Impact of Data Integrity Audits on Pharma Microbial QC Labs

David Jones
Rapid Micro Biosystems

Most people are aware of the requirements of the code of federal regulations 21 CFR Part 11 for computer software security, which have been a major pharmaceutical IT focus for approximately 10 years.

However, within the 21 CFR 11 requirements lurked another high-risk component: "Data Integrity". Simply put, Data Integrity ("DI") is the assurance that data records are accurate, complete, intact and maintained within their original context so as to make the data trustworthy.

In pharmaceutical QC labs, there are often many manual steps in the performance of a routine QC analytic test to release a product (Figure 1). High risk areas were associated with the amount of human input required and how closely that input was monitored and verified.

The phrase "Data Integrity" was until recently relatively unknown in pharma QC laboratories. The floodgates of regulatory observations and warning letters were opened by compliance audit findings with High Performance Liquid Chromatography (HPLC) systems where stored sample data was "reworked" to achieve compliant product, or where tests were repeated until a "good run" was obtained without documenting all runs. Figure 2 shows the trend in warning letters containing data integrity over the last few years.

Failure to comply with the requirements of the data integrity guidelines have had significant effects on sites that have holds put on product release or on-site qualification. The costs and delay to implement compliant processes are significant.

The basic principles of Data Integrity revolve around the simple acronym- ALCOA, These stand for:

- Attributable-traceable to a unique person,
- Legible- no pencil, no correction, no liquid fluid, no hidden field that won't allow access, no deletion, overwriting,
- Contemporaneous- no backdating, no prefilling, date and time,
- Original true copy-in paper world (analytical worksheet); e-world (FTIR-spectra, injection sequence, electronic backup copy of the source TF-IR spectra file, compare to the original electronic data confirming ALL metadata is in the electronic copy set,
- Accurate- verification and confirmation through QMS.

