http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-1-overview

Review and Critique of the MRHA Data Integrity Guidance for Industry — Part 1: Overview

Fri, 05/29/2015 - 10:41am

R.D. McDowall, Ph.D.

R.D. McDowall is Director of R D McDowall Ltd.This is the first of four articles looking at the new UK's Medicines and Heathcare products Regulatory Agency (MHRA) guidance for industry on data integrity.

Global Data Integrity Problems

Data integrity is a major regulatory topic that has been the subject of a number of articles by myself in *Scientific Computing* over the past few years focusing on chromatography data systems[1,2] and warning letters issued in mid 2013.[3,4] Data integrity is not just an Indian and Chinese problem, but a global issue, as many data integrity problems are based on poor and/or outdated working practices rather than a minority of cases involving data falsification.

Cases of fraud and falsification have occurred in the United States with Able Laboratories[5] and Leiner Health Products[6] over 10 years ago. As a result, the United States regulator, the Food and Drug Administration (FDA) has taken the lead, such as:

- Updated Compliance Program Guide 7346.832,[7] which has as objective 3 a data integrity audit of laboratory data.
- Trained their inspectors in data integrity, which means that there is now a focus on computerized systems and the data contained therein rather than paper output.
- There is level 2 guidance for some aspects of data integrity: shared user log-ins, why paper cannot be raw data from a computerized system, and using samples as SST injections.[8]

The European Medicines Agency has started posting GMP non-compliances online, where many cases of data integrity have been noted,[9] and Health Canada has now stated that GMP inspections will be unannounced due to data integrity issues that it has uncovered.[10]

MHRA Approach to Data Integrity

Our story begins in December 2013 when the MHRA gave the pharmaceutical industry an early Christmas present via their Web site.[11] This announcement stated that, from January 2014:

The MHRA is setting an expectation that pharmaceutical manufacturers, importers and contract laboratories, as part of their self-inspection programme must review the effectiveness of their governance systems to ensure data integrity and traceability.

This was an extension of self-inspections (internal audits) under Chapter 9 of EU GMP.[12] However, in addition to the pharmaceutical company itself, it was also an expectation that the data integrity of a company's suppliers (e.g. API suppliers, contract manufacturing and contract laboratories) are included in these assessments as well.

In March 2014, the MHRA wrote to suppliers of chromatography data systems to request a copy of their software and documentation in order to understand how each system worked. The unwritten lines of the letter I have surmised are so that the inspectors can identify how data can be falsified using a specific CDS application. It is not known how many suppliers responded with copies of their software.

The next stage of the story is that, in April 2014, MHRA and other European inspectors received training in data integrity from Monica Cahilly who is one of the trainers for the FDA on this subject.

In January 2015, MHRA released a guidance for industry on data integrity.[13] Following feedback from the industry, MHRA issued a second version in March 2015.[14] This is a good point, as it shows that they are willing to listen to the pharmaceutical industry.

The focus of these articles is an interpretation and critique of the second version of the MRHA data integrity guidance for laboratories working to European Union GMP regulations, such as analytical development in R&D and quality control in pharmaceutical manufacturing. In doing so, some of the main differences between the first and second versions of the document will be highlighted and discussed.

MHRA Data Integrity Guidance Overview

The guidance[14] consists of 16 pages and is divided into two main sections: discussion and definitions. The discussion section of three pages consists of an introduction, followed by topics on establishing data criticality and inherent integrity risk, and designing systems to assure data quality and integrity. The definitions section comprises 19 definitions combined with the MHRA expectation or guidance. It is the latter section, specifically the regulatory guidance or expectation, which provides most of the value in the document.

As we shall see, writing of the document in parts appears rushed, as there are some interesting areas, which I believe to be unintended, where there are mistakes and gaps in the definition portion of the document. Some of the original gaps have been corrected in the second version of the guidance, but some errors remain. The major issue I have with the definitions section is that it reads as a shopping list, rather than integrating and interleaving the definitions together to create a better idea of exactly what is required by the agency.

In summary, the MHRA guidance can be described as good, but not that good.

It is good in that it sets out definitions and regulatory expectations for data integrity and clarifies some points. It is not that good, as it fails to integrate the individual definitions and expectations into a meaningful description of what needs to be done to comply, mainly by lacking figures. However, these comments should not understate the fact that this is the first comprehensive guidance for industry on data integrity that has been issued by a regulatory authority. Although some may argue that the FDA's CPG 7346.832[7] should be the first such document, as it outlines a laboratory data integrity audit. However, this is intended for inspectors not industry. The MHRA document is a guidance for industry.

1 in 28?

One question that struck me reading the document for the first time was why has the MHRA taken the steps to publish this guidance for industry? The UK is one of the 28 member states of the European Union, and each member state has its own regulator responsible for inspections within its borders and for inspections outside the EU. However, the regulations and the majority of guidance documents or concept papers are usually issued by the European Medicines Agency (EMA), the pan European body responsible for regulations, product licensing, etcetera. What will one EU-competent authority achieve working on its own?

Introduction to the MHRA Guidance – Setting the Scene

The introduction to the MRHA guidance looks at the justification for data integrity and the first sentence sets the scene:

Data integrity is fundamental in a pharmaceutical quality system which ensures that medicines are of the required quality.[14]

It goes on to state that this guidance is complimentary to EU GMP. It also reiterates an MHRA expectation for a data governance system, which repeats their original 2013 approach.[11] It also warns companies not to return to paper, as this would be a breach of European Union directive 2001/83/EC[15] which, in article 23, requires companies to take account of scientific and technical progress.

Two changes have been made in the March 2015 version of the document in the introduction:

- The first is informational and refines the scope of the document to active substances (APIs) and dosage forms. Therefore, it excludes excipients from the scope of the guidance, presumably as these are lower risk.
- The second change is more far-reaching for regulated organizations. In the original version, the guidance stated that organizations are "*not expected to implement a forensic approach to data checking, ….*".
  However the revised version slips in four additional words to read "*not expected to implement a forensic approach to data checking on a routine basis, ….*".
  This changes the whole approach to data integrity: the original version wanted a system *to provide an acceptable state of control based on data integrity risk*. However, do we now need to have CSI on standby to rush in waving their torches around looking for clues whenever a data integrity alarm is raised?
  Perhaps a more rational approach is that we leave the forensics to the regular self-

inspections or for cause audits and the acceptable state of control to routine operations, such as the second-person checks of laboratory data and the reportable results?

Drowning in Integrity Definitions

The MHRA guidance document gives a definition of data integrity which is shown in Table 1 along with four other definitions (two from the FDA, one from National Institute of Science and Technology and one from the Institute of Electronic and Electrical Engineers – IEEE) of either data integrity or integrity. I have deliberately listed all five definitions in Table 1 to illustrate that different organizations, or even different divisions of the same regulatory organization, can have different approaches to the same subject.

