World Health Organization

DRAFT WORKING DOCUMENT FOR COMMENTS:

Guideline on data integrity

Please send your comments to **Dr Sabine Kopp**, Team Lead, Norms and Standards for Pharmaceuticals, Technical Standards and Specifications (kopps@who.int), with a copy to Ms Claire Vogel (vogelc@who.int) before 15 August 2020. Please use our attached Comments Table for this purpose.

Our working documents are sent out electronically and they will also be placed on the WHO Medicines website (http://www.who.int/medicines/areas/quality_safety/quality_assurance/guidelines/en/) for comments under the "*Current projects*" link. If you wish to receive all our draft guidelines, please send your email address to jonessi@who.int and your name will be added to our electronic mailing list.

SCHEDULE FOR DRAFT WORKING DOCUMENT QAS/19.819/Rev.1

[*Note from Secretariat*: in view of COVID-19, the schedule had to be adapted as face-to-face meetings were postponed and/or replaced by virtual meetings]

Guideline on data integrity

Description of activity	Date
Preparation of the document following recommendation of the Fifty-fourth WHO Expert Committee on Specifications for Pharmaceutical Preparations (ECSPP).	October 2019
Mailing of working document inviting comments, including to the Expert Advisory Panel on the International Pharmacopoeia and Pharmaceutical Preparations (EAP), and posting of the working document on the WHO website for public consultation.	November 2019– January 2020
Consolidation of comments received and review of feedback. Preparation of working document for discussion.	March – May 2019
Discussion of working document and feedback received during the public consultation and the informal Consultation on Good Practices for Health Products Manufacture and Inspection. In view of the logistical situation with regard to COVID-19 , the consultation was replaced by virtual meetings of a working group composed of inspectors from Brazil, China, India, Italy and South Africa, as well as UNICEF.	10 and 12 June 2020
Mailing of the revised working document inviting comments, including to the EAP, and posting the working document on the WHO website for the second round of public consultation.	June 2020
Consolidation of comments received and review of feedback. Preparation of working document for discussion.	September 2020
Presentation to the Fifty-fifth ECSPP meeting.	12-16 October 2020
Any other follow-up action as required.	

Guideline on data integrity

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

1. Introduction and background

1.1. Data governance and its related measures 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, commonly referred to as "ALCOA+".

1.2. 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) and good laboratory practice (GLP) 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 to meet regulatory requirements or are inappropriately managed and validated *(3,4)*; (v) inappropriate data flow (e.g. manual data transfer); and (vi) failure to adequately review and manage original data and records.

1.3. Data governance control strategies using quality risk management principles *(5)* are required to mitigate such risks. Examples of controls may include, but are not limited to:

- the establishment and implementation of a data integrity (DI) policy;
- the establishment and implementation of procedures that will facilitate compliance with DI 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 application of quality risk management (QRM) with the identification of all areas of risk to DI through data integrity risk assessment (DIRA) and the implementation of appropriate controls to eliminate or reduce risks to an acceptable level throughout the life-cycle of the data;
- ensuring sufficient resources are available to implement and complete a DI program and to monitor compliance with DI policies and procedures and processes, and to facilitate continuous improvement of both;

- the provision of necessary training for personnel in, for example, GxP, computerized systems and the principles of DI;
- the implementation and validation of computerized systems appropriate for their intended use, including all relevant DI requirements in order to ensure that the computerized system has the necessary controls to protect the electronic data *(3)*;
- 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 DI requirements.

2. Scope

2.1. This guideline provides information, guidance and recommendations to facilitate compliance with regulatory requirements related to DI documentation and record management.

2.2. The scope of this guideline is designated as "GxP" for pharmaceutical products. The principles could also be applicable to vector control products.

2.3. Where possible, this guideline has been harmonised with other published documents. This guideline should also be read with other WHO good practices guidelines and publications.

2.4. The principles of this guideline 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 documented in quality agreements.

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 – implicit basic ALCOA principles.

archiving. Archiving is the process of protecting records from the possibility of being further altered or deleted, and storing these records under the control of independent data management personnel throughout the required retention period. Archived records should include, for example, associated metadata and electronic signatures.

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.

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 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 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.

