Sudarshan Publications

Available online at https://s-epub.in/index.php/trp/index

Pharma Research

An International Peer Reviewed, Indexed Journal

10.62655/s-epub.2023.v15.i01.pp19-26

DATA INTEGRITY IN PHARMACEUTICAL QUALITY CONTROL LABORATORIES

JAKUB DVORAK[1]*, DIVYA SWAIN[2], TOMMASO BRAMBILLA[3]

Affiliation

[1]Department of Pharmaceutical Management, Vysocina Institute of Pharmaceutical Chemistry, Jihlava, Czech Republic
[2]Department of Regulatory Affairs, Sahyadri Institute of Pharmaceutical Education, Karad, India
[3]Department of Pharmaceutical Management, Cilento Institute of Pharmaceutical Sciences, Salerno, Italy

*Corresponding author: Jakub Dvorak, j.dvorak@vipc.cz
ORCID iD: Jakub Dvorak 0000-0001-2145-9353; Divya Swain 0000-0001-2435-7862; Tommaso Brambilla 0000-0001-0928-831X

Article info

Received date:26-05-2022
Accepted date:13-09-2022
Published date:21-01-2023
ISSN;0975-8216

Keywords:

Data integrity, Quality control laboratory, ALCOA, Audit trail, Good manufacturing practice

Abstract

Data integrity failures in pharmaceutical quality control laboratories have become one of the most frequent reasons for regulatory action against manufacturers. This review draws on published regulatory guidance, textbooks and commentaries to describe what data integrity means in practice, how the main regulators define it, and where laboratory systems typically fail. It covers paper and electronic records, chromatography data systems, out-of-specification handling, audit trail review, and the organisational factors that sit behind technical lapses. Most of the literature is guidance or opinion rather than empirical study, and measures of the effectiveness of controls are scarce. The review reports no new data and identifies this as the main gap.

1. INTRODUCTION

A release decision for a batch of medicine rests on laboratory results. If those results cannot be trusted, nothing built on them can be trusted either, whether it is a certificate of analysis, a stability claim or a shelf-life. That simple fact is why regulators treat the reliability of laboratory data as a matter of patient safety and not merely of paperwork [1, 2]. In the past decade, inspection findings have repeatedly involved deleted files, unofficial testing, shared passwords and results that appear to have been selected, and data integrity is now a standing theme in enforcement correspondence [3, 4].

The concept is not new. Good practice requirements for records have existed for decades, and the Barr Laboratories case in 1993 already established that retesting without scientific justification was not acceptable [5]. What changed was the spread

of computerised instruments. Laboratories moved from bound notebooks to chromatography data systems, spectrometer software and laboratory information systems, and each step introduced new ways in which data can be created, altered or lost without anyone noticing [6, 7]. Regulators responded by publishing detailed guidance, and the pharmaceutical quality system standards set out expectations for management as well [8, 9].

Quality control is where the stakes are highest because the laboratory sits at the junction between manufacturing and the patient. Production records show how a batch was made, but laboratory data show whether it meets its specification. A falsified or poorly controlled result can pass a defective batch into the supply chain or, equally harmful, condemn a good one [1, 4]. Pressure is also greatest here: release timelines, limited instrument time and a shortage of trained analysts are common features of busy laboratories, and they create the conditions in which shortcuts appear [4, 10].

This review focuses on the quality control laboratory, where release and stability testing take place and where data integrity risks are concentrated. It summarises how authorities define the term, how failures arise in everyday laboratory work, which technical and organisational controls have been proposed, and what the literature does not yet show. The material is drawn from regulatory documents, standards and published commentary. No inspection data or laboratory records are reported, and nothing here describes a particular facility. Terms are used as they appear in the source documents, and where they differ between regulators the difference is noted. The intended readers are laboratory managers, analysts and pharmacists in quality roles.

2. DEFINITIONS AND REGULATORY EXPECTATIONS

Data integrity is commonly defined as the extent to which data are complete, consistent and accurate throughout their life cycle. The United Kingdom medicines regulator describes it in these terms, and the United States Food and Drug Administration uses a similar formulation in its question-and-answer guidance [1, 2]. The World Health Organization has issued its own guidance on good data and record management, later revised, and the Pharmaceutical Inspection Co-operation Scheme has produced a detailed document intended for use by inspectors in many countries [3, 11, 12]. These documents agree far more than they differ.

The acronym ALCOA captures the core attributes: data should be attributable, legible, contemporaneous, original and accurate. It is usually credited to an FDA official of the 1990s, and later guidance extended it to ALCOA+ by adding complete, consistent, enduring and available [2, 13]. The practical meaning is straightforward. A result must show who generated it and when, it must be recorded at the time of the work and not reconstructed later, the first record must be kept, and the record must be readable for the whole retention period. Each attribute can be tested during an audit by tracing a reported result back to the raw data.

