According to the FDA, source data should be "ALCOA": attributable, legible, contemporaneous, original and accurate

ALCOA+

Desired state		
A	Attributable	Who performed an action and when? If a record is changed, who did it and why? Link to the source data
B	Legible	Data must be recorded permanently in a durable medium and be readable
C	Contemporaneous	The data should be recorded at the time the work is performed, and date-and-time stamps should follow in order
O	Original	Is the information the original record or a certified true copy?
A	Accurate	No errors or editing performed without documented amendments
+	Complete	All data including repeat or reanalysis performed on the sample
+	Consistent	Consistent application of data time stamps in the expected sequence
+	Enduring	Recorded on controlled worksheets, laboratory notebooks, or electronic media
+	Available	Available/accessible for review/audit for the lifetime of the record

Throwing People into the Works

Human error can disrupt even the best-planned and -implemented IT system. Leadership and organizational culture can have a positive effect on data integrity.

Software applications follow logical processes and thus generally produce a repeatable outcome from a given sequence of steps – although there are occasional exceptions to this where a fault condition arises at inconsistent intervals. A process of validation can be used to give a high degree of assurance that the application, when properly controlled and used, will consistently return the same result.

Throwing people into the works – people by nature being unpredictable and prone to variability in techniques and judgment – can disrupt even the best-planned and implemented information technology (IT) system.

In P. G. Wodehouse's 1934 novel *Right Ho, Jeeves*, the phrase "He should have had sense enough to see that he was throwing a spanner into the works" is used to describe a character who is deliberately causing disruption and disorder.

The monitoring of human-error rates can be a powerful indicator of a company's error culture.

A perfect example of this can be found in an April 2015 US Food and Drug Administration (FDA) Warning Letter:[1]

> *[T]he analyst at your firm altered the file name in the spectrophotometer containing the sample identification information for (b)(4) API lot # (b)(4), tested on April 2, 2014, to support the release of two previously manufactured lots, # (b)(4) and (b)(4). . . . This practice is unacceptable and raises serious concerns regarding the integrity and reliability of the laboratory analyses conducted by your firm.*

This statement clearly indicates an analyst deliberately falsified a result in a computerized system. (It should be recognized, however, that while some GxP data changes may not be the result of intentional falsification, they also lead to data-integrity issues.)

The importance of leadership

Management responsibilities

ISO 9001:2015[2] clearly identifies one of the key roles of management: ensuring the availability of resources. This is reaffirmed in many, if not all, GxP regulations around the world.

Applying this requirement to data integrity, management must:

- Provide sufficient competent people to complete the assigned tasks: Overworked people may feel pressured to maximize yield or productivity at the expense of data integrity.

- Provide sound, reliable equipment and instrumentation for production and quality personnel to achieve the expected throughput: Outdated equipment may neither provide the technological controls for data integrity nor produce accurate data. Frequent equipment downtime can increase pressure on the staff to seek alternative ways keep up with their workload.

- Maintain the facilities and operating environment in a fit state for their intended purposes: Lack of physical security and poor IT infrastructure can themselves jeopardize data integrity by allowing unauthorized access to a server room, for example, or by losing data from a local hard drive.

  These responsibilities are in addition to providing leadership in all matters of data integrity and compliance, as effective executive leadership is a critical component in maintaining a high level of data integrity. A corporation must emphasize the importance of data integrity to the organization through word and action, including embedding the quality requirements within the business process.

  Executive leadership must encourage right behaviors by prioritizing data integrity when setting objectives, performance targets and incentives.

  Leadership should drive a strategy that focuses on prevention, detection and response. The priority of effort for prevention should be greater than the priority of effort for detection; effort for detection should be greater than effort for response. This translates into:

- Select, install and configure systems that are capable of providing the technical controls essential to protecting data integrity, such as unique accounts, granular privileges and audit trails. (A more comprehensive discussion on technical controls and data integrity by design can be found in "An Ounce of Prevention.")

- Ensure that effective review processes are in place to detect any data-integrity issues throughout the operational life. (Detailed information on results review, audit-trail review, periodic review, data audits, etc., is covered in "Big Brother Is Watching.")

- On detection, ensure that the preventive actions implemented reduce or eliminate data-integrity risks by technical or design controls (preferred) and by influencing human behavior. (This is discussed in "Doing the Right Thing.")

Leadership must first accept that there have always been – and always will be – data-integrity issues on some level. Investigating and understanding the existing data-integrity issues within an organization is a strong foundation from which to begin the process of reducing such issues.

The MHRA Data Integrity Definitions and Guidance states the objective as being to "design and operate a system which provides an acceptable state of control based on the data integrity risk, and which is fully documented with supporting rationale."[3] Once a system with inherent controls has been put in place, detection is the next essential safeguard against the daily threats to data integrity. The reporting process for data-integrity problems must be understood from the top level all the way down to the line operators, and it must come with immunity from management censorship or retribution.

Metrics

Poorly chosen metrics can undermine integrity by encouraging the wrong behaviors and potentially providing the "pressure" element envisaged by Donald Cressey in his hypothesis on fraud[4] and pictorially represented in the "Fraud Triangle" (see Figure 1).

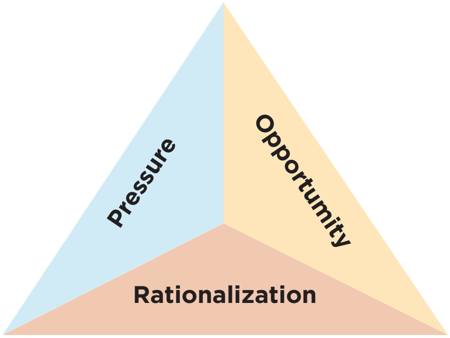


Figure 1 The fraud triangle

When such pressures are combined with the opportunity for data falsification presented by poor technical controls, it can be just a small step further for an employee to rationalize that altering the data is a minor misdemeanor and may even save the company money in the long term. At this point, the employee now has the motive (pressure) and ability (opportunity) to commit fraud and has even convinced himself or herself that it in the company's best interest to do so (rationalization) – when, in reality, fraud can only be detrimental to both the company and the employee.

As an example of pressure resulting from metrics, some companies may determine and monitor the throughput of the laboratory performing quality-control analyses. If the lab's performance is measured through the number of samples analyzed during a time period, then there is no pressure on the analysts relating to the pass or fail status of the samples analyzed. This prevents any temptation to "encourage" samples to pass but could give rise to poor-quality sample and column preparation as the analysts have no incentive to care about the result.

Redefining the metric as the number of passing samples in a time period, however, may provide substantial motivation for the analysts to make samples pass by whatever means they can in order to return a high efficiency, especially if there is potential for a pay rise or promotion linked to this.

A carefully chosen metric may involve the number of samples analyzed in a time period, but it would also need to factor in any incorrect test results as detected by second-person review or even repeat testing as part of an investigation.

Falsification for profit is discussed in more detail in "Doing the Right Thing," as is the use of positive metrics linked to rewards.

Cultural considerations

Cultural considerations can refer to a corporate culture (that is, the paradigm within which an organization operates) or a geographic culture (the moral and behavioral norm within a particular country or region).

Corporate culture

Corporate culture can vary widely, from a family-owned private company to a publicly traded corporation with an independent board of directors that comprises leading industry figures and subject-matter experts.

From a regulatory perspective, there is no difference: The expectation for data integrity and product quality remains the same. The publicly traded corporation may, however, by its very nature lend itself to significantly more transparency than the family-owned private company:

- The corporation may be subject to Sarbanes–Oxley or other financial audits that could identify any corporate culture of adverse data practices.
- There are no family loyalties and potentially fewer conflicts of interest involved in the corporation if an employee reports a data-integrity concern outside of his direct reporting structure.
- The corporate directors should consider the impact of any company activity on their individual industry reputations.

It should be noted, however, that a larger corporate business may suffer from:

- A level of inertia that must be overcome, especially when it is required to update the quality system and the way of working to mitigate (perceived or real) gaps in the quality system
- A lack of crossover knowledge, such as having more resources dedicated solely to "quality functions," but such specialism may restrict an understanding of laboratory processes

Small start-up companies, common in the fields of biotechnology, sensing, and software development, have their own unique challenges:

- Little or no segregation of duties – all personnel have multiple roles
- Minimal independence and impartiality of departments
- A reliance on improvisation and innovation to work around problems
- An immature, and possibly incomplete, quality management system
- Potentially less focus on specific industries (particularly in a software start-up)

A company looking to succeed and grow should be amenable to input and suggestions from its customers, including ways to strengthen its data-integrity approaches.

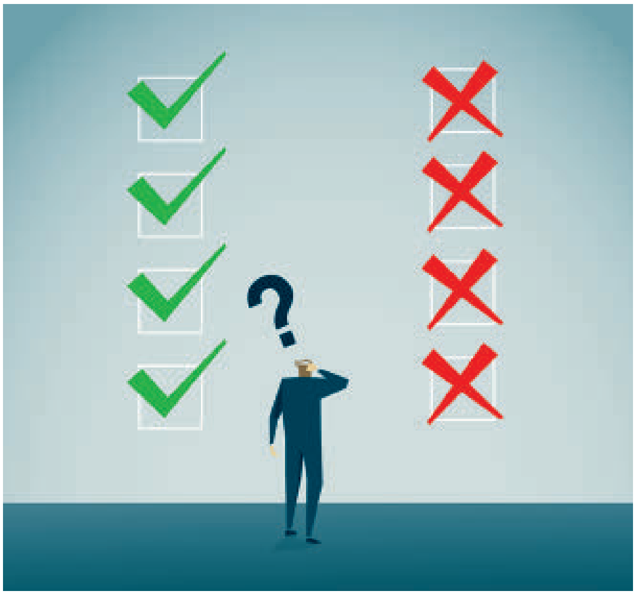
Geographic culture

Even in today's global society, geographic culture has a significant impact on site operations. There are many published works on geographic culture available; some of the cultural classifications in this section were taken from *The Culture Map*, by Erin Meyer.[5]

Cultures based on an egalitarian style with consensus decision making – as found, for example, in Scandinavian countries – may have a natural advantage in promoting data integrity. Openness and a willingness to discuss difficult situations can support an environment where failing results are seen as a group problem to be resolved with clearly documented corrective actions that mitigate the manufacturing or other root cause.

Similarly, people from cultures that tend toward direct negative feedback, such as in the Netherlands, will likely feel comfortable escalating an issue through the management structure.

In a more hierarchical society, especially one that intuitively uses indirect negative feedback, as might be found in highly traditional cultures like Japan or China, reporting an out-of-specification result could be seen as either a personal failing on the part of the analyst or even an implied criticism of the manufacturing department. Such cultures will have to invest significant effort to consciously overcome traditional thinking in order to achieve the openness around data integrity that is needed for compliance.



Human error

"Doing the Right Thing" focuses on intentionally fraudulent actions that undermine the integrity of data; it is, however, important to recognize that such actions are thankfully in the minority and that data is more often affected by genuine human error.

Minimizing human error

In his three-part article "Optimizing Human Performance," Gerry McAuley sees human error as indicative of failures in the systems and processes within the organization.[6–8] When transparent, open investigations are conducted to determine the true root cause – which may be a combination of failures across a number of individuals and processes – and followed up with effective solutions, the incidence of human error can be reduced.

McAuley proposes moving from the current and pervasive mindset that human errors should be dealt with by "reprimanding, retraining, adding extra lines to SOPs, and thinking people just need to read them" to a paradigm based on openness and a real understanding of people and behaviors and ultimately to a corporate culture where "individuals who try to hide, ignore, or respond inappropriately to perceived human errors are not able to exist in the business."