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).

metadata. Metadata are data about data that provide the contextual information required to understand those data. These include structural and descriptive metadata. Such data describe the structure, data elements, interrelationships and other characteristics of data. They also permit data to be attributable to an individual. Metadata necessary to evaluate the meaning of data should be securely linked to the data and subject to adequate review. For example, in weighing, the number 8 is meaningless without metadata, such as, the unit, milligram, gram, kilogram, and so on. Other examples of metadata include the time/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).

4. Data governance

4.1. Senior management is responsible for the establishment, implementation and control of an effective quality system and a data governance system by assuring that policies, training and technical systems are in place.

4.2. 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 at all levels of the organization.

4.3. Senior management should be accountable for the implementation of systems and procedures in order to minimise the potential risk to DI, 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)*.

4.4. There should be a written DI policy.

4.5. Data should be attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring and available. This is generally referred to as ALCOA+.

4.6. The quality system, including documentation such as procedures and formats for recording 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 data roles, responsibilities and accountability throughout the life cycle and consider the design, operation and monitoring of processes/systems to comply with the principles of DI, including control over intentional and unintentional changes to data.

4.8. Data governance systems should include e.g.:

- training in the importance of DI principles;

- the creation of an appropriate working environment;
- active encouragement of collecting feedback and continuous improvement; and
- the reporting of errors, unauthorized changes, omissions and undesirable results.

4.9. The data governance programme should include policies and procedures addressing data management. Elements of effective management governance should at least include:

- management oversight and commitment;
- the application of QRM;
- quality metrics and performance indicators;
- validation;
- change, incident and deviation management;
- security, cybersecurity, access and configuration control;
- database build, data collection, data review, blinded data, randomization;
- the tracking, trending, reporting of DI anomalies, and lapses or failures for further action;
- the prevention of commercial, political, financial and other organizational pressures;
- adequate resources, 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 DI, product quality and patient safety.

4.10. There should be a system for the regular review of documents and 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.

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

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

4.13. Significant DI lapses identified that may impact patient safety, product quality or efficacy, should be reported to the relevant medicine regulatory authorities.

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

4.15. Good documentation practices should be followed in order to ensure that all records are complete and in accordance with ALCOA+ principles.

4.16. 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:

- restricting the ability to change dates and times for recording events;
- using controlled documents and forms for recording GxP data;
- controlling the issuance of blank paper templates for data recording of GxP activities, with reconciliation and authenticity controls where required;
- defining access and privilege rights to automated systems, ensuring segregation of duties;
- enabling audit trails and restricting the ability to enable or disenable audit trails;
- having automated data capture systems and printers connected to equipment and instruments in production and quality control where possible;
- ensuring the proximity of printers to sites of relevant activities;
- design processes in a way to avoid the unnecessary transcription of data or unnecessary conversion from paper to electronic and vice versa; and
- ensuring access to original electronic data and metadata for personnel responsible for reviewing and checking data.

4.17. Systems, procedures and methodology used to record and store data should be periodically reviewed for effectiveness and updated, as necessary, in relation to new technology.

5. Quality risk management

5.1. The DIRA should be documented. This should cover systems and processes that produce data or, where data are obtained, data criticality and inherent risks.

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

5.3. DI risks should be assessed, mitigated, communicated and reviewed 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 DIRA has highlighted areas for remediation, 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 remediation 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 DI. Examples include the appropriate content and design of procedures, formats for recording, access control, the use of computerized systems and other means.

5.6. Controls should cover risks to data. Risks to data manipulation include deletion of, changes to, and exclusion of data or results from data sets without written justification, authorisation where appropriate, and detection.

5.7. In line with the current approach in GxP, this guideline recommends a documented risk-based approach over the life cycle of data considering data criticality. DIRA should be carried out in order to identify and assess areas of risk.

5.8. Efficient risk-based controls and the review of data and documents should be identified and implemented. The effectiveness of the controls should be verified.

6. Management review

6.1. There should be management oversight of quality metrics relevant to data governance.

6.2. Management should ensure that computerized systems are meeting regulatory requirements in order to ensure DI compliance and to avoid the acquisition of inadequate systems and software.

6.3. The effectiveness of the controls implemented should be measured against the quality metrics and performance indicators. These should include, for example:

- the tracking and trending of data;
- A review of audit trails in, for example, production, quality control, GLP, case report forms and data processing; and
- routine audits and/or self-inspections, including DI and computerized systems.

7. Outsourcing

7.1. The outsourcing of activities 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 DI requirements.

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 contracted organization that are relevant to the contract giver's product or services.