Table 1: Data Integrity and Integrity Definitions

Source	**Definition of Data Integrity or Integrity**
MHRA[14]	The extent to which all data are complete, consistent and accurate throughout the data lifecycle (data integrity).
FDA 1[16]	The degree to which a collection of data are complete, consistent and accurate (data integrity)
FDA 2[17]	Data, information and software are accurate and complete and have not been improperly modified (integrity)
NIST[18]	The property that data has not been altered in an unauthorized manner (data integrity). Data integrity covers data in storage, during processing, and while in transit
IEEE[19]	The degree to which a system or component prevents unauthorized access to, or modification of, computer programs or data (integrity)

What can we learn from these definitions of integrity and data integrity? Let us attempt to reconcile and combine them into a single approach for data integrity:

- Data must be complete, consistent and accurate (MHRA & FDA 1, 2).
- Data have a life cycle (MHRA, NIST).
- Data must not have been improperly modified (FDA, NIST).
- If using a computerised system the software should prevent unauthorised modification of data (FDA 2, IEEE).

The first three bullet points hold for manual processes, as well as hybrid and electronic computerized systems, and the fourth point covers hybrid and electronic systems.

Wrong Definition of Data and Integrity Criteria

In the definition section, data is defined as *information derived or obtained from raw data (e.g. a reported analytical result)*.[14] This definition is misleading. How can be data be defined as information? Data are processed and reduced to information, which itself can be further interpreted to produce knowledge. However you look at it, data can never be information. MHRA's own definition equated information as analytical results (i.e. a reduction of raw data).

In the regulatory expectation for data is the requirement to comply with ALCOA principles. Table 2 shows these criteria in the first five rows (Attributable, Legible, Contemporaneous, Original and Accurate). The first line of each criterion is the MHRA requirement, and underneath are my additions to them. However, when looking at data integrity, ALCOA principles, which were developed for paper records, are not sufficiently comprehensive. The GAMP Data Integrity SIG has adopted the EMA GCP[20] criteria for electronic source data, which are shown in Table 2 in the last four rows and summarized as ALCOA+. The four additional criteria are: Complete, Consistent, Enduring and Available. Therefore, the data definition and the regulatory expectation sections in the MHRA guidance need to be revised, in my opinion, to be comprehensive for paper, hybrid and electronic processes and systems.

Table 2: ALCOA+ Criteria for Data Integrity

Criterion	**Meaning**
Attributable	Attributable to the person generating the data (MHRA) Who acquired the data originally or performed an action subsequently to it and when?
Legible	Legible (MHRA) Can you read the data together with any metadata or all written entries on paper?
Contemporaneous	Contemporaneous (MHRA) Documented (on paper or electronically) at the time of an activity
Original	Original record or true copy (MHRA) Written observation or printout or a certified copy thereof Electronic record including metadata of an activity

Accurate	Accurate (MHRA) No errors in the original observation(s) No editing without documented amendments / audit trail entries by authorized personnel
Complete	All data from an analysis, including any data generated before a problem is observed, data generated after repeat part or all of the work or reanalysis performed on the sample. For hybrid systems, the paper output must be linked to the underlying electronic records used to produce it.
Consistent	All elements of the analysis, such as the sequence of events, follow on and data files are date (all processes) and time (when using a hybrid or electronic systems) stamped in the expected order
Enduring	Recorded on authorized media e.g. laboratory notebooks, numbered worksheets, for which there is accountability or electronic media Not recorded on the backs of envelopes, laboratory coat sleeves, cigarette packets or Post-It notes
Available	The complete collection of records can be accessed or retrieved for review and audit or inspection over the lifetime of the record.

The next part of this series will look at the data governance system.

References

1. R.D.McDowall, *Scientific Computing*, November 2007 http://www.scientificcomputing.com/articles/2007/09/four-cds-compliance-lessons?cmpid=horizontalcontent
2. R.D.McDowall, *Scientific Computing*, March – April 2011 http://www.scientificcomputing.com/articles/2011/05/ensuring-data-integrity-regulated-environment?cmpid=horizontalcontent
3. R.D.McDowall, *Scientific Computing*, July 2013 http://www.scientificcomputing.com/articles/2013/09/fdas-focus-laboratory-data-integrity-%E2%80%93-part-1
4. R.D.McDowall, *Scientific Computing*, August 2013 http://www.scientificcomputing.com/articles/2013/09/fdas-focus-laboratory-data-integrity-%E2%80%93-part-2
5. Able Laboratories 483 Inspectional Observations (July 2005)
6. Leiner Health Products FDA Warning Letter, (August 2007)

7. Compliance Program Guide (CPG) 7346.832 Pre Approval Inspections, FDA May 2010, effective May 2012
8. FDA Level 2 guidance Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance - Records and Reports. Question 7: In warning letters to firms, why has FDA objected to the practice of using actual samples to perform system suitability testing (sometimes also referred to as "trial," "test," or "prep" runs)? http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm124787.htm
9. EMA EudraGDMP database http://eudragmdp.ema.europa.eu/inspections/gmpc/searchGMPNonCompliance.dof
10. Health Canada database http://www.hc-sc.gc.ca/dhp-mps/compli-conform/gmp-bpf/index-eng.php
11. MHRA Expectation Regarding Self Inspection and Data Integrity, December 2013 The original web page was http://www.mhra.gov.uk/Howweregulate/Medicines/Inspectionandstandards/GoodManufacturingPractice/News/CON355490 This page has now been archived and a snapshot is available at http://webarchive.nationalarchives.gov.uk/20141205150130/http://www.mhra.gov.uk/Howweregulate/Medicines/Inspectionandstandards/GoodManufacturingPractice/News/CON355490
12. EU GMP Chapter 9 Self Inspections: http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
13. MHRA GMP Data Integrity Definitions and Guidance for Industry version 1, January 2015
14. MHRA GMP Data Integrity Definitions and Guidance for Industry version 2, March 2015
15. Medicinal Products for Human Use, EU Directive 2001/83/EU, Article 25 (2001)
16. FDA Glossary of Computer Systems Software Development Terminology (1995)
17. FDA Guidance on Content of Premarket Submissions for Management of Cybersecurity in Medical Devices (Oct 2014)
18. NIST SP 800 – 33 Underlying Technical Models for Information Technology Security, December 2001
19. IEEE Software Engineering Standard 610 (Glossary, 1990
20. Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials, European Medicines Agency, 2010

R.D. McDowall is Director of R D McDowall Ltd. He may be contacted at editor@ScientificComputing.com.

http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-2-data-governance-system

Review and Critique of the MRHA Data Integrity Guidance for Industry — Part 2: Data Governance System

Fri, 05/29/2015 - 11:27am

R.D. McDowall, Ph.D.