The monitoring of human-error rates can be a powerful indicator of the company's error culture, with a consistently high incidence of error changing little over time showing that mistakes are accepted as inevitable with no effort made to improve working practices.

Effective executive leadership is a critical component in maintaining a high level of data integrity

Effective mechanisms to reduce human-error rates include (most effective first):

Use people less: Increased use of direct interfaces between systems in place of human manual transcription should mean less human error.

Use people only for their strengths: Humans are very effective at monitoring multiple systems simultaneously, whereas it would require a highly complex automated system to achieve the same monitoring function. The data in Table A, however, shows that humans are naturally poor at manual data entry, so this should be avoided by implementing the direct interfacing of equipment and automated transfer of data.

Limit opportunities for human error: Use drop-down lists in place of free text entry, for example, so that searching for a particular product name will not fail due to a spelling error.

Human error rates

Professor Raymond Panko at the University of Hawaii has been collating data on human-error rates and has uploaded key figures to his website; a small selection of that data has been reproduced here. It should be noted that even a second-person review will not necessarily catch 100% of the errors present and so the actual error rate may be higher than quoted here (see Table A).

Interestingly, more recent data from Potter[9] seems to suggest that entering data in a more critical system – in-flight management, for example – does not lower error rates, as one might be expect given the perceived importance of the situation; it can actually give a worse error rate than situations without such pressure. Alternatively, the increased error rate could be attributed to less accurate keyboard input from users accustomed to word processing and spell-checking to correct errors compared to the necessity for high accuracy among professional typists using manual typewriters in the earlier studies (although spell-checking itself can create errors when it "corrects" a word erroneously and thus changes the meaning of the statement).

Table A Selected error rates in data entry

Scenario	Error Rate*	Researcher, Date
Expert typist	1%	Grudin, 1983
Student performing calculator tasks	1–2%	Melchers and Harrington, 1982
Entries in an aircraft flight management system, per keystroke; higher if heavy workload	10%	Potter, 1995

* Detected by second-person review

Table B Selected error rates in spreadsheet development

Summary	Error Rate*	Auditor, Date
50 spreadsheets audited; 0.9% of formula cells contained errors that would give an incorrect result	86%	Powell, Baker and Lawson, 2007
7 spreadsheets audited	86%	Butler, 2000
22 spreadsheets audited, only looking for major errors	91%	KPMG, 1998

* Percent of spreadsheets with detectable errors

A regulator does not distinguish between human error and data falsifications when assessing the impact of a data-integrity failure.

It should also be noted that Potter's study found that the error rate increased with a heavy workload, which reinforces the message in the section on Management Responsibility: It is essential to have sufficient staff to manage the workload and preserve data integrity.

Panko has further researched error rates in spreadsheet programming. In his article "What We Know About Spreadsheet Errors,"[10] he leverages experiences from financial spreadsheet audits by lead auditing companies to compile an error rate for spreadsheet development (see Table B).

While it may not be feasible for companies to audit all of their data entry in such a formal and controlled fashion using an outside company, careful tracking and trending of the findings from properly conducted root cause investigations should be able to provide some measurable metric around the incidence of human error within the company. This metric can then be monitored to measure the efficacy of data-integrity activities as part of the company's ongoing commitment to quality.

When discussing the incidence of genuine human error, it's important to note that a regulator does not distinguish between human error and data falsifications when assessing the impact of a data-integrity failure.

This is clearly evident in a January 2015 FDA Warning Letter:

> *In correspondence with the Agency, you indicate that no malicious data integrity patterns and practices were found. Also, you state that no intentional activity to disguise, misrepresent, or replace failing data with passing data was identified and no evidence of file deletion or manipulation was found. Your response and comments focus primarily on the issue of intent and do not adequately address the seriousness of the CGMP violations found during the inspection.*[11]

This statement shows that the FDA does not make allowances for how the data-integrity issues occur; it only cares that the issues have occurred and may impact product quality and patient safety.

Conclusion

Corporate leadership, corporate culture, and geographic culture all have a significant impact on the integrity of data. Strong corporate leadership should provide the paradigm to improve data integrity. Furthermore, implementing an effective framework of administrative safeguards and technical controls – examined in "An Ounce of Prevention" – should minimize genuine human error and ultimately reduce opportunities for deliberate falsification. ■

Charlie Wakeham and Thomas Haag

Implementing a Corporate Data Integrity Program

This article provides a condensed version of a presentation the author made at the ISPE Europe Annual Conference, 7-9 March 2016, in Frankfurt, Germany. Both the article and presentation are compiled from materials developed by the ISPE GAMP® Data Integrity Special Interest Group. Both also borrow from "Considerations for a Corporate Data Integrity Program," a recently published ISPE GAMP Community of Practice concept paper that shares implementation considerations based on the experiences of several companies, including successes and challenges. Although the specifics of each company's data-integrity program are different, the considerations described provide direction for creating a successful corporate data-integrity program.

A well-defined strategy is the cornerstone of any data-integrity program.

A well-defined strategy is the cornerstone of any data-integrity program. To design and implement a successful program you must have a keen understanding of your current state of affairs and business process knowledge; you must also make sure that those processes support your data-integrity requirements.

The assessment activities outlined below can serve as a basis for defining and establishing your strategy. The high-level plan presented here will define the approach, timeline, resource requirements, and rationale required to execute your data-integrity program. It may also provide a means to track progress for senior management reports, as well as a documented rationale and plan to outline your program and actions during audits and inspections. Finally, it outlines a method that can help align multisite activities and provide a holistic approach to compliance.

At a minimum, a well-defined strategy demonstrates your commitment to managing data-integrity issues within your company and creates a corporate governance oversight process.

Identifying and establishing executive sponsorship is crucial to getting support for your data-integrity program. The sponsor is responsible for the program's overall success, and will be required to set direction, define priorities, provide resources and break down organizational barriers. The sponsor will also help executives be aware of the four key benefits that a data-integrity program can deliver: financial, risk reduction, regulatory, and legal product liability.

What are the critical success factors?

Management accountability

While a successful data-integrity program requires cross-functional oversight and participation, management accountability at all levels of the corporation – from the CEO to operations floor supervision – plays a key role in ensuring data integrity. Managers should "walk the talk" and personify integrity in response to a failure. They should foster an environment in which employees are encouraged to identify and report data-integrity issues on the shop floor. They should never incentivize data falsification and should always discourage the "wanting-to-please" mentality that can lead to data corruption.

Accountable managers also provide the appropriate resources to ensure data integrity – including people, capable instruments and systems, along with sound and understandable business processes. They acknowledge that data-integrity issues will occur, and that human error contributes greatly to data integrity issues. And they drive a strategy that focuses on prevention, detection and response.

Knowledge sharing and training

As you roll out your data-integrity program, sharing and addressing a number of questions will help build a good data-integrity foundation across your organization. These include, but are not limited to:

- What does data integrity mean and how does it apply to my day-to-day business activities?
- What role do equipment qualification and computerized system validation play in data integrity?
- How does data integrity relate to 21 CFR Part 11 and EU GMP Annex 11?
- What are our roles and responsibilities? What are those of the regulatory agencies?
- When does data integrity start and when does it end?

It's important to make information readily available to all levels of the organization. Employees from the executive suite to the shop floor should have appropriate levels of knowledge and accountability about data-integrity requirements and expectations.

Establishing a data-integrity knowledge repository or knowledge base is a great way to provide historical and current information. Consulting subject matter experts both within and outside of your organization early in the process is crucial to establishing an appropriate knowledge foundation.

Data integrity should be inherent to your processes, so that it can provide a foundation for more focused training. Data handlers should be trained to understand that they are data-integrity stewards. They should understand the business processes and the data they generate. They are responsible for identifying and escalating concerns regardless of the effect on delivery, quotas, or timelines. Those in quality and compliance roles should have advanced training to ensure that data-integrity requirements are implemented within systems and processes, and that they support the business processes and business owners.

Are your controls in place?

Quality management system

Data integrity and data governance are an integral part of your quality system. It's appropriate to start with organizational and procedural controls, therefore, when designing a data-integrity program.

Does your quality management system (QMS) adequately address the regulatory requirements associated with data-integrity? An assessment will identify any procedural controls that might be lacking. Do adequate processes exist within the QMS to prevent, detect, report, and address data-integrity failures? Are the ALCOA+ requirements clearly addressed within the QMS? Are there adequately defined processes to generate and review data? And are there proper controls for the entire data life cycle? If you have a good and well-defined corporate QMS aligned with current GxPs, most of

Data integrity and data governance are an integral part of your quality system.

these items should be addressed and traceable to the appropriate regulation applicable to your business processes.

Organizational gaps are more likely to be identified as sites and local business areas define and execute their local procedures, however, a more detailed gap assessment may be required to truly understand the state of data-integrity controls in place at this level.

Corporate quality culture

This leads to another control you should assess and understand: corporate and quality culture.

Just as behaviors can promote appropriate actions and foster an environment that champions integrity, the opposite is equally true: Cost-saving measures may encourage password sharing due to limited user license purchases; poorly conducted investigations may blame human error or find no assignable cause. Changing a standard operating procedure (SOP) may be proposed as a preventive action, but all too often it can be ignored and not truly address the root issue.

Poorly chosen metrics can also undermine data integrity. Metrics that encourage pressure, opportunity, and rationalization can support fraudulent practice and may encourage data-integrity issues. Emphasizing speed rather than accuracy and quality, for example, can force employees to cut corners and focus on the wrong things.

Technology

As with organizational controls, you must also assess technical controls, which include your equipment and computer systems. Are these properly qualified and/or validated to ensure data integrity? All too often, systems are not qualified, designed, or configured to ensure data integrity. System access and security should be properly defined and audit trails properly utilized to review, detect, report, and address data integrity issues.

Compliance

Understanding how organizational and technical controls are executed and applied in your business processes is critical. An audit or self-assessment process should monitor compliance with your QMS and the regulatory requirements of your business. A quick measure of data-integrity compliance can be taken with a review of the self-assessment, internal audits, and third-party reports and observations associated with these activities. What types of data integrity issues exist? Are there repeat findings related to data-integrity issues? Are there systemic issues and do they stem from a corporate or quality culture issue?

Of course it is only possible to review this data if these self-assessment and audit processes are designed and able to identify data-integrity risks and gaps. They should utilize forensic audit techniques and focus on data-integrity compliance issues. This will be critical to the long-term monitoring and overall effectiveness of your program; it will also help ensure you are identifying and addressing data-integrity issues before regulatory inspections find them.

If you are fortunate enough to have received an inspection visit from a regulatory agency that has implemented forensic data-integrity inspection techniques, you will be able to use the results of that visit as yet another indication of your acceptable state of control of data-integrity risks. Otherwise, a review of regulatory observations from other companies can

Worth the effort?

This question is often asked as companies determine how to address data integrity within their organizations. The MHRA GMP Data Integrity Definition and Guidance for Industry March 2015 provides some interesting perspectives related to this question.

It states that “Data Integrity is fundamental in a pharmaceutical quality system which ensures that medicines are of the required quality.” It goes on to say that “The data governance system should be integral to the pharmaceutical quality system … ” So there is clearly an expectation that companies address data integrity and data governance in their pharma quality system because it is fundamental to ensuring product quality.

Does this mean that companies must implement elaborate and highly resourced programs to address data integrity? The MHRA guidance further states that “The effort and resources assigned to data governance should be commensurate with the risk to product quality and should be balanced with other quality assurance resource demands.” So the effort and resources should be aligned with the risk and with other quality demands.

