Waters
THE SCIENCE OF WHAT'S POSSIBLE.®

Why is Electronic CDS Data a Major Data Integrity Concern for Regulators?

Tools and advice on electronic Data Integrity and how it specifically applies to chromatography systems and the challenges they present in audit situations

Heather Longden, Senior Marketing Manager of Informatics and Regulatory Compliance
Waters Corporation, Milford, MA, USA

INTRODUCTION

Pharmaceutical manufacturers are bound by regulatory agencies to follow and employ current Good Manufacturing Practices (cGMPs) for the preparation and analysis of drug products. Additionally, they have significant responsibility to demonstrate, document and file regulatory information before releasing new products to the market following Good Laboratory Practices (GLPs) and for proving clinical safety and efficacy following Good Clinical Practices (GCPs).

Analytical techniques, such as chromatography, are extensively used for measuring and quantifying components in a mixture, supporting many claims of product quality required by these GxPs. The chromatography data systems (CDS) used to capture, process and document the data have highlighted specific concerns about suspected regulatory and quality issues at some labs because the applications provided important benefits in terms of time-stamped, automated audit trails, change histories and (where used) secure electronic signatures. These technologies make data falsification more difficult and more traceable than with paper records; however, the added complexity and volume of available metadata presents its own challenges when devising comprehensive review processes.

What follows is a look at how chromatography data systems address specific concerns and challenges when demonstrating Data Integrity to an auditor or regulator.

WHY IS DATA INTEGRITY OF PARTICULAR CONCERN TODAY?

Data Integrity is not a new concern. It has been a regulatory expectation since written, and then printed, records were the norm. Today, however, the extent of metadata in electronic records is on a completely different scale; it provides significantly more evidence of a user's behavior than what would have been easily apparent in a written laboratory report.

Tools found in chromatography data systems should provide regulators additional confidence in the Data Integrity. However, as auditors and quality groups are learning how to read the metadata stored in electronic records, they are also highlighting potentially suspicious practices or those that cannot be readily explained. This is the source of today's strong focus on Data Integrity.

Unfortunately, agencies have lost trust that analysts always behave with honesty and integrity based on the additional information uncovered in the electronic records. They are now hoping that a lab's quality department will take advantage of this useful metadata to manage users' behavior and prevent falsified or even simply "polished" data. Regulatory agencies expect the quality unit and reviewers to monitor the data reported and to ensure that "testing into compliance" is not occurring.

WHAT IS DATA INTEGRITY?

Data Integrity refers to the accuracy and consistency of data, facts and statistics over a product's lifecycle. Data Integrity ensures recoverability, searchability and traceability of any original records.

While sof tware and built-in technical controls are key parts of Data Integrity, humans are the most critical variable because they create, review and approve the data.
This can be seen significantly in chromatography versus other analytical or measurement techniques that are used to create data. Chromatographic analysis relies heavily on analysts' accurate adherence to procedures while preparing samples, standards and mobile phases and ensuring the instrument and chemistries are set up correctly before analysis, as well as scientifically evaluating and potentially reprocessing the data post acquisition, before the final results can be relied upon.

The human component relies on many aspects, including:

- A culture for Data Integrity
- Governance of data and quality focused review processes
- Data uniquely associated with specific users
- Users having the skill and the training to do the job in the most accurate way possible
- Safeguards against fraud

Analysts executing poor quality separation methods require additional manual steps to generate meaningful and consistent results. Therefore, to minimize the need for human intervention, laboratories should ensure the reliability and robustness of their separations. Analytical methods must be properly validated for accuracy, precision and robustness, while chromatographic instruments should be constantly evaluated for system suitability and robustness. Instruments must be regularly maintained as well as adequately qualified or calibrated throughout their use. Standards and reagents require accurate preparation in addition to high quality and reliable suppliers. Validated and documented procedures must be in place to minimize the potential for human error (malicious or unintentional).

COMPUTERIZED SYSTEMS

At the request of regulators, Data Integrity controls are now expected to be built into chromatographic data collection applications and systems. Laboratory procedural controls should be in place for computer system validation, data traceability and periodic review of data handling. It is expected that software applications should only be run on a qualified network, should include a disaster recovery plan as well as backup and restore processes and all these aspects should be part of the validation process.

It is clear that computerized systems improve traceability and provide the capability to prevent and detect undesirable user actions by including more controls and documentation. Some basic tools for quality assurance (QA), quality auditors, and regulators include:

- Access levels
- System polices
- Audit trails

QUALITY DATA REVIEW

Because of the tools offered by compliant-ready applications, it is critical that quality reviews, as well as inspections, focus on original electronic data in their original dynamic form. Related metadata, used to determine the trustworthiness of those data, are often missing from printed reports. This missing information may result in misleading interpretations leading to quality risks. Regulators are also hiring investigators or auditors with laboratory backgrounds who understand the systems, and some are learning how a good well-controlled laboratory should function, from the laboratories that they visit.

