Waters
THE SCIENCE OF WHAT'S POSSIBLE.®

# Data Integrity for Your Laboratory Computerized Systems

Lynn Archambault, Informatics and Regulatory Compliance
Product Marketing Manager, Waters Corporation

## INTRODUCTION

Electronic data and computerized systems introduce new challenges to maintaining data integrity. With increased scrutiny from regulatory authorities, companies need to have a thorough understanding of their regulated systems and how to assess them for gaps. This whitepaper provides insight on the evolution of data integrity for computerized systems and the current regulatory focus including specific examples of how to maintain control over Empower® Chromatography Software data and how the Laboratory Analytics application provides additional capabilities for oversight of laboratory data, instruments, and workflows that cannot be easily obtained in Empower.

## DATA INTEGRITY: COMPANY CONCERNS AND CULTURE

Data integrity is the current buzz word for regulated companies and it is important to understand how data integrity is defined in order to deal with the problems that surround it. The Medicines & Healthcare products Regulatory Agency (MHRA) guidance document, released in March 2015, is a brief document covering the topic well. This document provides a short definition of data integrity: *The extent to which all data for its entire lifecycle is complete, consistent, and accurate.* The World Health Organization also has a definition of data integrity which uses the MHRA definition but adds to it using the principles of ALCOA: Attributable (who acquired the data or performed an action), Legible (can you read and understand the data entries), Contemporaneous (documented at the time of the activity), Original (first recorded observation), and Accurate (reflects what took place). As companies move to more automated laboratories with computerized systems, most of the fundamental ALCOA principles are satisfied by an automatically generated audit trail providing the who, the when, accuracy, originality, and legibility, all on the system with the raw data. As these issues begin to resolve, companies are starting to transition to other things that they need to be aware of. For example, it's important to remember that electronic records need to be complete, consistent, and enduring. Data integrity is important because it is a company's responsibility to ensure that its product(s) have the correct identity, purity, strength, quality, and safety for patients. The only way that a company is able to prove to a regulatory body and to the public that a product is manufactured and released as it is intended is by having the data to support their claim. When the regulatory agency performs an audit they can see all the good data that supports a product. Unfortunately agencies have investigated further and found reasons to question the data that supports that product or found additional data that is not reported. In most cases a company will have all the good data available as it supports their product, but what a regulatory agency wants to see is what happens when something goes wrong. Was a CAPA created, was a root cause found, and was the issue rectified? They want to ensure control is maintained at all times so that a product is the same

every time it is manufactured, no matter the location, facility, staff, date, or time.

Transitioning from written and printed records to electronic records helps to resolve the fundamental principles of ALCOA but it's still easy to get overwhelmed. Consequently it's important to take the time to plan for what to do and how the transition will be accomplished. Current processes and procedures should be reviewed to determine if there are proactive controls that could be implemented to reduce errors and risks with minimal impact to the business to make sure that the data is correct, accurate and enduring. All companies should take the time to empower employees to make sure that they understand how their actions can affect product quality. It is also important to understand what a company can reasonably accomplish with its existing resources, to decide what items can be implemented either short or a long term and to understand any constraints.

Data integrity as it relates to a company culture is also extremely important. Compliance doesn't mean that it has to be difficult and cumbersome; a company needs to ensure that compliance is part of their business goal so that employees are not tempted to avoid the correct process. Employees shouldn't be punished for revealing that something went wrong, otherwise when something does go wrong, they may not speak up because they're afraid of repercussions. A culture that promotes open communication is better able to resolve and address problems early on so that nothing gets swept under the rug, ultimately resulting in data integrity problems.