It also states that “As such, manufacturers and analytical laboratories are not expected to implement a forensic approach to data checking on a routine basis, but instead design and operate a system which provides an acceptable state of control based on data integrity risk, and which is fully documented with supporting rationale.” The emphasis is on designing and implementing a system to provide an acceptable state of control based on data integrity risk.

The MHRA guidance also says that “consideration should be given to the organizational (e.g., procedural) and technical (e.g., computer system access) controls applied to different areas of the quality system” and the “effort and resources … be commensurate with its criticality in terms of impact to product quality attributes.”

provide insight into current trends and concerns. Data-integrity observations issued for one site are potential indicators of issues in other sites. You should determine if similar issues exist within your sites and develop action plans to close any gaps. There is no faster way to lose the trust of a regulatory agency than to have the same issues identified at multiple sites within your organization, as it highlights not only the possibility of a systemic issue, but corporate and quality culture issues as well.

Define your metrics

Defining and establishing appropriate data-integrity program metrics are necessary for two reasons: First, it ensures a positive return on investment. Senior management that invests time, money, and resources into a program expects a return on that investment, otherwise why invest in the first place? Second, metrics also measure the success of the program and demonstrate progress against goals.

In early stages of the program, reporting of data-integrity issues will increase with increased awareness and improved detection, which may skew the metrics. It is important to manage this "bad news" and continue to foster an environment of open reporting. A program-reporting process will also bolster success. Your plan should define the reporting expectations to senior management, area business leadership, the program team, and those on the shop floor. It is an opportunity to share metrics and progress to date, as well show progress against the plan. It also identifies and communicates issues and provides a mechanism to agree on next steps.

Audit your processes

Audit processes also are critical to the success of the program. Multiple types of audits should be conducted, including, but not limited to:

- Initial gap assessment or audit of nonconformance
- Periodic audit of long-term data archives
- Supplier qualification
- Closeout gap assessment or full audit following program completion
- Ongoing internal quality audits of established data integrity controls to ensure continuing effectiveness and compliance

These will provide critical information to set a baseline and measure success, as well as highlight possible gaps, corrections, and additions to your project scope. For initial and closeout assessments, consider using an independent auditor. (This does not necessarily mean an outside expert, but someone independent of the internal core team.)

Conduct review processes

A final key to successful implementation of a data-integrity program is defining and implementing a robust review process, including result reviews and periodic reviews.

Result review: Results review is defined as the review of individual results or sets of results conducted prior to making an accept/reject decision about product or data quality. It should compare results against specifications, limits, and acceptance criteria. It should also evaluate the completeness and correctness of metadata. The review process makes room for judgment about the accuracy and integrity of any manually entered values; it also reviews information associated with any decisions or actions taken.

The reviewer shoudl assess and understand the effect that any manual adjustments or alterations of the data or metadata might have on the results or product decision, and should also be aware of any changes to method versions used in creation of the result. The reviewer should also assess the results' conformity to sound scientific practice and documented procedures. Increased review rigor should be applied for manual adjustments and/or results that barely meet specifications.

Three Key Factors to Consider

I am frequently asked "How much effort is required to implement a corporate data integrity program?"

MHRA GMP Data Integrity Definition and Guidance for Industry March 2015 makes it clear:

> *The degree of effort and resources applied to the organizational and technical control of data lifecycle elements should be commensurate with its criticality in terms of impact to product quality attributes.*

My typical answer is, "It depends," because three factors should be considered to develop your initial data-integrity strategy and define your corporate program:

First: What were the outcomes of the gap assessments and audits of your organizational controls (i.e., your QMS and procedures)? If significant gaps exist, then a greater effort will be required to address the integrity risks. This may also result in the creation of site and/or local procedures to implement the new controls and processes.

Second: What were the outcomes of the gap assessment and audits of the technical controls associated with equipment and computer systems? This could result in updates, reconfiguration, or even replacement of a number of systems, all of which must be qualified and/or validated. Depending on the extent of the changes to these systems, the amount of effort and resources required will vary by projects and/or system.

Third: Are there gaps associated with business processes and their execution? These are typically identified by conducting a detailed business process review and gap assessment with the people responsible. Changing business processes is not always easy, especially when they have been in place for a significant period of time. Mitigating gaps may require changes in procedure and the organization's quality and business culture. Training may be required to support the changes. As always, management accountability and support is critical and will have a direct influence on successful implementation, especially when dealing with multiple sites.

Understanding how organizational and technical controls are executed and applied in your business processes is critical.

The result review should not overlook the audit trail review,[1] which provides the most effective means of assessing data integrity. Unfortunately, in some cases the audit trail is not easily accessible or permanently associated with the result, making the review difficult to complete and data-integrity issues difficult to detect.

Appropriate and accessible audit trails can prevent and detect data-integrity issues, but reviewing the audit trail and metadata associated with the volume of results generated in today's business processes can present logical and resource challenges. Technology controls implemented within many systems, however, have provided a means to review by exception.

This applies risk-based methodology to data review based on alerts highlighting a subset of results that require additional detailed review; these may be results and data that are within but close to the specification limit, have been manually manipulated (i.e., integration), or have been reprocessed. These types of systems also require validation to verify and document the alert functionality.

Periodic reviews: Computer systems require periodic reviews to ensure they continue to operate in a matter consistent with their intended use and remain in a validated state consistent with that use. GAMP® 5 is a great resource that outlines the concepts of periodic review. From a data-integrity perspective, periodic review should include evaluation of any changes to system configuration that could affect data integrity. It should also focus on any data deletions, including what was deleted, why, and by whom. In addition, the review should target system administration activities and user accounts, especially accounts disabled following unsuccessful login attempts.

Other periodic review activities include SOP review to ensure that appropriate data integrity controls are addressed, system validation records are current and reflect the intended use of the system, required SOP records are maintained, change control process is functioning properly, and system performance is not affected negatively by the intended use of the system. ■

Michael Rutherford

An Ounce of Prevention

The administrative and technical controls needed to mitigate risks to data integrity prove Ben Franklin's maxim that "an ounce of prevention is worth a pound of cure."

Computerized systems' functionality is based on a combination of hardware, software, processes, personnel and environment. When such systems are used for the collection, storage, sharing, use and archiving of regulated data, the following guiding principles will apply:

- Data should be collected, stored, shared and used only for legitimate business purposes.
- Data should be collected, stored, shared and used in a secure manner.
- Any data that is to be shared externally must be transferred by secure means.
- Active, responsible data stewards should be assigned to all critical data.
- Users should only have access to the data needed to do their jobs and should be granted access levels commensurate with the requirements of their jobs.
- Data, as well as any associated metadata that provides content and meaning to the data, must be retained for the relevant retention period.

- Any change to critical original data must be recorded in an audit trail. This should capture who made the change, the old and new values, the date and time of the change and the reason for the change.
- All data users should be appropriately trained on requirements related to data collection, storage, sharing, integrity and use.
- When the same information is available from multiple systems, the authoritative source of the information should be documented in the quality system and some effective mechanism put in place to ensure that the other systems are updated and remain consistent with the authoritative source.
- Specific definitions for and the purpose of data collected in electronic systems must be clear to users.
- Consistency checks should be implemented within and between records.
- Quality oversight of data processes is essential.
- All computer systems used for data collection, storage, sharing, use, and archiving must be validated for their intended use.

Administrative safeguards

Administrative safeguards consist of administrative controls – generally documented in policies and procedures – to manage the selection, development, implementation and maintenance of security measures to protect GxP data, and to manage the conduct of the workforce in relation to the protection of that data. Appropriate controls must be established for all phases of the data life cycle, from initial creation through processing (including any modification, deletion, transformation, or migration), use, retention, archiving and retrieval.

Policies and procedures

Pharmaceutical companies typically have numerous policies and procedures that impact data integrity in some way, including, but not limited to, the following examples:

- Good documentation practices
- Data life cycle approach
- Computerized systems validation
- Risk management
- Security management, including access management
- System administration
- Change control, especially manual direct database updates by information technology (IT)
- Incident response and management
- Backup and restore, including monitoring for backup errors
- Disaster recovery and business-continuity planning
- Archiving and record retention
  - Retrieval and readability checks on archived data and metadata, including audit trails
  - Archived data security and data-integrity controls
- Review processes, including audit-trail review (see “Big Brother Is Watching”)

Security management process and access management

Many of the system-specific administrative security controls addressed here have an indirect but significant impact on data integrity. While overall responsibility for the controls lies with the business (as the data owners), the actual implementation of some controls may rely on IT or technical organizations.

> **Appropriate controls must be established for all phases of the data life cycle**

Wherever possible, logical security controls for a computer system should be based on technical rather than procedural controls. Security and access controls include, but are not limited to:

- Securing communication of user credentials, such as:
  - Ensuring that you are talking with the correct person before sharing verbal information
  - Verifying the user’s identity in person
  - Using one communication channel for the password and a different one for the username
  - Limiting the amount of time that the initial password is valid
- Ensuring appropriate approval(s) for access
- Identifying unique users to ensure nonrepudiation of changes and/or electronic signatures
- Accessing roster review for user accounts
- Removing access or privileges in a timely manner when they are no longer needed
  - Note: An automated method based on job moves provides a much higher degree of assurance than procedural or manual control
- Enacting good password practices such as using pass phrases that are at least eight characters long and include letters, numbers, and special characters; prohibiting password sharing
- Modifying all default passwords (especially for system-administrator accounts)
- Password expiration interval
- Creating challenge questions for reauthentication, such as password resets
- Establishing an appropriate automatic logoff interval (inactivity timeout) within the application
- Ensuring that appropriate role-based security is established
- Segregating business-related duties. For example: System-administrator access data deletion and/or system configuration changes should not be assigned to individuals with a direct interest in the data (anyone who generates, reviews or approves data). Where this is unavoidable due to personnel limitations, mitigating controls should be implemented, such as dual user accounts with different privileges
- Segregating IT-related duties: ensure that the same individual cannot request, grant and approve access for themselves, for example
- Following the least-privilege rule to ensure that each users has only enough privileges to allow him or her to fulfill his or her job function
- Limiting the number of IT system administrators to the minimum possible, taking into account the size and nature of the organization
- Monitoring of turning on/off system audit trails
- Limiting access to other system configuration parameters that could affect data integrity, such as user account modification and number of failed access attempts before lockout
- Configuring security notifications from the application to a designated authority
- Reviewing access logs
- Accessing roster review for nonuser accounts, such as administrator accounts

- Implementing time synchronization and time-clock security controls
- Assuring that vendor-provided software is maintained at a release that is supported by the vendor. This ensures that the latest security patches and service packs can be applied as soon as reasonably possible to close known security risks.

Contracts and other arrangements

From contract manufacturing or laboratory services to outsourcing IT or using "as a service" options such as SaaS, PaaS or IaaS (software, platform, or infrastructure as a service), these service providers have a potential impact on data integrity that must be evaluated, controlled, and mitigated. From the providers' willingness to be audited through the completion of audits or assessments prior to supplier selection, and throughout the ongoing service engagement, data integrity should be an area of absolute focus. Both IT controls and business processes must be reviewed to ensure that appropriate controls are in place to guarantee data integrity. Establishing clear requirements related to data security and integrity in the contract and/or quality agreement provides a baseline for ongoing monitoring to ensure that expectations will be met. Record-retention periods and access requirements (including system availability) must be clearly defined and achieved.

