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

Table of Contents

1. Background

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

The way in which regulatory data, which supports human, animal and environmental safety assessment, is generated, handled, reported 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, automation of systems and use of remote technologies and the increased complexity in ways of working, for example, via third party service providers. Systems can range from manual processes with paper records to the use of fully computerised systems. However, the main purpose of the regulatory requirements remains the same: having confidence in the quality and the integrity of the data and being able to reconstruct activities performed during the conduct of non-clinical safety studies.

2. Introduction

2.1 This document provides guidance for test facilities that conduct GLP studies.

2.2 The guidance aims to promote a risk-based approach to the management of data which includes data risk, criticality and lifecycle. Users of this paper need to understand their data flows (as a lifecycle) to identify data which is likely to impact on GLP compliance. In turn, this will allow the identification of the most effective and efficient risk-based controls.

2.3 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. The data should be collected and maintained in a secure manner, such that is attributable, legible, contemporaneously recorded, original or a true copy and accurate. 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 chapters 7.5 and 7.6). Data integrity provides control over the data (i.e., whether it can be trusted), whereas data quality refers to the assurance that data produced are generated in compliance with applicable standards and can be used for its intended purpose.

2.4 The principles described in this guidance are primarily focused on electronic data, but these could equally be applied to the control of other data types and formats.

2.5 The guidance refers to the acronym ALCOA being Attributable, Legible, Contemporaneous, Original, and Accurate. ALCOA has historically been regarded as the attributes of data quality which 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) lifecycle.

2.6 This guidance should be read in conjunction with OECD Documents No 1 (*OECD Principles on Good Laboratory Practice*), 15 (*Establishment and Control of Archives that Operate in Compliance with the Principles of GLP*) and 17 (*Application of GLP Principles to Computerised Systems*) and applicable national regulations. The references to the GLP Principles about data integrity can be found in chapters 1.2.2.f, 1.4.3, 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 both this document and Document No 17 then reference is made within the text.

3. GLP Responsibilities on 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 and they are responsible for the quality of their data.

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 study report, raw data and supporting material are archived. The study director should also ensure all the data needed to reconstruct their studies is available and ensure that all systems apparatus used are fit for purpose.

Archivist

Archivist is the individual responsible for the management, operations and procedures for archiving in accordance with established Standard Operating Procedures (SOPs), and the Principles of GLP (including archiving of data, physically or 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 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); and
- 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).

Quality Assurance Personnel

Quality Assurance Personnel should conduct inspections to determine if all studies are conducted in accordance with the Principles of GLP (including periodic review or surveillance of data governance measures).

4. The principles of data integrity

4.1 TFM should ensure that systems implemented within the test facility produce data that are attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring and available in all its forms, i.e. paper and electronic.

4.2 Arrangements in place within the test facility with respect to people, systems and facilities should be designed, operated and where appropriate adapted to support a suitable working environment, i.e. creating the right environment to enable data integrity controls to be effective.

4.3 Approaches used for the management of data governance (e.g. standard operating procedures) should always be endorsed by TFM. The effectiveness of the data governance approach should be monitored and assessed on a regular basis.

4.4 TFM is expected to implement a fully documented system that provides an acceptable state of control based on the data integrity risk with supporting rationale. An example of a suitable approach is to perform a data integrity risk assessment where the processes that produce, modify 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 are also considered acceptable.

4.5 Systems should maintain appropriate levels of control and wider data governance measures should ensure that the quality assurance programme can detect opportunities for data integrity failures within the test facility's systems.

4.6 The effort and resource applied to assure the integrity of the data should be commensurate with the risk and impact of a data integrity failure.

4.7 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.

5. Establishing data criticality and inherent integrity risk

5.1 Data, depending on what it is, have varying importance on safety decisions. Data criticality may be determined by considering how the data impacts on the aims of studies as detailed in the study plan and how ultimately the data influences the regulatory decision-making process.

