Annex 4

Guideline on data integrity

This document replaces the WHO *Guidance on good data and record management practices* (Annex 5, WHO Technical Report Series, No. 996, 2016) (*1*).

1. Introduction and background 137
2. Scope 137
3. Glossary 138
4. Data governance 140
5. Quality risk management 144
6. Management review 145
7. Outsourcing 146
8. Training 146
9. Data, data transfer and data processing 147
10. Good documentation practices 148
11. Computerized systems 149
12. Data review and approval 152
13. Corrective and preventive actions 152
References 153
Further reading 153
Appendix 1 Examples in data integrity management 155

1. Introduction and background

1.1. In recent years, the number of observations made regarding the integrity of data, documentation and record management practices during inspections of good manufacturing practice (GMP) (*2*), good clinical practice (GCP), good laboratory practice (GLP) and Good Trade and Distribution Practices (GTDP) have been increasing. The possible causes for this may include (i) reliance on inadequate human practices; (ii) poorly defined procedures; (iii) resource constraints; (iv) the use of computerized systems that are not capable of meeting regulatory requirements or are inappropriately managed and validated (*3*, *4*); (v) inappropriate and inadequate control of data flow; and (vi) failure to adequately review and manage original data and records.

1.2. Data governance and related measures should be part of a quality system, and are important to ensure the reliability of data and records in good practice (GxP) activities and regulatory submissions. The data and records should be 'attributable, legible, contemporaneous, original' and accurate, complete, consistent, enduring, and available; commonly referred to as "ALCOA+".

1.3. This document replaces the WHO *Guidance on good data and record management practices* (Annex 5, WHO Technical Report Series, No. 996, 2016) (*1*).

2. Scope

2.1. This document provides information, guidance and recommendations to strengthen data integrity in support of product quality, safety and efficacy. The aim is to ensure compliance with regulatory requirements in, for example clinical research, production and quality control, which ultimately contributes to patient safety. It covers electronic, paper and hybrid systems.

2.2. The guideline covers "GxP" for medical products. The principles could also be applied to other products such as vector control products.

2.3. The principles of this guideline also apply to contract givers and contract acceptors. Contract givers are ultimately responsible for the integrity of data provided to them by contract acceptors. Contract givers should therefore ensure that contract acceptors have the appropriate capabilities and comply with the principles contained in this guideline and documented in quality agreements.

2.4. Where possible, this guideline has been harmonised with other published documents on data integrity. This guideline should also be read with other WHO good practices guidelines and publications including, but not limited to, those listed in the references section of this document.

3. Glossary

The definitions given below apply to the terms used in these guidelines. They may have different meanings in other contexts.

ALCOA+. A commonly used acronym for "attributable, legible, contemporaneous, original and accurate" which puts additional emphasis on the attributes of being complete, consistent, enduring and available throughout the data life cycle for the defined retention period.

Archiving. Archiving is the process of long-term storage and protection of records from the possibility of deterioration, and being altered or deleted, throughout the required retention period. Archived records should include the complete data, for example, paper records, electronic records including associated metadata such as audit trails and electronic signatures. Within a GLP context, the archived records should be under the control of independent data management personnel throughout the required retention period.

Audit trail. The audit trail is a form of metadata containing information associated with actions that relate to the creation, modification or deletion of GxP records. An audit trail provides for a secure recording of 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.

Backup. The copying of live electronic data, at defined intervals, in a secure manner to ensure that the data are available for restoration.

Certified true copy or true copy. A 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.

Data. All original records and true 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 GMP activity and which

allow full and complete reconstruction and evaluation of the GMP activity. Data should be accurately recorded by permanent means at the time of the activity. Data may be contained in paper records (such as worksheets and logbooks), electronic records and audit trails, photographs, microfilm or microfiche, audio or video files or any other media whereby information related to GMP activities is recorded.

Data criticality. This is defined by the importance of the data for the quality and safety of the product and how important data are for a quality decision within production or quality control.

Data governance. The sum total of arrangements which provide assurance of data quality. These arrangements ensure that data, irrespective of the process, format or technology in which it is generated, recorded, processed, retained, retrieved and used will ensure an attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring and available record throughout the data life cycle.

Data integrity risk assessment (DIRA). The process to map out procedures, systems and other components that generate or obtain data; to identify and assess risks and implement appropriate controls to prevent or minimize lapses in the integrity of the data.

Data life cycle. All phases of the process by which data are created, recorded, processed, reviewed, analysed and reported, transferred, stored and retrieved and monitored, until retirement and disposal. There should be a planned approach to assessing, monitoring and managing the data and the risks to those data, in a manner commensurate with the potential impact on patient safety, product quality and/or the reliability of the decisions made throughout all phases of the data life cycle.