Documentation and data management

The process used to store and manage documentation and data and the repository in which the data resides can have a significant impact on data integrity. Documents or other types of data files stored and managed in a regulated electronic document management system or other validated electronic-record computer system that leverages a relational database can be controlled on a much higher level than documents or other types of data files managed in a file share on a server using a manual process. This is explicitly supported in the MHRA GMP Data Integrity Definitions and Guidance for Industry March 2015,[1] which states:

> *There is an inherently greater data-integrity risk with flat files (e.g., when compared to data contained within a relational database), in that these are easier to manipulate and delete as a single file. [Emphases in the original]*

Paper-based records

It should be noted that while paper-based records have been used for much longer than electronic records, they share many of the same concerns, such as how to ensure the record remains:

- Legible: Considerations around fading ink and the well-known issues with thermal printouts
- Available: How to protect the record during long-term archiving: Will the paper degrade? Is it threatened by moisture or pest species?
- Retrievable: How to quickly locate one record among the many thousands retained; advantages and disadvantages of onsite vs. offsite storage

Paper records, of course, have an additional issue in that they lack the independent audit trail that can accompany an electronic record; this means that it is not possible to identify record backdating, repeat testing of the sample, or whether all results have been retained.

For certain records, there is clear regulatory guidance that the electronic record must be retained, as the paper record is not sufficient. The US Food and Drug Administration (FDA), for example, makes a very clear statement about chromatographic data:

> *For High Performance Liquid Chromatography (HPLC) and Gas Chromatography (GC) systems (and other computerized systems involving user inputs, outputs, audit trials, etc.), the predicate rules, such as 21 CFR 211.68 and 21 CFR 211.180(d), require the electronic records themselves to be retained and maintained in accordance with those regulations. . . . [T]he printed chromatograms used in drug manufacturing and testing do not satisfy the predicate rule requirements in 21 CFR Part 211. . . . [T]he electronic record must be maintained and readily available for review by, for example, QC/QA personnel or the FDA investigator.*[2]

Technical controls

Technical controls should be introduced to mitigate the risks associated with human actions in the design of the original system. However, because technical controls are designed, tested, and implemented by humans, it is important to recognize that the controls may themselves have design flaws.

Computerized systems are often generalized as IT systems or software solutions. In fact, a computerized system is not limited to software but should be considered as a business process supported by the use of IT solutions. The key here is that business process comes before the technical solution – never the other way around.

Business process comes before the technical solution – never the other way around

Data integrity by design

The integrity of regulated data should be safeguarded in three spaces: during collection and processing, when transferring between systems and in storage.[3] Evaluating the risks to data integrity at each stage of the data flow in a business process can identify opportunities to improve data integrity by intelligent system design; transcription errors, for example can be eliminated from a workflow by directly interfacing the source and target systems such that the data is transferred electronically using a validated process. Transmission security across an open network can be strengthened by using integrity controls such as a checksum and encryption processes. Highly critical data editing or deletion functions can be additionally secured and justified using transactional safeguards such as password authentication at the time of execution, and the recording of an explanation for the action via free text or (preferably) preconfigured reason. A user interface that highlights potential data-integrity issues, such as manually integrated results or repeat samples, assists by focusing review efforts on the results with the highest risk.

Technical controls have an important advantage over human controls. The repeatable and reliable behavior of any validated IT system (whether it is a distributed clinical database or a manufacturing executing system) can be designed, tested, operated, and maintained in such a way that data integrity is ensured and well documented. However, it is also true that even the best systems – in terms of implementation, efficiency, and quality – could not ensure data integrity without qualified data stewards. Data

Introducing humans to a validated IT system creates a more complex and unpredictable interaction

stewards are the guardians of data integrity; their role is to speak up when something is amiss, and they should not fear the repercussions typically associated with slowing down the process to achieve better quality. The repetitive and sometimes heavy lifting of data should be left to validated IT systems, allowing data stewards to concentrate on more valuable and creative endeavors, such as monitoring data across multiple systems and identifying any patterns in the data or process. They are only human, after all.

Physical safeguards

Physical security begins with restricting site access to authorized visitors only.

IT architecture can be selected to improve data integrity by eliminating hard drives within the laboratory with all of the risks inherent in physical access to the storage media. A system based on a client/server architecture provides the ability to isolate the physical data location (the server) in a dedicated, climate-controlled environment (server room) with additional physical security ensuring that only a very select subset of authorized personnel are able to access the server room. Controlling which terminals (clients) can be used for specific functionality can reduce the likelihood of inappropriate system usage. In a distributed system, for example, it may be reasonable to have client terminals throughout the whole site, but those in the warehouse could be prevented from initiating a packaging run on the production line.

Finally, careful consideration should be given to the storage media used for backup and long-term archival storage as the data on it must remain accessible, secure, and protected for an extended period. This may require a process for transferring the archived data to new media due to "shelf-life" limits, a newer system and/or a change of technology.

Computerized system validation

Introducing humans to a validated IT system creates a more complex and unpredictable interaction which, when refined and documented, becomes process. The marriage of the trained human user armed with an efficient process to a validated IT system produces the computerized system. The life cycle for a computerized system is a continuum from the initial idea for the system to its final decommissioning; it must address the potential need for the data to live on after system decommissioning to satisfy record-retention requirements.

Computerized system validation (CSV) is a process that is applied to provide verifiable objective evidence that a system meets predetermined specifications, governed by clearly documented procedures and used only by individuals with appropriate expertise and training. System access should be limited to only those personnel with a legitimate business reason for accessing the system, and granular levels of privileges should be further used to limit personnel access to specific functionality or data within the system, according to their job roles. CSV ensures that a system comprising people, process, procedures, hardware, software, operating system and networks is fit for its intended purpose.

Ensuring data integrity in a GxP system is extended, but not guaranteed, by CSV. It may be necessary to accommodate vendor solutions that have data-integrity gaps in the technical controls; these must be mitigated as part of the validation process. CSV only ensures that a system is fit for its intended purpose; it cannot absolutely prevent data-entry error or intentional misuse of the system.

Computerized systems that handle GxP-relevant data must be validated to ensure that health authority requirements for good (manufacturing, laboratory or other) practices are met, noting that the Code of Federal Regulations Title 21, Part 11,[4] PIC/S GMP Guide[6] and/or the EudraLex Annex 11[5] (along with equivalent regulations for other countries) provide specific requirements around the use of regulated electronic data, records, and signatures.

Conclusion

An effective and well-maintained framework of administrative safeguards and technical controls can "remove temptation" when it comes to falsifying data. The controls can eliminate obvious opportunities for misdeeds and encourage correct use of the system. It is acknowledged, however, that intelligent, skilled people may well be capable of circumventing even sophisticated controls. "Big Brother Is Watching" examines the review processes designed to monitor for evidence of wrongdoing and discusses training approaches to reinforce awareness of data integrity in a GxP environment. ■

Charlie Wakeham and Thomas Haag

How Good Is Your Data?

New methods can increase data integrity in the lab

Chances are the integrity of your data is at risk.

Surprised? Data-intensive science is becoming far more mainstream in daily laboratory operations, and the laboratory has also become a strategic source of scientific evidence to support daily manufacturing and research operations in almost all pharmaceutical operations. Yet in 2013 the US Food and Drug Administration (FDA) reported that laboratory processes and deficiencies associated with laboratory controls are among its top three regulatory observations. The same report also cited a 50 percent increase in warning letters related to aspects of data integrity.

Data integrity is the assurance that data records are accurate, complete, intact and maintained within their original context, including their relationship to other data records. Ensuring data integrity means protecting original data from accidental or intentional modification, falsification, malicious intent (fraud) or even deletion (data loss).

Data integrity is not a new phenomenon; its basic principles have been described in several ISPE Good Automated Manufacturing Practice (GAMP®)

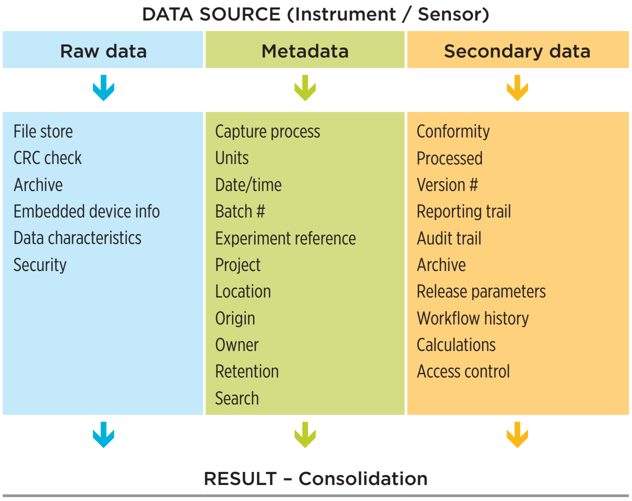
guidelines.[1] This article summarizes some recent regulatory findings and highlights how organizations can reduce data-integrity inconsistencies.

Changing the emphasis

In a data-integrity-focused audit, emphasis shifts from information based solely on technical and scientific contexts to evidence proving that the final analytical results are not false. As regulators increase their focus on data integrity and reliability, auditors are conducting examinations with multiple regulations and standards in mind, including pharmaceutical quality/ manufacturing standards,[2] good laboratory practices, GAMP, good clinical practices and application integrity policy in addition to FDA-recognized standards.

According to the FDA, source data should be "attributable, legible, contemporaneous, original and accurate" (ALCOA), and must meet regulatory requirements for record keeping. "ALCOA+" refers to additional terms included by the European Medicines Agency's concept paper on electronic data[3] in clinical trials ("Desired state" table, page 42). It is highly recommended that this concept be used.

Figure 1 Informatics data journey



DATA SOURCE (Instrument / Sensor)

Raw data	Metadata	Secondary data
File store CRC check Archive Embedded device info Data characteristics Security	Capture process Units Date/time Batch # Experiment reference Project Location Origin Owner Retention Search	Conformity Processed Version # Reporting trail Audit trail Archive Release parameters Workflow history Calculations Access control

RESULT – Consolidation

Informatics data journey

When samples are analyzed, several types of scientific data are created in the laboratory. They can be categorized in three different classes (Figure 1):

Raw data: Created in real time, this is all data on which quality decisions are based.[4] Raw data files can be unstructured, and are often based on a proprietary, vendor-defined format.

Metadata: This "data about the data" is used for cataloging, describing and tagging data resources. Metadata adds basic information, knowledge and meaning, and helps organize electronic resources, provide digital identification, and supports resource archiving and preservation.

Secondary or processed data: This is raw data transformed by scientific methodologies such as spectroscopy, chromatography, etc. To maintain data integrity, altering methods to reprocess will require a secured audit trail functionality, data, and access security.

Data integrity is the assurance that data records are accurate, complete, intact and maintained within their original context

QbD decreases variability

Corrective and preventive action (CAPA) is one of the four elements that support a proactive continuous improvement process within the product life cycle approach. Today's CAPA systems are good, but they focus on a traditional reactive approach. The ICH Q10 guideline[5] recommends a much more proactive approach to make biopharmaceutical manufacturing simple, sustainable, and more robust. Modern laboratory informatics platforms such as a laboratory information management system (LIMS), electronic lab notebook or laboratory execution system will significantly improve the use of previous knowledge created in laboratories. Scale-up information, clinical research, translational medicines and failed reactions during discovery may well contribute to a better understanding of the drug substance than we have anticipated.

Self-documenting processes

Automating metadata capture is very effective for maintaining data integrity and has been adopted by many industries. Scientific laboratories, however, lag behind. More than 75 percent of laboratory experiments or analysis starts with some kind of manual process, such as weighing. The majority of results are still written down on a piece of paper or are retyped into a computer or tablet. Self-documenting processes capture metadata automatically without human interaction, eliminating transcription errors and avoiding the need to retype data.