5.2 The risks to data are reflected in its potential to be deleted, amended or excluded without authorisation along with a lack of opportunity for detection of such activities and events. The risks to data may be increased by complex or inconsistent processes, with open-ended and subjective outcomes compared to simple tasks that are undertaken consistently, are well defined and have a clear objective.

5.3 Data may be captured and recorded by:

(i) documentation, on paper, of a manual observation or of an activity;
(ii) electronic means, using equipment that range from simple instruments through to complex highly configurable computerised systems;
(iii) using a hybrid system where both paper-based and electronic records constitute the original record;
(iv) other means such as photography, imagery, chromatography plates, etc.

Paper

Data recorded manually on paper may require independent verification based on a data integrity risk assessment or by other requirements. Examples include contemporaneous (or timely) second person verification of data entry or cross-checks of related information sources (for example, equipment logbooks, test system data, etc.).

Electronic media

The risks related to data integrity may depend on the degree to which equipment or computerised systems that capture, generate or use data can be configured, and the potential for manipulation during transfer of data between computerised systems.

Hybrid

A hybrid system may include combinations of paper (or other non-electronic media) and electronic records.

Where hybrid systems are required to be used, this should be clearly documented as to what constitutes the whole data set. All records that are defined by the data set should be reviewed by an appropriately qualified person and retained.

Other

Where the data are captured by a photograph or imagery (or other media), the requirements for storage of that format throughout its lifecycle should follow the same considerations as for the other formats, considering any additional controls required for that format. Where the original format cannot be retained due to degradation issues, alternative mechanisms for recording (e.g. photography or digitisation) and subsequent storage may be considered, and the selection rationale documented (e.g. thin layer chromatography).

Where raw data are supported by electronic communication methods, for example email (see also section 7.13), electronic messaging and social media networks, consideration should be given to mechanisms for the collation (including ensuring that the records are complete, and integrity is intact), retention and subsequent archiving of such records. Where possible these should be retained in their original format but if this is not possible then TFM should implement processes for faithful transcription in a retainable format.

5.4 A data integrity risk assessment should consider all factors required to follow a process or perform a function. It is expected to consider not only the computerised system in isolation but also all supporting activities and functions such as: guidance, process, 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.

5.5 Where the data integrity risk assessment has highlighted areas for remediation prioritisation of actions, including acceptance of an appropriate level of residual risk should be documented, communicated to TFM, and kept under review. In situations where long-term remediation actions are identified, risk-reducing short-term measures should be implemented to provide acceptable data governance until a more permanent solution is implemented.

6. Designing systems and processes to assure data integrity.

6.1 Systems and processes, designed to maintain data integrity may include (list is not exhaustive):

- Accessibility of templates or forms at locations where activities take place so that data are recorded promptly, accurately and legibly.
- Access to blank paper proformas for raw/source data recording should be controlled where this is appropriate. Reconciliation, or the use of controlled books with numbered pages, may be necessary to prevent recreation of a record.
- User access rights that prevent (or audit trail with suitable controls, if prevention is not possible) unauthorised data entries or amendments. Use of 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.
- The provision of a work environment (such as adequate space, sufficient time for tasks, and properly functioning equipment) that permit performance of tasks and recording of data as required.
- Access to original records for staff performing data review activities.
- Access to data systems for quality assurance personnel performing audits.
- Reconciliation of controlled printouts.
- Sufficient training in data integrity principles provided to all appropriate staff (including TFM).
- Subject matter experts are included in the risk assessment process.

6.2 The use of scribes to record 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 etc.

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 wherever possible, although it is accepted that this countersigning step will be retrospective. The process for supervisory (scribe) documentation completion should be described in an approved procedure, which should also specify the activities to which the process applies.

7. Definition of terms and interpretation of requirements

7.1 Data

Facts, figures and statistics collected together for reference or analysis. All original records and true or verified copies of original records, including source data and metadata and all subsequent transformations and reports of these data, which are generated or recorded at the time of the GLP activity and allow full and complete reconstruction and evaluation of the GLP activity.