## DATA REVIEW

While printed copies may still have a place in some laboratories, electronic records really need to be maintained. The FDA posted responses to frequently asked questions and has said that printed chromatograms alone are not sufficient because they do not contain the metadata that goes into creating the chromatogram. Therefore reviewers need to review the electronic data in its entirety because the process is impractical to deal with using paper documents. Since there are technical controls in place for software systems, that are not available for printed copies, the data is more secure. All of the information can be reviewed within the system providing better assurance and understanding of the whole picture. Resistance may be encountered during the transition to electronic reviews in places where paper

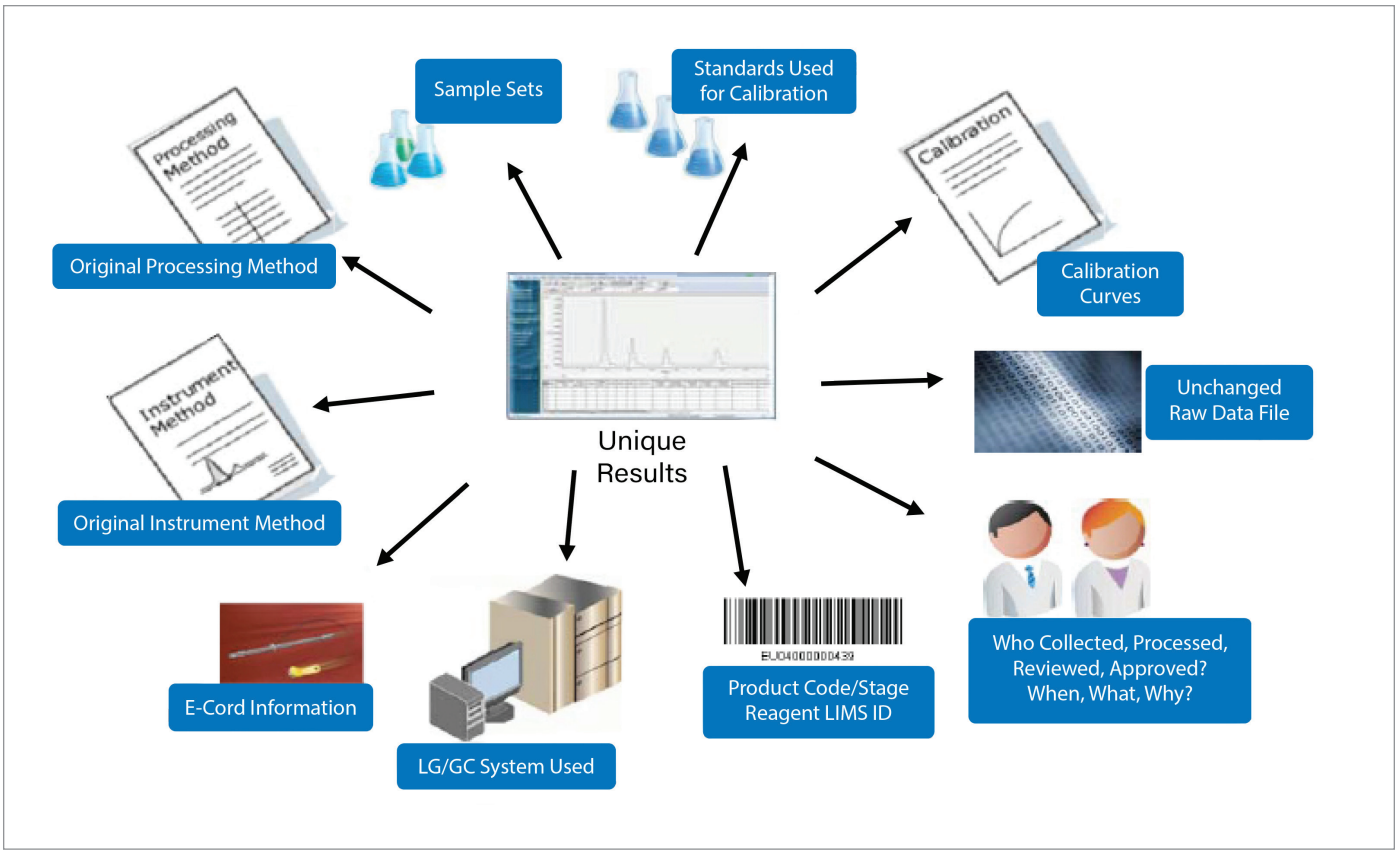


*Figure 1. Empower: Data and metadata is linked.*

reviews have historically been acceptable. As illustrated in Figure 1, to assist with electronic reviews, chromatography data systems (CDSs) like Waters® Empower uses a relational database connecting the results to all of the associated metadata. The review of associated data can be performed by performing a series of software selections and using the 'view as' option to review methods, audit trails, calibration curve information, sign offs, etc. For newer users this process may be overwhelming and it is for this reason the Result Audit Viewer was developed, as shown in Figure 2. The Result Audit Viewer (available in Empower 3 FR 2) is a single place to simplify the review process for all the audit trails, methods and method differences between versions. Electronic data review is important not just because it's a requirement, but because it's the only way to understand the process of what happened by who, when and why. A