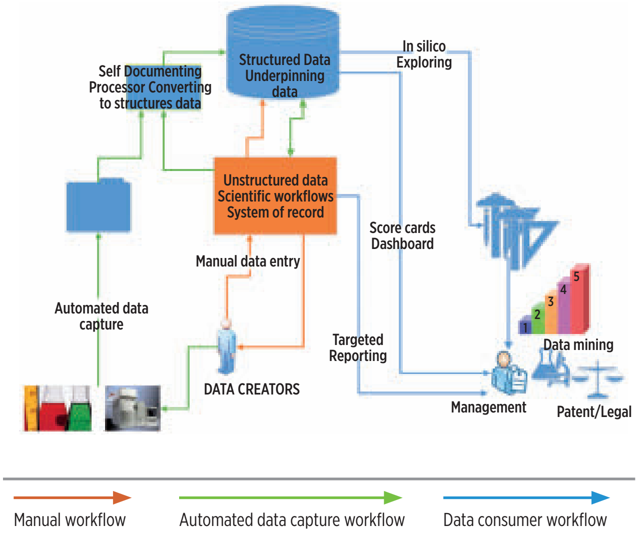
Single point of truth

To avoid data-integrity challenges, it is crucial to have a master copy of the data – a single source of data used across multiple systems, applications and/or processes. To achieve a single point of "truth" and significantly reduce data integrity challenges in the laboratory, we need to understand the key differences between spreadsheets and databases.

Red flags
Alteration of raw, original data and records
Multiple analyses of assay with the same sample **without adequate justification**
Manipulation of a **poorly defined** analytical procedure and associated data analysis in order to obtain passing results
Backdating stability test results to meet the required commitments
Creating acceptable test results **without performing** the test
Using test results from previous batches to **substitute testing** for another batch

Source: FDA

Figure 2 Embed data consumers' compliance requirements – adopt life cycle mindset



Best practices

Define a **single** point of truth for (meta)data

Reduce, automate, and **simplify** workflow complexities

Stop spreadsheet madness

Implement **self-documenting** processes at the source

Utilize **best-practice** analysis protocols

Adopt and use data **industry standard**s and processes

Embed data consumers (compliance) requirements

Avoid custom software extensions

The perception that a spreadsheet can act as a database is wrong. The primary function of a spreadsheet is to manipulate, calculate and visualize data, whereas the primary function of a database is to store and retrieve data in a structured manner. A spreadsheet has serious drawbacks when used for data storage: It cannot enforce relationships, there are no multi-user capabilities and it offers no data validation or protection against data corruption.

Workflow complexities

Simplifying scientific processes would significantly reduce challenges in data integrity. Although our industry is trying to harmonize scientific processes, other regulated industries are ahead in this field., There are, however, signals that our industry is recognizing the need. For example, balance and titrator suppliers have increased the value of their instruments by implementing approved and validated methods and industry best-practice workflows in their firmware. Almost all major suppliers allow methods to be implemented directly into balances and other wet chemistry instruments.

Simplifying scientific processes would significantly reduce challenges in data integrity

Figure 3 Benefits of open data standards

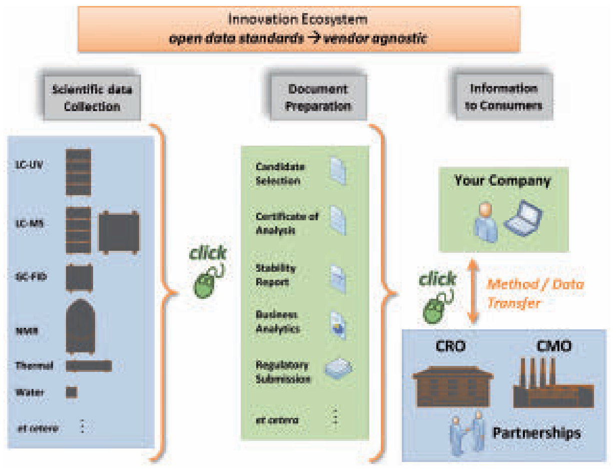


Image used with permission from Allotrope Foundation © 2016.

Integrating LIMS processes with enterprise workflows can also significantly reduce the probability of data-integrity issues. Process harmonization will initially increase the validation burden, but the effort will pay off in the long-term, significantly reducing the amount of potential data-integrity failure points and boosting efficiency for laboratory staff and management.

The final example is mapping the entire laboratory workflow and related operations, from sample receipt to release of results, which can consolidate operational workflow. The net effect may significantly reduce validation effort and decrease data integrity risks.

One serious concern is the lack of data standards in the scientific community. Without standards, data integrity will remain challenging and auditing and verifying is an expensive exercise. ■

Peter Boogaard

Big Brother Is Watching

Corporate training can be considered a "human" control for preventing data-integrity problems. Reinforce "right behavior" with ongoing training and monitor effectiveness with review processes.

The foundation for a high level of data integrity is knowing and understanding what data integrity is, the importance it has for a company and the personal role each employee has in protecting it.

Companies should also recognize regulatory authorities' increasing awareness and expectations for data integrity. While this is not new, the focus on and approach to managing and inspecting it are changing. As technologies, electronic systems and business models modernize, the industry must understand how to manage data in a changing environment.

Data-integrity training

Data-integrity problems can affect a company's reputation and profitability. To avoid the problems associated with data-integrity breaches, a "speak up/quality first" culture must be endorsed by company management, as we discussed in "Throwing People into the Works," and data-integrity training should be implemented from senior executives down to the line-operator level.

At the line-operator level, data integrity should be inherent in the process and not compromised to meet delivery timelines. Key data handlers should be formally trained to understand their roles in maintaining the integrity of the data they handle: They are data stewards, responsible for highlighting and escalating any concerns about data and quality regardless of the effect on production quota or deadlines. Training should not only ensure a common understanding of data, data integrity, falsification and data life cycles, but should also emphasize electronic good documentation practices, also referred to as good data management.

Foundational data-integrity training is only part of the bigger data-integrity picture, however. An additional, deeper understanding of technical expectations and requirements, inspection and auditing techniques and process governance are required to establish holistic data integrity for those with data steward or quality assurance (QA) responsibilities.

It is a regulatory expectation that companies understand their data life cycles and how data flows through their processes and systems. Personnel in roles that own these processes and systems (such as business-process and system owners) must understand their responsibilities in maintaining data integrity. These could include:

- Understanding how and where the data is used and its effect on product quality and patient safety
- Knowing what other review processes and data stewards are involved in each data flow, particularly those downstream of the system
- In-depth knowledge of system functionality with the most potential impact on data integrity and how to detect such activity

Personnel in QA and compliance roles must also have an advanced understanding of data-integrity requirements to ensure that these requirements are implemented within the systems and processes, and to support the business-process and system owners.

A corporation's data-integrity training program should be both general and specific. It should target the correct audiences and consider the specific scale of the corporation. In a large pharmaceutical company, high-level training for all employees might be at a foundational level, but the content and focus may be quite different for different functions. (The consequences of a data-integrity issue will be very different for a line operator compared to the operations director, for example.) This training approach might be ineffective for small and/or startup companies, however. In those cases it might be more effective to roll out both foundational and detailed training simultaneously.

Training on the general principles of data integrity could be complemented by more detailed, contextual training appropriate for data stewards who play a direct role in data handling. The specific training provided for such persons (including quality and compliance personnel) must extend beyond

People might cause data-integrity problems, but they are also superior to machines when it comes to detecting them

the general requirements and definitions of data integrity. This role-based training should focus on critical thinking and auditing techniques and could include specific-use cases related to the roles. Data-integrity training for laboratory auditors and process owners, for example, might include a comprehensive review of US Food and Drug Administration (FDA) warning letters that describe data-integrity observations in laboratory settings and practical exercises around examining audit trails.

Review processes

People might cause data-integrity problems, but they are also superior to machines when it comes to detecting integrity issues. Software applications can generate an audit trail, but only a human can decide “Was that integration parameter change a scientifically valid one?” For this reason, review processes remain in the human domain. Review processes can be discrete or continuous, one-off or repeated, and scheduled or unscheduled. In the sections below, different types of review processes and their timing are discussed.

Result review

Result review is defined here as a review of individual results, or sets of results, that is done prior to making the accept/reject decision about the product or data quality. To make that decision effectively, it is essential that the result review:

- Compares the result against specifications/limits
- Evaluates the completeness and correctness of the metadata supporting the result
- Determines the accuracy and integrity of any manually entered values
- Reviews any decisions or actions taken
- Understands any manual adjustment or alteration of the data or metadata
- Investigates any changes to the method versions used in the creation of the result
- Assesses conformity to sound scientific practice and documented procedures

Where there is a data audit trail that is easily accessible and permanently associated with the result, a review is likely to be the most effective route to assessing the integrity of the data.

The MHRA GMP Data Integrity Definitions and Guidance for Industry March 2015[1] states that:

> *Audit trail review should be part of the routine data review/approval process, usually performed by the operational area which has generated the data (e.g., laboratory).*

A data audit trail review should be conducted by someone who really understands the business process supported by the system and therefore understands the impact of the actions recorded in the data audit trail.

Result review should involve increased rigor of focus for results that involve manual adjustment and/or “just passing” results; an application offering the ability to highlight such results automatically provides an additional level of efficiency and assurance and may allow for the review-by-exception approach to data review.

Review by exception

Review by exception applies a risk-based approach to data review. In an environment where hundreds or even thousands of results are generated daily, if an equal amount of time is devoted to reviewing each result, by simple mathematics that amount of time will be very small. For just 100 samples, even spending as little as 2 minutes per result means more than 3 hours’ review time daily from each reviewer on those 100 samples – and more than one level of review may be required. Realistically, it is just not possible to review each result and its history effectively in 2 minutes.

Where the process or application permits, review by exception creates alerts to highlight a subset of results requiring additional effort, such as those:

- Within but close to the limit of the specification
- That have been manually integrated
- Where manually entered critical data have been changed
- That have been reprocessed

A detailed result review (as discussed above) is then conducted on this subset of results to understand what has been changed and why in order to decide whether to approve or reject the results. The remainder of the results, where the result is well within specification and no changes or adjustments have been made, can then be approved with a minimal level of review. A company wishing to operate review by exception has the responsibility to determine and document what that minimal level of review is, and to justify it during a regulatory inspection. Some level of validation will be required to document and verify the alert functionality.

Audit trail review

There was much debate within the industry in 2011 when the revised EudraLex Volume 4, Annex 11[2] stated that “audit trails must be regularly reviewed.” In reality, audit trail reviews were a regulatory expectation as far back as FDA Warning Letter 06-ATL-09 in 2006,[3] which stated:

> *Although the audit function is discussed in your procedures [for a chromatography data system], there is no specific requirement regarding any review of the audit trails, and your records failed to include documentation that a second person had conducted such a review. In fact, our investigator was told that no such audit had ever been performed. However, a second person must review these audit trails, particularly given the lack of controls for preventing data manipulation. Such an audit may well have detected the data manipulation which was occurring at your facility.*

This has been further reinforced more recently in warning letters 10-NWJ-03 in 2010[4] (“your firm’s review of laboratory data does not include a review of an audit trail or revision history to determine if unapproved changes have been made”) and 320-12-08 in 2012[5] (“your SOP does not have provisions for any audit trail reviews to ensure that deletions and/or modifications do not occur”).

Audit trail review offers a means to detect data-integrity issues but also functions as a deterrent. This is reflected in the National Institute of Standards and Technology Special Publication 800-12: “Introduction to Computer Security”:

> *Audit trails are a technical mechanism that help managers maintain individual accountability. By advising users that they are personally accountable for their actions, which are tracked by an audit trail that logs user activities, managers can help promote proper user behavior. Users are less likely to attempt to circumvent security policy if they know that their actions will be recorded in an audit log.*[6]

The last article in this series, “Doing the Right Thing” (p. 60), discusses other behavioral controls that can be combined with the audit trail review to discourage inappropriate activities.

