The degree of metadata missing reduces the ability to interpret the data.

Audit trail

The audit trail is a form of metadata that contains information associated with actions that relate to the creation, modification or deletion of electronic data. An audit trail provides an automated secure way of recording life cycle details such as creation, additions, deletions or alterations of information in an electronic record without obscuring or overwriting the original record. An audit trail facilitates the reconstruction of the history of such events relating to the record, including the 'who, what, when and why' of the action.

3.2. Data structure

Data can have different structures.

Static format

A static record format, such as a paper or electronic record, is one that is fixed and allows no interaction between the user and the record content. For example, all paper records are static records. Electronic records that do not contain any link to other records that allow interaction are also static records. A printout from a basic electronic balance, where no electronic data is stored, is an example of a static record from an electronic system.

Dynamic format

Records in a dynamic state are mostly electronic records that allow for an interactive relationship between the user and the record content. Examples of a dynamic format include chromatography data maintained as electronic records to allow the user to zoom on the baseline, to view the integration more clearly, or to have direct access via electronic links to the sequence of analysis, the table of results, the audit trails and the methods of acquisition and integration. Records electronically signed are also dynamic records as they contain a link with the authentication of the signature.

File structure

The way in which most of the electronic data are structured within the GLP environment will depend on what the data will be used for and the end user will almost always have this dictated to them by what software / computerised system is available.

Flat files

A flat file consists of a single table of data, has no internal hierarchy and allows the user to specify data attributes i.e. its data structure is self-contained and limited.

Flat files can be thought of as being similar to the files in a file cabinet drawer, a collection of single records each containing standalone data. The most commonly known flat file would be a .csv or .xls file or a Microsoft Word™ text only document.

Relational databases

Relational databases are a collection of tables linked together using a common piece of data, such as a study number, and can be arranged to highlight specific information for *ad hoc* queries. A relational database is a scalable and query friendly tool that provides the ability to capture a wide variety of data types. Relational databases are usually not used to record raw data.

Relational databases store different components of associated data and metadata in different places. Each individual record is created and may be retrieved by compiling the data and metadata for review using a database reporting tool.

For example, electronic records in a database format allows the user to track, trend and query data.

3.3. Electronic signature

An electronic signature is a signature in digital form that represents the hand-written ('wet') signatory.

Different types of systems exist from simple ones (e.g. internal user identification with password) to complex systems of signatures (e.g. with an external, certified electronic signature service that provides with timestamp and encrypted information behind the signature). To be considered as an electronic signature in legal terms, the associated level of control required is defined where relevant by local regulation.

3.4. Data integrity

Data integrity is the degree to which data are complete, consistent, accurate, trustworthy and reliable and that these characteristics of the data are maintained throughout the data life cycle. Assuring data integrity requires appropriate quality and risk management systems, including adherence to sound scientific principles, good documentation practices and training of personnel.

3.5. Data quality

Data quality is the assurance that the data produced are generated according to applicable standards and fit for intended purpose. Data quality is assured by appropriate study design that accurately and scientifically addresses the experimental question and hypotheses being studied and by the availability of adequate resources. Data quality affects the value and overall acceptability of the data in regard to decision-making or onward use.

3.6. Data life cycle

The data life cycle includes all phases in the life of the data from generation and recording through processing (including analysis, transformation or migration), use, data retention, archive, retrieval and destruction.

- Data approval: Data approval is the act of authorising data after collection, processing or verification to record that data are suitable for their intended use.
- Transcription: Transcription is the process where data are manually copied from a source into another record of data set.

  Transcription can occur when:

  - the same information is recorded in different records (for example, the date of arrival of the test item is recorded in multiple records such as logbooks or proformas);
  - data are entered into a computerised system for calculations. Transcription of manual records into an electronic system constitutes an example of a hybrid system.

- Data processing: Data processing is a sequence of operations performed on data in order to extract, present, calculate or obtain derived data in a defined format. Examples might include calculations in a spreadsheet, statistical analysis of individual test system data to present trends, or conversion of a raw electronic signal to a chromatogram and subsequently a calculated numerical result.
- Data migration: Data migration is the process of moving electronic data between different data storage types, computerised systems, or simply the transition of data from one format to another. This may include changing the format of data, but not the content or meaning, to make it usable or visible on an alternative computerised system.
- Computerised system transaction: A computerised system transaction is a single operation or sequence of operations performed as a single logical unit of work. The operation(s) that constitute(s) a transaction may not be saved as a permanent record on durable storage until the user commits the transaction through a deliberate act (e.g. pressing a save button, see also "data approval"), or until the system forces the saving of data.
- Data retention: Data retention is the storage of data which may be for the purpose of archiving (protected data for long-term storage) or back-up (electronic data or for the purposes of disaster recovery).
- Back-up: A data back-up is a copy of current data, metadata and system configuration settings maintained for the purpose of recovery including disaster recovery.