R.D. McDowall is Director of R D McDowall Ltd.This is the second of a four-part series reviewing and critiquing the recent Medicines and Healthcare products Regulatory Agency (MHRA) guidance for industry document on data integrity.[1] The first part of the series[2] provided a background to the guidance document and discussed the introduction to the document. In this part, we will look at the MHRA requirement for a data governance system.

Data Governance System

The MHRA guidance document defines a data governance system as:

The sum total of arrangements to ensure that data, irrespective of the format in which it is generated, is recorded, processed, retained and used to ensure a complete, consistent and accurate record throughout the data lifecycle.[1]

Let us explore what this should entail. First, no other regulatory agency is requiring organizations to have a data governance system. However, this is a good idea given the number of issues involving data integrity that have been found recently. The rationale for this is based on MHRA's interpretation of ICH Q10 on Pharmaceutical Quality Systems (PQS),[3] which is incorporated in Part 3 of EU GMP[4] and that of EU GMP Chapter 1 on PQS,[5] which is based in part on ICH Q10.

Under the clause 1.8 for GMP for medicinal products it states

> (vi) Records are made, manually and/or by recording instruments, during manufacture which demonstrate that **all the steps required by the defined procedures and instructions were in fact taken** and that the quantity and quality of the product was as expected.

There is a similar requirement for quality control laboratories in clause 1.9 which states:

> (iv) Records are made, manually and/or by recording instruments, which **demonstrate that all the required sampling, inspecting and testing procedures were actually carried out**. Any deviations are fully recorded and investigated;

I believe that it is on these two clauses that MHRA bases the interpretation for a data governance system. As required by EU GMP Chapter 4, records are evidence that instructions have been executed correctly.[6] However, it is a long stretch from sections 1.8 and 1.9 of EU GMP to a data governance system. In contrast, FDA has a least burdensome approach to the interpretation of their medical device regulations,[7] in a risk-based world, should this not be the way forward?

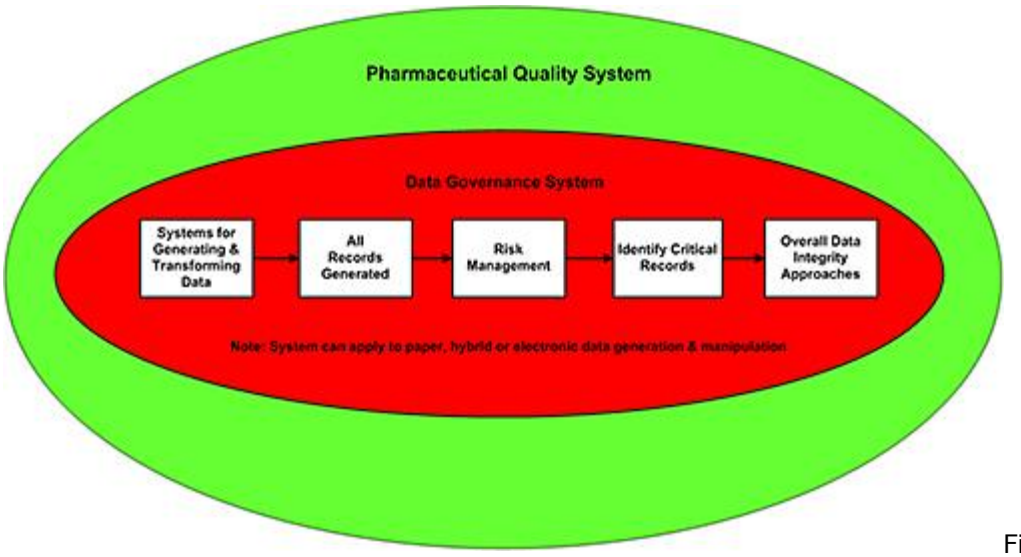

Figure 1: A Data Governance System within an overall Pharmaceutical Quality SystemAs shown diagrammatically in Figure 1: a data governance system can operate within the overall pharmaceutical quality system, as there are data integrity requirements contained in EU GMP Annex 11[8] for computerized systems, as well as requirements in Chapters 4 and 6 for paper, hybrid and electronic records.[6,9] The data governance system covers all processes involved in generating data and records during the course of pharmaceutical supply, manufacturing, testing and release. The controls to be applied to individual records, especially critical ones, are determined by risk management, which will outline the data integrity approaches for each system.

More detail is provided on the data governance structure by the MHRA in the definitions section of the guidance:[1]

- *Data governance should address data ownership throughout the lifecycle, and consider the design, operation and monitoring of processes / systems in order to comply with the principles of data integrity including control over intentional and unintentional changes to information.*
- *Data Governance systems should include staff training in the importance of data integrity principles and the creation of a working environment that encourages an open reporting culture for errors, omissions and aberrant results.*
- *Senior management is responsible for the implementation of systems and procedures to minimise the potential risk to data integrity, and for identifying the residual risk, using the principles of ICH Q9. Contract Givers should perform a similar review as part of their vendor assurance programme*[1]

From this, we can derive the following elements of a data governance system, which are listed below and shown linked in Figure 1:

- management responsibilities
- risk assessment
- data owners, who can be equated to the process owners of computerised systems under Annex 11[8] and the responsibilities combined
- policies and procedures
- training, including data integrity
- creating a no-blame culture around data integrity