Presenting both the good as well as the "less-than-perfect data" is necessary to demonstrate that errors are not ignored or dismissed, specifically for reanalysis and reprocessing. A proper process must be followed for a lab error investigation to determine if the root cause could be assigned to a mistake in the analysis. Only then can repeat testing be performed. If no lab error is clearly identified, a full out-of-specification (OOS) investigation should be initiated to determine the cause of a product quality failure.

GUIDANCE DOCUMENTS

Regulators need to trust the data they are presented with as this is what they rely on most to ensure the quality of work performed when they are not in the laboratory. As a result, many guidances have been written about Data Integrity and, although written by several different agencies and industry groups, they are well aligned (Figure 1).

Both final and draft guidance documents indicate that data must be ALCOA:

- **A**ttributable to a particular user
- **L**egible (clear and concise data entries)
- **C**ontemporaneous (recorded at the time of the activity)
- **O**riginal (the first recorded observation or a verified true copy of the original observation)
- **A**ccurate (scientifically valid and error-free)

In addition, data must be (+):

- Complete (including any repeat processing)
- Consistent
- Enduring
- Available

The challenge for chromatographic analysis is its complexity. As instrumentation becomes more sophisticated, printouts only summarize the data (in static form) and are not a complete representation of the original (dynamic) electronic record. Printed chromatograms do not satisfy the GMP requirements that any printed record should be a true, accurate and complete copy of every item stored as part of the electronic record.[1]

Figure 1. Data Integrity Guidance.

REGULATORY CONCERNS FOR DATA INTEGRITY

Failure to establish that lab records include complete data is a GMP violation of 211.194(a). Firms must keep all data associated with an analysis and all calculations performed whether they were correct or incorrect and whether they needed to be repeated or invalidated.

European Union (EU) non-conformance reports include observations of a) manipulation of laboratory data, b) the opportunity to manipulate data based on missing technical controls, and c) incomplete data review processes, which should be able to intercept manipulated data (Figure 2).

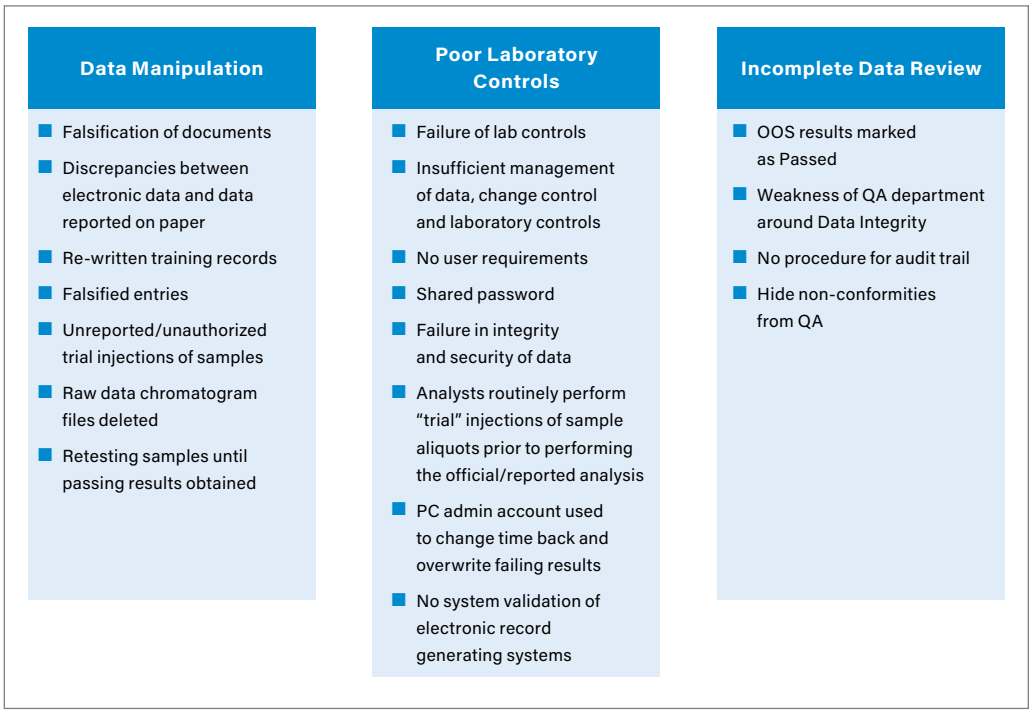

Figure 2. Summary of EU Non-Conformances.