  Back-up allows for provisions made for the recovery of data files or software, for the restart of processing, or for the use of alternative computer equipment following a system failure or disaster.
- Archive: Archive means a designated area or facility (e.g. cabinet, room, building or computerised system) for the secure storage and retention of records and materials.

3.7. Data governance

Data governance is the sum total of arrangements to ensure that data (irrespective of the format in which they are captured, generated, recorded, processed, retained, archived and used) are attributable, legible, contemporaneous, original (or verified copy), accurate, complete, consistent, enduring and accurate (ALCOA+) throughout their life cycle.

These arrangements can consist of a single standalone system or across a combination of systems within a test facility.

4. GLP responsibilities for data, from generation to archive

Study Personnel

All study personnel are responsible for recording raw data promptly and accurately and in compliance with the Principles of GLP.

Study Director

The study director should ensure that:

- all raw data are fully documented and recorded;
- computerised systems used in the study have been validated, including requirements associated with data integrity; and
- after completion (including termination) of the study, the study plan, the final report, raw data and supporting material are archived so that all the material, including data, needed to reconstruct the study remain available.

Archivist

The archivist is the individual responsible for the management, operations and procedures for archiving in accordance with the Principles of GLP, including archiving of data, physically and electronically.

Test Facility Management

Test Facility Management (TFM) is responsible for the organisation and functioning of the facility where data are generated. TFM should:

- ensure that a sufficient number of qualified personnel, appropriate facilities, equipment, and materials are available for the timely and proper conduct of the study, including resources to ensure data governance;
- ensure the maintenance of a record of the qualifications, training, experience and job description for each professional and technical individual;
- ensure that personnel clearly understand the functions they are to perform and, where necessary, provide training for these functions, including training on data integrity;
- ensure that appropriate and technically valid standard operating procedures (SOPs) are established and followed, and approve all original and revised SOPs, including those relating to the data governance system;
- ensure that an individual is identified as responsible for the management of the archives, including data, paper and electronic archiving;
- establish procedures to ensure that computerised systems are suitable for their intended purpose, and are validated, operated and maintained in accordance with the Principles of GLP, including functionalities associated with data integrity;
- implement systems that comply with current regulatory expectations; and
- ensure that residual risks associated with data integrity are identified and mitigated.

Quality Assurance Personnel

Quality Assurance (QA) Personnel should conduct inspections to determine if all studies are conducted in accordance with the Principles of GLP. This may include data collection, data capture systems, implemented data governance measures and associated SOPs and should be included in the QA Programme of the test facility.

5. Principle actions to ensure data integrity

1. TFM should ensure that systems implemented within the test facility produce data that are attributable, legible, contemporaneous, original, accurate, complete,

consistent, enduring and available (ALCOA+) in all its forms, i.e. paper and electronic. The study director should verify that the implemented systems are fit for the integrity of the study data.

2. TFM is expected to implement a fully documented system with supporting rationale that provides an acceptable state of control based on the data integrity risk. An example of a suitable approach is to perform a data integrity risk assessment where the processes that produce, process and/or store data are mapped out and each of the formats and their controls are identified and the data criticality, inherent risks and appropriate mitigations documented. Other documented approaches to the identification and control of data integrity risks can be acceptable.
3. Arrangements in place within the test facility with respect to organisation and personnel, systems and facilities should be designed, operated and where appropriate adapted to support a suitable working environment, i.e. providing an appropriate environment to enable the function of effective data integrity controls.
4. Data governance must be applied across the whole data life cycle to provide assurance of data integrity. Data governance should address data ownership and accountability and consider the design, operation and monitoring of processes/systems in order to comply with data integrity requirements, including control over all changes to data. Data governance systems should also ensure that data are readily available and accessible. Electronic data should be available in human-readable form.
5. The approaches used for the management of data governance should use risk management techniques to detect risks for data integrity failures within the test facility's systems, to minimise the potential risk to data integrity and to identify any residual risk. Approaches used for the management of data governance (e.g. SOPs) should always be approved by TFM. The effectiveness of the data governance approach should be monitored and assessed on a regular basis as defined by TFM.
6. TFM is expected to ensure appropriate resources and training. Data governance systems should include staff training in the importance of data integrity concepts and the creation of a working environment that enables transparency, and actively encourages reporting of errors, omissions and aberrant results.
7. The risks to data are reflected in their potential to be deleted either unintentionally or intentionally, amended, altered or excluded without authorisation or without the ability to detect such activities and events. The risks to data may be increased by complex or inconsistent or missing processes, with open-ended and subjective outcomes. Simple, well-defined tasks that are undertaken consistently and have a clear objective should be established to mitigate such risks.
8. A data integrity risk assessment (or equivalent) should consider all factors required to follow a process or perform an activity. TFM should nominate personnel to conduct the risk assessment and it is advised to be performed by a multidisciplinary team of subject matter experts that may include members with knowledge of the process, study directors, specialists in information technology (IT), QA and all other relevant functions. It is expected to consider not only the system in isolation but also all supporting activities and functions such as regulations, processes, interfaces to other systems, human intervention, training and quality systems. Automation or the use of a validated system may lower but not eliminate the risk to data integrity. Where there is human intervention, particularly influencing how or what data are recorded or reported, there may be an increased risk from poor