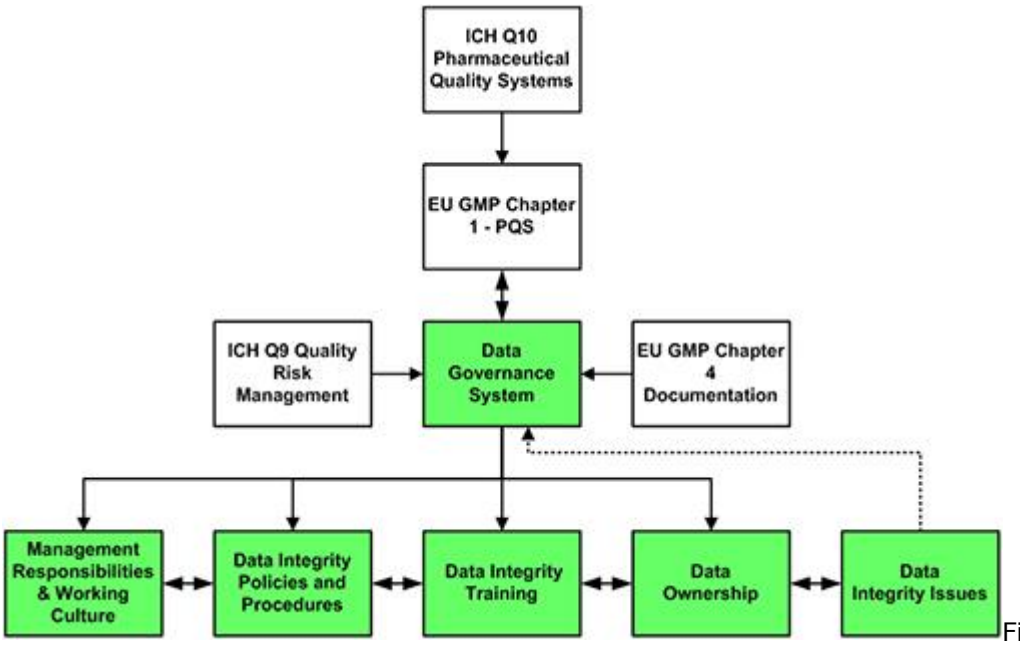

Figure 2: A suggested Data Integrity Governance Structure based on the MHRA Guidance Taking the criteria that were abstracted above from the MHRA guidance above, we can interweave them within an existing pharmaceutical quality system shown in Figure 2. Note, as shown in Figure 2, the five areas of the data governance system are not standalone silos, but interact with each other.

- **Management Responsibilities**: Senior management now has overall responsibility for quality and compliance with GMP within the PQS, as defined in EU GMP Chapter 1,[5] what needs to be added are the additional responsibilities for data integrity within each senior manager's functional area.[10] The responsibilities are to ensure that data are acquired, secured, transformed and reported in accordance with defined procedures and that deviations will be documented and investigated. Typically, these responsibilities, but not the accountability, will be devolved to the data owners of specific processes and computerized systems.
- **Working Culture and Data Integrity Issues**: This is the most important area that senior management can foster. What is required is the creation and maintenance of an open and no-blame culture to enable staff to raise data integrity issues. Part of this culture is the ability of staff to raise data integrity issues without fear of retribution via reporting mechanism to senior management.
- **Policies, Procedures and Training**: Procedures for ensuring data integrity for all activities (both GMP and non-GMP to avoid dual standards) followed by training in these procedures for all staff is essential.[11,12] Data integrity must be included in the regulatory requirement for on-going GMP training to reinforce the message. Part of the policies and procedures is the requirement for risk assessment. This needs to be undertaken to determine the impact and criticality of the records generated by each system to determine the controls. Then, via a gap and plan process, assess the existing controls in place to determine what, if any, additional controls are required to ensure data integrity. The GAMP good practice guide on Compliant Part 11 Records and Signatures[13] already has a list of controls to protect electronic records, and this could be adapted by organizations to include paper records as well.
- **Data Ownership**: There is a requirement for a data owner under the MHRA guidance. Rather that create another role, I would suggest that, for computerized systems, the existing process owner in the laboratory for each system should also be responsible for the integrity of data generated and managed within their systems. However, there are potential problems — what happens if data are transferred manually to a spreadsheet for further

calculations or are transferred from one system to another electronically — is the same person the data owner? However, if the responsibilities of the process owner and data owner are combined, the issue should be resolved for the majority of processes and systems.

In addition to the MRHA document, there is an extreme example of a data governance system in operation today, and that is documented in the Ranbaxy consent decree that the company and the FDA agreed upon in January 2012.[14] This established the post of Chief Data Integrity Officer reporting to the Board with a number of tasks to carry out to resolve the long standing falsification issues that had arisen over the previous four to five years. Part of the setup was the establishment of a whistleblowing phone line that any company employee can call without fear of retribution. I am not advocating such a governance structure, as the Ranbaxy approach has been defined to correct falsification carried out over some time. What is required is to integrate the data governance within the pharmaceutical quality system as shown in Figure 1.

However, after writing this section, I am still reminded that this is a single inspectorate within the European Union — how effective will this request for a data governance system be? Why is the EU not acting in unison?

Summary

In this part of the review of the MHRA data integrity guidance, we have focussed in the data governance system promoted by the UK regulator. There is a basis for this when interpreting EU GMP Chapter 1, and an outline of the elements for such a data governance system are presented and discussed. In the next part of this review and critique series, we will look at data criticality and a data life cycle.

References

1. MHRA GMP Data Integrity Definitions and Guidance for Industry version 2, March 2015
2. R.D.McDowall, *Scientific Computing*, Part 1 http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-1-overview
3. ICH Q10 Pharmaceutical Quality System, 2006
4. EU GMP Part 3 ICH Q10 Pharmaceutical Quality System http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
5. EU GMP Chapter 1 Pharmaceutical Quality System http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
6. EU GMP Chapter 4 Documentation, 2011 http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
7. FDA The Least Burdensome Provisions of the FDA Modernization Act of 1997: Concept and Principles; Final Guidance for FDA and Industry 2002 http://www.fda.gov/RegulatoryInformation/Guidances/ucm085994.htm
8. EU GMP Annex 11 Computerised Systems, 2011 http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
9. EU GMP Chapter 6 Quality Control, 2014 http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
10. J.Avellanet and E.Hitchings, Pharmaceutical Engineering, in press
11. EU GMP Chapter 2 Personnel (Principle and clause 2.11) http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
12. 21 CFR 211, Current Good Manufacturing Practice for Finished Pharmaceutical Products §211.25
13. GAMP Good Practice Guide Compliant Part 11 Records and Signatures, ISPE Tampa FL, 2005

14. Consent Decree of Permanent Injunction, Ranbaxy Laboratories and Food and Drug Administration, Case 1:12-cv-00250-JFM, United States Court District of Maryland, January 2012

R.D. McDowall is Director of R D McDowall Ltd. He may be contacted at editor@ScientificComputing.com.

http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-3-data-criticality-and-data-life-cycle

Review and Critique of the MRHA Data Integrity Guidance for Industry — Part 3: Data Criticality and Data Life Cycle

Fri, 05/29/2015 - 12:21pm

R.D. McDowall, Ph.D.

R.D. McDowall is Director of R D McDowall Ltd.This is the third of a four-part series reviewing and critiquing the recent Medicines and Healthcare products Regulatory Agency (MHRA) guidance for industry document on data integrity.[1] The first part of the series[2] provided a background to the guidance document and discussed the introduction to the document. The second part reviewed and discussed the data governance system.[3] In this part, we will look at data criticality and the data life cycle.