Data should be:

A - attributable to the person generating the data
L – legible and permanent
C – contemporaneous
O – original record (or certified true copy)
A - accurate

Data governance measures should also ensure that data are complete, consistent, enduring and available throughout the lifecycle, where:

Complete – the data must be whole, a complete set

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

Enduring – lasting throughout the data lifecycle

Available – readily available for inspection purposes

Data generated should be identified at the time of recording by the individual(s) responsible for direct data entries. Computerised system design should always provide for the retention of full audit trails to show all changes to the data without obscuring the original data. 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. Reason for changes should be given.

7.2 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. 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.

Raw data generated during the conduct of the study should be recorded directly, promptly and accurately. Where data are generated as a result of direct computer input (e.g. typing a value) by

an individual then this should be identified at the time of data input by the person responsible for the entry.

Raw data must permit the full reconstruction of the study activities. Where this has been captured in a dynamic state and generated electronically, paper copies cannot usually be considered as 'raw data' due to the absence of associated metadata. Information that is captured in a dynamic state should remain available in that state.

In the case of basic electronic equipment which 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 basic 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 as electronic data.

In all definitions, the term 'data' includes raw data.

7.3 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 LIMS, etc.).

7.4 Metadata

Metadata are any information used for the identification, description, and relationships of electronic records or their elements. Metadata give data meaning, provides context, defines structure, and enables retrievability across systems, and usability, authenticity, and auditability across time.

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

7.5 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. The data should be collected and maintained in a secure manner, such that they are attributable, legible, contemporaneously recorded, original or a true copy and accurate. Assuring data integrity requires appropriate quality and risk management systems, including adherence to sound scientific principles and good documentation practices.

7.6 Data Quality

Data quality is the assurance that the data produced are generated according to applicable standards and fit for intended purpose in regard to the meaning of the data and the context that supports it.

Data quality affects the value and overall acceptability of the data in regard to decision-making or onward use.

7.7 Data Governance

Data governance is the sum total of arrangements to ensure that data (irrespective of the format in which they are generated, recorded, processed, retained and used) are complete, consistent and accurate records throughout its lifecycle.

Data governance should address data ownership and accountability throughout the lifecycle, 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 include staff training in the importance of data integrity concepts and the creation of a working environment that enables visibility, and actively encourages reporting of errors, omissions and aberrant results.

TFM should be accountable for the implementation of systems and procedures to minimise the potential risk to data integrity, and for identifying the residual risk, using risk management techniques.

Data governance systems should also ensure that data are readily available and directly accessible on request from national competent (monitoring or receiving) authorities. Electronic data should be available in human-readable form.

7.8 Data Lifecycle

Data lifecycle 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 governance, as described in the previous section, must be applied across the whole data lifecycle to provide assurance of data integrity. Data should be retained securely, under the control of the archivist, including where relevant an appropriate electronic repository whether this is on the original system, subject to suitable controls or in a stand-alone electronic archive.

7.9 Recording and Collection of Data

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

The selected method should ensure that data of appropriate accuracy, completeness, content and meaning are collected and retained for its intended use. Systems that generate dynamic data should allow the dynamic nature of the data to be retained. Retention of static (printed/manual) data generated from dynamic data is not appropriate.

As data are required to allow the full reconstruction of activities the amount and the resolution of data to be collected should be justified.

If used, blank forms (including, but not limited to, worksheets and laboratory notebooks) should be controlled by documented procedures. For example, numbered sets of blank forms may be issued as appropriate and should be reconciled upon completion of all issued forms. Similarly, bound paginated notebooks, stamped for official use by a document control group, allow for detection of unofficial notebooks as well as of any gaps in notebook pages.

7.10 Data Transfer/Migration

Data transfer or migration is the process of moving data between different data storage types, computerised systems, or simply the transition of data from one state 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.