Regulators are often starting from the assumption that data is not being captured and reported with honesty and integrity. It then becomes the job of the laboratory to prove otherwise. Key inspection themes are outlined in Figure 3.

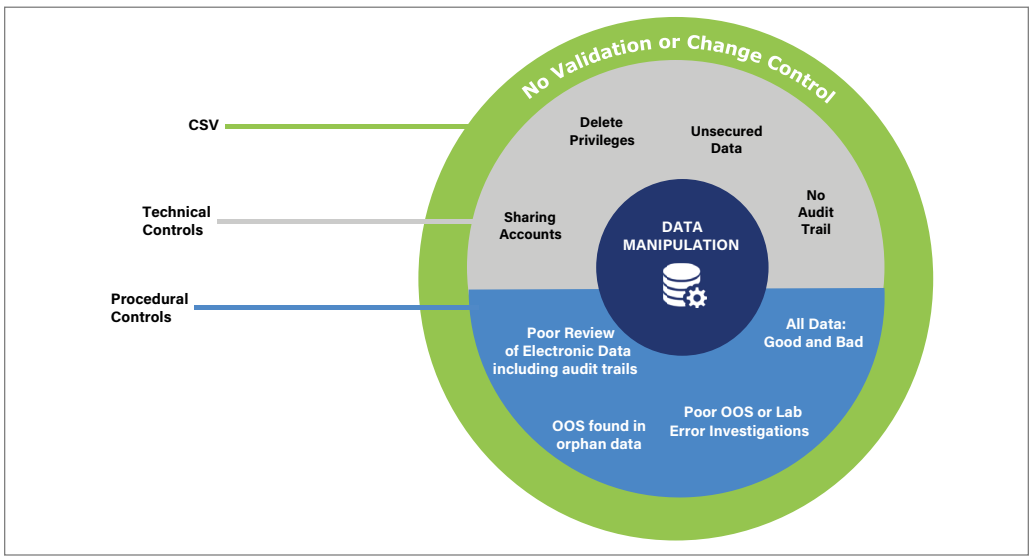

Figure 3. Inspection themes.

One way to prove integrity is through technical controls. If systems do not allow users to delete data, it becomes easier to prove that data could not have been erased. Shared accounts are also problematic for demonstrating unquestionable accountability for data creation or modification. Many laboratories are still using instrumentation with software that has no audit trails, which is a failure to meet the technical controls requirements set out in 1997.

Additionally, managers should be sure that simply hiding or ignoring data is not occurring, specifically when a run must be repeated. This might include a defined investigative process and proper scientific justification for invalidation of any data. FDA and other agencies provide detailed guidances on these expectations.

To ensure drug quality, regulatory agencies will look at and expect in-house quality units to continuously observe all reported and non-reported electronic data (orphan data):

- Are analysts cherry picking only the good results?
- Are samples being "tested into compliance" or polished to meet specification?
- Is data secure?
- Is there hidden or deleted data?

This problem is tied to OOS results, which may be either ignored or invalidated without proper justification and then simply retested. In these cases, the data review often does not include the original and all versions of results. Moreover, orphan data captured to a "test" folder without proper scientific invalidation could cause suspicion as deliberately cherry picking or making the results look better.

Properly looking for the root causes of invalidated results, whether for "in specification" or OOS results, and eliminating that root cause problem, will subsequently reduce the need for any future repeat testing. Root causes that can be addressed to prevent future failures and reprocessing include:

- Poorly developed or validated analytical methods
- Inconsistent column separation performance
- Sample, standard, reagent or mobile phase preparation errors
- Instrument failures
- Analyst error

SPECIFIC CONCERNS ABOUT THE CHROMATOGRAPHIC PROCESS: REPEAT INJECTIONS AND TEST INJECTIONS

Guidances suggest that reanalyzing or reintegrating a sample should never be required; however, tests fail for a variety of reasons such as instrument failure, lack of system equilibration, improper/expired columns, or a mistake. When a mistake is made, there is often pressure to rectify or hide the problem.

Justifications such as "I'll be fired if I admit my mistakes," "I have no time to do an OOS investigation," and "No one will notice if I'm clever about covering it up" are probably the biggest reasons why analysts attempt to hide errors in their lab from their own quality units.

It warrants repeating that there must be a scientific reason for reanalyzing samples. This should be documented in a deviation report (or similar document) and regulators are concerned if only the repeat sample set is reported. If the data is documented as a repeat, regulators/auditors want to see the original data and the scientific justification for the repeat.