Dynamic data. Dynamic formats, such as electronic records, allow 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 or reviewer (with appropriate access permissions) to reprocess the data and expand the baseline to view the integration more clearly.

Electronic signatures. A signature in digital form (bio-metric or non-biometric) that represents the signatory. In legal terms, it is the equivalent of the handwritten signature of the signatory.

Good practices (GxP). An acronym for the group of good practice guides governing the preclinical, clinical, manufacturing, testing, storage, distribution

and post-market activities for regulated pharmaceuticals, biologicals and medical devices, such as GLP, GCP, GMP, good pharmacovigilance practices (GVP) and good distribution practices (GDP).

Hybrid system. The use of a combination of electronic systems and paper systems.

Medical product. A term that includes medicines, vaccines, diagnostics and medical devices.

Metadata. Metadata are data that provide the contextual information required to understand other data. These include structural and descriptive metadata, which describe the structure, data elements, interrelationships and other characteristics of data. They also permit data to be attributable to an individual. Metadata that are necessary to evaluate the meaning of data should be securely linked to the data and subject to adequate review. For example, in the measurement of weight, the number 8 is meaningless without metadata, such as, the unit, milligram, gram, kilogram, and so on. Other examples of metadata include the time or date stamp of an activity, the operator identification (ID) of the person who performed an activity, the instrument ID used, processing parameters, sequence files, audit trails and other data required to understand data and reconstruct activities.

Raw data. The original record (data) which can be described as the first-capture of information, whether recorded on paper or electronically. Raw data is synonymous with source data.

Static data. A static record format, such as a paper or electronic record, that is fixed and allows little or no interaction between the user and the record content. For example, once printed or converted to static electronic format chromatography records lose the capability of being reprocessed or enabling more detailed viewing of baseline.

4. Data governance

4.1. There should be a written policy on data integrity.

4.2. Senior management should be accountable for the implementation of systems and procedures in order to minimise the potential risk to data integrity, and to identify the residual risk using risk management techniques such as the principles of the guidance on quality risk management from WHO (*5*) and The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) (*6*).

4.3. Senior management is responsible for the establishment, implementation and control of an effective data governance system. Data governance should be embedded in the quality system. The necessary policies, procedures, training, monitoring and other systems should be implemented.

4.4. Data governance should ensure the application of ALCOA+ principles.

4.5. Senior management is responsible for providing the environment to establish, maintain and continually improve the quality culture, supporting the transparent and open reporting of deviations, errors or omissions and data integrity lapses at all levels of the organization. Appropriate, immediate action should be taken when falsification of data is identified. Significant lapses in data integrity that may impact patient safety, product quality or efficacy should be reported to the relevant medicine regulatory authorities.

4.6. The quality system, including documentation such as procedures and formats for recording and reviewing of data, should be appropriately designed and implemented in order to provide assurance that records and data meet the principles contained in this guideline.

4.7. Data governance should address the roles, responsibilities, accountability and define the segregation of duties throughout the life cycle and consider the design, operation and monitoring of processes/systems to comply with the principles of data integrity, including control over authorized and unauthorized changes to data.

4.8. Data governance control strategies using quality risk management (QRM) principles (5) are required to prevent or mitigate risks. The control strategy should aim to implement appropriate technical, organizational and procedural controls. Examples of controls may include, but are not limited to:

- the establishment and implementation of procedures that will facilitate compliance with data integrity requirements and expectations;
- the adoption of a quality culture within the company that encourages personnel to be transparent about failures, which includes a reporting mechanism inclusive of investigation and follow-up processes;
- the implementation of appropriate controls to eliminate or reduce risks to an acceptable level throughout the life cycle of the data;
- ensuring sufficient time and resources are available to implement and complete a data integrity programme; to monitor compliance

with data integrity policies, procedures and processes through e.g. audits and self-inspections; and to facilitate continuous improvement of both;

- the assignment of qualified and trained personnel and provision of regular training for personnel in, for example, GxP, and the principles of data integrity in computerized systems and manual/paper based systems;
- the implementation and validation of computerized systems appropriate for their intended use, including all relevant data integrity requirements in order to ensure that the computerized system has the necessary controls to protect the electronic data (*3*); and
- the definition and management of the appropriate roles and responsibilities for contract givers and contract acceptors, entered into quality agreements and contracts including a focus on data integrity requirements.

4.9. Data governance systems should include, for example:

- the creation of an appropriate working environment;
- active support of continual improvement in particular based on collecting feedback; and
- review of results, including the reporting of errors, unauthorized changes, omissions and undesirable results.