Data transfer/migration procedures should include a rationale and be robustly designed and validated to ensure that data integrity is maintained during the data lifecycle. Careful consideration should be given to understanding the data format and the potential for alteration at each stage of data generation, transfer, migration and subsequent storage.

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

(See also section 2.8 of OECD Document 17)

7.11 Data Processing

A sequence of operations performed on data in order to extract, present or obtain information in a defined format. Examples might include 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.

There should be adequate traceability of any user-defined parameters used within data processing activities to the raw data, including attribution to who performed the activity examples include the selection and application of chromatography integration parameters or selection of gating parameters for analysis of a flow cytometry assay.

The original raw data and audit trails 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 to ensure that the processing parameters are not being manipulated to achieve a more desirable end point.

7.12 Invalidating Data

Data may only be invalidated where it can be demonstrated through sound scientific or technical justification that the data are not representative of the quantity measured, sampled or acquired. For example, failure during system suitability checks.

In all cases, this justification should be documented and considered during data review and reporting. All data (even if invalidated) should be retained with the original data set and be available for review in a format that allows the validity of the decision to exclude the data to be confirmed.

7.13 Original Record and True Copy

7.13.1. Original record (see also raw data definition)

The first or source capture of data or information e.g. original paper record of manual observation, or electronic raw data file from a computerised system and all subsequent data required to fully reconstruct the conduct of the GLP activity. Original records can be Static or Dynamic.

Static Records

A static record format, such as a paper or electronic record, is one that is fixed and allows little or no interaction between the user and the record content. For example, a paper record of a clinical observation of an animal, dated and signed is a static record.

Dynamic Records

Records in a dynamic state, such as electronic records, allow for an interactive relationship between the user and the record content. For example, electronic records in database formats allow the user to track, trend and query data; chromatography records maintained as electronic records allow the user (with proper access permissions) to reprocess the data and expand the baseline to view the integration more clearly or to view links between data points, audit trails and technical parameters.

Chromatography records maintained as electronic records contain instrument and experimental information

Email

Relevant email that allows verification of activities and responsibilities should be retained in their original format. Email could have different roles in the reconstruction of the GLP study. An assessment should be conducted to define if emails and their attached documents are raw data. If emails are raw data, consideration should be given to mechanisms to maintain the attributability and integrity of the email (such as ensuring that the sender and receiver can be determined alongside appropriate dates and times. Processes for the collation of email should be established (including ensuring that the records are complete and integrity is intact), for example, attachments should remain associated with the corresponding email and email ‘chains’ are preserved.

Manual Observations

Where the data obtained requires manual observation to record it, such as a measurement of a weight or pH value where the instrument does not have any means of output, the process

should be risk assessed and additional actions might be considered to demonstrate data integrity (such as image capture or presence of a witness).

7.13.2 True Copy (synonymous with 'verified copy' as used in the GLP principles)

A true copy (irrespective of the type of media used) of the original record that has been verified (i.e., by a 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.

A true copy may be stored in a different format or document type to the original record if required. If the true copy is electronic it must retain the metadata and audit trail 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 true copy, may be reconstructed.

Original records and true copies must preserve the integrity (accuracy, completeness, content and meaning) of the record. True copies of original records may be retained in place of the original record (e.g. scan of a paper record), 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 competent authorities require that original records are retained in the format in which they were generated.

It should be possible to create a true copy of electronic data, including relevant metadata, for the purposes of review, backup and archival. Consideration should be given to the dynamic functionality of a 'true copy' throughout the retention period (see 'archive').

Data must be retained in a dynamic form where this is critical to its integrity or verification for example moving images used to demonstrate an activity cannot be reduced to a single static image. If the computer 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 computer 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 original 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. It would also require a validated means to verify that the printed records were an accurate representation.

7.14 Transcription

Transcription is the process where a written copy of an original record is made. Transcription should be avoided wherever possible. Where manual transcriptions occur, it is advisable that they are verified by a second person or validated system. The original records should be regarded as raw data and should be retained.