organisational controls or data verification due to overreliance on the system's validated state.

9. Where the data integrity risk assessment (or equivalent) has highlighted areas for remediation, then the prioritisation of actions, including acceptance of an appropriate level of residual risk, should be documented by the designated team and communicated for approval to TFM. Periodic reviews of the risk assessment should be performed to take into account the implemented actions and the possible changes in processes. In situations where long-term remediation actions are identified, risk-reducing short-term measures should be identified, documented, communicated for approval to TFM and implemented to provide an acceptable level of control in data governance until a more permanent solution is implemented.
10. Regulatory decision-making requires study data to be relevant and reliable. Data criticality may be determined by considering how the data impacts on the objectives, validity and GLP compliance of a study.
11. The effort and resource applied to assure data integrity should be commensurate with the risk and the impact of the associated data integrity failure.
12. Test facilities should be aware that appropriate data integrity controls are necessary for computerised systems as well as paper-based manual systems, although the controls may not be the same. Hybrid systems may be used if their ability to ensure data integrity is demonstrated (see also in section 7.3 “Review of data from hybrid systems”).

6. Data integrity requirements through the data life cycle

6.1. General requirements on data

Test facilities should have an appropriate level of process understanding and technical knowledge of systems used for data recording, including their capabilities, limitations and vulnerabilities.

The provision of a work environment that permits performance of tasks and recording of data as required is essential. Examples include adequate space, sufficient time for tasks and properly functioning equipment.

The following requirements are applicable to all data.

Data should be:

A - attributable to the person generating/modifying/reviewing the data

L - legible

C - contemporaneous

O - original record (or verified copy of it)

A - accurate

Data governance measures should also ensure that data are complete, consistent, enduring and available throughout the life cycle (ALCOA+), where:

Complete - the data must be whole, a complete set

Consistent - the data must be self-consistent and free from self-contradiction

Enduring - permanent, lasting throughout the data life cycle

Available - readily available

Data generated should be identified at the time of recording by the individual(s) responsible for the data entry.

Computerised system design should always provide for the retention of full audit trails to show all changes to the data without obscuring the original record. It should be possible to associate all changes to data with the person having made those changes and the date they were made, for example, by use of a data audit trail or equivalent mechanisms, or timed and dated (electronic) signatures. Reason for changes must be given.

6.2. Generation, capture or recording of raw data

Raw data generated during the conduct of the study should be recorded directly, promptly, legibly and accurately. All the raw data should be signed and dated, either electronically or on paper or on other media. Where raw data are generated as a result of direct computer input (e.g. typing a value), raw data should be identified by the identity of the person responsible for the recording and by the time of entry.

When the original electronic captured data are not considered as the raw data, this should be justified and documented.

Manual recording

Data recorded manually may require independent verification based on a data integrity risk assessment or by other requirements. Examples can include contemporaneous (or timely manner) second person verification of data entry or cross-checks of related information sources (for example, equipment logbooks, test system data, etc.) or data review. The level of control should be commensurate with the identified risk of error in the manual recording.

Manual observations should be directly and simultaneously recorded by the observer. If there is an exceptional need to confirm the manual observations (e.g. because of its high level of criticality on the validity of the study), additional actions might be considered to demonstrate data integrity (such as image capture or presence of a witness to confirm the observation). Records of the additional actions undertaken by the observer, and where relevant a witness, must be kept as additional data with the raw data recorded by the observer.

The use of scribes to contemporaneously record the activity on behalf of another operator can be considered where justified, for example:

- The act of contemporaneous recording compromises the activity (e.g. documenting test item preparation under sterile conditions by study personnel).
- In-life examination of test systems.

The recording by the second person should be contemporaneous with the task being performed and the records should identify both the study personnel performing the task and the person completing the record. The study personnel performing the task should countersign the record when possible to formalise the fact they performed the action (not the acceptance of the recorded data). The process for scribe documentation completion should be described in SOP, which should also specify the activities to which the process applies.