Audit trail mechanisms in clinical and pharmacovigilance systems are the norm in both configurable and nonconfigurable software. Here, audit trails may be regarded as forensic tools to aid investigation when the integrity of a record is questioned; until then, it may be sufficient just to review the audit trail configuration to verify that:

- It is turned on and has not been turned off since the last review
- It is configured to capture the required metadata
- Ability to change the audit trail configuration (including system clock) is subject to the proper segregation of duties

Reviews of system audit trails and logbooks are a more pressing concern in laboratory environments and manufacturing sites, however, where the sophistication of the interfacing systems can limit the ease of transmission between them. Suggestions of what to look for within the system audit trail (as distinct from the data audit trail) are discussed in “Doing the Right Thing.”

Audit trail review offers a means to detect data-integrity issues but also functions as a deterrent

Periodic review

During the system’s periodic review, the following could be evaluated within the audit trail as part of monitoring human behavior and the effectiveness of the technical controls:

- Changes to system configuration that could impact data integrity controls
- Data deletion: What was deleted and why? If data was deleted as part of an archiving process, verify that the archived data is still accessible
- Account disabling due to successive failed logons: Look for repeat offenders and any timing patterns

Such a review process may only be practical in a system where the audit trail can be filtered. The practicalities and benefits of audit trail reviews are examined in the 2015 article by Perez et al., “A Risk-Based Approach to Audit Trails,”[7] and will not be duplicated in this discussion.

Personnel records and system-administrator logs can be reviewed for ongoing assurance of data integrity by:

- Checking the active user account list to ensure that only current personnel retain access to the system
- Confirming via training records that all active personnel are adequately trained to operate the system
- Ensuring that system/database backups are happening on the defined schedule, the integrity of the backup is verified and trial restoration of the system occurs periodically in a documented manner

Other periodic activities involve the review of standard operating procedures (SOPs), system records, SOP records, change control and system performance. These are essential for ongoing compliance, but they are out of the scope of this article. Periodic review and SOPs are covered in practical detail in the GAMP® Good Practice Guide *A Risk-Based Approach to Operation of GxP Computerized Systems*.[8]

Periodic review should be performed at a defined interval based on the GxP criticality of the system. Review frequency may be increased where issues have been found in system operation or in previous periodic reviews; similarly, a consistent lack of issues may provide justification to formally document and apply a decreased review frequency.

Data audit

A range of data audit activities can be undertaken as part of the scheduled periodic review process, unscheduled as part of an investigation or even in preparation for a regulatory inspection or customer audit.

One effective exercise could be to conduct a mock inspection of a specific data-handling process, where the entire data flow would be explained as if it were being presented to a regulatory official. This will highlight any confusion about where the data resides and how it passes from one system to another; it may identify areas of weakness in the system(s).

Another exercise could be to pick a single result and trace it back to the raw data, including any laboratory notebook entries. Verify the data integrity and audit trail at each step and demonstrate that all raw data, paper or electronic, is readily retrievable, fully supports the final result and is consistent with any summary data filed with the regulatory agencies as part of an application for approval.

Repeating the exercise in the opposite direction – to verify that all data has been processed and reported and to confirm that there is no orphan data that could indicate trial injections or other malpractices – is equally important.

Further proactive data audit activities could be based on the regulators' own guidance; the FDA Compliance Program Guidance Manual on preapproval inspections,[9] for example, suggests that inspectors should:

- Review data on finished product stability, dissolution, content uniformity and active pharmaceutical ingredient impurity
- Determine if data was not submitted to the application that should have been
- Look for invalidated out-of-specification results and assess whether it was correct to invalidate them
- Seek out inconsistencies in manufacturing documentation, such as identification of actual equipment used

Review process documentation

Within regulated industries, simply completing an action is not sufficient; there must be some documented evidence of when it was completed and by whom. The MHRA GMP Data Integrity Definitions and Guidance for Industry March 2015 requires that:

> *There should be evidence available to confirm that review of the relevant audit trails have taken place. When designing a system for review of audit trails, this may be limited to those with GMP relevance (e.g., relating to data creation, processing, modification, and deletion).*[1]

Reviewing audit trail entries associated with results (i.e., data audit trail) may be governed by a Review of GxP Data SOP and documented by some statement along the lines of "By approving this report I certify that I have reviewed the data, metadata, manually entered values and the audit trail records associated with this data, in accordance with Review SOP XXX." This statement could be included in the signature process for the electronic record and be visible on the printed and displayed report.

The MHRA guidance goes on to state:

> *QA should also review a sample of relevant audit trails, raw data, and metadata as part of self-inspection to ensure ongoing compliance with the data governance policy/procedures.*[1]

Data-integrity problems can affect a company's reputation and profitability

Such a review may occur during the periodic review or be triggered as part of an investigation into a data integrity noncompliance. The issue around reviewing system audit trails – those that capture all logon/logoff activities, system-configuration changes, etc. – is really about scale, and there are implications to be considered for a sampling-based approach to audit trail review (especially where filtering is not available to focus on GxP critical entries. With known system implementations of up to 2,000 users in a global organization, the quantity of entries in an audit trail can dwarf any human effort to review them. Again, this is dealt with more fully in Perez et al.; the point here is whatever the approach – filtered, sampled, or reviewed only as part of an investigation – the approach, the justification for the approach and the completed review process should be formally documented in a manner likely to be acceptable to a regulator.

Conclusion

Happily, the Big Brother state detailed with horrifying clarity in George Orwell's book *1984* did not come to pass. For ensuring data integrity, however, some of the book's concepts of retraining and monitoring are essential, although thankfully at a less drastic level. There is clear regulatory evidence that ongoing monitoring of technical controls via review processes is required to demonstrate that data integrity has been evaluated, achieved and protected. ■

Charlie Wakeham and Thomas Haag

Doing the Right Thing

Tools and techniques encourage positive responses

Good behaviors can promote and encourage integrity within a company, but negative behaviors and measurements can damage integrity. One example of a damaging behavior is a company attempting to save costs by not buying enough user licenses for an application, thus forcing user-account sharing. As a result, system activity cannot be reliably and independently attributed to a single individual.

Poorly conducted investigations often blame the human factor or find no assignable cause. A change to standard operating procedures (SOPs) may be proposed as a preventive action. In reality, human behavioral controls such as SOPs can easily be ignored, and the process may be adversely affected, giving rise to data integrity issues. These behavioral fails can only be detected later, after the harm has occurred. This preventive action will therefore likely fail to guard against similar issues in the future.

Outside of the pharmaceutical industry, falsification and fraud occurred in respected financial institutions such as JP Morgan (2003) and Credit Suisse Group (2007–2008). The article "Compliance Alone Won't Make Your Company Safe"[1] discusses the premise that good people can still behave inappropriately and that creating a "policeman culture" of enforcing rules and procedures may discourage generally honest employees from admitting that they wandered away from the straight and narrow or inadvertently made a mistake.

Personnel involved in pharmaceutical manufacture, development, testing, etc., typically have a strong scientific or engineering background: "If it can be calculated, measured, or analyzed, then it is tangible and will be accepted." In this fact-based environment, the complex interaction of soft skills needed to direct people's behavior and responses is easily overlooked to the detriment

Any person making critical product-quality decisions must be free from commercial, marketing or financial pressure that could influence his or her decision

of data integrity. A properly applied combination of leadership direction, motivation, metrics and independent controls can be used to direct and reward the right behaviors, fostering data integrity.

Behaviors

Improvisation

In "Throwing People into the Works," improvisation was mentioned briefly in the context of small or startup companies, but improvisation is a mindset that can be widespread in any company or country where insufficient or inappropriate resources are a way of life.

Improvisation is the ability to work around a lack of people or absent or damaged equipment, and even a lack of training, to "get the job done somehow." The downside to a culture of improvisation is that SOPs or other controls will not be followed, and the integrity of any data produced by such means is therefore highly suspect. This reinforces yet again the importance of management provision for sufficient and suitable resources.

The scientific and engineering mindset of people in skilled professions can also create a culture in which any rule or impediment will be seen as a challenge to be gotten around: "Ah, but in that case I could ... " and this is more difficult to mitigate. "Big Brother Is Watching" emphasized the importance of training to reinforce the "right behavior" as one defense against this puzzle-solving mentality, but the "six sources of influence" discussed later in this article may prove more effective overall compared to training alone.

Impartiality

Any person making critical product-quality decisions must be free from commercial, marketing or financial pressure that could influence his or her decision.

For example, a quality control (QC) lab supervisor who reports to the operations department may be at risk of undue pressure to pass batches even if he or she has valid concerns about the test results. Good practice would recommend reporting through the independent quality assurance department.

Falsification for profit

Greed has been the motivator in a number of high-profile company fraud cases in recent years. Corporate-scale data-integrity fraud has included such extremes as performing bioequivalence testing on the branded product but presenting the results as those for the generic version; more common and widespread fraudulent activities may include:

- Unofficial testing to see if the sample will pass before running the "official" sample for the batch record. Some examples are US FDA warning letters 320-14-08,[2] 320-14-01,[3] 320-14-005,[4] and UK MHRA Non-Conformance Report 8913/378537-0004 NCR.[5]
- Concealing, destroying, or overwriting raw data and samples. Some examples are US FDA warning letters 320-14-08,[2] 320-15-07,[11] 320-14-11,[6] and Italian Medicines Agency Non-Conformance Report IT/GMP/NCR/INT/1-2014.[7]
- Renaming or misrepresenting results from a passing batch in support of other batches: US FDA Warning Letter 320-15-09[8] and the Trade and Industry Inspection Agency of State of Lower Saxony – Oldenburg, Germany, DE-MI-04 2011029 are examples.
- Manually manipulating chromatography integrations to alter the result: US FDA Warning Letter 320-15-04[9] is an example

The extent of falsification achievable by an individual depends on a combination of their motivation and their seniority within the organization, counteracted by the efficacy of administrative and technical controls to prevent such falsification (see Table A).

The extent and impact of falsification is greatly magnified if collusion is involved. A senior QC manager has the power to direct his or her staff to collude for falsification, resulting in systemic fraud within the laboratory, whereas an individual analyst can only try to persuade a coworker to try to falsify data and inherently runs the risk of being reported to management for inappropriate behavior. Geographic and corporate cultures may also influence the ease with which collusion may occur; strongly hierarchical cultures may be more susceptible to collusion instigated at a senior level as these cultures inherently discourage any disagreement with authority figures. (Cultural considerations are discussed in more detail in the first article in this series.)

Understanding effective risk controls

In formal risk methodology,[12] there are the following risk treatment options:

Avoid: Stop the activity or do it in a different way that eliminates the risk

Reduce (also termed "mitigate"): Adopt measures to reduce the likelihood of occurrence or reduce the severity of harm or increase the probability of detection

Retain: Accept a low level of residual risk

Transfer: Transfer the risk creating activity (more practical for physical risk than data-integrity risk)

Table A: Potential for falsification as a function of motivation and seniority

Motivation		Seniority	
		Lab technician	Operations director
	Greed	Data integrity issues may now have become quite sophisticated within the lab domain: ▪ Variety of saved test methods used for a range of known scenarios to effect the desired result ▪ Pool of "good projects" from which data is copied in place of new testing Falsification may be routinely happening to maximize lab or technician throughput in exchange for financial incentives or career advancement.	Data integrity issues may constitute systemic, corporate fraud, where: ▪ All raw materials are used and all finished goods are released, irrespective of quality ▪ The company benefits from significant savings on staff and equipment achieved through reduced quality and development testing In this scenario, the finished goods may be highly unsafe and ineffective, but all focus is on operating profits and bonuses.
	Fear	Data integrity issues here are likely to be on an individual sample or test level, and may take the form of: ▪ Test method or parameters altered to influence the result ▪ Test samples destroyed ▪ Test samples substituted to ensure a passing result Falsification is occurring when samples fail, because the management culture does not promote honesty and cares only about passing results.	Data integrity issues may be focused around production yield, such as: ▪ Pressuring the quality department to release borderline product ▪ Understating rejected batches or having them mixed with passing batches during rework Falsification is aimed at hiding poor performance from the shareholders, and is endemic throughout the production environment.

