# DATA INTEGRITY
# (EMA – AUGUST 2016)

**Date : ** 17AUG2016

## Data integrity

Data integrity enables good decision-making by pharmaceutical manufacturers and regulatory authorities.It is a fundamental requirement of the pharmaceutical quality system described in EU GMP chapter 1, applying equally to manual (paper) and electronic systems.

Promotion of a quality culture together with implementation of organisational and technical measures which ensure data integrity is the responsibility of senior management. It requires participation and commitment by staff at all levels within the company, by the company's suppliers and by its distributors.

Senior management should ensure that data integrity risk is assessed, mitigated and communicated in accordance with the principles of quality risk management. The effort and resource assigned to data integrity measures should be commensurate with the risk to product quality, and balanced with other quality assurance resource demands. Where long term measures are identified in order to achieve the desired state of control, interim measures should be implemented to mitigate risk, and should be monitored for effectiveness.

The following questions and answers describe foundational principles which facilitate successful implementation of existing guidance published by regulatory authorities participating in the PIC/S scheme. It should be read in conjunction with national guidance, medicines legislation and the GMP standards published in Eudralex volume 4.

The importance of data integrity to quality assurance and public health protection should be included in personnel training programmes.

WHO - Annex 5: guidance on good data and record management practices

## 1. How can data risk be assessed?

Data risk assessment should consider the vulnerability of data to involuntary or deliberate amendment, deletion or recreation. Control measures which prevent unauthorised activity and increase visibility / detectability can be used as risk mitigating actions.

Examples of factors which can increase risk of data integrity failure include complex, inconsistent processes with open-ended and subjective outcomes. Simple tasks which are consistent, well-defined and objective lead to reduced risk.

Risk assessment should include a business process focus (e.g. production, QC) and not just consider IT system functionality or complexity. Factors to consider include:

- Process complexity
- Process consistency, degree of automation /human interface
- Subjectivity of outcome / result
- Is the process open-ended or well defined

This ensures that manual interfaces with IT systems are considered in the risk assessment process. Computerised system validation in isolation may not result in low data integrity risk, in particular when the user is able to influence the reporting of data from the validated system.

## 2. How can data criticality be assessed?

The decision which data influences may differ in importance, and the impact of the data to a decision may also vary. Points to consider regarding data criticality include:

- What decision does the data influence?

For example: when making a batch release decision, data which determines compliance with critical quality attributes is of greater importance than warehouse cleaning records.

- What is the impact of the data to product quality or safety?

For example: for an oral tablet, active substance assay data is of greater impact to product quality and safety than tablet dimensions' data.

## 3. What does 'Data Lifecycle' refer to?

'Data lifecycle' refers to how data is generated, processed, reported, checked, used for decision-making, stored and finally discarded at the end of the retention period.

Data relating to a product or process may cross various boundaries within the lifecycle, for example:

- IT systems
  - ➢ Quality system applications
  - ➢ Production
  - ➢ Analytical
  - ➢ Stock management systems
  - ➢ Data storage (back-up and archival)
- Organisational
  - ➢ Internal (e.g. between production, QC and QA)
  - ➢ External (e.g. between contract givers and acceptors)
  - ➢ Cloud-based applications and storage

## 4. Why is 'Data lifecycle' management important to ensure effective data integrity measures?

Data integrity can be affected at any stage in the lifecycle. It is therefore important to understand the lifecycle elements for each type of data or record, and ensure controls which are proportionate to data criticality and risk at all stages.

## 5. What should be considered when reviewing the 'Data lifecycle'?

The 'Data lifecycle' refers to the:

- Generation and recording of data
- Processing into usable information
- Checking the completeness and accuracy of reported data and processed information

- Data (or results) are used to make a decision
- Retaining and retrieval of data which protects it from loss or unauthorised amendment
- Retiring or disposal of data in a controlled manner at the end of its life

'Data Lifecycle' reviews are applicable to both paper and electronic records, although control measures may be applied differently. In the case of computerised systems, the 'data lifecycle' review should be performed by business process owners (e.g. production, QC) in collaboration with IT personnel who understand the system architecture. The description of computerised systems required by EU GMP Annex 11 paragraph 4.3 can assist this review. The application of critical thinking skills is important to not only identify gaps in data governance, but to also challenge the effectiveness of the procedural and systematic controls in place.

Segregation of duties between data lifecycle stages provides safeguards against data integrity failure by reducing the opportunity for an individual to alter, mis-represent or falsify data without detection.

Data risk should be considered at each stage of the data lifecycle review.