Access to the current version of templates or forms used to record the raw data, should be available at locations where activities take place so that data can be recorded promptly. The number of used templates compared to the number of available copies should be controlled to avoid duplication and to support the identification of data integrity issues, such as the detection of recreation or transcription of a record. If templates or forms to record data are available by printing, the number of printouts should be controlled.

Risk assessment should identify the level of control needed and the absence of full control and reconciliation should be justified by risk assessment to determine why some situations are exempt from this requirement.

The use of blank paper proformas for raw data recording should be limited and controlled but should also be available to allow the contemporaneous recording of unexpected events. The reconciliation between the available sets of blank forms at the beginning and upon completion of all issued forms should be implemented. The use of paginated books can be an appropriate solution, so that the deletion of pages could be detected. Risk assessment should identify the level of control needed and the absence of full control and reconciliation should be justified.

Nevertheless, the system implemented for controlling access to forms should allow an easy availability of the proper document to avoid the potential use of improper recording of data on an unapproved form and any subsequent transcription.

Data generated as a direct computer input should be identified at the time of data input by the individual(s) responsible for direct data entries.

For electronic data, access to applications should not hamper the contemporaneous recording of data. User access rights should prevent unauthorised data entries.

Automatic recording

External devices or system interfacing methods that eliminate manual data entries and human interaction with the computerised system, such as barcode scanners, ID card readers, or printers, can be used when validated.

The risks related to data integrity may depend on the degree to which equipment or computerised systems that automatically capture, record, or generate data can be configured and validated, and the potential for manipulation or loss of data within the system.

Hybrid systems

In the case of basic electronic equipment that does not store electronic data or provides only a printed data output (e.g. certain balances or pH meters), then the printout can constitute the raw data.

Where the electronic equipment does store electronic data but only holds a certain volume before overwriting it, all efforts should be made to extract and control the data and metadata as electronic data. Printing it to paper if immediately signed and dated or transforming it into another format is acceptable if no information is lost. Data (including metadata) in their retained format, should be verified prior to deletion from electronic equipment.

Other media

Data can be captured by a photograph or imaging methodologies and technologies (or other media), the requirements for traceability of the recording stay the same.

Recording in flat files

Most flat files do not allow the traceability of the identity of the person recording the data and the date and time of the record. Some flats files may carry basic metadata relating to file creation and date of the last amendment but do not provide an adequate data audit trail. Flat files should generally not be used for direct data capture or storing raw data.

Where the use of flat files is necessary, and control of the data cannot be achieved by an alternative method, then risk mitigations must be established that take into account the use of such files. Examples of possible mitigations could include encryption, document location access controls, or technical safeguards that can detect modifications made to the file outside of the originating software.

6.3. Metadata

For raw data to have full meaning the data requires metadata and should be considered as part of the data (see also section 6.13 "Data audit trail").

Metadata should be generated contemporaneously with the data and should be retained with the associated data.

6.4. Electronic Signatures

An electronic signature should be equivalent to the handwritten signature of the signatory and may be used to signify approval, authorisation or verification of specific data entries.

In order to ensure data integrity, the use of electronic signatures should be appropriately controlled with consideration given to:

- how the signature is attributable to an individual and to the purpose it is being used for (e.g. approval, verification, acknowledgement);
- how the act of signing is recorded within the system so that it cannot be altered or manipulated without invalidating the signature or status of the entry;
- how the time and date of the signature is recorded along with the name of the owner and the meaning of the signature;
- how the record of the signature will be associated with the entry made and how this can be verified; and
- how the security of the electronic signature is ensured i.e. so that it can only be applied by the owner of that signature.

An inserted image of a signature or a footnote indicating that the document has been electronically signed (where this has been entered by a means other than the validated electronic signature process) is not sufficient.

If, in connection with an electronic signature functionality, a traditional authentication consisting of a user ID and a secret password is replaced by biometric authentication (e.g. fingerprint, hand, face or iris scanner), the implemented solution should be thoroughly validated and documented.

(See also section 3.9 of OECD Document No 17 (OECD, 2016[4])*)*

6.5. Generation of verified copies

A verified copy (irrespective of the type of media used) of data should be confirmed (i.e. documented with dated signature or by generation through a validated process) to have the same information, including data that describe the context, content, and structure, as the original. Original and verified copies must preserve the integrity (accuracy, completeness, content and meaning) of the data.

Verification must be attributable to the individual who performs the verification. The date (and time if relevant) of the generation of the verified copy should be retained with the relevant copy.