Test injections may be viewed with concern if they routinely use sample preparations to ensure systems are ready for use. While it is scientifically sound that no analysis should be initiated until chromatographic systems are functioning properly, test injections from samples should not be used for this purpose. This could potentially raise a regulatory issue and suspicion of pretesting or unofficial testing of the sample. Also, analysts sometimes try to justify a failed series of injections as simply a test of the system. An independent solution or a well-characterized secondary standard, for instance, is a better choice for test injections or "system readiness checks."

If system suitability is not met, ideally the run should be aborted to ensure questionable data is not produced or collected. Alternatively, it may be sufficient to ensure that any data collected is not processed if it could not be trusted due to a system suitability failure. One way to minimize the occurrence of failing chromatographic systems is to ensure that both equipment and methods exceed robustness expectations. This would reduce analytical runs that need to be repeated.

SPECIFIC CONCERNS ABOUT THE CHROMATOGRAPHIC PROCESS: REINTEGRATION OF CHROMATOGRAMS

Documentation of why an analyst reprocesses chromatograms should be available. This might be simply recorded in the comments that form part of the required audit trail, or it may need additional documentation. However, reviewers and QA must appreciate that it is unrealistic to expect chromatography to integrate perfectly the first time every time. Unless the laboratory has very clean, robust and well resolved chromatograms, it is perfectly normal to require some optimization of integration or identification parameters for each day's run. If a laboratory gets perfect integration right the first time for all chromatograms, it may raise suspicion. If the data looks too good to be true, then it probably is.

Multiple integration attempts could indicate deliberate polishing or manipulation or at least give rise to questions, specifically if the sample or run failed in the original integration and passed when reintegrated.

Reviewing audit trails and original processed data is the only way to determine if reprocessing was scientifically required or conducted for another reason.

Automated processing (i.e., leveraging the algorithms and integration parameters in the processing method) is only an approximation of the peak integration that a good chromatographer would manually assign, leveraging their own scientific knowledge and experience. Preference may be to use software for convenience and speed of processing results with some idea it creates consistency. However, automation does not bestow a higher level of quality on the integration.

Processing parameters must often be adapted by analysts to get the most accurate peak integration, especially if the run includes very disparate levels of component concentrations. A single accurate set of parameters to automatically process the entire data set cannot be easily derived. In such cases, manual integration may be required for individual runs to ensure accurate integration.

The alternative practice of optimizing integration parameters to a new version of method, for each and every sample, is rarely viewed as good practice. Confidence that calibration standards, system suitability chromatograms and sample analyses are all processed using the same set of processing parameters is expected. Some CDS applications, such as Empower, will rely on the assumption that standards and samples will be processed using the same version of the processing method.

Saving each version of results is a key element that the FDA guidance includes. Each reprocessing or reintegration is part of the GxP record, and should be reviewed to ensure that subsequent iterations were not processed to polish or hide OOS results. It may also be possible in the CDS to obscure from the analyst the effects of integration changes to calculated values so as not to influence the placement of baselines, either automatically or manually.

Forcing lab processes that only allow automated processing of chromatograms will result in staff spending large portions of their day programming integration events to ensure that the resulting peaks are integrated correctly, with no obvious indication of manual intervention. Complex parameters and timed events in an automated integration process can ultimately be equivalent to manual integration (such as the "forced peak start" event). The degree of manipulation that can be done under the auspices of an "automated method" might be as customized as a manual integration activity could produce. In this case, the degree of human intervention is of a similar level, and yet the casual reviewer will not easily see how manipulative the analyst has been. Clearly and transparently using manual integration may well result in higher level of quality.

The placing of baselines, specifically for unresolved peaks, should always follow expectations consistent with the method as it was validated. Each day's analysis will not be identical to the previous day. Therefore, a clear procedure for adapting the integration to the raw data should be expected with appropriate levels of oversight.

A quality method with good resolution enables the analyst to have a processing method that performs integration reproducibly the first time. Training on how to use the integration parameters is essential as well as having well-understood procedures for processing and reprocessing chromatograms including good examples of the expected integration for each individual method (product or analyte specific). Reviewers should pay special attention to data that is reprocessed, whether with automated algorithms, with highly customized integration events, or manually.

TRACEABILITY

Audit trails should be included in the electronic meta data and be an integral part of the review process. It provides history and supports trust for the results being reviewed. The level of review and oversight that audit trails provide also deters analysts from using shortcuts in the system or manipulating the data.

Current chromatography data systems offer an internal database, which is an important traceability tool. Chromatography systems equipped with Waters® Empower® Software can link all aspects of metadata together into a traceable solution to ensure that metadata links can never be broken (Figure 4).