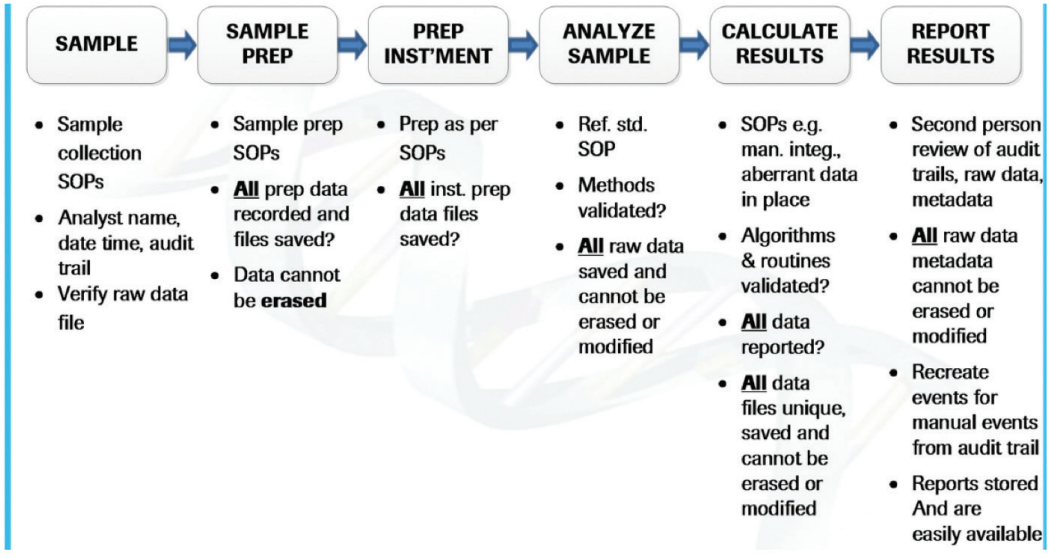

Figure 1. Sample Process Steps in the Pharmaceutical QC Lab

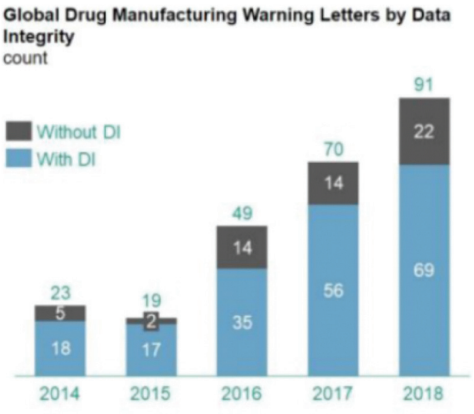

Figure 2. Warning Letter Trends 2014-2018; Source: FDAZilla

- § 211.68 requires that "backup data are exact and complete" and "secure from alteration, inadvertent erasures, or loss" and that "output from the computer … be checked for accuracy"
- § 212.110(b) requires that data be "stored to prevent deterioration or loss"
- §§ 211.100 and 211.160 require that certain activities be "documented at the time of performance" and that laboratory controls be "scientifically sound"
- § 211.180 requires that records be retained as "original records," or "true copies," or other "accurate reproductions of the original records"
- §§ 211.188, 211.194, and 212.60(g)require "complete information," "complete data derived from all tests," "complete record of all data," and "complete records of all tests performed"
- §§ 211.22, 211.192, and 211.194(a)require that production and control records be "reviewed" and that laboratory records be "reviewed for accuracy, completeness, and compliance with established standards"
- §§ 211.182, 211.186(a), 211.188(b)(11), and 211.194(a)(8) require that records be "checked," "verified," or "reviewed")

As noted in several of the warning letters, any failure to meet these cGMP regulations can amount to a data integrity failure and a violation of cGMP regulations.

Figure 3. Examples of Data Integrity references from CFR

All of these have requirements within the codes of Federal Regulations (CFR) that have been in effect for many years for "pen and paper" recording but have now moved into the electronic era. Example CFRs are shown in Figure 3.

There are also many global regulations and guidance's available on this topic:

- US FDA Data Integrity and Compliance With CGMP, draft Guidance for Industry April 2016 [https://tinyurl.com/ycuzvzgu]
- MHRA 'GXP' Data Integrity Guidance and Definitions March 2018 [https://tinyurl.com/y82jlq5z]
- PIC/S PI 041-1 (Draft 2) Good practices for data management and integrity in regulated GMP/GDP environments 10 August 2016 [https://picscheme.org/en/news?itemid=33]
- WHO TRS 996 Annex 05 'Guidance on good data and record management practices' 2016 [https://tinyurl.com/jmep977] and 'Guidance on good data and record management practices' September 2015 QAS15-624 [https://tinyurl.com/y726b37n]
- CH E6(R2) [www.ich.org]

Not all of these documents agree or are very explicit. As a result the PDA created a Technical Report to provide industry with a representative guidance, Technical Report No. 80: Data Integrity Management System for Pharmaceutical Laboratories, released in August 2018. Within this document, reference is made to the specific risks for the Microbiology QC lab:

- 7.2 Risk-Based Mitigation
  - » Currently, a high percentage of tests conducted in microbiological laboratories are observational, that is, the results (such as a colony count) are viewed and manually recorded on a paper document or in a computer record. Absent an easy, reliable method to verify the recorded data, some laboratories require microbiologists to use second-person verification (e.g., supervisor) by physical examination of the test plates. Further, the second person verification could be performed as a discreet step prior to approval of the data or combined with the data-approval step.
  - » Risk factors for collection, control, and verification of microbiology data are reduced with computer interface technology, such as automated plate readers or rapid methods that produce an electronic record that is retrievable, and relatively tamper proof or digitally time-and-date stamped photography equipment. This can include automation and the use of advanced methods with a validated data recording (for example, ATP bioluminescence platform) system and audit trail capabilities.

This set of data integrity risks is also echoed in the most recent MHRA guidance "MHRA GXP Data Integrity Guidance and Definitions: Revision 1 Mar 2018:

- 6.11.1. Original record
  - » Where the data obtained requires manual observation to record (for example results of a manual titration, visual interpretation of environmental monitoring plates) the process should be risk assessed and depending on the criticality, justify if a second contemporaneous verification check is required or investigate if the result could be captured by an alternate means.
- 6.11.2 True Copy
  - » "Where manual transcriptions occur, these should be verified by a second person or validated system."

Recently auditors have been applying the requirements that all manually recorded information needs to be verified by a second signatory, the so called "4 eyes" observation. Auditors are more likely to visit the micro lab and pull plates from the contaminated trash and verify the counts match the records that happened historically. Within the micro testing process, there are many steps that open questions of error or fraud: did the dish go in the air sampler? Did the sample

get incubated for the correct time at the correct temperature? Did the operator read the plate and record the correct result? Did the recorded result get transferred to the batch record?

The norm for microbiological testing is the visual inspection of broth for the detection of microbial growth or the enumeration of microbial plate counts. Typically, the results are recorded on paper worksheets or more recently into electronic notebooks. The transcribed data, though not always the original microbial cultures, are checked by a second person for completeness and the absence of recording and arithmetical errors, then entered into LIMS and approved. These laboratory procedures place a high reliance on the integrity, experience and training of the individual analyst, the ability of the peer reviewer or supervisory staff to detect poor data integrity, and the culture of the company to set the highest standard. Unfortunately, the procedures are susceptible to falsification or human error.