An electronic verified copy of data recorded in paper format can be generated, provided that there is a documented process, in place to ensure that the outcome is a verified copy.

6.6. Correction or amendment of data

Any change in the raw data should be made so as not to obscure the previous entry, should indicate the reason for change and should be dated and signed or initialled by the individual making the change.

For data generated as a direct computer input, computerised system design should always provide for the retention of full audit trails to show all changes to the data without obscuring the original record. It should be possible to associate all changes to data with the persons having made those changes, for example, by use of timed and dated (electronic) signatures (see also section 6.13 “Data audit trail”). Reason for changes should be given and recorded.

6.7. Transcription

Transcriptions should be avoided as they increase the risks of errors and inconsistencies. Where transcriptions cannot be avoided, they should be verified by a second person or operated by a validated system. The original records should be regarded as raw data and should be retained.

6.8. Invalidating or Excluding Data

Data may only be invalidated or excluded where it can be demonstrated through sound scientific or technical justification or logical sense that the data are not representative of the recorded event. The rejection of analytical results due to equipment malfunction, or the invalidation of a clinical observation monitored from a dead animal are relevant examples.

Investigations to find the cause of the generation of data that must be invalidated or excluded are essential. In all cases, the justification of the invalidation or exclusion should be documented and considered during data review and reporting. For common cases (e.g. incoherent analytical results for a single sample, or failure to meet acceptance criteria), the rules to exclude or invalidate data should be defined in advance in the study plan or in SOPs. All data (even if invalidated) should be retained with the data set and be available for review in a format that allows the validity of the decision to invalidate or exclude the data to be confirmed.

6.9. Data Processing

There should be adequate traceability of any user-defined parameters within data processing activities, including attribution to who performed the activity. Examples include

calculations or (with proper access permissions) the selection and application of chromatography integration parameters or selection of gating parameters for analysis of a flow cytometry assay. Processing data rules should be clearly defined and controlled by SOPs.

The raw data and available audit trails of the process should be retained. Retained records should allow reconstruction of all data processing activities regardless of whether the output of that processing is subsequently reported. If data processing has been repeated with progressive modification of processing parameters, this should be visible with documented justification to ensure that the processing parameters are not being manipulated to achieve a more desirable end point.

6.10. Data Migration

Data migration procedures should include a rationale and be robustly designed and validated to ensure that data integrity is maintained during the data life cycle. Careful consideration should be given to understanding the data format and the potential for alteration at each stage of data generation, migration and subsequent storage. Measures to ensure and demonstrate that data are not altered during each step of the process should be in place.

The challenges of migrating data are often underestimated, particularly regarding maintaining the full meaning and integrity of the records, including associated metadata.

In case of migration from a party (the "sender") to another (the "receiver"), the data, and the associated metadata, date/time of migration, expected format and specification of a transfer protocol or agreement used to migrate the data should be defined before migration. Mechanisms of communication and coordination between the sender and the receiver should be in place to ensure that the received data have the same attributes as the sent data.

(See also section 2.8 of OECD Document No 17 (OECD, 2016[4])*)*

6.11. Relational Database

Retrieval of information from a relational database requires a database reporting tool or the original application that created the record.

Amendments to data should not be performed directly into the database fields but should be via the originator software package, so that appropriate audit trail entries and controls remain in place. Nevertheless, if a data change by a system administrator is required directly into the database, this should be justified, controlled, documented, have the study director's approval and the process should be described in an SOP.

Access rights for database entry or amendment should be controlled, and consistent with the requirements for computerised system user access/system administrator roles (see also section 8.2 "Computerised system access and roles").

6.12. Computerised System Transactions

A computerised system transaction where a parameter must be within a defined limit, range, or distribution to ensure quality of the data should be considered as critical. Computerised systems should be designed to ensure that the execution of such transactions are recorded contemporaneously. Where transactional systems are used, the combination of multiple unit operations into a combined single transaction should be avoided (e.g. multiple data entry before saving), and the time intervals before saving of data should be

minimised. Systems should be designed to require saving data to permanent memory before prompting users to make changes. Exceptions to these requirements should be justified.

TFM should define during the development of the system (e.g. via the user requirements specification) what critical transactions are linked to that system based on the functionality and the level of risk associated with the system. Critical transactions should be documented with process controls that consider system design (prevention), together with monitoring and review processes. Oversight of activities should alert to failures that are not addressed by the process design. Surveillance activities of critical transactions should be considered as part of the QA programme.