clearly outlined process and subsequent SOPs makes it easier to spot outliers and perform further investigations. Outlining a procedure in how to perform data review should involve assessment of the data lifecycle process (acquisition through reporting). Reviewing the process is typically facilitated by starting at either the end result and work backwards to acquisition or starting at acquisition and work towards the end result. For example, try to highlight the key triggers that have the most impact to the quality of your data. Maybe during review it is important to look at final results (summaries, averages, CofA) and work back through the data to understand how this information was calculated so you can focus on if there were any changes to these areas during your review. Assess what is critical to your integration, to sample set meta data, and key items in the audit trails. Once you identify where your highest risks are these are the areas to focus on with more scrutiny during data review. Try to isolate forensic investigation to only instances when it is required. A list of warning signs can be defined e.g. abnormal re-integration, an excess number of multiple results, unusual metadata changes or results that only just meet specification. A regular review is more likely to catch mistakes or deviations. This is a proactive approach to make sure that all the information is accurate allowing you to make a good, quality decisions on whether a product is suitable for patients.

## AUDIT TRAIL REVIEW

The European Commission document, Annex 11: Computerized Systems states that audit trails need to be convertible to a generally intelligible form and regularly reviewed: not a forensic approach but instead a risk-based approach. The point of an audit trail review is similar to reviewing paper documents; when data is crossed out and