Within the pharmaceutical industry, it is common to immediately try to control the probability of occurrence as the risk treatment, most often by implementing people-based controls. As discussed throughout this series, however, people are the wild card in data integrity – the major source of variation – so it seems illogical to rely on them to be the controls.

Effective risk controls:

- Do not rely solely on people's actions
- Are built-in
- Are easy to comply with
- Are well communicated and understood
- Are supported and enforced by management
- Have backups/contingencies
- Make errors/failures clearly visible
- Fail over to a safety condition

Controls that rely on people to consistently perform an action the right way out of many possible ways are ineffective. Simply writing an instruction into an SOP may have little or no long-term effect on the probability that someone will do something the wrong way. Single training events may affect the probability of correct performance in the short term after the training, but will have minimal influence in the long term as people move within the organization and old habits reassert themselves.

Six sources of influence

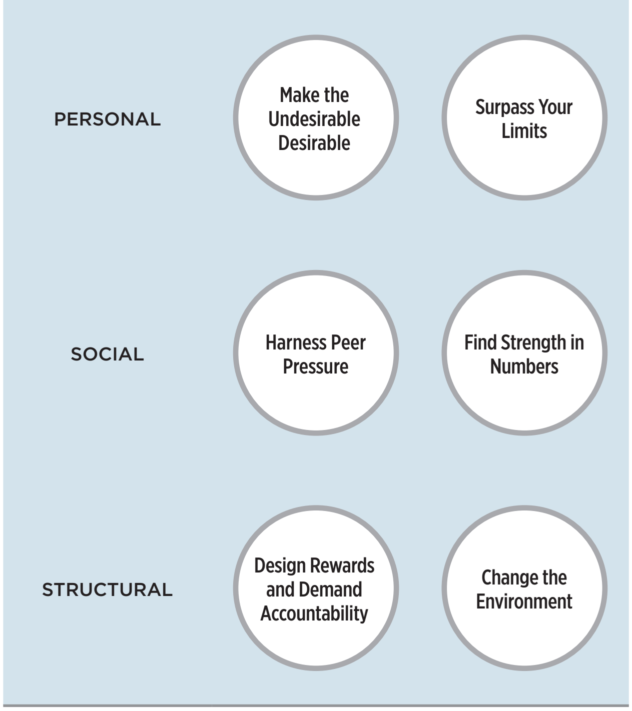
In *Influencer*,[10] Joseph Grenny and his colleagues propose a model for influencing behavior and attitudes. In the example below, this model has been applied to data integrity in a hypothetical QC laboratory (see Figure 1).

Make the undesirable desirable

This first step is about helping employees find the personal motivation to care about data integrity. Analysts testing many samples daily and under pressure to deliver throughput may easily lose sight of what the sample means: It's not a tick in a box or a statistic; it's vital information about whether a lifesaving medication will work.

Connecting the behavior to an outcome has a powerful impact. If possible, find out whose neighbor, child or parent relies on that medication and (with permission) use that person (our "real-life patient") to make it personal for all the lab staff. Spin the story. Add a picture and some background about the real-life patient: What are his or her hobbies? Does he or she have kids? Pets? Now, finding a failing sample is not a blot on the analyst's day; it's an important victory keeping this real person safe and healthy so he or she can continue sailing/studying environmental science at college/playing with his or her grandchildren.

Figure 1: Grenny's six sources of influence



PERSONAL	Make the Undesirable Desirable	Surpass Your Limits
SOCIAL	Harness Peer Pressure	Find Strength in Numbers
STRUCTURAL	Design Rewards and Demand Accountability	Change the Environment

> The extent and impact of falsification is greatly magnified if collusion is involved

This type of motivation allows the analyst to find intrinsic satisfaction in the right behavior; his or her diligence during testing has safeguarded the health of a patient who is personally connected to him or her.

Surpass your limits

It's easy to fall into having "just enough" knowledge to get by in a job. In this model, analysts are encouraged to spend time each day – as little as 10 minutes can have an effect – honing their chromatographic and application knowledge. Analysts can improve their understanding of integration, consider ways to improve sample preparation, or begin defining custom field calculations that could be used to eliminate calculation errors and the need to copy data into a spreadsheet for analysis.

It is essential that the lab manager provide strong support for this kind of self-improvement by helping analysts set short-term goals to measure the improvement and providing praise for the achievement.

Harness peer pressure

Here, the social aspects begin to affect behavior. Within the lab, there is likely to be one or more analysts or scientists to whom the others turn for advice and assistance. This person is the opinion leader, and he or she is vital to the success of the data-integrity initiative.

A message about data integrity, written by senior management, distributed throughout the company and read aloud by the lab manager is just a string of words falling on uninterested ears.

A respected colleague (the opinion leader) who really appreciates why data integrity is important and will set the example in the lab makes data integrity part of the lab environment. Once data integrity is embraced as the way to work, then peer pressure will keep everyone focused on integrity.

Every work area has the undiscussable. This is the "elephant in the room" – the topic that nobody is brave enough to raise. An example may be a dosage of a particular product that is always close to the impurity limit because it's made on an older and somewhat outdated manufacturing line. Whatever the topic, bring it into the open. Discuss it at the daily standup or weekly meeting. Acknowledge the problem, and share the experiences (and the frustration) of borderline sample results. Reinforce the need to scrupulously follow the sample preparation SOP, keep to the current method version, and never, ever just "tweak" the integration to turn that "just failed" into "close enough to pass." Making the problem open and shared takes away the temptation to nudge the sample into passing.

Strength in numbers

Research studies have proven repeatedly that groups perform significantly better than individuals. In the new culture of openness within the lab, it is now time to make data integrity a collective rather than individual responsibility. Analysts should be encouraged to work together to identify potential risks to data integrity and propose mitigations to those risks.

Teams of two or more can be formed to conduct data-integrity reviews (such as data audits, as discussed in "Big Brother Is Watching") periodically as an internal audit or on the first batch of a new product line. After all, who knows better where the data-integrity "holes" may hide than an analyst?

Increasing knowledge and confidence around data integrity will, in turn, lead to continual improvement in the overall integrity of the laboratory data.

Design rewards and demand accountability

It is important to note that the reward step happens late in the influencing process and not as the prime motivator as so many corporate leaders believe.

The dangers of inappropriately selected metrics were discussed briefly at the beginning of this series. The aim is to reward the right behavior rather than rewarding results. (Remember the samples analyzed per time period and all the inherent pitfalls associated with that metric?) Rewards should be small and symbolic rather than substantial enough to fuel a greed motivation.

Our team of analysts conducting the data audits who find a recurring flaw in the sample receipt register could get rewarded with, for example, an extended break to enjoy a round of specialty coffees bought by the company. Remember: Praise and recognition from their peers and their manager can mean just as much as the reward itself.

Good behaviors can promote and encourage integrity within a company, but negative behaviors and measurements can damage integrity

Change the environment

Earlier in this article, we looked at effective controls and found those to be controls that were built in and easy to follow. This influence model makes the same point: If the system is set up to make it easy to do the right thing, then people will do the right thing.

Creating approved methods for instrument control, data processing, and reporting all combine to make tasks quick and simple for the analyst – while ensuring that he or she is doing them correctly. Creating custom field calculations to eliminate calculation errors and getting sample weights read into the system electronically to eliminate transcription errors significantly strengthen data integrity by not only reducing the probability of error but also removing the simplest means for an analyst to falsify the sample weight or the concentration of active ingredient.

Using a combination of software applications, hardware interfaces and workflow design, it is possible to create an environment that, by its very nature, drives data integrity.

Conclusion

Over the course of this series, we have looked at leadership and culture; physical, administrative, and technical controls; training and monitoring; and now behaviors and positive influences. In recognizing the complexity of the problem of protecting data integrity, we have come to understand that there cannot be a single one-size-fits-all solution. Integrity is threatened by both human error and human greed, but greed will be more damaging to data integrity and will affect a greater number of records. We saw that the US FDA does not make allowances for how the data-integrity issues occur – whether by genuine error or deliberate falsification; it only cares that the issues have occurred and may impact product quality and patient safety. We looked at the audit findings from regulators around the world, inspecting to their own national regulations, and saw that they are focusing on and identifying the same data-integrity concerns, such as unofficial testing and failure to keep raw data.

This harmonization of inspection approach among regulators provides the common goal for all regulated companies, but ultimately it is the people factor within those companies that determines whether the goal is attained; all too often, data integrity is not consistently achieved and maintained. It will take a combination of making falsification so utterly unacceptable as to be unthinkable and increasing the probability of detection to the extent that it's simply not even worth it to restore confidence in data integrity across our industry. ■

Charlie Wakeham and Thomas Haag

A Special Interest Group (SIG) for Data Integrity

Launched in January 2014, the sponsor of the Data Integrity GAMP SIG, Mike Rutherford, had signed up some 50 members before the announcement at ISPE's 2013 Annual Meeting. The group now boasts more than 100 members, a sign, says Mike, of the topic's importance in the pharmaceutical manufacturing industry. "The group is working with Board member Chris Reid to make the SIG and ISPE-centric activity that reaches beyond GAMP."

In 2014, the SIG set four overarching objectives:

1. Understand existing and future regulatory expectations, guidance and enforcement strategies.
2. Identify and propose appropriate data integrity control strategies for critical data and key quality attributes throughout the life cycle that also address data management from the operational through to the record retention phase.
3. Provide tools to align requirements with a product's life cycle.
4. Provide a pragmatic and tangible framework for managing data integrity risks across the industry.

In the two years since its formation the SIG has generated presentations on how to identify and mitigate data integrity risk; identified which global GxP regulations and guidances are linked to data integrity; and developed a prototype tool with hundreds of these references which, while available only to GAMP SIG members today, may be rolled out to the broader membership in the future.

Goals for 2016 are three-fold:

1. Develop a GAMP guide on electronic records and data integrity that will include current thinking on governance. A session will take place at the 2016 Annual Meeting in Atlanta, this September, with the guide targeted for publication by Q1 2017.
2. Develop a GPG on how to apply the GAMP guide, as well as one that focuses on pragmatic solutions.
3. Create content for ISPE conferences, such as the Europe Annual Conference just held in Frankfurt, Germany, the 2016 Annual Meeting, and the upcoming GAMP regional conference in Copenhagen. They are also supporting the development of a Data Integrity Workshop at the ISPE/FDA GMP Conference in June of this year.

As a topic that is the focus of regulatory agencies around the world, data integrity "is something you absolutely need to be thinking," says Mike. ISPE is devoting much effort to it and solutions will continue to evolve for this business problem.