4.10. The data governance programme should include policies and procedures addressing data management. These should at least where applicable, include:

- management oversight and commitment;
- the application of QRM;
- compliance with data protection legislation and best practices;
- qualification and validation policies and procedures;
- change, incident and deviation management;
- data classification, confidentiality and privacy;
- security, cybersecurity, access and configuration control;
- database build, data collection, data review, blinded data, randomization;
- the tracking, trending, reporting of data integrity anomalies, and lapses or failures for further action;

- the prevention of commercial, political, financial and other organizational pressures;
- adequate resources and systems;
- workload and facilities to facilitate the right environment that supports DI and effective controls;
- monitoring;
- record-keeping;
- training; and
- awareness of the importance of data integrity, product quality and patient safety.

4.11. There should be a system for the regular review of data for consistency with ALCOA+ principles. This includes paper records and electronic records in day-to-day work, system and facility audits and self-inspections.

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

4.13. Where weaknesses in data integrity are identified, the appropriate corrective and preventive actions (CAPA) should be implemented across all relevant activities and systems and not in isolation.

4.14. Changing from paper-based systems to automated or computerised systems (or vice-versa) will not in itself remove the need for appropriate data integrity controls.

4.15. Records (paper and electronic) should be kept in a manner that ensures compliance with the principles of this guideline. These include but are not limited to:

- ensuring time accuracy of the system generating the record, accurately configuring and verifying time zone and time synchronisation, and restricting the ability to change dates, time zones and times for recording events;
- using controlled documents and forms for recording GxP data;
- defining access and privilege rights to GxP automated and computerized systems, ensuring segregation of duties;
- ensuring audit trail activation for all interactions and restricting the ability to enable or disable audit trails (Note: 'back-end' changes and 'hard' changes, such as hard deletes, should not be allowed). Where audit trials can be disabled then this this action should also appear in the audit trail;

- having automated data capture systems and printers connected to equipment and instruments in production (such as Supervisory Control and Data Acquisition (SCADA), Human Machine Interface (HMI) and Programme Logic Control (PLCs) systems), in , quality control, and in clinical research (such as Clinical Data Management (CDM) systems), where possible;
- designing processes in a way to avoid the unnecessary transcription of data or unnecessary conversion from paper to electronic and vice versa; and
- ensuring the proximity of an official GxP time source to site of GxP activity and record creation.

4.16. Systems, procedures and methodology used to record and store data should be periodically reviewed for effectiveness. These should be updated throughout the data life cycle, as necessary, where new technology becomes available. New technology implementation must be evaluated before implementation to verify the impact on data integrity.

5. Quality risk management

Note: documentation of data flows and data process maps are recommended to facilitate the assessment, mitigation and control of data integrity risks across the actual and intended data process(es).

5.1. Data Integrity Risk Assessment (DIRA) should be carried out in order to identify and assess areas of risk. This should cover systems and processes that produce data or, where data are obtained and inherent risks. The DIRAs should be risk-based, cover the life cycle of data and consider data criticality. Data criticality may be determined by considering how the data is used to influence the decisions made. The DIRAs should be documented and reviewed, as required, to ensure that it remains current.

5.2. The risk assessments should evaluate, for example, the relevant GxP computerised systems, supporting personnel, training, quality systems and outsourced activities.

5.3. DI risks should be assessed and mitigated. Controls and residual risks should be communicated. Risk review should be done throughout the document and data life cycle at a frequency based on the risk level, as determined by the risk assessment process.

5.4. Where the risk assessment has highlighted areas for remedial action, the prioritisation of actions (including the acceptance of an appropriate level of residual risk) and the prioritisation of controls should be documented and communicated. Where long-term remedial actions are identified, risk-reducing short-term measures should be implemented in order to provide acceptable data governance in the interim.

5.5. Controls identified may include organizational, procedural and technical controls such as procedures, processes, equipment, instruments and other systems in order to both prevent and detect situations that may impact on data integrity. Examples include the appropriate content and design of procedures, formats for recording, access control, the use of computerized systems and other means.

5.6. Efficient risk-based controls should be identified and implemented to address risks impacting data integrity. Risks include, for example, the deletion of, changes to and exclusion of data or results from data sets without written justification, authorisation where appropriate, and detection. The effectiveness of the controls should be verified (see Appendix 1 for examples).

6. Management review

6.1. Management should ensure that systems (such as computerized systems and paper systems) are meeting regulatory requirements in order to support data integrity compliance.

6.2. The acquisition of non-compliant computerized systems and software should be avoided. Where existing systems do not meet current requirements, appropriate controls should be identified and implemented based on risk assessment.

6.3. The effectiveness of the controls implemented should be evaluated through, for example:

- the tracking and trending of data;
- a review of data, metadata and audit trails (e.g. in warehouse and material management, production, quality control, case report forms and data processing); and
- routine audits and/or self-inspections, including data integrity and computerized systems.