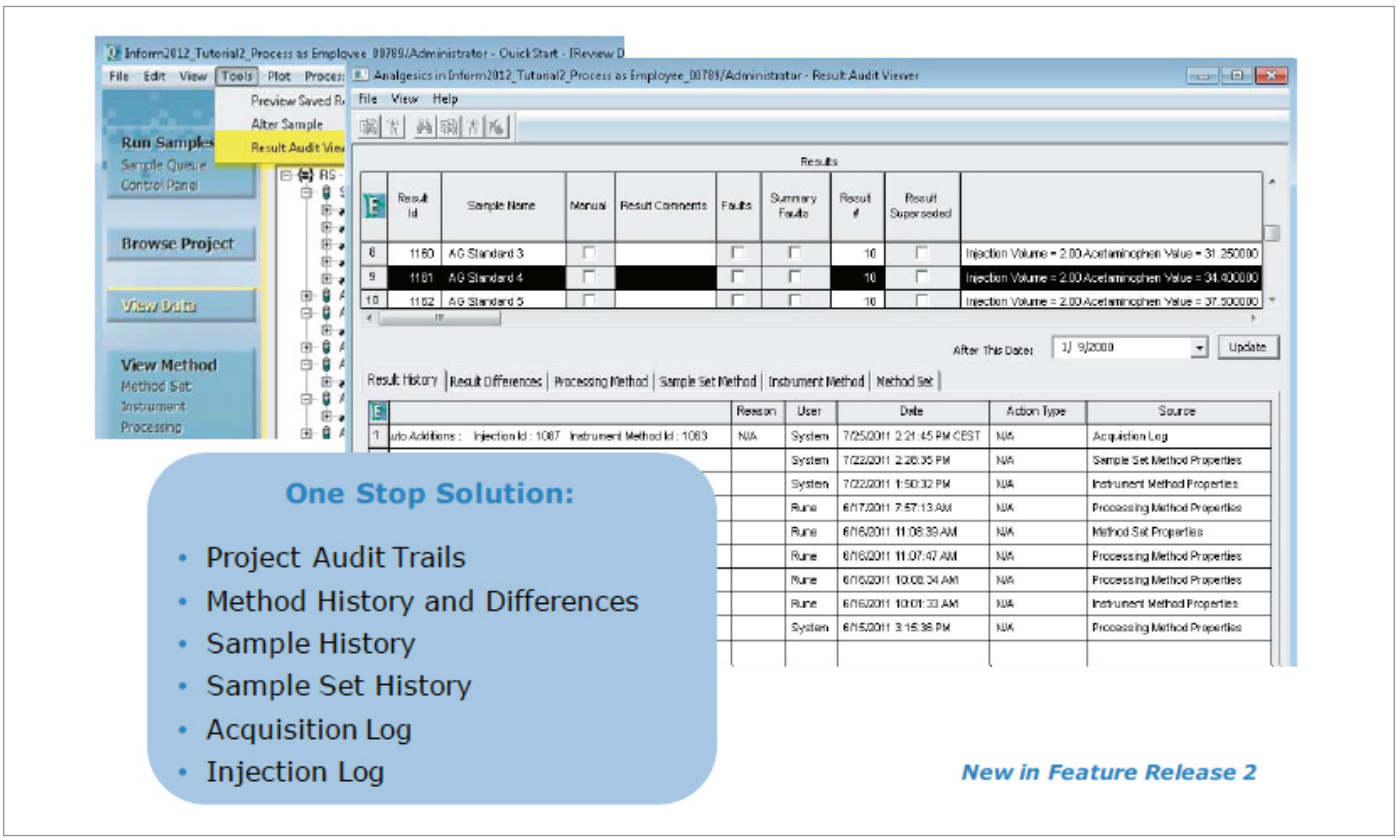


*Figure 2. Result audit viewer tool.*

corrected, it is initialed and dated, and a reason added. The same should be true for electronic records and audit trail review with the focus on important GMP changes. There are a number of items that are captured in the audit trail so it is important to understand the process and outline the highest risks. For example it may be important to review manually entered information like sample weights or dilutions, to verify that custom fields were not altered, or that the analyst did not change the method. Clarifying these points in a review SOP document helps to keep all individuals aware of what is expected and typically performed. Capturing that audit trail review was performed can be as simple as having a statement in your report that says: *I sign this data to attest I have performed/reviewed/approved this data according to SOPxyz per my outlined role and responsibilities.* This keeps audit trail review from being held to a higher standard just because it is in a computerized system. For example a user in a lab will only attest that they performed certain tasks in accordance with their SOPs, such as pipetting a certain volume of sample, and this personal declaration should be acceptable for data review as well because the individual is accountable for their actions.

Frequent reviews can uncover frequent reoccurring issues, like a processing method that is frequently changed. If something happens repeatedly it may be related to a lack of robustness of the method or an analyst training issue. In the absence of software tools, a comprehensive assessment like the FDA is asking for, can be a tedious manual process. Figure 3 shows Laboratory Analytics, a separate application from Empower, can expose manual integration events in your entire Empower System that has taken place. Having this information, such as knowledge of when manual integration events took place in every project in Empower in a single dashboard, makes it much easier to perform investigations, reviews, decisions and changes.

**"Data Integrity Dashboard"**

*Keyword Analysis*

Provides the ability to search using key words, with respect to Sample set name, sample name, project name, and custom field values.