Establishing Data Criticality and Inherent Integrity Risk

This section of the guidance first discusses the data governance system that we discussed above and then moves on to look at data generation. The spectrum of data generation is purported to be represented by Figure 1 in the guidance, which turns out to be a diagram drawn by Monica Cahilly during the April 2014 training of the MHRA inspectors.[4] However, the diagram is, to my mind, only focused on instruments and computer systems, and I have drawn up a more detailed description of what should be presented in a data integrity guidance, see Figure 1.

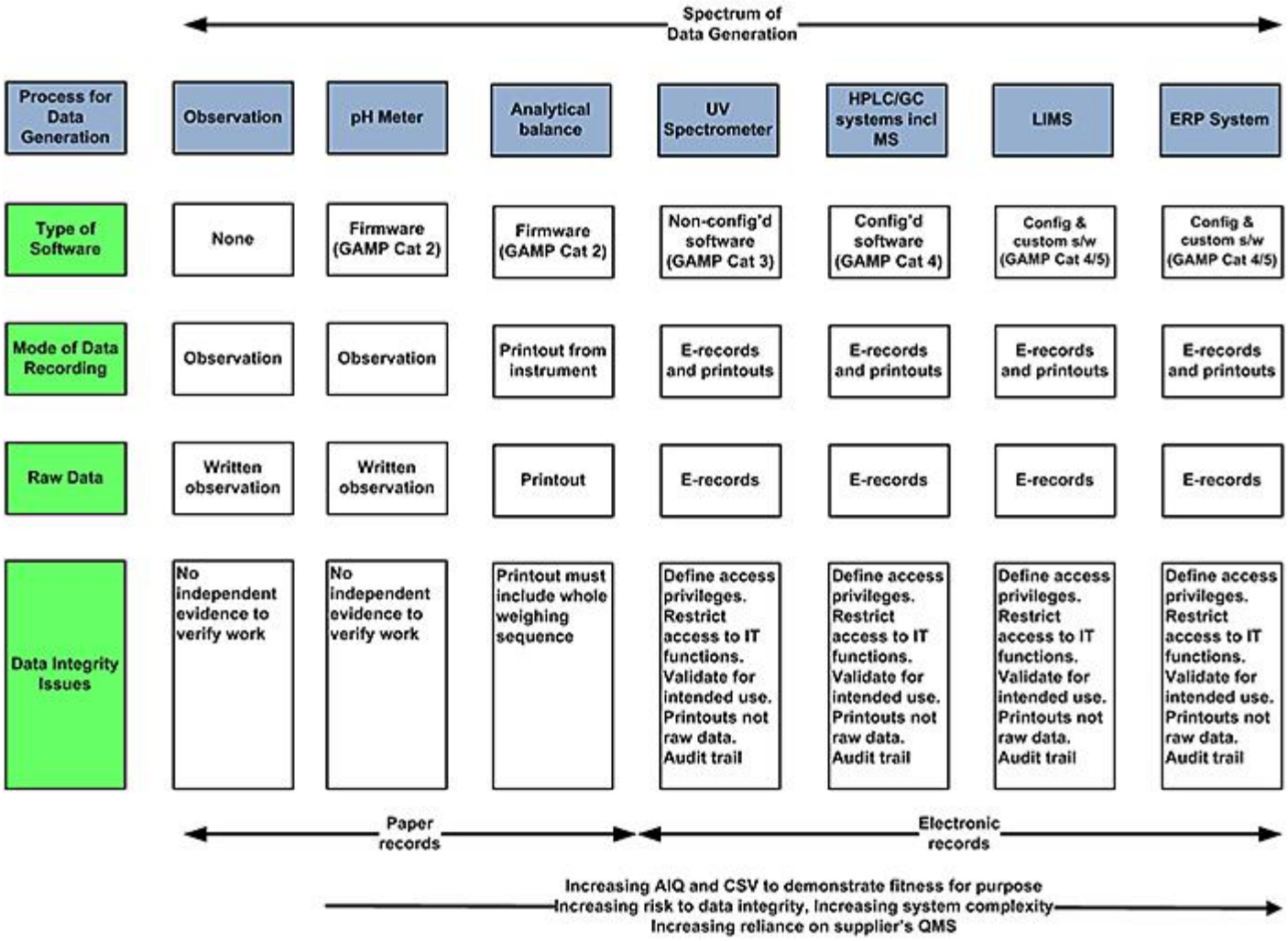

Figure 1: Spectrum of data generation processes in a laboratory highlighting data integrity issuesOn the horizontal axis at the top of the figure are the different processes that can be used in a laboratory environment to generate data; these vary from observation to simple instruments, such as balances and pH meters, to chromatography data systems through to LIMS (laboratory information management systems) and ERP (enterprise resource planning) systems. The vertical axis consists of the attributes of each process, such as whether software is used and, if so, the GAMP classification, the mode of data recording, the raw data produced and the main data integrity issues of each process. Note that Figure 2 quotes firmware as Category 2 software, although this has been discontinued in GAMP version 5,[5] it equates to Group B instruments in USP <1058> on Analytical Instrument Qualification (AIQ).[6] When mapping USP <1058> groups versus GAMP software categories,[7] if Category 2 software were reinstated, there would be equivalence between Category 2 software and Group B instruments.

The first three processes from observation to analytical balance have paper records, and the remaining four items have electronic records as raw data. Dependent upon how the latter four computerized systems are used, they can either be hybrid or electronic by using electronic signatures. Furthermore, the pH meter and analytical balance are discussed here from the perspective of being standalone instruments rather than being interfaced to a LIMS or ELN (electronic laboratory notebook). The problem with the MHRA figure is that it focusses only on instruments and computerized systems and does not consider data gathered by observation.

Figure 1 also shows that, for analytical instruments and laboratory computerized systems, the following items hold true:

- Going from left to right, there is increasing complexity.
- Increasing amounts of AIQ and / or CSV are required to demonstrate fitness for intended use as one goes from a simple instrument to a complex computerized system.
- There is increasing risk to data integrity from either inadvertent acts by users or deliberate falsification going from left to right.
- There is increasing reliance of a laboratory on a supplier's quality management system the further to the right one goes.