7.3. Where data and document retention are contracted to a third party, particular attention should be paid to understanding the transfer, storage and restoration of data held under that agreement, as well as controls to ensure the integrity of data over their life cycle. This includes

data in motion and data at rest. Tools should be identified to ensure data integrity, for example, encryption.

7.4. No activity, including outsourcing of databases, should be sub-contracted to a third party without the prior approval of the contract giver. This should be stated in the contractual agreements where appropriate.

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

8. Training

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

8.2. Personnel should agree to abide by DI principles and should be made aware of the potential consequences in cases of non-compliance.

8.3. Personnel should be trained in good documentation practices and measures to prevent and detect DI issues. Specific training may be required in cases where computerized systems are used in the generation, processing, interpretation and reporting of data and where risk assessment has shown that this may be required. Such training should include, for example, evaluating the system security, back-up, configuration settings and reviewing of electronic data and metadata, such as audit trails and logs, for individual computerized systems used in the generation, processing and reporting of data.

9. Data and data transfer

9.1. Data may be recorded manually reflecting an observation, result or other data and information on paper, or electronically by using equipment and instruments including those linked to

computerised systems. A combination of manual and electronic systems may also be used, referred to as a "hybrid system".

9.2. The same considerations for DI apply to data sets such as photographs, videos, DVDs, imagery and chromatography plates. There should be a documented rationale for the selection of such a method.

9.3. Risk-reducing supervisory measures 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 acquisition and processing methods) should ensure that data meet ALCOA+ principles. Where results or data are processed using a different method/parameters, then the acquisition method should be recorded. Audit trails with the required details should allow for reconstruction of all data processing and administrative activities.

9.6. Data transfer should not result in any changes to the content or meaning of the data. The transfer should be tracked in the audit trail or by other suitable means.

9.7. Data transfer should be validated and computerized interfaces tested, especially systems which map and or transform data moving between computerized systems.

10. Good documentation practices

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

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, or other 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 available.

10.4. Specific controls should be implemented in order to ensure the integrity of 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;
- the use of permanent, indelible ink;
- no use of pencil or erasers;
- the use of single-line cross-outs to record changes with the identifiable person who made the change, date and reason 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 of sequentially numbered copies of blank forms with authenticity controls; 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. The computerized system selected should be suitable and validated for its intended use.

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. Suitably configured and validated, software should be used where instruments and equipment with computerised systems are used. The validation should cover the design, implementation

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

11.5. The appropriate controls of detection for lapses in DI principles should be in place. Technical controls should be used whenever possible. Additional controls should be implemented where stand-alone systems with a user-configurable output is used, for example, Fourier-transform infrared spectroscopy (FTIR) and UV spectrophotometers. Examples of detection and prevention mechanisms may include, but are not limited to, instrument usage logbooks, electronic audit trails, and external software to lockdown the personal computer workstation.

11.6. Critical records or data, including metadata, should be reviewed and retained according to risk assessment. Reduced effort and/or frequency should be justified.

Access and privileges

11.7. There should be a documented system in place that defines the access and privileges of users of computerized systems. There should be no discrepancy between paper records and electronic records, including the creation and inactivation of users.

11.8. Access and privileges should be in accordance with the role and responsibility of the individual with the appropriate controls to ensure DI (e.g. no modification, deletion or creation of data outside the allocated responsibility).

11.9. 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. Minimally, activity logging for such accounts and the review of logs by designated roles should be conducted in order to ensure appropriate oversight.

11.10. 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)*.

Audit trail

11.11. 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.

11.12. All 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 periodical verification that the audit trail remained enabled throughout the data life cycle.

11.13. 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.

11.14. Routine data review should include a review of audit trails. Evidence of the reviews should be maintained.

Electronic signatures

11.15. Each electronic signature should be appropriately controlled. An electronic signature should be:

- validated;

- attributable to an individual;
- free from alteration and manipulation; and
- date- and time-stamped, where appropriate.

11.16. 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 review and approval

11.17. There should be a documented procedure for the routine and periodic review, as well as the approval of data.

11.18. 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.

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

Data backup, retention and restoration

11.20. Data should be retained 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.

11.21. Data and records should be kept in a secure area which provides appropriate protection. Access should be controlled.