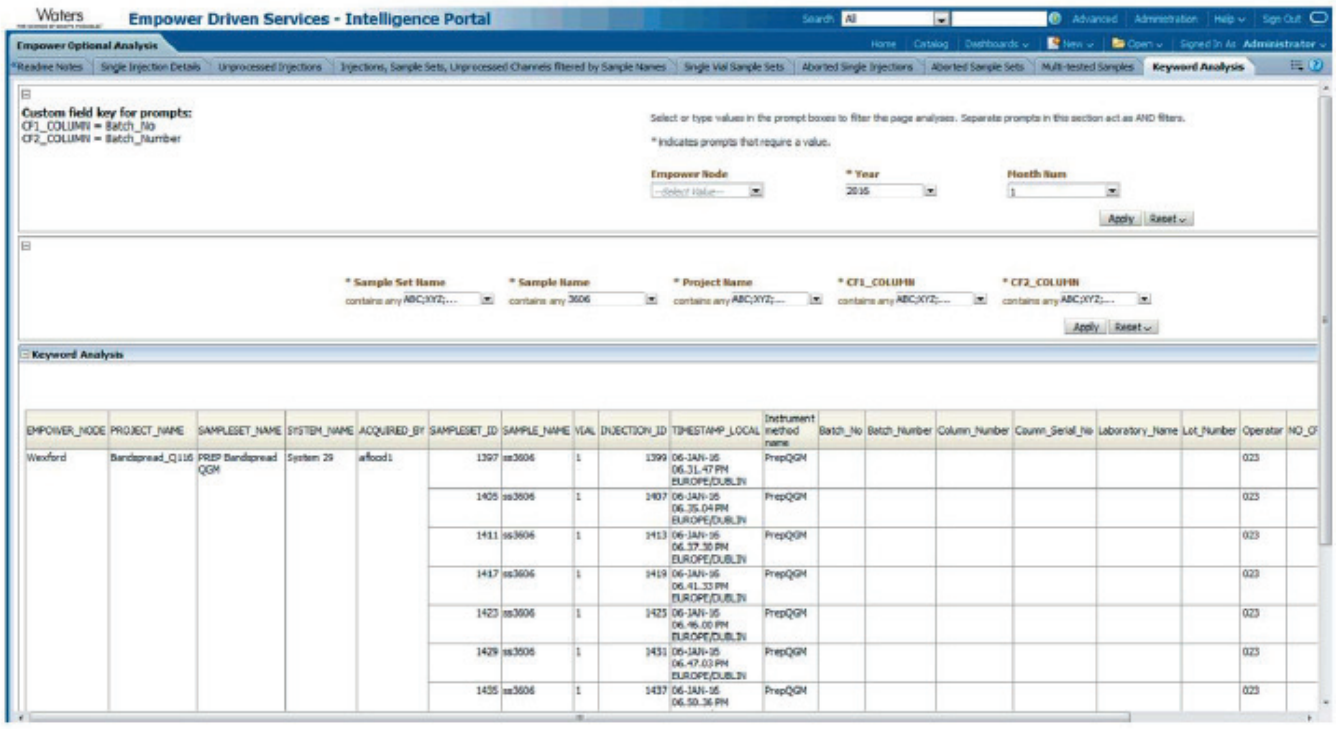


*Figure 3. How to document data review including audit trails.*

## MULTIPLE RESULTS

It's important to understand what's hiding behind a paper report. In some cases, a sample set can be processed to create a result set and then re-processed to get a second set of results, and maybe a third time with an additional set of results, etc. When possible it is very helpful to have result sets to keep everything together making it easier to report and review. It is very helpful to have a processing SOP in place. This SOP provides everyone (users, reviewers and auditors) the ability to understand the workflow. This allows reviewers to understand if results should be investigated further or not. It is important to remember that multiple processing is not wrong; a procedure that only allows data to be processed once just isn't practical, since an analyst might then sit in front of a screen for hours before processing data if they only have once chance to do it correctly without having to write a deviation. Integration parameters are important but not all individuals that are reviewing data understand the process as well as those in the laboratory and it is always helpful for everyone to understand not just the values but what good integration looks like for a method. It should also be recognized that manual integration isn't evil. Yes, it could be used to manipulate data but processing methods can be adjusted to manipulate data as well. Analysts and reviewers should understand what good integration looks like for your data so your data is always reported and reviewed consistently.