Let us look at four examples of data gathering from Figure 1:

- **Observation**: Manual observations may be found in many laboratories for tests such as color or odor of samples, as well as recording data from some instruments as shown in the first column on the left of Figure 3. As noted here, the data integrity issue is that there is no independent evidence to verify that the value or result recorded is correct, has suffered from a transcription error (value only) or has been falsified. Therefore, each process using observation only needs to be risk assessed to determine the criticality of the data being generated: for example, is an odor determination the same criticality as the pH determination of HPLC mobile phase?
- **Instrument**: The example used in Figure 1 is an analytical balance with a printer. Given the importance of accurately measuring reference materials and samples and the impact that a balance can have on a regulated laboratory, it is important that the integrity of measurement is maintained. At a minimum, a printer is essential for an analytical balance, as the MHRA guidance makes clear[1] and discussed later in this paper. However we need to consider more detail: what data need to be recorded when making a weighing measurement? In my view, the printer needs to record the weights captured during any weighing operation e.g. weight of weighing vessel, tared weight and the weight of material.
- **Hybrid System**: The hybrid system, typified by a UV spectrometer using GAMP Category 3 software, is the worst of both worlds, as the laboratory has to manage and co-ordinate two different and incompatible media types: paper records and electronic records. The issues are that paper cannot be defined as raw data as noted by the EU and FDA.[8,9] Note that the FDA level 2 guidance[9] is a much better discussion of why paper cannot be raw data. Other data integrity issues are that configuration of the software must be recorded, including definitions of user types and the access privileges for each type, and validation of this configured software for the intended use. Many hybrid systems consist of the instrument connected to a standalone workstation, where there are potential issues of access to the operating system, clock, the data files themselves via the OS and effective and validated backup and recovery.[10] This situation is specifically commented in the MHRA guidance in the definitions section.[1] Systems using the operating system to store the data files in open access directories can suffer from the stupidity of operators performing unintended deletions, as well as attempts at falsification from individuals. However the use of a database should protect data from many falsification attacks. But, in reality, data need to be acquired and stored securely in the network when using flat file systems.
- **Electronic System**: Using a chromatography data system with GAMP category 4 software with electronic signatures as an example. In this instance, the raw data are electronic records with electronic signatures. To ensure data integrity, the application has to be configured for security and access control (definition of user types and access privileges) and also for the use of electronic signatures. Data are acquired to the network and are secured with a database. Validation for intended use will demonstrate that the configured systems works. The audit trail documents changes made by authorized individuals. The issue now is the separation of system administration roles from that of the use of the system by chromatographers.

Note that this approach can only be a generalization: know your instrument or system and how it operates is the key maxim here. For example, modern balances can have clocks, and their screens can access software such as electronic laboratory notebooks or LIMS acting as terminals, as well as an analytical instrument. Simply having a balance connected to such an application may not be enough — where is the time and date stamp applied in such cases: at the balance or in the software application? Can anybody change the clock in the balance and impact the time stamp in the application?

The Data Life Cycle

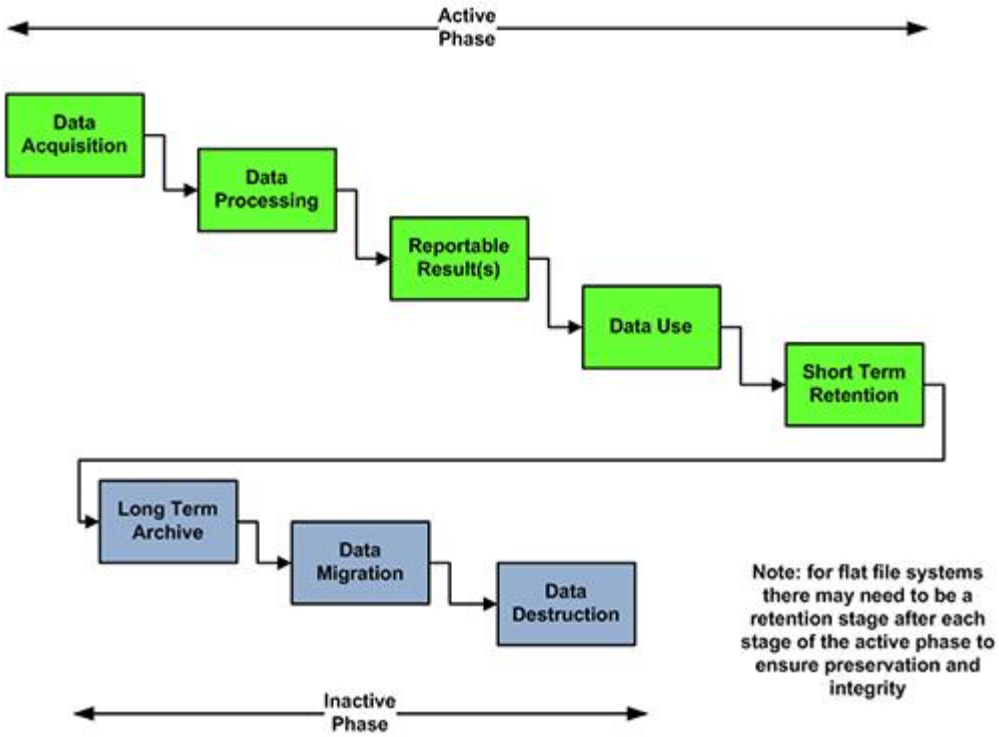

Figure 2: A suggested Data Life CycleThe MHRA definition of data expects a data lifecycle but does not give any clue about what one should be. In the absence of guidance, here is my suggestion of what such a data life cycle could be as shown in Figure 2. Firstly, there are two phases of a data life cycle for laboratory data: an active phase and an inactive phase.

The active phase of the data life cycle consists of the following activities:

- **Data acquisition:** the process of controlling and recording the observation or generating the data from the analytical procedure
- **Data processing:** interpretation or processing of the original data
- **Generate reportable result:** calculation of the reportable result for comparison versus specification
- **Information and Knowledge Use:** use of the result for the immediate purpose, but also over a longer time for trending
- **Short Term Retention:** storage of the data and information in a secure but accessible environment for any further use e.g. complaints, investigations, as well as audits / inspections

Note that, for many laboratory computerized systems where electronic records are stored in flat files within the operating system, there may needs to be a retention process performed after each stage of the active phase to ensure preservation of the record and the integrity.

The inactive phase of the data lifecycle consists of the following stages:

- **Long-term Archive:** movement of the records into a secure archive for long-term retention
- **Data Migration:** if necessary or required, there may be one or more migrations of data from one system / repository to another over the retention period
- **Data / Record Destruction:** when the retention period has elapsed, then a formal process to destroy the data / records should be executed, providing that there is no litigation pending.

However, this life cycle does not account for any other use of the data e.g. trending over time or product quality reviews where the information generated during an analysis is used as the input data for generation of additional information or knowledge abstraction.

Summary

In this part of the MHRA data integrity guidance, we have looked at the data risk and criticality via different ways of generating data from observation to an electronic computerized system using electronic signatures. In addition, we have considered a data lifecycle and looked at some of the issues surrounding this. In the last part of the series we will look at the section on system design, discuss a few of the definitions that constitute the bulk of the guidance document and summarize the guidance document.