7.15 Computer System Transactions:

A computer system transaction is a single operation or sequence of operations performed as a single logical 'unit of work'. The operation(s) that makes 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), or until the system forces the saving of data.

The metadata (e.g. username, date, and time) are not captured in the system audit trail until the user saves the transaction to durable storage. In computerised systems, an electronic signature may be required for the record to be saved and become permanent.

A critical step is computer system transaction where a parameter must be within a defined limit, range, or distribution to ensure the quality of the data. Computer systems should be designed to ensure that the execution of critical steps is recorded contemporaneously. Where transactional systems are used, the combination of multiple unit operations into a combined single transaction should be avoided, 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.

The test facility should define during the development of the system (e.g. via the user requirements specification) what critical steps are appropriate for that system based on the functionality of the system and the level of risk associated with the system. Critical steps 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 steps should be considered as part of the quality assurance programme.

7.16 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 GLP records. An audit trail provides a secure automated way of recording life-cycle details such as creation, additions, deletions or alterations of information in a record, either paper or electronic, without obscuring or overwriting the original record. An audit trail facilitates the reconstruction of the history of such events relating to the record regardless of its medium, including the 'who, what, when and why' of the action.

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 and original 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, where applicable, 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 be switched on. Any personnel involved in a study (study directors, heads of analytical departments, analysts 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 recording any changes made.

The relevance of data retained in audit trails should be considered by the test facility to permit robust data review/verification. It is not necessary for audit trail review to include every system activity (e.g. user log on/off, keystrokes etc.) and should be achieved by review of appropriately designed and validated system reports.

Where relevant audit trail functionality does not exist (e.g. within legacy systems) and another technical solution is not easily available alternative levels of control can be implemented. Alternative levels of control may be achieved by, for example the use of logbooks or monitoring of access. Alternative controls should be proven to be effective and defined within an SOP.

Where an add-on software or a compliant system does not currently exist, this should be justified by evidence that a compliant solution is being worked upon and what mitigation temporarily supports the continued use.

Routine data review should include a documented audit trail review as appropriate as determined by a risk assessment. When designing a system for review of audit trails, this may be limited to those with GLP relevance (e.g. relating to data creation, processing, 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. Review of data from hybrid systems should be clearly defined and described so that it is possible to determine the actual data sources reviewed. See also 'data governance'.

Where systems do not meet the audit trail and individual user account expectations, demonstrated progress should be available to address these shortcomings. This should either be through add-on software which provides these additional functions or by an upgrade to a compliant system. Where remediation has not been identified or implemented in a timely manner, a deficiency may be cited.

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

7.17 Electronic Signatures

A signature in digital form that represents the hand-written ('wet') signatory. The electronic signature should be equivalent in legal terms 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 record of the signature will be associated with the entry made and how this can be verified; and

- the security of the electronic signature i.e. so that it can only be applied by the 'owner' of that signature.

Where a paper or pdf copy of an electronically signed document is produced, the metadata associated with the electronic signature should be maintained.

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. Where a document is electronically signed then the metadata associated with the signature should be electronically retained. A document that is signed electronically is only valid when retained electronically unless the paper print out retains all associated metadata.

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.

For printed copies of electronically signed documents refer to True Copy section.

(see also section 3.9 of OECD Document 17)

7.18 Data Review and Approval

The approach to reviewing specific record content, such as critical data fields and metadata such as cross-outs on paper records and audit trails in electronic records, should meet all applicable regulatory requirements and be risk-based.

There should be a procedure that describes the process for the review and approval of data. Data review should also include a review of relevant metadata, including audit trails or elements of them. Data review should be documented and may be risk based. The record of the review should include any issues found, the date that review was performed and the signatures of those performing the review.

A procedure should describe the actions to be taken if data review identifies an error or omission. This procedure should enable data corrections or clarifications to provide visibility of the original record, and audit trailed traceability of the correction, using ALCOA principles (see 'data' definition).