View filters can be particularly useful in determining what to actually review more in depth. View filters can provide additional information and they are easy to populate in different projects so they are available for all users. Figure 4 highlights their utility. Adding the number of results stored to the channels tab shows if there are unprocessed raw data. In the results tab by adding a result number, the number of results stored, and the number of sign offs any results that were not reported or how many times particular samples were processed can be quickly observed, getting a good picture of how well the general process was followed. Laboratory Analytics provides this information and oversight for all projects in Empower. For example, as can be seen in Figure 5, implementing a procedure for sign off reduced the review time in one department which was previously the bottle neck in the process. These numbers are quickly generated and are helpful not just for effectiveness reviews but for general business oversight and understanding.

Filter By: Review | Edit View | Update | Max Rows: 1000

Sample Sets | Injections | **Channels** | Methods | Result Sets | Results | Peaks | Fractions | Sign Offs | Curves | View Filters | Custom Fields

| | SampleName | Vial | Injection | Sample Type | Date Acquired | Channel | Channel Description | # of Results Stored |
|---|---|---|---|---|---|---|---|---|
| 1 | IQ Blank Injection | 1 | 1 | Unknown | 9/14/2015 11:48:57 AM EDT | W2996 250.0nm-1.2 | W2996 PDA 250.0 nm at 1.2 | 1 |
| 2 | External Unk. 4 | 7 | 1 | Unknown | 9/17/1997 5:37:56 PM EST | 486 | 254nm | 3 |
| 3 | External Unk. 3 | 6 | 1 | Unknown | 9/17/1997 5:30:59 PM EST | 486 | 254nm | 3 |
| 4 | External Unk. 2 | 5 | 1 | Unknown | 9/17/1997 5:24:04 PM EST | 486 | 254nm | 3 |
| 5 | External Unk. 1 | 4 | 1 | Unknown | 9/17/1997 5:17:07 PM EST | 486 | 254nm | 3 |
| 6 | External Std. 3 | 3 | 1 | Standard | 9/17/1997 5:10:10 PM EST | 486 | 254nm | 3 |
| 7 | External Std. 2 | 2 | 1 | Standard | 9/17/1997 5:03:14 PM EST | 486 | 254nm | 3 |
| 8 | External Std. 1 | 1 | 1 | Standard | 9/17/1997 4:56:03 PM EST | 486 | 254nm | 3 |
| 9 | PQ Unk. 4 | 7 | 1 | Unknown | 9/17/1997 5:37:56 PM EDT | 486 | 254nm | 0 |

Filter By: Review | Edit View | Update | Max Rows: 1000

Sample Sets | Injections | Channels | Methods | Result Sets | **Results** | Peaks | Fractions | Sign Offs | Curves | View Filters | Custom Fields | Audit Trails

| | SampleName | Vial | Injection | Date Processed | Processing Method | Result Id | Result # | # of Results Stored | Number of Sign Offs | Channel Id | Processed By |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Internal Std. 1 | 30 | 1 | 12/16/2015 8:01:29 AM EST | Internal Std Processing Method | 1240 | 1 | 3 | 0 | 1182 | System |
| 2 | Internal Std. 1 | 30 | 1 | 12/16/2015 8:01:38 AM EST | Internal Std Processing Method | 1258 | 2 | 3 | 0 | 1182 | System |
| 3 | Internal Std. 1 | 30 | 1 | 12/16/2015 8:01:49 AM EST | Internal Std Processing Method | 1267 | 3 | 3 | 0 | 1182 | System |
| 4 | Internal Std. 2 | 31 | 1 | 12/16/2015 8:01:30 AM EST | Internal Std Processing Method | 1242 | 1 | 3 | 0 | 1188 | System |
| 5 | Internal Std. 2 | 31 | 1 | 12/16/2015 8:01:50 AM EST | Internal Std Processing Method | 1269 | 3 | 3 | 0 | 1188 | System |
| 6 | Internal Std. 2 | 31 | 1 | 12/16/2015 8:01:41 AM EST | Internal Std Processing Method | 1260 | 2 | 3 | 0 | 1188 | System |