7. Outsourcing

7.1. The selection of a contract acceptor should be done in accordance with an authorized procedure. The outsourcing of activities, ownership of data, and responsibilities of each party (contract giver and contract accepter) should be clearly described in written agreements. Specific attention should be given to ensuring compliance with data integrity requirements. Provisions should be made for responsibilities relating to data when an agreement expires.

7.2. Compliance with the principles and responsibilities should be verified during periodic site audits. This should include the review of procedures and data (including raw data and metadata, paper records, electronic data, audit trails and other related data) held by the relevant contract accepter identified in risk assessment.

7.3. Where data and document retention are contracted to a third party, particular attention should be given to security, transfer, storage, access and restoration of data held under that agreement, as well as controls to ensure the integrity of data over their life cycle. This includes static data and dynamic data. Mechanisms, procedures and tools should be identified to ensure data integrity and data confidentiality, for example, version control, access control, and encryption.

7.4. GxP activities, including outsourcing of data management, should not be sub-contracted to a third party without the prior approval of the contract giver. This should be stated in the contractual agreements.

7.5. All contracted parties should be aware of the requirements relating to data governance, data integrity and data management.

8. Training

8.1. All personnel who interact with GxP data and who perform GxP activities should be trained in relevant data integrity principles and abide by organization policies and procedures. This should include understanding the potential consequences in cases of non-compliance.

8.2. Personnel should be trained in good documentation practices and measures to prevent and detect data integrity issues.

8.3. Specific training should be given in cases where computerized systems are used in the generation, processing, interpretation and reporting of data and

where risk assessment has shown that this is required to relevant personnel. Such training should include validation of computerized systems and for example, system security assessment, back-up, restoration, disaster recovery, change and configuration management, and reviewing of electronic data and metadata, such as audit trails and logs, for each GxP computerized systems used in the generation, processing and reporting of data.

9. Data, data transfer and data processing

9.1. Data may be recorded on paper or captured electronically by using equipment and instruments including those linked to computerised systems. A combination of paper and electronic formats may also be used, referred to as a "hybrid system".

9.2. Data integrity consideration are also applicable to media such as photographs, videos, DVDs, imagery and thin layer chromatography plates. There should be a documented rationale for the selection of such a method.

9.3. Risk-reducing measures such as scribes, second person oversight, verification and checks should be implemented where there is difficulty in accurately and contemporaneously recording data related to critical process parameters or critical quality attributes.

9.4. Results and data sets require independent verification if deemed necessary from the DIRA or by another requirement.

9.5. Programmes and methods (such as processing methods in sample analysis (see also Good Chromatography Practices, TRS 1025) should ensure that data meet ALCOA+ principles. Where results or data are processed using a different method/parameters, then each version of the processing method should be recorded. Data records, content versions together with audit trails containing the required details should allow for reconstruction of all data processing in GxP computerized systems over the data life cycle.

9.6. Data transfer/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, transfer and subsequent storage. The challenges of migrating data are often underestimated, particularly regarding maintaining the full meaning of the migrated records.

Data transfer should be validated. The data should not be altered during or after it is transferred to the worksheet or other application. There should be an audit trail for this process. The appropriate quality procedures should be followed if the data transfer during the operation has not occurred correctly. Any changes in the middle layer software should be managed through the appropriate Quality Management Systems (*7*).

10. Good documentation practices

Note: The principles contained in this section are applicable to paper data.

10.1. Good documentation practices should be implemented and enforced to ensure compliance with ALCOA+ principles.

10.2. Data and recorded media should be durable. Ink should be indelible. Temperature-sensitive or photosensitive inks and other erasable inks should not be used. Where related risks are identified, means should be identified in order to ensure traceability of the data over their life cycle.

10.3. Paper should not be temperature-sensitive, photosensitive or easily oxidizable. If this is not feasible or limited, then true or certified copies should be generated.

10.4. Specific controls should be implemented in order to ensure the integrity of raw data and results recorded on paper records. These may include, but are not limited to:

- control over the issuance and use of loose paper sheets at the time of recording data;
- no use of pencil or erasers;
- use of single-line cross-outs to record changes with the identifiable person who made the change, date and reason for the change recorded (i.e. the paper equivalent to an electronic audit trail);
- no use of correction fluid or otherwise, obscuring the original record;
- controlled issuance of bound, paginated notebooks;
- controlled issuance and reconciliation of sequentially numbered copies of blank forms with authenticity controls;
- maintaining a signature and initial record for traceability and defining the levels of signature of a record; and
- archival of records by designated personnel in secure and controlled archives.

11. Computerized systems

(Note. This section highlights some specific aspects relating to the use of computerized systems. It is not intended to repeat the information presented in the other WHO guidelines here, such as the WHO Guideline on computerized systems (3), WHO Guideline on validation (2) and WHO Guideline on good chromatography practices (7). See references.)