References

1. MHRA GMP Data Integrity Definitions and Guidance for Industry version 2, March 2015
2. R.D.McDowall, Scientific Computing, Part 1 http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-1-overview
3. R.D.McDowall, Scientific Computing, Part 2 http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-2-data-governance-system
4. M Cahilly, Personal Communication
5. Good Automated Manufacturing Practice (GAMP) guidelines Version 5, ISPE, Tampa Florida, 2008
6. United States Pharmacopoeia, <1058>, Analytical Instrument Qualification
7. L. Vuolo-Schuessler, M. E. Newton, P. Smith, C.Burgess, and R.D. McDowall, Pharmaceutical Engineering 46, (1), 46 – 56, 2014
8. EU GMP Chapter 4 Documentation, 2011 http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
9. FDA Level 2 guidance Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance - Records and Reports. Question 3: How do the Part 11 regulations and "predicate rule requirements" (in 21 CFR Part 211) apply to the electronic records created by computerized laboratory systems and the associated printed chromatograms that are used in drug manufacturing and testing? http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm124787.htm

10. EU GMP Annex 11 Computerised Systems, 2011
http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm

R.D. McDowall is Director of R D McDowall Ltd. He may be contacted at editor@ScientificComputing.com.

http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-4-system-design-definitions-and-overall-assessment

Review and Critique of the MRHA Data Integrity Guidance for Industry — Part 4: System Design, Definitions and Overall Assessment

Fri, 05/29/2015 - 1:56pm

R.D. McDowall, Ph.D.

R.D. McDowall is Director of R D McDowall Ltd This is the fourth and final part of a series reviewing and critiquing the recent Medicines and Healthcare products Regulatory Agency (MHRA) guidance for industry document on data integrity.[1] The first part of the series[2] provided a background to the guidance document and discussed the introduction to the document. The second part reviewed the data governance system,[3] and the third part discussed data criticality and the data lifecycle.[4] This part reviews the system design, some of the definitions, and finishes with an overall assessment of the guidance.

Designing Systems to Assure Data Quality and Integrity

This portion of the MHRA guidance[1] consists of two sections. The first is a list of bullet points for the design of systems, and the second is a discussion on scribes for documenting GMP activities.

Turning to the first section, my view is that many of the bullet points are poorly written, with some basic errors. Below are the bulleted points from the MHRA document, and underneath each one are my comments and critique:

- *Access to clocks for recording timed events*.
  This is a poorly written item, as it implies that anyone or any system can access a clock, be it for a manual process or a computer process. But this is anybody's guess. I think that this is intended for computerized systems rather than manual processes, unless there is a test such as loss on drying (LOD). What this point should make is that an application needs access to the system clock to provide the time and date stamp for events within it. By implication, this means that workstations should be networked to ensure that the time stamp can maintain the accuracy from the time server that is linked to a trusted time source, so that manual intervention is not required. However, the main issue is that access to the system clock must be restricted to authorized individuals to prevent time traveling and data falsification.
- *Control over blank paper templates for data recording*.
  Perhaps a better phrasing of this requirement is to be found in the FDA 1993 guide on Inspection of Pharmaceutical Quality Control Laboratories: *We expect raw laboratory data to be maintained in bound (not loose or scrap sheets of paper) books or on analytical sheets for which there is accountability, such as pre-numbered sheets*.[5] Far more succinct and to the point.
- *User access rights which prevent (or audit trail) data amendments*.
  Perhaps a better way to express this is that user types / roles need to be defined and documented along with the corresponding access privileges per role. In addition, access privileges that enable a user to modify or delete records need to be justified. Where the access privileges allow either data modification or data deletion, these need to be monitored by the audit trail in the application.
- *Automated data capture or printers attached to equipment such as balances*.
  Put at its most basic: inspectors do not trust people to make manual observations of critical

data from analytical balances. They want independent verification of the weights of reference standards and samples used in analytical procedures. Indeed, standalone balances without printers may have been acceptable 30 years ago, but no longer due to cases of data falsification, but also human error. Analytical balances with a printer are now the status quo, see the discussion in Part 3 of this series.[4] However, what about other instruments such as a pH meter for checking mobile phases or that buffers have been made up correctly — is a printer necessary? Enter stage left a risk assessment!

- *Proximity of printers to relevant activities.*
  This applies mainly to hybrid systems as, if data are acquired, processed and reported electronically with electronic signatures, the need for a printer in proximity to the activity diminishes.
- *Access to raw data for staff performing data checking activities*.
  This is similar to the FDA GMP requirement for complete data and for the second person review to see all data generated in the course of an analysis.[3,9]

On the second section, my advice on scribes in a normal laboratory environment is don't use them, as this would cause more compliance problems than it solves. Furthermore, there is no equivalent position from the FDA on the subject.

Definitions and Expectations Associated with Data

There are 19 definitions in the MHRA document,[1] this critique will only focus on three of them: raw data, metadata and data. The problem with these three is that we have a surfeit of data and little information about how they link together and this, I would suggest, is a major omission from the guidance: figures are better for putting context around some of the key definitions. For the purposes of simplicity, I have not included the regulatory expectations, although the criteria for data integrity (ALCOA+) was discussed earlier in this article.[1]

Table 1 lists the three MRHA definitions for raw data, metadata and data from the guidance document.[1] These definitions are presented in the document, but are not really linked with what happens in practice in the laboratory. The principle of EU GMP Chapter 4 is more informative: *Records include the raw data which is used to generate other records*.[6] Therefore, by regulatory definition, we need to consider far more than just raw data and the associated contextual metadata, but also the processed or interpreted data derived from them, as well as the generation of the reportable result. Furthermore, as mentioned earlier, data cannot be considered as information. As such, the MHRA definitions should be revised again to reflect these concerns.

Table 1: MHRA Definitions for Raw Data, Metadata and Data

Word	**MHRA Definition[1]**
Raw Data	• Original records and documentation, retained in the format in which they were originally generated (i.e. paper or electronic), or as a 'true copy'. • Raw data must be contemporaneously and accurately recorded by permanent means. • In the case of basic electronic equipment which does not store electronic data, or provides only a printed data

	output (e.g. balance or pH meter), the printout constitutes the raw data.
Metadata	• Metadata is data that describes the attributes of other data, and provide context and meaning. • Typically, these are data that describe the structure, data elements, inter-relationships and other characteristics of data. • It also permits data to be attributable to an individual.
Data	• Information derived or obtained from raw data (e.g. a reported analytical result)

What do these definitions mean in practice? Let us look at three options shown in Figure 1:

- a paper-based test using observation with documentation by writing in a laboratory notebook
- a test conducted using a hybrid system
- and, finally, one using electronic workflows and electronic signatures

How do these different tests link with the three definitions? These three tests are also broken down in Table 2 into raw data with the associated metadata, processed data, information and knowledge. The latter two topics are either misunderstood in the MHRA guidance (information) or not mentioned (knowledge) which is subject to a separate paper on the subject.[7]

The first example is an observation of a test for example color or odor, this is written into a laboratory notebook or a controlled sheet. The second example is a hybrid system where observations generate electronic records and the metadata are written, as well as contained within the application, generated data are manually typed into a spreadsheet for calculation of the reportable result. The last example is an electronic system where all activities are contained within the application and underlying database. The reportable result is electronically signed by the tester and the reviewer.