*Figure 4. View Filters: Oversight for what you need to know?*

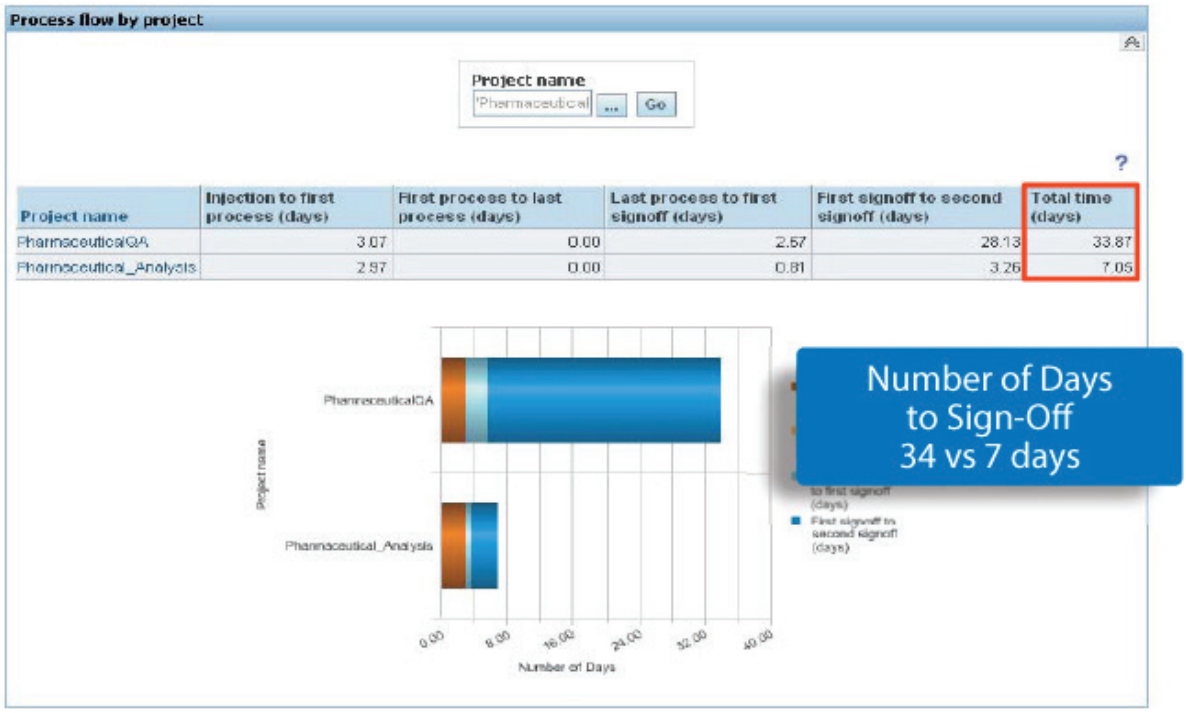


*Figure 5. Oversight of the laboratory workflow.*

**Provides the ability to:**

- Identify a list of injections/samples that are not associated to a sample set
- Filtered by Node, Year, and Month (Year is a required filter)
- View associated information, including custom field information