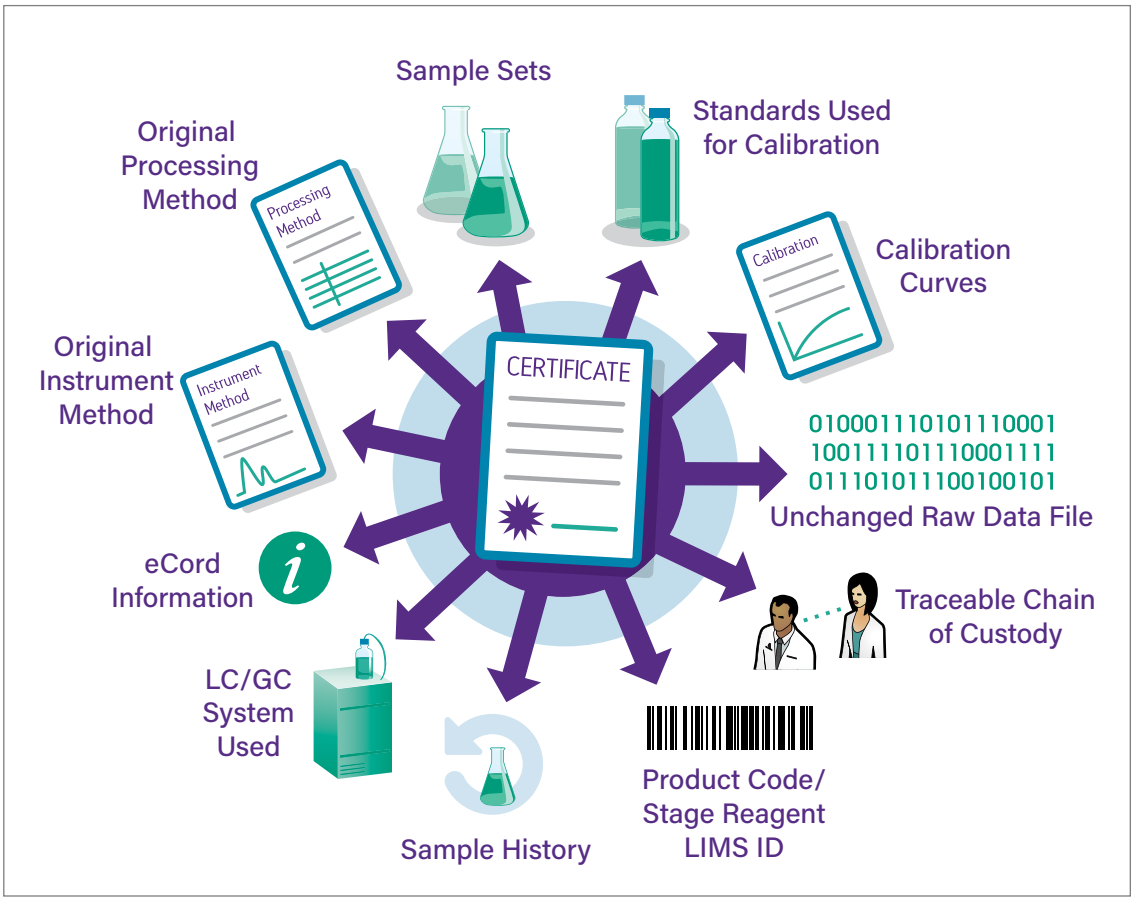

Figure 4. Empower traceability.

SUMMARY

Chromatography data systems capture important information (or metadata) for electronic records including audit trails which leverage time stamps and change histories. To ensure product quality, the metadata should be regularly reviewed by quality control staff to manage users' behavior to prevent generation of falsified data – either maliciously or inadvertently.

Establishing a culture where laboratory staff are empowered to raise and act upon concerns about product quality issues, analytical method improvements or workflow enhancements is essential. Equally, imposing unreasonable barriers to analytical work, in an automatic, immediate reaction to regulatory observations, might simply tempt staff to find alternative ways to achieve their work goals. Companies need to balance critical compliance measures against the practicality of the implementation and the needs of the business to ensure consistent quality of analytical results.

Reference

1. FDA, "Data Integrity and Compliance With CGMP Guidance for Industry," April 2016, https://www.fda.gov/downloads/drugs/guidances/ucm495891.pdf.

Waters

THE SCIENCE OF WHAT'S POSSIBLE.®

Waters Corporation
34 Maple Street
Milford, MA 01757 U.S.A.
T: 1 508 478 2000
F: 1 508 872 1990
www.waters.com