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).

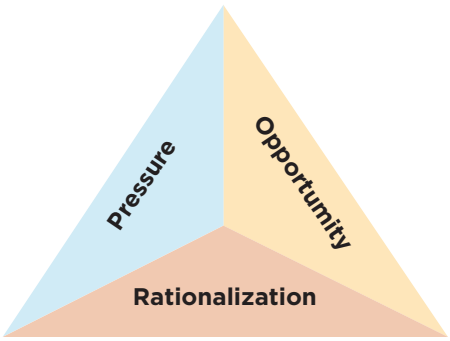


Figure 1 The fraud triangle

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.

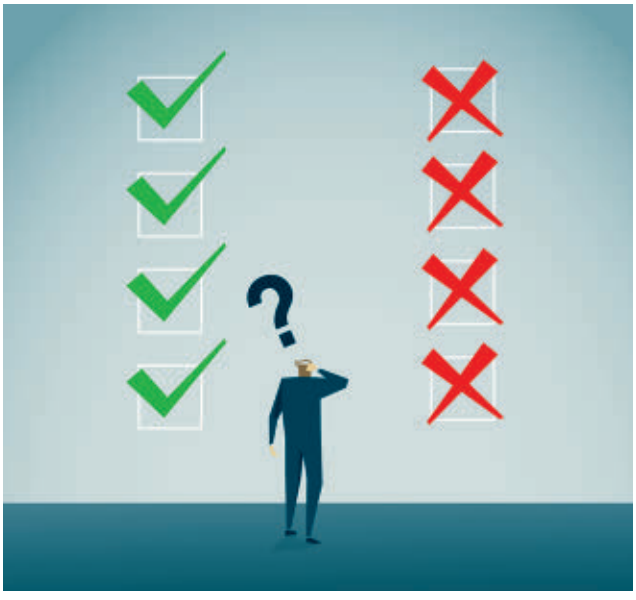
Geographic culture

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.



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

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 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.