11.1. Each computerized system selected should be suitable, validated for its intended use, and maintained in a validated state.

11.2. Where GxP systems are used to acquire, record, transfer, store or process data, management should have appropriate knowledge of the risks that the system and users may pose to the integrity of the data.

11.3. Software of computerized systems, used with GxP instruments and equipment, should be appropriately configured (where required) and validated. The validation should address for example the design, implementation and maintenance of controls in order to ensure the integrity of manually and automatically acquired data; ensure that Good Documentation Practices will be implemented; and that data integrity risks will be appropriately managed throughout the data life cycle. The potential for unauthorized and adverse manipulation of data during the life cycle of the data should be mitigated and, where possible, eliminated.

11.4. Where electronic instruments (e.g. certain pH meters, balances and thermometers) or systems with no configurable software and no electronic data retention are used, controls should be put in place to prevent the adverse manipulation of data and to prevent repeat testing to achieve the desired result.

11.5. Appropriate controls for the detection of lapses in data integrity principles should be in place. Technical controls should be used whenever possible but additional procedural or administrative controls should be implemented to manage aspects of computerised system control where technical controls are missing. For example, when stand-alone computerized systems with a user-configurable output are used, Fourier-transform infrared spectroscopy (FTIR) and UV spectrophotometers have user-configurable output or reports that cannot be controlled using technical controls. Other examples of non-technical detection and prevention mechanisms may include, but are not limited to, instrument usage logbooks and electronic audit trails.

Access and privileges

11.6. There should be a documented system in place that defines the access and privileges of users of systems. There should be no discrepancy between paper records and electronic records where paper systems are used to request changes for the creation and inactivation of users. Inactivated users should be retained in the system. A list of active and inactivated users should be maintained throughout the system life cycle.

11.7. Access and privileges should be in accordance with the role and responsibility of the individual with the appropriate controls to ensure data integrity (e.g. no modification, deletion or creation of data outside the defined privilege and in accordance with the authorized procedures defining review and approval where appropriate).

11.8. A limited number of personnel, with no conflict of interest in data, should be appointed as system administrators. Certain privileges such as data deletion, database amendment or system configuration changes should not be assigned to administrators without justification – and such activities should only be done with documented evidence of authorization by another responsible person. Records should be maintained and audit trails should be enabled in order to track activities of system administrators. As a minimum, activity logging for such accounts and the review of logs by designated roles should be conducted in order to ensure appropriate oversight.

11.9. For systems generating, amending or storing GxP data, shared logins or generic user access should not be used. The computerised system design should support individual user access. Where a computerised system supports only a single user login or limited numbers of user logins and no suitable alternative computerised system is available, equivalent control should be provided by third-party software or a paper-based method that provides traceability (with version control). The suitability of alternative systems should be justified and documented (*8*). The use of legacy hybrid systems should be discouraged and a priority timeline for replacement should be established.

Audit trail

11.10. GxP systems should provide for the retention of audit trails. Audit trails should reflect, for example, users, dates, times, original data and results, changes and reasons for changes (when required to be recorded), and enabling and disenabling of audit trails.

11.11. All GxP relevant audit trails should be enabled when software is installed and remain enabled at all times. There should be evidence of enabling the audit trail. There should be periodic verification to ensure that the audit trail remains enabled throughout the data life cycle.

11.12. Where a system cannot support ALCOA+ principles by design (e.g. legacy systems with no audit trail), mitigation measures should be taken for defined temporary periods. For example, add-on software or paper-based controls may be used. The suitability of alternative systems should be justified and documented. This should be addressed within defined timelines.

Electronic signatures

11.13. Each electronic signature should be appropriately controlled by, for example, senior management. An electronic signature should be:

- attributable to an individual;
- free from alteration and manipulation
- be permanently linked to their respective record; and
- date- and time-stamped.

11.14. An inserted image of a signature or a footnote indicating that the document has been electronically signed is not adequate unless it was created as part of the validated electronic signature process. The metadata associated with the signature should be retained.

Data backup, retention and restoration

11.15. Data should be retained (archived) in accordance with written policies and procedures, and in such a manner that they are protected, enduring, readily retrievable and remain readable throughout the records retention period. True copies of original records may be retained in place of the original record, where justified. Electronic data should be backed up according to written procedures.

11.16. Data and records, including backup data, should be kept under conditions which provide appropriate protection from deterioration. Access to such storage areas should be controlled and should be accessible only by authorized personnel.

11.17. Data retention periods should be defined in authorized procedures.

11.18. The decision for and manner in which data and records are destroyed, should be described in written procedures. Records for the destruction should be maintained.

11.19. Backup and restoration processes should be validated. The backup should be done routinely and periodically be restored and verified for completeness and accuracy of data and metadata. Where any discrepancies are identified, they should be investigated and appropriate action taken.