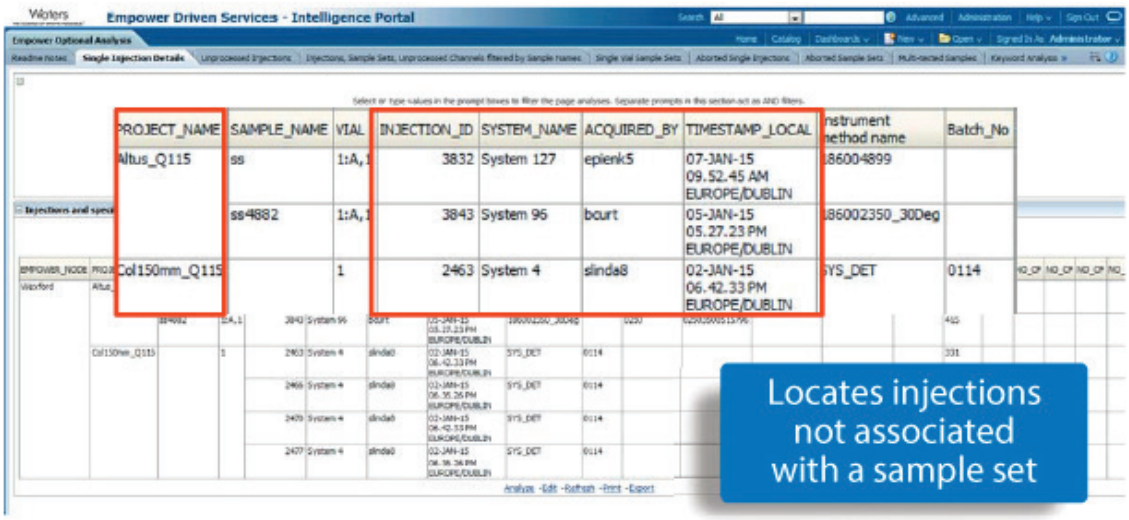


*Figure 6. "Data Integrity Dashboard" – Single injection details.*

## ORPHAN DATA

Orphan data is data that's not processed, reported or accounted for, and auditors will want to know why it exists. A well-defined SOP with technical controls in place will prevent someone from acquiring sample data in a project and then re-running the sample officially in a different project. Technical controls that include project access and project creation therefore become increasingly important. A way to limit orphan data is to have more control over providing the privileges for project folder creation and access. Project hierarchy and how a project is set up for archival purposes for the consistency of its life cycle along with the ability to retrieve it when it's needed will also play a role. The new Laboratory Analytics Data Integrity Dashboard example shown in Figure 6 can help to indicate possible orphan data. Examples include:

- Single injections
- Sample sets from a single vial
- Unprocessed channels
- Injection/sample set abort activity
- Multiple injections of the same sample when the sample name is the same

Information from these dashboards can then be used as guidance of where to look in Empower and investigate for further information about the history behind this data.

## PERIODIC EVALUATION AND INSTRUMENT MANAGEMENT

Periodic evaluation requires compiling a lot of information to show that a system is validated and in a state of control, and incident management requires that a root cause of any critical incidents should be identified. Both cases can be very tedious and labor intensive. The use of electronic systems can be a big advantage to help access information and compile it efficiently.

Periodic review is a way to assess any gaps or deficiencies in the system. Since processes, procedures, and methods are always evolving it is easier to continually assess and make minor adjustments than to leave something for several years and have to make major changes. A periodic review will facilitate continual improvements and is also a good time to review the system audit trails and search for any key changes, following an SOP. Depending on criticality, GxP use, and system maturity a periodic review timeframe may change (based upon risk). Laboratory Analytics can be used as a tool to compile information for CAPA effectiveness reviews, method lifecycle management, identification of projects or users for an internal audit, or compile information to assist in the identification of a root cause.

## HISTORICAL DATA

Planning for the future is important but Laboratory Analytics also provides ability to mine old data, for a comprehensive assessment that regulators are typically looking for. The ability to data mine all data using the same tools and dashboards allows a quick assessment of whether or not CAPA initiatives, Lean/Six Sigma implementations, or process improvements have had an impact.

## CONCLUSION

Data integrity involves technical controls, procedural controls, educated employees, and a company culture that acknowledges and rectifies gaps. Empower provides a number of build in features to help with your day-to-day activities to assist in data integrity. Regulatory agencies are pushing for a more comprehensive assessment to provide assurance that the issue is isolated or implemented corrective actions were successful. Laboratory Analytics can help to extract data and remodel it into a structure providing better access to needed information more quickly and efficiently. Laboratory Analytics is another application that searches and filters information of interest but all the data is maintained securely in Empower. While Laboratory Analytics does not replace day-to-day review it enhances oversight and capabilities to search across many projects and millions of rows of data for more comprehensive assessment on the integrity of your data and business process.

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