## 6. 'Data lifecycle': What risks should be considered when assessing the generating and recording of data?

The following aspects should be considered when determining risk and control measures:

- How and where is original data created (i.e. paper or electronic)
- What metadata is associated with the data, to ensure a complete, accurate and traceable record, taking into account ALCOA principles. Does the record permit the reconstruction of the activity
- Where is the data and metadata located
- Does the system require that data is saved to permanent memory at the time of recording, or is it held in a temporary buffer

In the case of some computerised analytical and manufacturing equipment, data may be stored as a temporary local file prior to transfer to a permanent storage location (e.g. server). During the period of 'temporary' storage, there is often limited audit trail provision amending, deleting or recreating data. This is a data integrity risk. Removing the use of temporary memory (or reducing the time period that data is stored in temporary memory) reduces the risk of undetected data manipulation.

- Is it possible to recreate, amend or delete original data and metadata;

Controls over paper records are discussed elsewhere in this guidance.

Computerised system controls may be more complex, including setting of user privileges and system configuration to limit or prevent access to amend data. It is important to review all data access opportunities, including IT helpdesk staff, who may make changes at the request of the data user. These changes should be procedurally controlled, visible and approved within the quality system.

- How data is transferred to other locations or systems for processing or storage;

Data should be protected from possibility of intentional or unintentional loss or amendment during transfer to other systems (e.g. for processing, review or storage). Paper records should be protected from amendment, or substitution. Electronic interfaces should be validated to demonstrate security and no corruption of data, particularly where systems require an interface to present data in a different structure or file format.

Does the person processing the data have the ability to influence what data is reported, or how it is presented.

## 7. 'Data lifecycle': What risks should be considered when assessing the processing data into usable information?

The following aspects should be considered when determining risk and control measures:

- How is data processed;

Data processing methods should be approved, identifiable and version controlled. In the case of electronic data processing, methods should be locked where appropriate to prevent unauthorised amendment.

- How is data processing recorded;

The processing method should be recorded. In situations where raw data has been processed more than once, each iteration (including method and result) should be available to the data checker for verification.

- Does the person processing the data have the ability to influence what data is reported, or how it is presented;

Even 'validated systems' which do not permit the user to make any changes to data may be at risk if the user can choose what data is printed, reported or transferred for processing. This includes performing the activity multiple times as separate events and reporting a desired outcome from one of these repeats.

Data presentation (e.g. changing scale of graphical reports to enhance or reduce presentation of analytical peaks) can also influence decision making, and therefore impact data integrity.

## 8. 'Data lifecycle': What risks should be considered when checking the completeness and accuracy of reported data and processed information?

The following aspects should be considered when determining risk and control measures:

- Is original data (including the original data format) available for checking;

The format of the original data (electronic or paper) should be preserved, and available to the data reviewer in a manner which permits interaction with the data (e.g. search, query). This approach facilitates a risk-based review of the record, and can also reduce administrative burden for instance utilising validated audit trail 'exception reports' instead of an onerous line-by-line review.

- Are there any periods of time when data is not audit trailed;

This may present opportunity for data amendment which is not subsequently visible to the data reviewer. Additional control measures should be implemented to reduce risk of undisclosed data manipulation.

- Does the data reviewer have visibility and access to all data generated;

This should include any data from failed or aborted activities, discrepant or unusual data which has been excluded from processing or the final decision-making process. Visibility of all data provides protection against selective data reporting or 'testing into compliance'.

- Does the data reviewer have visibility and access to all processing of data;

This ensures that the final result obtained from raw data is based on good science, and that any data exclusion or changes to processing method is based on good science. Visibility of all processing information provides protection against undisclosed 'processing into compliance'.

## 9. 'Data lifecycle': What risks should be considered when data (or results) are used to make a decision?

The following aspects should be considered when determining risk and control measures:

- When is the pass / fail decision taken;

If data acceptability decisions are taken before a record (raw data or processed result) is saved to permanent memory, there may be opportunity for the user to manipulate data to provide a satisfactory result, without this change being visible in audit trail. This would not be visible to the data reviewer.

This is a particular consideration where computerised systems alert the user to an out of specification entry before the data entry process is complete (i.e. the user 'saves' the data entry), or saves the record in temporary memory.

## 10. 'Data lifecycle': What risks should be considered when retaining and retrieving data to protect it from loss or unauthorised amendment?

The following aspects should be considered when determining risk and control measures:

- How / where is data stored;

Storage of data (paper or electronic) should be at secure locations, with access limited to authorised persons. The storage location must provide adequate protection from damage due to water, fire, etc.

- What are the measures protecting against loss or unauthorised amendment;

Data security measures should be at least equivalent to those applied during the earlier Data lifecycle stages. Retrospective data amendment (e.g. via IT helpdesk or data base amendments) should be controlled by the pharmaceutical quality system, with appropriate segregation of duties and approval processes.