12. Data review and approval

12.2. There should be a documented procedure for the routine and periodic review, as well as the approval of data. Personnel with appropriate knowledge and experience should be responsible for reviewing and checking data. They should have access to original electronic data and metadata.

12.3. The routine review of GxP data and meta data should include audit trails. Factors such as criticality of the system (high impact versus low impact) and category of audit trail information (e.g. batch specific, administrative, system activities, and so on) should be considered when determining the frequency of the audit trail review.

12.4. A procedure should describe the actions to be taken where errors, discrepancies or omissions are identified in order to ensure that the appropriate corrective and preventive actions are taken.

12.5. Evidence of the review should be maintained.

12.6. A conclusion, where required, following the review of original data, metadata and audit trail records should be documented, signed and dated.

13. Corrective and preventive actions

13.1. Where organizations use computerized systems (e.g. for GxP data acquisition, processing, interpretation, reporting) which do not meet current GxP requirements, an action plan towards upgrading such systems should be documented and implemented in order to ensure compliance with current GxP.

13.2. When lapses in GxP relevant data regarding data integrity are identified, a risk-based approach may be used to determine the scope of the

investigation, root cause, impact and CAPA, as appropriate. Health authorities, contract givers and other relevant organizations should be notified if the investigation identifies a significant impact or risk to, for example, materials, products, patients, reported information or data in application dossiers, and clinical trials.

References

1. Guidelines on good manufacturing practices for pharmaceutical products: main principle. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: forty-eighth report. Geneva: World Health Organization; 2013: Annex 2 (WHO Technical Report Series, No. 986; https://www.who.int/medicines/areas/quality_safety/quality_assurance/TRS986annex2.pdf?ua=1, accessed 4 May 2020).
2. Good manufacturing practices: guidelines on validation. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations; fifty-third report. Geneva: World Health Organization; 2019: Annex 3 (WHO Technical Report Series, No. 1019; http://digicollection.org/whoqapharm/documents/s23430en/s23430en.pdf, accessed 5 May 2020).
3. Good manufacturing practices: guidelines on validation. Appendix 5. Validation of computerized systems. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: fifty-third report. Geneva: World Health Organization; 2019: Annex 3 (WHO Technical Report Series, No. 1019; https://www.who.int/medicines/areas/quality_safety/quality_assurance/WHO_TRS_1019_Annex3.pdf?ua=1, accessed 4 May 2020).
4. Guidelines on quality risk management. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: forty-seventh report. Geneva: World Health Organization; 2013: Annex 2 (WHO Technical Report Series, No. 981; https://www.who.int/medicines/areas/quality_safety/quality_assurance/Annex2TRS-981.pdf, accessed 4 May 2020).
5. ICH harmonised tripartite guideline. Quality risk management Q9. Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceutical for Human Use; 2005 (https://database.ich.org/sites/default/files/Q9%20Guideline.pdf, accessed 12 June 2020).
6. Good chromatography practices. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: fifty-fourth report. Geneva: World Health Organization; 2020: Annex 4 (WHO Technical Report Series, No. 1025; https://www.who.int/publications/i/item/978-92-4-000182-4, accessed 12 June 2020).
7. MHRA GxP data integrity guidance and definitions; Revision 1: Medicines & Healthcare Products Regulatory Agency (MHRA), London, March 2018 (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/687246/MHRA_GxP_data_integrity_guide_March_edited_Final.pdf, accessed 12 June 2020).

Further reading

- Data integrity and compliance with CGMP guidance for industry: questions and answers guidance for industry. U.S. Department of Health and Human Services, Food and Drug Administration; 2016 (https://www.fda.gov/files/drugs/published/Data-Integrity-and-Compliance-With-Current-Good-Manufacturing-Practice-Guidance-for-Industry.pdf, accessed 15 June 2020).

- Good Practices for data management and integrity in regulated GMP/GDP environments. Pharmaceutical Inspection Convention and Pharmaceutical Inspection Co-operation Scheme (PIC/S), November 2018 (https://picscheme.org/layout/document.php?id=1567, accessed 15 June 2020).
- Baseline guide Vol 7: risk-based manufacture of pharma products; 2nd edition.
- ISPE Baseline® Guide, July 2017. ISPEGAMP® guide: records and data integrity; March 2017.
- Data integrity management system for pharmaceutical laboratories PDA Technical Report, No. 80; August 2018.
- ICH harmonised tripartite guideline. Pharmaceutical Quality System Q10. Geneva: International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceutical for Human Use; 2008 (https://database.ich.org/sites/default/files/Q10%20Guideline.pdf, accessed 2 October 2020).