Many companies are reacting to the "4 eyes" requirement intended to mitigate the opportunity for falsification or human error. However, there are many interpretations of the requirement. Implementations range from 2 analysts reading every plate and dealing with enumeration differences between the analysts, to random checks, to formal checks but at prescribed 3 to 6 month intervals. The variation in approaches may lead to even closer scrutiny by an auditor as inconsistency of approach leads to loss of confidence.

The most conservative approach would involve removal of the human variable. While this is very difficult for the sample collection the incubation, enumeration and data transfer to a LIMS database can easily be automated. Automation can verify whether samples have not been taken early so missing sample points can be re-taken rather than wait 5-7 days to find data is missing that is required for batch release.

The Growth Direct™ system facilitates compliance data integrity through automation of the enumeration and reporting phases. Sample worklists can be down loaded from a LIMS platform (e.g. MODA for EM, or a general system such as Labware) to the Growth Direct system and collected samples validated against the list upon loading. The automation of the incubation process ensures the correct incubation time and temperatures including change over for serial incubation are accurately controlled. Interim plate reading through the incubation period allows faster reaction times in the event that a sample demonstrates an action or alert level excursion. On completion of the incubation, the results can be approved on the Growth Direct prior to the printout of reports if working in standalone mode, or dropped into the LIMS system for electronic approval. In LIMS mode, complete security is obtained from start of incubation to formal batch record.

With the introduction of an automated system the security of both raw data and metadata are of key importance. Original records and documentation are retained in the format in which they were originally generated (i.e. paper or electronic) as a 'true copy'. Raw data must be contemporaneously and accurately recorded by permanent means. In the case of basic electronic equipment such as a balance or pH meter which provides only a printed data output and does not store electronic data, the printout constitutes the raw data. Metadata is data that describe the attributes of other data, providing context and meaning. Typically, these are data that describe the structure, data elements, interrelationships and other characteristics of data. It also permits data to be attributable to an individual.

There is global ambiguity as to what constitutes "raw data" or what does not. McDowall and Burgess have suggested a definition for a "Primary Analytical Record", which contains raw data [LCGC Europe, Volume 28, Issue 11, November 2015, Pages 621-626]. A more comprehensive concept may be the term "Primary Analytical Record" which is not a single record but a collection of data, metadata, information, and knowledge. In the absence of an internationally agreed definition for raw and/or complete data, Rapid Micro Biosystems recommends using the concept of Primary Analytical Record as it best describes not merely the data, but also the processes required to obtain such an analytical record.

The Growth Direct contains two servers: the System Manager server to control the mechanical movements and holds the SQL database, and the Growth Analyzer server to process the images and create the CFU result which is transferred to the System Manager for output.

At each time point, the system camera takes three images (at three height levels above the membrane to ensure all processed objects are in focus) at each time point. The images are processed to create a "mask" of the image which is stored on the Growth Analyzer server hard drive as the "time point history" (TPH) file. The mask contains all the key parameters detected on the membrane surface, baseline, object position, brightness, etc. The Growth Analyzer software calculates and outputs the Colony Forming Unit (CFU) count at that time point to the system server SQL database, then to the user interface screen for user information. The full .tif images at each time point are temporarily saved to the Growth Analyzer server hard drive; however, they are never used again by the system.

At the next time point the same process is performed, the system uses the new "mask" with the new data and the previous "mask" to update the count and relevant metadata. At any moment there are two masks, current and previous, held in the TPH file on the Growth Analyzer server hard drive for each sample. The system server hard drives are RAID (Random Access of Independent Disks) configured; so, there are always two copies of the calculated result information, i.e. CFU, on the system. The Growth Analyzer server is not RAID configured so only one copy exists.

At the last time point the final CFU count is obtained, which is then stored to the SQL database and produces the final CFU result for the sample. As the images are very large, up to 145 TB/year/system for a fully utilized system, they are overwritten on the Growth Analyzer server when the allowed storage area is full.

When the final CFU result is calculated the final "mask" is also deleted. This equates to a Sample Result record in the RMB terminology. A System Result is a CFU count for a specific cassette.