6.13. Data Audit Trail

Where computerised systems are used to capture, process, modify, report, store or archive data electronically, system design should always provide for the retention of audit trails to show all changes to, or deletion of the data while retaining previous data. It should be possible to associate all data and changes to data with the persons making those changes, and changes should be dated and time stamped (time and including, where applicable, the time zone). The reason for the change should also be recorded. The items included in the audit trail should be those of relevance to permit reconstruction of the process or activity.

Audit trails should always be switched on during GLP activities. Any personnel with a direct interest in the data (study directors, heads of analytical departments, study personnel etc.) should not have the ability to amend or switch off the audit trail functionality. Where a system administrator amends or switches off the audit trail functionality, the audit trail should record this automatically and it should also be recorded automatically when the audit trail functionality is switched on again.

Where relevant audit trail functionality does not exist or systems do not meet the audit trail and individual user account expectations (e.g. within legacy systems), demonstrated progress should be available to address these shortcomings. This should either be through add-on software that provides these additional functions or by an upgrade to a compliant system. Remediation has to be identified and implemented in a timely manner.

If a system has no audit trail capability and review of available systems cannot identify alternatives and technological adaptations or additions to the existing system (i.e. remediation is not possible), this should be justified by evidence that a compliant solution is being worked upon and what mitigation activities, such as alternative level of control, temporarily supports the continued use. Alternative levels of control may be achieved by, for example, the use of manual logbooks or the definition of strict restricted access rights to the system. The printouts of the data could be also considered if integrity of data, including metadata, is ensured. Alternative controls measures should be proven to be effective, risk-based, defined within an SOP and periodically reviewed for reassessment.

Some GLP Compliance Monitoring Authorities may not accept systems without audit trail functionality including those with alternative control measures.

(See also section 3.4 of OECD Document No 17 (OECD, 2016[4])*)*

6.14. Data retention

Data required to allow the full reconstruction of activities of the studies should be collected and retained. Data should be retained with the associated metadata when applicable. Derived data should be retained with their raw data when necessary for the reconstruction of the study.

Data and document retention arrangements should ensure the protection of records from intended or unintended alteration or loss. Secure controls must be in place to ensure the data integrity of the record throughout the retention period.

The selected method for retention should ensure that data of appropriate accuracy, completeness, content and meaning are collected and retained for its intended use.

Retention of dynamic data

Information that is captured in a dynamic state should remain available in that state. For example, video recordings used to demonstrate an activity cannot be reduced to a single static image or to a series of single images.

Computerised systems that generate dynamic records should allow the dynamic nature of the data to be retained.

It may be a challenge to print on paper dynamic records without losing the interactive relationship between the user and the record content.

Any printouts should comprise of all associated available metadata, and should keep the link that binds them to the raw data. For example, if the associated metadata are printed in another page from the raw data, the integrity of the link is not ensured and the relationship to the raw data questionable.

When electronic raw data cannot be converted to verified copies (e.g. to prints on paper or pdf) without a loss of information (e.g. associated metadata), they should remain available in the original state.

If the computerised system cannot be maintained, e.g. if it is no longer supported, then records will be archived according to a documented archiving strategy prior to decommissioning the computerised system. It is conceivable for some data generated by electronic means to be retained in an acceptable paper or electronic format, where it can be justified that a static record maintains the integrity of the raw data. However, the data retention process must be shown to include verified copies of all raw data, metadata, relevant audit trail and result files, any variable software/system configuration settings specific to each record, and all data processing runs (including methods and audit trails) necessary for reconstruction of a given raw data set.

When printing on paper is the chosen solution, it would require a validated means to verify that the printed records were an accurate representation of the data set.

All information should be retained. Any loss of information should be identified and the risk on the integrity of the data set should be assessed and documented.

Retention of electronic signature

An electronically signed document is generally a dynamic record. Where a document is electronically signed, the metadata associated with the signature (i.e., printed name of the signer, meaning of signature, and date and time of the signature) should be electronically retained. A document that is signed electronically is only valid when retained electronically unless the paper print out or the pdf copy retains all the traceability to the signers identity, date and time and meaning of signature.

Retention of electronic communications

Electronic communications are another example of records in a dynamic state.

Where data are supported by electronic communication methods such as email and electronic messaging (e.g. allowing verification of GLP activities and responsibilities), processes for ensuring retention and the collation of electronic communications should be established (including ensuring that the records are complete and integrity is intact). Such mechanisms should be designed to maintain the attributability and integrity of those relevant electronic communications such as ensuring that the sender and receiver can be determined alongside appropriate dates and times. Any attachments should remain associated with the corresponding message and message chains should be preserved.

Where possible, these should be retained in their original format but if this is not possible, TFM should implement processes for faithful transcription and verification in a retainable format.

The printout on paper or the migration in a flat pdf file of electronic communication cannot ensure the required integrity.