Appendix 1

Examples in data integrity management

This Appendix reflects on some examples in data integrity management in order to support the main text on data integrity. It should be noted that these are examples and are intended for the purpose of clarification only.

Example 1: Quality risk management and data integrity risk assessment

Risk management is an important part of good practices (GxP). Risks should be identified and assessed and controls identified and implemented in order to assist manufacturers in preventing possible DI lapses.

As an example, a Failure Mode and Effects Analysis (FMEA) model (or any other tool) can be used to identify and assess the risks relating to any system where data are, for example, acquired, processed, recorded, saved and archived. The risk assessment can be done as a prospective exercise or retrospective exercise. Corrective and preventive action (CAPA) should be identified, implemented and assessed for its effectiveness.

For example, if during the weighing of a sample, the entry of the date was not contemporaneously recorded on the worksheet but the date is available on the print-out from a weighing balance and log book for the balance for that particular activity. The fact that the date was not recorded on the worksheet may be considered a lapse in data integrity expectations. When assessing the risk relating to the lack of the date in the data, the risk may be considered different (lower) in this case as opposed to a situation when there is no other means of traceability for the activity (e.g. no print-out from the balance). When assessing the risk relating to the lapse in data integrity, the severity could be classified as "low" (the data is available on the print-out); it does not happen on a regular basis (occurrence is "low"), and it could easily be detected by the reviewer (detection is "high") – therefore the overall risk factor may be considered low. The root cause as to why the record was not made in the analytical report at the time of weighing should still be identified and the appropriate action taken to prevent this from happening again.

Example 2: Good documentation practices in data integrity

Documentation should be managed with care. These should be appropriately designed in order to assist in eliminating erroneous entries, manipulation and human error.

Formats

Design formats to enable personnel to record or enter the correct information contemporaneously. Provision should be made for entries such as, but not limited to, dates, times (start and finish time, where appropriate), signatures, initials, results, batch numbers and equipment identification numbers. When a computerized system is used, the system should prompt the personnel to make the entries at the appropriate step.

Blank sheets of paper

The use of blank sheets should not be encouraged. Where blank sheets are used (e.g. to supplement worksheets, laboratory notebooks and master production and control records), the appropriate controls have to be in place and may include, for example, a numbered set of blank sheets issued which are reconciled upon completion. Similarly, bound paginated notebooks, stamped or formally issued by designated personnel, allow for the detection of unofficial notebooks and any gaps in notebook pages. Authorization may include two or three signatures with dates, for example, "prepared by" or "entered by", "reviewed by" and "approved by".

Error in recording data

Care should be taken when entries of data and results (electronic and paper records) are made. Entries should be made in compliance with good documentation practices. Where incorrect information had been recorded, this may be corrected provided that the reason for the error is documented, the original entry remains readable and the correction is signed and dated.

Example 3: Data entry

Data entry includes for example sample receiving registration, sample analysis result recording, logbook entries, registers, batch manufacturing record entries and information in case report forms. The recording of source data on paper records should be done using indelible ink, in a way that is complete, accurate, traceable, attributable and free from errors. Direct entry into electronic records should be done by responsible and appropriately trained individuals. Entries should be traceable to an individual (in electronic records, thus having an individual user access) and traceable to the date (and time, where relevant). Where appropriate, the entry should be verified by a second person or entered through technical means such as the scanning of bar-codes, where possible, for the intended use of these data. Additional controls may include the locking of critical data entries after the data are verified and a review of audit trails for

critical data to detect if they have been altered. The manual entry of data from a paper record into a computerized system should be traceable to the paper records used which are kept as original data.

Example 4: Dataset

All data should be included in the dataset unless there is a documented, justifiable, scientific explanation and procedure for the exclusion of any result or data. Whenever out of specification or out of trend or atypical results are obtained, they should be investigated in accordance with written procedures. This includes investigating and determining CAPA for invalid runs, failures, repeats and other atypical data. The review of original electronic data should include checks of all locations where data may have been stored, including locations where voided, deleted, invalid or rejected data may have been stored. Data and metadata related to a particular test or product should be recorded together. The data should be appropriately stored in designated folders. The data should not be stored in other electronic folders or in other operating system logs. Electronic data should be archived in accordance with a standard operating procedure. It is important to ensure that associated metadata are archived with the relevant data set or securely traceable to the data set through relevant documentation. It should be possible to successfully retrieve all required data and metadata from the archives. The retrieval and verification should be done at defined intervals and in accordance with an authorized procedure.

Example 5: Legible and enduring

Data and metadata should be readable during the life cycle of the data. Electronic data are normally only legible/readable through the original software application that created it. In addition, there may be restrictions around the version of a software application that can read the data. When storing data electronically, ensure that any restrictions which may apply and the ability to read the electronic data are understood. Clarification from software vendors should be sought before performing any upgrade, or when switching to an alternative application, to ensure that data previously created will be readable.