"What's important for members to understand is they needn't panic." ■

References

Throwing People into the Works

1. US Food and Drug Administration. Warning Letter 320-15-09. 6 April 2015. www.fda.gov/ICECI/EnforcementActions/WarningLetters/ucm443247.htm.
2. International Organization for Standardization. ISO 9001:2015: Quality Management Systems – Requirements. www.iso.org/iso/catalogue_detail?csnumber=62085
3. UK Medicines and Healthcare Products Regulatory Agency. "MHRA GMP Data Integrity Definitions and Guidance for Industry March 2015." www.gov.uk/government/uploads/system/uploads/attachment_data/file/412735/Data_integrity_definitions_and_guidance_v2.pdf.
4. Cressey, Donald R. *Other People's Money: A Study in the Social Psychology of Embezzlement*. Glencoe, Illinois: Free Press, 1953.
5. Meyer, Erin. *The Culture Map: Breaking through the Invisible Boundaries of Global Business*. Philadelphia: Public Affairs, 2014.
6. McAuley, Gerry. "Optimizing Human Performance, Part I." *BioPharm International* 27, no. 7. 1 July 2014. www.biopharminternational.com/optimizing-human-performance-part-i.
7. ——— . "Optimizing Human Performance, Part II: A Road Worth Traveling." *BioPharm International* 27, no. 8. 3 September 2014. www.biopharminternational.com/optimizing-human-performance-part-ii-road-worth-traveling.
8. ——— . "Optimizing Human Performance: A Road Worth Traveling, Part 3." *BioPharm International* 27, no. 9. 1 July 2014. www.biopharminternational.com/optimizing-human-performance-road-worth-traveling-part-3.
9. Potter, H. "Needed: A Systematic Approach for a Cockpit Automation Philosophy." *Proceedings of the Workshop on Flight Crew Accident and Incident Human Factors*. 21–23 June 1995. Washington, DC: US Federal Aviation Administration, Office of System Safety.
10. Panko, Raymond R. "What We Know about Spreadsheet Errors." *Journal of End User Computing* 10, no. 2 (Spring 1998): 15–21. Revised May 2008. http://panko.shidler.hawaii.edu/SSR/Mypapers/whatknow.htm.
11. US Food and Drug Administration. Warning Letter 320-15-06. 30 January 2015. www.fda.gov/ICECI/EnforcementActions/WarningLetters/2015/ucm432709.htm.

Implementing a Corporate Data Integrity Program

1. UK Medicines and Healthcare Products Regulatory Agency. "MHRA GMP Data Integrity Definition and Guidance for Industry March 2015." www.gov.uk/government/uploads/system/uploads/attachment_data/file/412735/Data_integrity_definitions_and_guidance_v2.pdf
2. Avellanet, John, and Eve Hitching, "Considerations for a Corporate Data Integrity Program" ISPE GAMP Community of Practice Concept Paper.
3. Wakeham, Charlie, Eve Hitching, and Thomas Haag. Special Report on Data Integrity. *Pharmaceutical Engineering* 36, no. 2 (March-April 2016).

An Ounce of Prevention

1. UK Medicines and Healthcare Products Regulatory Agency. "MHRA GMP Data Integrity Definitions and Guidance for Industry March 2015." www.gov.uk/government/uploads/system/uploads/attachment_data/file/412735/Data_integrity_definitions_and_guidance_v2.pdf.
2. US Food and Drug Administration. "Questions and Answers on Current Good Manufacturing Practices, Good Guidance Practices, Level 2 Guidance – Records and Reports." Question 3. www.fda.gov/drugs/guidancecomplianceregulatoryinformation/guidances/ucm124787.htm#3.
3. Lopez, Orlando. "A Computer Data Integrity Compliance Model. *Pharmaceutical Engineering* 35, no.2 (March/April 2015): 79–87.
4. Code of Federal Regulations. Title 21, Part 11: "Electronic Records; Electronic Signatures." http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=11
5. European Commission. Health and Consumers Directorate-General. EudraLex, Volume 4, Annex 11: "Computerised Systems." http://ec.europa.eu/health/files/eudralex/vol-4/annex11_01-2011_en.pdf.
6. Pharmaceutical Inspection Co-Operation Scheme. "PIC/S GMP Guide." http://www.picscheme.org/publication.php?id=4

How Good Is Your Data?

1. International Society for Pharmaceutical Engineering. "GAMP® Good Practice Guides." www.ispe.org/gamp-good-practice-guides
2. US Food and Drug Administration. "Pharmaceutical Quality/Manufacturing Standards (CGMP)." http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064971.htm
3. European Medicines Agency. "Reflection Paper on Expectations for Electronic Source Data and Data Transcribed to Electronic Data Collection." EMA/INS/GCP/454280/2010. 9 June 2010. http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/08/WC500095754.pdf
4. European Commission. Health and Consumers Directorate-General. "Documentation." Chapter 4 of EudraLex Volume 4, Good Manufacturing Practice (GMP) Guidelines. 30 June 2011. http://ec.europa.eu/health/files/eudralex/vol-4/chapter4_01-2011_en.pdf
5. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline. "Pharmaceutical Quality System: Q10." 4 June 2008. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q10/Step4/Q10_Guideline.pdf

Big Brother Is Watching

1. UK Medicines and Healthcare Products Regulatory Agency. "MHRA GMP Data Integrity Definitions and Guidance for Industry March 2015." www.gov.uk/government/uploads/system/uploads/attachment_data/file/412735/Data_integrity_definitions_and_guidance_v2.pdf.
2. European Commission. Health and Consumers Directorate-General. EudraLex, Volume 4, Annex 11: "Computerised Systems." http://ec.europa.eu/health/files/eudralex/vol-4/annex11_01-2011_en.pdf.
3. US Food and Drug Administration. Warning Letter 06-atl-09. 28 September 2006. www.fda.gov/ICECI/EnforcementActions/WarningLetters/2005/ucm076083.htm.
4. –––. Warning Letter 10-NWJ-03. 14 January 2010. www.fda.gov/ICECI/EnforcementActions/WarningLetters/ucm197966.htm.
5. –––. Warning Letter 320-12-08. 23 February 2012. www.fda.gov/ICECI/EnforcementActions/WarningLetters/2012/ucm294321.htm.
6. Gutman, Barbara, and Edward A Roback. "An Introduction to Computer Security: The NIST Handbook." National Institute of Standards and Technology Special Publication 800-12. October 1995. http://csrc.nist.gov/publications/nistpubs/800-12/handbook.pdf.
7. Perez, Randy, Chris Reid, and Sion Wyn. "A Risk-Based Approach to Audit Trails." *Pharmaceutical Engineering* 35, no. 2 (March/April 2015). www.pharmaceuticalengineering.org.
8. International Society for Pharmaceutical Engineering. GAMP® Good Practice Guide. *A Risk-Based Approach to Operation of GxP Computerized Systems – A Companion Volume to GAMP® 5*. January 2010. www.ISPE.org.
9. US Food and Drug Administration. Compliance Program Guidance Manual 7346.832. "Pre-Approval Inspections." 12 May 2010. http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/Manufacturing/QuestionsandAnswersonCurrentGoodManufacturingPracticescGMPforDrugs/ucm071871.pdf.

Doing the Right Thing

1. De Cremer, David, and Bjarne Lemmich. "Compliance Alone Won't Make Your Company Safe." *Harvard Business Review*, 18 May 2015. https://hbr.org/2015/05/compliance-alone-wont-make-your-company-safe.
2. US Food and Drug Administration. Warning Letter 320-14-08. 7 May 2014. www.fda.gov/ICECI/EnforcementActions/WarningLetters/2014/ucm397054.htm.
3. ——— . Warning Letter 320-14-01. 25 November 2013. www.fda.gov/ICECI/EnforcementActions/WarningLetters/2013/ucm376913.htm.
4. ——— . Warning Letter 320-14-005. 6 March 2014. www.fda.gov/iceci/enforcementactions/warningletters/2013/ucm390278.
5. UK Medicines and Healthcare Products Regulatory Agency. Statement of Non-Compliance with GMP. Report UK GMP 8913 Insp GMP 8913/378537-0004 NCR. www.pharmacompass.com/assets/pdf/edqm/A1348.pdf.
6. US Food and Drug Administration. Warning Letter 320-14-11. 16 June 2014. www.fda.gov/ICECI/EnforcementActions/WarningLetters/2014/ucm401451.htm.
7. Italian Medicines Agency. Statement of Non-Compliance with GMP. Report IT/GMP/NCR/INT/1-2014. www.pharmacompass.com/assets/pdf/news/N1.pdf.
8. US Food and Drug Administration. Warning Letter 320-15-09. 20 April 2015. www.fda.gov/ICECI/EnforcementActions/WarningLetters/ucm443247.htm.
9. ——— . Warning Letter 320-15-04. 19 December 2014. www.fda.gov/ICECI/EnforcementActions/WarningLetters/ucm427976.htm.
10. Grenny, J. *Influencer*. McGraw-Hill, 2008.
11. US Food and Drug Administration. Warning Letter 320-15-07. 27 February 2015. http://www.fda.gov/iceci/enforcementactions/warningletters/2015/ucm436268.htm
12. Institute of Risk Management. "A Risk Management Standard." 2002. https://www.theirm.org/media/886059/ARMS_2002_IRM.pdf

About the authors

Peter Boogaard is founder of Industrial Lab Automation. His company provides services to address harmonization, integration, and consolidation of business to help clients enable cross-functional collaboration among research, development, quality assurance, and manufacturing corporations. Peter has extensive experience in laboratory management. He has published papers in international magazines and contributes to several industry advisory boards, is an active member of ISPE, and organizes the annual Paperless Lab Academy congress (www.paperlesslabacademy.com). A Dutch citizen, Peter studied analytical chemistry in Delft, Netherlands.

Thomas Haag has been working at Novartis for more than 13 years in the areas of Clinical Data Management, Systems Validation, and Quality Assurance. He is an experienced project manager, business analyst, and quality assurance engineer with a focus on computerized systems validation and data integrity. Haag, a trained artist, started his career in commercial industrial product development and eventually moved into information technology, application training, and software project management. He lives at the New Jersey Shore with his family and his golden retriever, Alice. He enjoys snowboarding, gardening, and good conversation.

Christopher Reid is CEO of Integrity Solutions Limited, a provider of quality and compliance solutions to regulated companies globally with locations in the UK, Tokyo, and North Carolina. His responsibilities include all aspects of business management, financial management, resourcing, and product/service development. Chris currently works with leading global organisations developing and implementing quality and compliance solutions including defining and implementing strategic quality initiatives, implementing corporate quality policies and standards, skills development, and system validation. Chris is a member of ISPE's International Board of Directors, European Forum, European Leadership Team, Co-Chair of the Knowledge Network Council, and member of the Global and European GAMP® Steering Committees. He has contributed to the development of GAMP 5 and a variety of GAMP® Good Practice Guides. Chris holds a BSc (Hons) in Computing Science from Staffordshire University.

Michael Rutherford is a Consultant – Business Systems Support – Laboratory and Quality Systems, Medicines Development Unit, at Eli Lilly and Company. Mike is a recognized global technical expert in ER/ES, data integrity, computer systems validation and quality, laboratory automation, and laboratory informatics. For the last five years he has been responsible for the technical, administrative, and operational leadership for multiple laboratory and quality related computerized systems, as well as information protection projects across the Lilly MDU. Mike has been involved with ISPE and GAMP leadership since 2003 and currently serves as the Global Chair of the GAMP COP, the past Chair of the GAMP Americas COP, and the sponsor of numerous GAMP Special Interest Groups, including those on data integrity and the cloud. He was the recipient of the 2014 ISPE Max Seales Yonker Member of the Year Award

Charlie Wakeham has more than 15 years' industry experience developing and validating computerized systems for regulated production and laboratory environments. An active GAMP® member for more than a decade, she has contributed to several Good Practice Guides and many ISPE conferences. She is a Regional CSV Consultant at Waters Corporation, where she works directly with customers in the Asia–Pacific region to provide efficient and practical assistance with the validation of the company's laboratory Informatics computerized systems.