Retention of verified copies

Verified copies from electronic dynamic records (generated by migration) should be retained in dynamic state, so that the verified copy could include the metadata required to ensure that the full meaning of the data (e.g. date formats, context, layout, electronic signatures and authorisations) is kept and its history, including the creation of the verified copy, may be reconstructed.

Verified copies may be retained in place of the original, provided that a documented system is in place to verify and record the integrity of the copy. Consideration should be given to any risk associated with the destruction of original records. It should be recognised that some regulatory authorities require originals to be retained.

Retention of data from hybrid systems

Where hybrid systems are required to be used, this should be clearly documented as to what constitutes the whole data set and SOP should define which records should be retained.

Retention of data on other media

Where data are captured by a photograph or imaging methodologies and technologies (or other media), the requirements for storage of that format throughout its life cycle should follow the same considerations as for all data, considering any additional controls required for that format. Where the original format cannot be retained due to degradation issues, alternative mechanisms for recording including verification of the faithfulness of the process (e.g. photography or digitalisation) and subsequent storage may be considered, and the selection rationale documented.

6.15. Back-up

Mechanisms for ensuring that back-ups have completed successfully should be considered. The systems used should be validated and each back-up can be verified to ensure that it has functioned correctly e.g. by confirming that the data size and other copied properties matches that of the original record.

Back-up and recovery processes for electronic data should be tested where appropriate. Such as when changes occur to either the process or tools or applications used during back-up or restore. Moreover, the sustainability of some electronic media used for back-up (such as CDs, DVDs, etc.) needs to be verified periodically.

The back-up procedures should be described in an SOP, and back-up activities should be documented.

Back-ups for recovery purposes do not replace the need for archiving of data and metadata for the purposes of reconstruction of the study activity.

6.16. Archive

Data should be archived securely, under the control of the unique archivist, including, where relevant, an appropriate electronic repository whether this is on the original system or elsewhere, subject to suitable controls or in a stand-alone electronic archive.

All archive sites (physical as well as electronic) associated with the archived data should be identified and documented.

The Principles of GLP for archiving must be applied consistently to electronic and non-electronic data. It is therefore important that electronic data are stored with the same levels of access control and indexing as non-electronic data.

Archived records may be the original record and/or a verified copy (see also in section 6.14 "Retention of verified copies") and should be protected such that they cannot be altered or deleted without detection.

Archive arrangements must be designed to permit retrieval and readability of data and metadata throughout the required retention period.

When legacy systems can no longer be supported, consideration should be given to the importance of the data, and if required, to maintaining the software for data accessibility purposes. This may be achieved by maintaining software in a virtual environment. Where this is not possible, data should be migrated before archiving in a controlled, tested, and verified way to a system that can continue to be accessed. Migration to an alternative file format that retains the verified copy attributes of the data may be necessary with increasing age of the legacy data.

Where migration with full original record functionality is not technically possible, selection from the options available would have to be based on risk and importance of data over time. The migration file format should be selected taking into account the balance of risk between long-term accessibility versus the possibility of reduced dynamic data functionality (e.g. data interrogation, trending, re-processing etc.). It is recognised that the need to maintain accessibility may require migration to a file format that loses some attributes and/or dynamic data functionality. It is the TFM's responsibility to assess the impact of such losses and maintain the link between the readable audit trail or electronic signatures and the audited data to an acceptable level.

(See also section 3.11 of OECD Document No 17 (OECD, 2016[4])*)*

7. Data review

7.1. General considerations

Data review consists of appropriate verifications of critical data for quality control which can be conducted by study directors or other personnel.

The objectives of data review are:

- to detect any deletion, amendment, alteration or exclusion;
- for the study directors, to check that all raw data generated are fully documented and recorded; and
- to assess the efficiency of data governance measures by review of a complete data set generated through processes throughout the data life cycle.

To be effective, the level of data review and the scope of it should be defined by a risk assessment. Identified critical data should be reviewed through the critical steps of their data life. Data review should also include a review of relevant metadata, including audit trails or elements of them.

Data review should be documented. The record of the review should include any deviations to the Principles of GLP, study plans or SOPs detected by the review, the date that review was performed and the signatures of those performing the review.

There should be a procedure that describes the process for the data review. A procedure should also describe the actions to be taken if data review identifies deviations. This procedure should enable data corrections or clarifications to provide visibility of the original record, and audit trailed traceability of the correction.

Many software packages allow configuration of customised reports to support data review. Changes to report configuration should be controlled to prevent unauthorised changes. The system should be validated and where relevant the report outputs should be verified.

Note: The data review conducted by QA aims to support the statement that the reported results accurately and completely reflect the raw data of the studies. It may also be effective when auditing data integrity governance procedures. The level of review should be linked with the criticality of the data.