Many software packages allow configuration of customised reports. Key actions may be incorporated into such reports with report validation proving this is done accurately. Validation and locking of automated reporting tools and reports may also be used reduce the checks required to verify the integrity of the data.

7.19 Computerised System User Access/System Administrator Roles

Full use should be made of 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. Test facilities 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. no modification of data outside the application 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 which 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. Increased data review is likely to be required for hybrid systems because they are vulnerable to non-attributable data changes. It is expected that test facilities should be implementing systems which comply with current regulatory expectations.

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 potentially a conflict of interest so that they can make 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, data review or approval).

Where an independent system administrator cannot be assigned (e.g. in very 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 and will be subject to review during competent authority inspection.

(see also sections 1.3.1 and 3.7 of OECD Document 17)

7.20 Data Retention

Data retention may be for the purpose of archiving (protected data for long-term storage) or backup (data for the purposes of disaster recovery).

Data and document retention arrangements should ensure the protection of records from deliberate or inadvertent alteration or loss. Secure controls must be in place to ensure the data integrity of the record throughout the retention period. See also data transfer/migration.

Data (or a true copy thereof) generated in paper format may be retained for example by using a validated scanner, provided that there is a process in place to ensure that the outcome is a true copy (see also 'true copy').

The procedures for destruction of data should consider data criticality and where applicable legislative retention requirements. The destruction of data should only be carried out with the approval of the TFM and if applicable, the Sponsor)

7.21 Backup

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.

A copy of current (editable) data, metadata and system configuration settings (variable settings which relate to a record or analytical run) maintained for the purpose of recovery including disaster recovery before archiving.

Backup and recovery processes for electronic data should be periodically tested. Each back up should be verified to ensure that it has functioned correctly e.g. by confirming that the data size transferred matches that of the original record.

The backup procedures should be described in an SOP, and backup activities should be documented and signed and dated by the responsible personnel.

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

7.22 Archive

A designated area or facility (e.g. cabinet, room, building or computerised system) for the secure storage and retention of records and materials.

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, indexing and expedient 'retrieval' as non-electronic data.

Archived records may be the original record or a 'true copy' and should be protected such that they cannot be altered or deleted without detection.

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

When legacy systems can no longer be supported, consideration should be given to maintaining the software for data accessibility purposes as long as reasonably practicable. This may be achieved by maintaining software in a virtual environment.

Migration to an alternative file format which retains as much as possible of the 'true copy' attributes of the data may be necessary with increasing age of the legacy data. Where migration with full original data 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. See also 'data transfer/migration'.

(see also section 3.11 of OECD Document 17)

7.23 File Structure

The way in which most of the 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. There is an inherently greater data integrity risk with flat files when compared to data contained within a relational database in that they are easier to manipulate and delete as a single file.

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.

7.23.1 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. The resulting file contains records but limited metadata.

Flat files may carry basic metadata relating to file creation and date of the last amendment but may not provide an audit trail. When creating flat file reports from electronic data the metadata and audit trails relating to the generation of the raw data may be lost, unless these are retained as a 'true copy'.

In terms of systems with flat files, the risks can be reduced if the flat file can be encrypted and/or the system can use a technical safeguard to detect if the file was modified outside the application. Additional security controls can be implemented to limit access to the flat file location (e.g. file repository in a server with control access or block the access to the operating system level in standalone systems).

Consideration also needs to be given to the 'dynamic' nature of the data, where appropriate (see 'original and true copy' definition).

7.23.2 Relational Database

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

Relational databases are inherently more secure than a flat file, as the data are held in a file format which preserves the relationship between data and metadata.

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

Amendments to data should not be performed via the 'back door' i.e. directly into the database fields but should be via the originator software package wherever possible so that appropriate audit trail entries and controls remain in place. If a 'back end' data change is required then this should be carefully controlled, well documented and have TFM approval.

Access rights for database entry or amendment should be controlled, and consistent with the requirements for computerised system user access/system administrator roles.