- Is data backed up in a manner permitting reconstruction of the activity;

Back-up arrangements should be validated to demonstrate the ability to restore data following IT system failure. In situations where metadata (including relevant operating system event logs) are stored in different file locations from raw data, the back-up process should be carefully designed to ensure that all data required to reconstruct a record is included.

Similarly, 'true copies' of paper records may be duplicated on paper, microfilm, or electronically, and stored in a separate location.

- What are ownership / retrieval arrangements, particularly considering outsourced activities or data storage;

A technical agreement should be in place which addresses the requirements of Part I Chapter 7 and Part II Section 16 of the GMP guide.

## 11. 'Data lifecycle': What risks should be considered when retiring or disposal of data in a controlled manner at the end of its life?

The following aspects should be considered when determining risk and control measures:

- The data retention period

This will be influenced by regulatory requirements and data criticality. When considering data for a single product, there may be different data retention needs for pivotal trial data and manufacturing process / analytical validation data compared to routine commercial batch data.

- How data disposal is authorised

Any disposal of data should be approved within the quality system and be performed in accordance with a procedure to ensure compliance with the required data retention period.

## 12. Is it required by the EU GMP to implement a specific procedure for data integrity?

There is no requirement for a specific procedure, however it may be beneficial to provide a summary document which outlines the organisations total approach to data governance.

A compliant pharmaceutical quality system generates and assesses a significant amount of data. While all data has an overall influence on GMP compliance, different data will have different levels of impact to product quality.

A quality-risk management (ICH Q9) approach to data integrity can be achieved by considering data risk and data criticality at each stage in the Data lifecycle. The effort applied to control measures should be commensurate with this data risk and criticality assessment.

The approach to risk identification, mitigation, review and communication should be iterative, and integrated into the pharmaceutical quality system. This should provide senior management supervision and permit a balance between data integrity and general GMP priorities in line with the principles of ICH Q9 & Q10.

### 13. How are the data integrity expectations (ALCOA) for the pharmaceutical industry prescribed in the existing EU GMP relating to active substances and dosage forms published in Eudralex volume 4?

The main regulatory expectation for data integrity is to comply with the requirement of ALCOA principles. The table below provide for each ALCOA principle the link to EU GMP references (Part I, Part II and Annex 11):

| | **Basic Requirements for Medicinal Products (Part I): Chapter 4[1] / Chapter 6[2]** | **Basic Requirements for Active Substances used as Starting Materials (Part II) : Chapter 6[3] / Chapter 5[4]** | **Annex 11 (Computerised System)** |
|---|---|---|---|
| **A**ttributable (data can be assigned to the individual performing the task) | [4.20, c & f], [4.21, c & i], [4.29, e] | [6.14], [6.18], [6.52] | [2], [12.4], [15] |
| **L**egible (data can be read by eye or electronically and retained in a permanent format) | [4.1], [4.2], [4.7], [4.8], [4.9], [4.10] | [5.43] [6.11], [6.14], [6.15], [6.50] | [7.1], [9], [10], [17] |
| **C**ontemporaneous (data is created at the time the activity is performed) | [4.8] | [6.14] | [12.4], [14] |
| **O**riginal (data is in the same format as it was initially generated, or as a 'verified copy', which retains content and meaning) | [4.9], [4.27], [Paragraph "Record"] | [6.14], [6.15], [6.16] | [8.2], [9] |
| **A**ccurate (data is true / reflective of the activity or measurement performed) | [4.1], [6.17] | [5.40], [5.45], [6.6] | [Paragraph "Principles"],[5], [6], [10], [11] |

[1] Chapter 4 (Part I): Documentation
[2] Chapter 6 (Part I): Quality Control
[3] Chapter 5 (Part II): Process equipment (Computerized system)
[4] Chapter 6 (Part II): Process equipment

### 14. How should the company design and control their paper documentation system to prevent the unauthorised re-creation of GMP data?

The template (blank) forms used for manual recordings may be created in an electronic system (Word, Excel, etc.). The corresponding master documents should be approved and controlled electronically or in paper versions. The following expectations should be considered for the template (blank) form:

- have a unique reference number (including version number) and include reference to corresponding SOP number
- should be stored in a manner which ensures appropriate version control
- if signed electronically, should use a secure e-signature

The distribution of template records (e.g. 'blank' forms) should be controlled. The following expectations should be considered where appropriate, based on data risk and criticality:

- enable traceability for issuance of the blank form by using a bound logbook with numbered pages or other appropriate system. For loose leaf template forms, the distribution date, a sequential issuing number, the number of the copies distributed, the department name where the blank forms are distributed, etc. should be known
- Distributed copies should be designed to avoid photocoping either by using a secure stamp, or by the use of paper colour code not available in the working areas or another appropriate system.

### 15. What controls should be in place to ensure original electronic data is preserved?

Computerised systems should be designed in a way that ensures compliance with the principles of data integrity. The system design should make provisions such that original data cannot be deleted and for the retention of audit trails reflecting changes made to original data.