On the Growth Direct, the sample result is a CFU count and therefore the "Raw Data" for the system. Interim or temporary data are encrypted

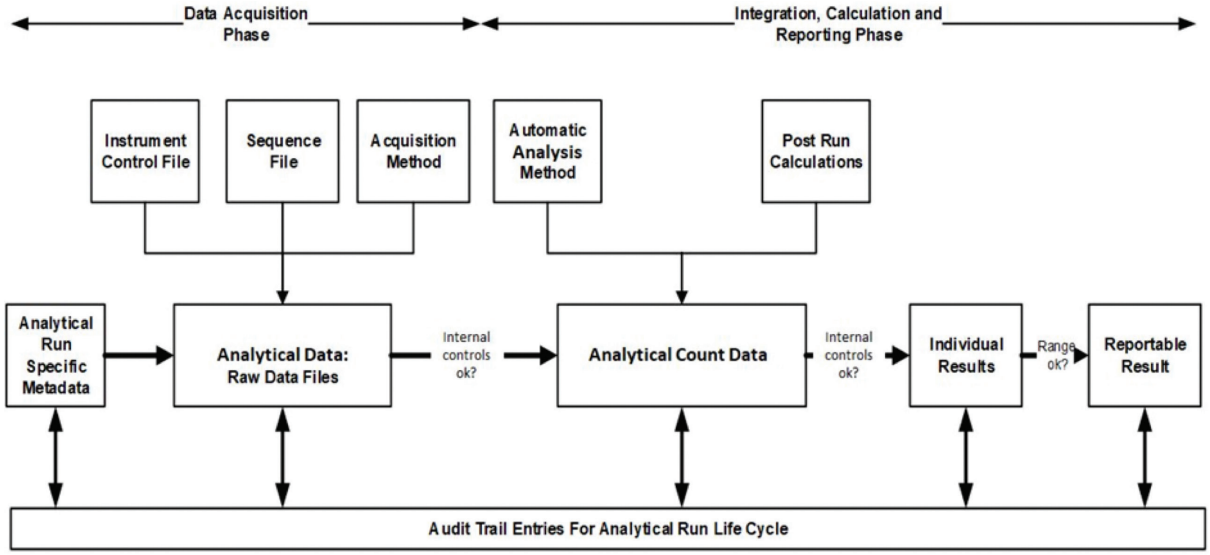

Figure 3. Examples of Data Integrity references from CFR

and overwritten. The deletion of the images could be misinterpreted as deletion of "raw data" within the context of GCP. This deletion however is considered acceptable in the light of the integrity of the software that calculates and reports the results. During the validation of the Growth Direct, the accuracy of the Growth Analyzer algorithms are assessed to verify that the organisms found in the facility are accurately detected and enumerated by the system. This is classically performed during the PQ phase. Manual counts are made of the organisms growing on the cassette after incubation and enumeration on the Growth Direct system. Equivalence between the Gold Standard of the Human Analyst and the Growth Direct confirmed accuracy of enumeration. Successful validation of the Growth Analyzer algorithms negates the need for storage or manual interrogation of the images, as the algorithms always work the same way for any sample cassette and have been proven accurate.

The Growth Direct system provides greater assurance of reporting a correct CFU count result than the manual method. Regulators demand that where an automated system replaces a manual or paper system, the automated system must be at least equivalent to the system it replaces. This mimics the current manual system where the CFU count is recorded and the plate disposed.

The CFU counts and metadata can be, viewed, printed, backed up or exported to LIMS, but cannot be deleted or modified by the user. The original metadata related to each result is stored with the result in the sequel database. The database cannot be accessed by the user other than to extract the required CFU information for printing or export to LIMS etc. via the User Interface. Not even the System Administrator can delete records from the SQL database. If the data set is deemed unacceptable, a comment can be added to the result report during the data approval process.

Summary

With increased regulatory scrutiny of the quality of data being produced by microbial QC labs under the requirements of the Data Integrity guidelines, more companies are looking to automate as much of the process as possible.

Preventing incorrect or fraudulent QC Microbiology data is not just about complying with regulations and preventing observations by the regulatory bodies and the associated remediation costs. Improving data integrity is ultimately about improving the quality of our drug supply and the in-process and facility control of the production of those medicines. The FDA acknowledges this and has been looking for ways to not only penalize companies who do not comply with regulations but is also looking for ways to incentivize drug companies with mature quality systems . Potential ways to reward companies is to use a rating system that could be used to inform drug purchasers of the quality management maturity of the facilities making the drugs they are looking to purchase. The implementation of an automated QC microbiology solution such as the GrowthDirect system would allow companies to demonstrate to the FDA, other regulatory bodies, and ultimately drug purchasers that their drugs are manufactured in a facility with a mature and robust quality management system.

References

1. https://www.fda.gov/news-events/fda-voices-perspectives-fda-leadership-and-experts/help-reduce-drug-shortages-we-need-manufacturers-sell-quality-not-just-medicine