Table 2: Records Associated with Manual Observation and Hybrid and Electronic Systems

Record	**Observation**	**Hybrid**	**Electronic**
Raw Data	• Written record	• Electronic files of the analysis	• Electronic files of the analysis
Metadata	• Further written data about the sample and analysis e.g.	• Electronic Files for control of the instrument, data acquisition,	• Electronic Files for control of the instrument, data acquisition,

	batch, test, analyst, date, etc.	interpretation and reporting of data. • Identification of who tested the sample, etc. • Audit trail entries of data changes • Further written data about the sample and analysis e.g. batch, test, etc.	interpretation and reporting of data. • Identification of who tested the sample, etc. • Audit trail entries of data changes • Further written data about the sample and analysis e.g. batch, test, etc.
Processed data		• Entry into spreadsheet for further calculation of individual values and reportable result • Spreadsheet file • Spreadsheet printout	• Interpretation of the raw data • Further metadata and audit trail entries
Information		• Individual values of aliquots	• Individual values of aliquots
Knowledge	• Reportable result • Handwritten signatures of tester and reviewer	• Printout of reportable result • Handwritten signatures of tester and reviewer • Linkage to underlying instrument raw data and spreadsheet file	• Reportable Result • Electronic signatures of tester and reviewer • Linkage to all data and metadata via application database

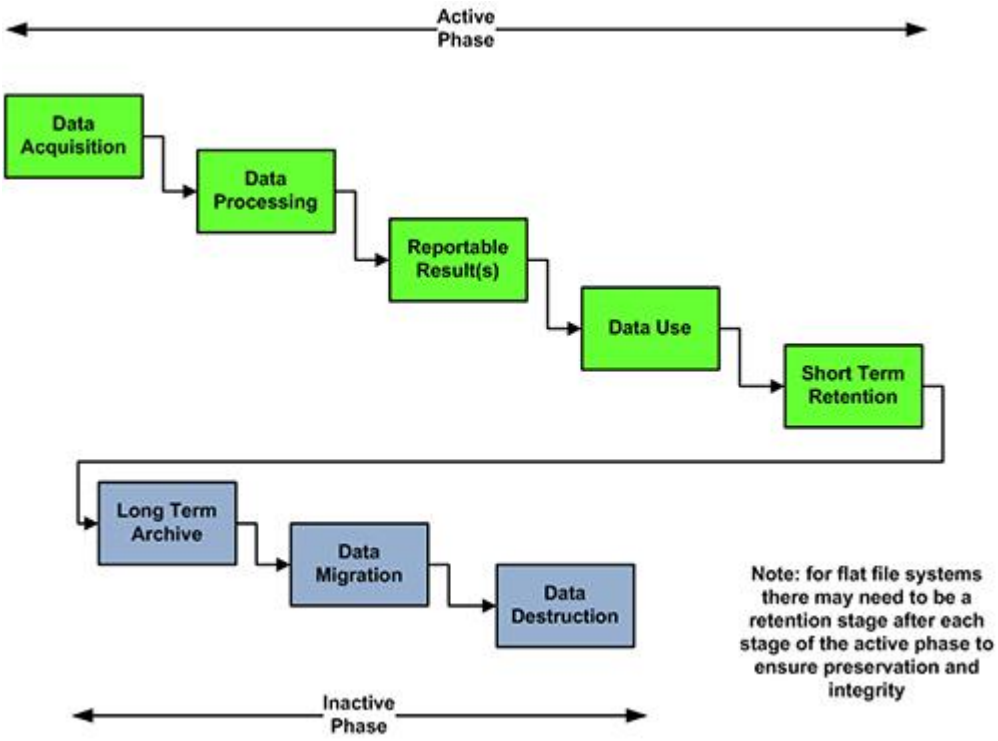

Figure 1: Three options for data generation in a regulated laboratoryThe aim of Table 2 and Figure 1 is to illustrate that simply presenting a series of definitions, even with regulatory expectations, is not enough. Context and explanation is all, and figures help understanding. In the MHRA document, Figures 2 and 3 show how to and how not to record data contemporaneously for a manufacturing system, the same approach should have been taken with many of the other definitions, as a picture is worth a thousand words.

Overall Assessment

My overall assessment of the MHRA data integrity guidance[1] is that it is good, but not good enough, and needs improvement as we have discussed in this series of articles.

It is good and provides a risk-based approach. There is more information on the data governance system than is provided on the MHRA Web site when first announced in 2013.[8] It also identifies the responsibilities of data owners and senior management in relation to data integrity. However, the guidance still is in need of improvement, such as it confuses data with information. A figure is needed to link together several related definitions. The section on design controls is poorly written and needs expansion to clarify what is required. As MHRA has shown a willingness to listen to comments from industry and has updated the document in a short time frame, my hope is that these articles along with other comments provide additional input to the review process.

References

1. MHRA GMP Data Integrity Definitions and Guidance for Industry version 2, March 2015
2. R.D.McDowall, Scientific Computing, Part 1 http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-1-overview?cmpid=horizontalcontent
3. R.D.McDowall, Scientific Computing, Part 2 http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-

integrity-guidance-industry-%E2%80%94-part-2-data-governance-system?cmpid=horizontalcontent
4. R.D.McDowall, Scientific Computing, Part 3 http://www.scientificcomputing.com/articles/2015/05/review-and-critique-mrha-data-integrity-guidance-industry-%E2%80%94-part-3-data-criticality-and-data-life-cycle?cmpid=horizontalcontent
5. FDA Guide to the Inspection of Pharmaceutical Quality Control Laboratories, 1993 http://www.fda.gov/ICECI/Inspections/InspectionGuides/ucm074918.htm
6. EU GMP Chapter 4 Documentation, 2011 http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
7. EU GMP Annex 11 Computerised Systems, 2011 http://ec.europa.eu/health/documents/eudralex/vol-4/index_en.htm
8. C Burgess and R D McDowall. LC-GC Europe, scheduled September 2015.
9. MHRA data integrity expectation Dec 2013 Web reference archived!

R.D. McDowall is Director of R D McDowall Ltd. He may be contacted at editor@ScientificComputing.com.