### 16. Why is it important to review electronic data?

In the case of data generated from an electronic system, electronic data is the original record which must be reviewed and evaluated prior to making batch release decisions and other decisions relating to GMP related activities (e.g. approval of stability results, analytical method validation etc.). In the event that the review is based solely on printouts there is potential for records to be excluded from the review process which may contain un-investigated out of specification data or other data anomalies. The review of the raw electronic data should mitigate risk and enable detection of data deletion, amendment, duplication, reusing and fabrication which are common data integrity failures.

*Example of an inspection citing:*

Raw data for HPLC/GC runs which had been invalidated was stored separately to the QC raw data packages and had not been included in the review process.

In the above situation, the procedure for review of chromatographic data packages did not require a review of the electronic raw data or a review of relevant audit trails associated with the analyses. This lead to the exclusion of records from the review process and to lack of visibility of changes made during the processing and reporting of the data. The company was unable to provide any explanation for the data which had been invalidated.

### 17. Is a risk-based review of electronic data acceptable?

Yes. The principles of quality risk management may be applied during the review of electronic data and review by exception is permitted, when scientifically justified.

Exception Reporting is used commonly as a tool to focus the review of electronic data such as (but not limited to) electronic batch records. Exception reporting rapidly highlights to the reviewer one of the most critical elements of batch review, i.e. the exceptions. The level of review of the full electronic batch record can vary based on the exceptions as well as the level of confidence and experience with a particular process. Appropriate testing and validation must be completed for the automated system and the output Batch Exception Report to ensure its functionality meets the business and regulatory requirements as per GMP.

### 18. What are the expectations for the self-inspection program related to data integrity?

Ongoing compliance with the company's data governance policy/procedures should be reviewed during self-inspection, to ensure that they remain effective. This may also include elements of the Data lifecycle discussed in Q3-Q9.

### 19. What are my company's responsibilities relating to data integrity for GMP activities contracted out to another company?

Data integrity requirements should be incorporated into the company's contractor/vendor qualification/assurance program and associated procedures.

In addition to having their own data governance systems, companies outsourcing activities should verify the adequacy of comparable systems at the contract acceptor. The contract acceptor should apply equivalent levels of control to those applied by the contract giver.

Formal assessment of the contract acceptors competency and compliance in this regard should be conducted in the first instance prior to the approval of a contractor, and thereafter verified on a periodic basis at an appropriate frequency based on risk.

### 20. How can a recipient (contract giver) build confidence in the validity of documents such as Certificate of Analysis (CoA) provided by a supplier (contract acceptor)?

The recipient should have knowledge of the systems and procedures implemented at the supplier for the generation of the CoA. Arrangements should be in place to ensure that significant changes to systems are notified and the effectiveness of these arrangements should be subjected to periodic review.

Data related to activities which are outsourced are routinely provided as summary data in a report format (e.g. CoA). These summary documents are reviewed on a routine basis by the contract acceptor and therefore the review of data integrity at the contract acceptor site on a regular periodic basis (e.g. during on-site audit) takes on even greater significance, in order to build and maintain confidence in the summary data provided.

**21. What are the expectations in relation to contract calibration service providers who conduct calibrations on-site and/or off-site? Are audits of these companies premises required?**

Using the principles of QRM to assess data criticality and risk, the company should include assessment of data governance systems implemented by the service provider when making decisions on service contracts. This may be achieved by on-site audit or desk-based assessment of information submitted by the service provider.

**22. What is expected of my company in the event that one of my approved contractors (e.g. active substance manufacturer, finished product manufacturer, quality control laboratory etc.) is issued with a warning letter/statement of non-compliance concerning data integrity, from a regulatory authority?**

It is considered that the company should evaluate the risk to its products manufactured/released using the principles of quality risk management. Risk assessments should be made available to Inspectors, on request.

Depending on the outcome of the risk assessment, appropriate action should be taken which may entail delisting the contractor from the approved contractor list. In the event that abnormal disruption in supply may result from a contractor compliance situation, relevant regulatory authorities should be consulted in this regard.

**23. Where does my company's responsibility begin and end in relation to data integrity aspects of the supply chain for medicinal products?**

All actors in the supply chain play an important part in overall data integrity and assurance of product quality.

Data governance systems should be implemented from the manufacture of starting materials right through to the delivery of medicinal products to persons authorised or entitled to supply medicinal products to the public.

Relative responsibilities and boundaries should be documented in the contracts between the relevant parties. Final responsibility of ensuring compliance throughout the supply chain rests with batch certifying QP.

**SOURCE (17AUG2016):**

http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/q_and_a/q_and_a_detail_000027.jsp&mid=WC0b01ac05800296ca#section16