7.2. Review of data audit trail

It is not necessary for audit trail review to include every system activity.

The relevant data among all the retained data in audit trails should be identified to permit robust data review/verification. The review should be conducted according to a documented risk-based process identifying the criticality of the data subject to the review and the criticality of transactions identified through the data flow. The review may be achieved by direct access to the system audit trail or by use of appropriately designed and validated system reports.

Routine data review should include a documented audit trail review as determined by the risk assessment. When designing a system for review of audit trails, this may be limited to those activities with GLP relevance (e.g. relating to data creation, processing, compliance with procedures, modification and deletion etc.). Audit trails may be reviewed as a list of relevant data, or by an ‘exception reporting' process. An exception report is a validated search tool that identifies and documents predetermined ‘abnormal’ data or actions, which requires further attention or investigation by the data reviewer.

Reviewers should have sufficient knowledge and system access to review relevant audit trails, raw data and metadata.

7.3. Review of data from hybrid systems

Increased data review is likely to be required for hybrid systems because they are vulnerable to non-attributable data changes. All records from hybrid systems that are

defined by the data set should be reviewed by a qualified person. The level of this control should be adapted to the processes used in the hybrid system. Review of data from hybrid systems should be clearly defined and described so that it is possible to determine the actual data sources reviewed.

8. Access to data

8.1. General considerations

Access rights to data and records should be always created based on the risk assessment of each phase of the data lifecycle.

Access right should be defined to allow the personnel to fulfil their GLP responsibilities.

Access to records for personnel performing data review activities should be maintained.

The necessary access (including to records, audit trails and system functionality), permissions and training should be available to support QA inspection to verify if all studies are conducted in compliance with the Principles of GLP.

8.2. Computerised system access and roles

User access

Full use should be made of access controls to ensure that personnel have access only to functionality that is appropriate for their job and study role, and that actions are attributable to a specific individual. TFM must be able to demonstrate the access levels granted to individual staff members and ensure that historical information regarding user access level is available. Where the system does not capture these data, then a paper record should be available. Controls should be applied to both the operating system and application levels. Individual login at operating system level may not be required if appropriate controls are in place to ensure data integrity (e.g. individual login at application level should be sufficient if modification of data outside the application is not possible).

For systems generating, amending or storing GLP data, shared logins or generic user access should not be used. Where the computerised system design supports individual user access, this function must be used. This may require the purchase of additional licences.

Systems that are not used in their entirety for GLP purposes but do have elements within them, such as approved suppliers, stock status, location and transaction histories that are GLP applicable require appropriate assessment.

It is acknowledged that some computerised systems support only a single user login or limited numbers of user logins. Where no suitable alternative computerised system is available, equivalent control may be provided by third-party software or a paper-based method of providing traceability (with version control). The suitability of alternative systems should be justified and documented.

System administrator access

System administrator access should be restricted to the minimum number of people possible taking account of the size and nature of the test facility. The generic system administrator account should not be available for routine use. Personnel with system administrator access should log in with unique credentials that allow actions in the audit

trail(s) to be attributed to a specific individual. The intent of this is to prevent giving access to users with a potential conflict of interest to prevent unauthorised changes that would not be traceable to that person.

System administrator rights (permitting activities such as data deletion, database amendment or system configuration changes) should not be assigned to individuals with a direct interest in the data (data generation, amendment, deletion, review or approval). Any changes to study data performed by a system administrator must only be done after receiving prior permission from the study director.

Where an independent system administrator cannot be assigned (e.g. in small test facilities), a similar level of control may be achieved using dual user accounts with different privileges with all changes performed under system administrator access subject to appropriate review and approval.

The individual should log in using the account with the appropriate access rights for the given task e.g. a laboratory technician performing data checking should not log in as system administrator where a more appropriate level of access exists for that task. The suitability of such an arrangement should be periodically reviewed.

(See also sections 1.3.1 and 3.7 of OECD Document No 17 (OECD, 2016[4])*)*

References

OECD (1997), *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 1: OECD Principles on Good Laboratory Practice (revised in 1997)*. [1]

OECD (2007), *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 15: Advisory Document of the Working Group on Good Laboratory Practice : Establishment and Control of Archives that Operate in Compliance with the Principles of GLP*. [2]

OECD (2014), *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 16: Advisory Document of the Working Group on Good Laboratory Practice : Guidance on the GLP Requirements for Peer Review of Histopathology*. [3]

OECD (2016), *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring Number 17: Advisory Document of the Working Group on Good Laboratory Practice : Application of GLP Principles to Computerised Systems*. [4]