Other risks include the fading of microfilm records, the decreasing readability of the coatings of optical media such as compact disks (CDs) and digital versatile/video disks (DVDs), and the fact that these media may become brittle.

Similarly, historical data stored on magnetic media will also become unreadable over time as a result of deterioration. Data and records should be stored in an appropriate manner, under the appropriate conditions.

Example 6: Attributable

Data should be attributable, thus being traceable to an individual and where relevant, the measurement system. In paper records, this could be done through the use of initials, full handwritten signature or a controlled personal seal. In electronic records, this could be done through the use of unique user logons that link the user to actions that create, modify or delete data; or unique electronic signatures which can be either biometric or non-biometric. An audit trail should capture user identification (ID), date and time stamps and the electronic signature should be securely and permanently linked to the signed record.

Example 7: Contemporaneous

Personnel should record data and information at the time these are generated and acquired. For example, when a sample is weighed or prepared, the weight of the sample (date, time, name of the person, balance identification number) should be recorded at that time and not before or at a later stage. In the case of electronic data, these should be automatically date- and time-stamped. In case hybrid systems are to be used, including the use for an interim period, the potential and criticality of system breaches should be covered in the assessment with documented mitigating controls in place. (The replacement of hybrid systems should be a priority with a documented CAPA plan.) The use of a scribe to record an activity on behalf of another operator should be considered only on an exceptional basis and should only take place where, for example, the act of recording places the product or activity at risk, such as, documenting line interventions by aseptic area operators. It needs to be clearly documented when a scribe has been applied.

> *"In these situations, the recording by the second person should be contemporaneous with the task being performed, and the records should identify both the person performing the task and the person completing the record. The person 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 that specifies the activities to which the process applies.*" (Extract taken from the Medicines & Healthcare Products Regulatory Agency (MHRA) *GxP data integrity guidance and definitions* (*10*).)

A record of employees indicating, their name, signature, initials or other mark or seal used should be maintained to enable traceability and to uniquely identify them and the respective action.

Example 8: Changes

When changes are made to any GxP result or data, the change should be traceable to the person who made the change as well as the date, time and reason for the change. The original value should not be obscured. In electronic systems, this traceability should be documented via computer generated audit trails or in other metadata fields or system features that meet these requirements. Where an existing computerized system lacks computer-generated audit trails, personnel may use alternative means such as procedurally controlled use of log-books, change control, record version control or other combinations of paper and electronic records to meet GxP regulatory expectations for traceability to document the what, who, when and why of an action.

Example 9: Original

The first or source capture of data or information and all subsequent data required to fully reconstruct the conduct of the GxP activity should be available. In some cases, the electronic data (electronic chromatogram acquired through high-performance liquid chromatography (HPLC)) may be the first source of data and, in other cases, the recording of the temperature on a log sheet in a room – by reading the value on a data logger. This data should be reviewed according to the criticality and risk assessment.

Example 10: Controls

Based on the outcome of risk assessment which should cover all areas of data governance and data management, appropriate and effective controls should be identified and implemented in order to assure that all data, whether in paper records or electronic records, will meet GxP requirements and ALCOA+ principles. Examples of controls may include, but are not limited to:

- the qualification, calibration and maintenance of equipment, such as balances and pH meters, that generate printouts;
- the validation of computerized systems that acquire, process, generate, maintain, distribute, store or archive electronic records;
- review and auditing of activities to ensure that these comply with applicable GxP data integrity requirements;
- the validation of systems and their interfaces to ensure that the integrity of data will remain while transferring between/among computerized systems;
- evaluation to ensure that computerized systems remain in a validated state;
- the validation of analytical procedures;

- the validation of production processes;
- a review of GxP records;
- ensuring effective review and oversight of the Batch Release Systems and processes by using different oversight and review techniques to ensure that data have not changed since the original entry; and
- the investigation of deviations, out of trend and out of specifications results.

Example 11: Accuracy

Points to consider for assuring accurate GxP records:

- the entry of critical data into a computer by an authorized person (e.g. entry of a master processing formula) requires an additional check on the accuracy of the data entered manually. This check may be done by independent verification and release for use by a second authorized person or by validated electronic means. For example, to detect and manage risks associated with critical data, procedures would require verification by a second person;
- validation and control over formulae for calculations including electronic data capture systems;
- ensuring correct entries into the laboratory information management system (LIMS) such as fields for specification ranges;
- other critical master data, as appropriate. Once verified, these critical data fields should normally be locked in order to prevent further modification and only be modified through a formal change control process;
- the process of data transfer between systems should be validated;
- the migration of data including planned testing, control and validation; and
- when the activity is time-critical, printed records should display the date and time stamp.