11.22. Retention periods should be defined in authorized procedures.

11.23. Records reflecting documented reasons for the destruction of data should be maintained.

559 11.24. Backup and restoration processes should be validated. The backup should be done and
560 periodically restored and verified for completeness and accuracy of data and metadata.
561 Where any discrepancies are identified, they should be investigated.
562

563 # 12. Corrective and preventive actions

564

565 12.1. Where organizations use computerized systems (e.g. for GxP data acquisition, processing,
566 interpretation, reporting) which do not meet current GxP requirements, a workplan towards
567 upgrading such systems should be documented and implemented in order to ensure
568 compliance with current GxP.
569

570 12.2. When GxP lapses in DI are identified, a risk-based approach may be used to determine the
571 scope of the investigation, root cause, impact and CAPA, as appropriate. Health authorities,
572 contract givers and other relevant organizations should be notified if the investigation identifies
573 a significant impact or risk to, for example, materials, products, patients, reported information
574 or data in application dossiers, and clinical trials.
575

576

References

1. Guidance on good data and record management practices. In: WHO Expert Committee on Specifications for Pharmaceutical Preparations: fiftieth report. Geneva: World Health Organization; 2016: Annex 5 (WHO Technical Report Series, No. 996; https://www.who.int/medicines/publications/pharmprep/WHO_TRS_996_annex05.pdf?ua=1, accessed 12 June 2020).

2. 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).

3. 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).

4. 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).

5. 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).

6. 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).

7. 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).

8. 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.

641 • Data integrity management system for pharmaceutical laboratories PDA Technical Report, No.
642 80; August 2018.
643
644
645
646
647

Annex 1. Examples in data integrity management

This Annex reflects on some examples in data integrity (DI) management in order to support the main text on DI. 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 manufacturing practices (GMP). 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 DI, 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 at the right time. Provision should be made for entries such as, but not limited to, dates, time (start, finish, where appropriate), signatures, initials, results, batch numbers and equipment identification numbers. The system should prompt the personnel to make the entries at the appropriate step.

Blank forms

The use of blank forms should not be encouraged. Where blank forms 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 forms issued which are reconciled upon completion. Similarly, bound paginated notebooks, stamped or formally issued by a 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 examples such as 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 in indelible ink 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 into a computerized system should be traceable to the paper records used.

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 should not be found 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 data and datasets from the archives. This includes metadata. This should be done in accordance with a procedure and verified at defined intervals.

Example 5: Legible and enduring

Data and metadata should be readable during the life cycle of the data. 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. 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 that captures user identification (ID), date and time stamps and the electronic signature must 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. The use of hybrid systems is discouraged but where legacy systems are awaiting replacement, upgrade or connection to upper level systems, documented mitigating controls should be 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)*.)

Example 8: Changes

When changes are made to any result or data, the change should be traceable to the person who made the change and the date, time and reason for the change. 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

Original data include the first or source capture of data or information and all subsequent data required to fully reconstruct the conduct of the GxP activity (*see the definition of raw data*). In some cases, the electronic data (electronic chromatogram acquired through high-performance liquid chromatography (HPLC)) may be the original 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 – may be considered the original data. Original data should be reviewed according to the criticality and risk assessment. Proof of review should be presented (e.g. as a signature (reviewed by:) and date of the review). For electronic records, this is typically signified by electronically signing the electronic data set that has been reviewed and approved. Written procedures for data review should clarify the meaning of the review and approval signatures in order to ensure that the personnel concerned understand their responsibility as reviewers and approvers to assure the integrity, accuracy, consistency and compliance with established standards of the electronic data and metadata subject to review and approval. Written procedures for data review should define the frequency, roles and responsibilities and approach to review of meaningful metadata, such as audit trails. These procedures should also describe how aberrant data are to be handled if found during the review. Personnel who conduct such reviews should have adequate and appropriate training in the review process as well as in the software systems containing the data subject to review.

Example 10: Controls

Based on the outcome of the data integrity risk assessment (DIRA) (which should cover all areas of data governance and data management), the 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 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 or archive electronic records;
- the validation of systems in order to ensure that the integrity of data will remain while transmitting between/among computerized systems;
- the validation of analytical procedures;
- the validation of production processes;
- a review of GxP records; and
- the investigation of deviations, out of trend and out of specifications results.

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;
- formulae for calculations entered into spreadsheets;
- master data entered into the laboratory information management system (LIMS) such as fields for specification ranges used to flag out of specification values on the certificate of analysis;
- 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.

***