Regulators outside the United States and Europe have taken similar positions. The WHO guidance applies to manufacturers supplying procurement programmes, and the European Medicines Agency publishes questions and answers that clarify expectations on topics such as shared accounts, data retention and the use of paper printouts [3, 12, 14]. Mutual recognition among inspectorates means that a serious finding in one country is often noticed in others, and a company with a poor record may face import alerts or suspended certificates of suitability elsewhere [11, 14]. For multinational companies, harmonising laboratory practices on the most demanding standard is the practical course.

The legal basis in the United States lies in the current good manufacturing practice regulations for finished pharmaceuticals, which require complete records of all laboratory data, including all tests and retests, and in the regulation on electronic records and electronic signatures, which requires controls such as secure computer-generated audit trails and unique user identification [15, 16, 17]. The corresponding European

provisions are in the EU GMP guide, particularly the chapters on documentation and quality control and the PIC/S good practices guide for computerised systems [18, 19, 20]. International standards add further expectations: ICH Q7 for active ingredients, EU GMP Chapter 1 for the quality system and ISO/IEC 17025 for testing laboratories [9, 21, 22].

Inspectors tend to test a laboratory by following a single batch from the certificate back to the raw data. They compare the injection sequence with the reported results, look for gaps in the numbering, and ask to see the audit trail for selected runs [11, 14]. Mismatches between what was run and what was reported are the most direct signal of trouble. Knowing this, laboratories can carry out the same trace themselves in internal audits, and many do so on a rotating basis as part of self-inspection [2, 11]. The exercise reveals gaps in procedure faster than a review of the written policy.

Several features set the guidance documents apart from older rules. They put the responsibility for data governance on senior management, and not only on analysts, and they expect risk-based thinking about which data and systems matter most [9, 23]. They also assume that electronic data and their metadata are the primary record. Printing a chromatogram and filing it is no longer enough if the underlying electronic file is not retained, because the paper copy cannot show deleted injections or altered integration events [1, 11, 24]. Documents of this kind also stress that the retention period for electronic records is the same as for paper, which has consequences for archiving and migration when software is replaced.

A short note on terminology avoids confusion. Metadata are the data that describe other data, for example the time stamp, user and instrument identity attached to a chromatogram. Raw data are the original records of observation. A dynamic record can be queried or processed, while a static one is fixed. The distinction explains why a PDF export from an electronic system is not the same as the system's record, and why guidance asks for the dynamic record to stay accessible [11, 14, 24]. Metadata of this kind are often the first thing lost when data are exported or migrated between systems.

3. WHERE LABORATORY DATA GO WRONG

Published analyses of regulatory actions show that failures are rarely subtle. Common observations include sharing of login credentials, analysts with administrator rights who can delete or edit files, instruments left to run without any audit trail enabled, and results stored on removable media or in unvalidated spreadsheets [4, 24, 25]. In several instances, inspectors found evidence of trial injections, where a sample was run and discarded before the official analysis, or reprocessing of chromatograms until a passing result appeared. These behaviours are usually deliberate, but not always malicious; some begin as ways of coping with workload and pressure [25].

Out-of-specification results are a particular hazard. The FDA guidance on investigating such results requires a laboratory phase and, if no assignable cause is found, a wider investigation, and it discourages repeated retesting until a result falls within the limits [26]. The Barr ruling made the same point legally. In daily practice the pressure to release product makes the temptation to discard an unwelcome value understandable, and laboratories that have no clear rules on invalidating results, or that allow analysts to decide privately, are exposed [5, 26]. The pattern is the same everywhere: a small shortcut, tolerated once, becomes routine.

Manual integration deserves its own mention. Peak integration in chromatography is partly a matter of judgement, but altering baselines after seeing a result is a classic source of concern. Good practice is to set integration parameters in a validated method, to restrict manual changes to justified cases, to record the reason and to retain both original and changed versions, with a reviewer checking the audit trail [6, 23, 24]. Chromatography data systems differ widely in how readily they enforce these rules, which is why system selection and configuration are quality decisions as well as technical ones [6, 23].

Sample handling comes before any instrument is switched on, yet it is easily

overlooked in integrity programmes. Samples must be logged on receipt, stored under defined conditions, and traceable to the batch and the pull point. Weighing, dilution and preparation steps are often recorded by hand, and the weights printed by a balance are part of the raw data and should be kept with the worksheet [15, 24, 27]. Stability studies extend the problem across years, since pull dates, chamber records and repeated testing must stay linked and retrievable long after the original analyst has moved on [24, 27].

Manual and microbiological tests are not exempt. Titrations, colony counts, pH readings and visual inspections are recorded on paper, and the temptation to adjust a reading or transcribe a rounded value is as real as with electronic data. Controlled, pre-numbered worksheets, entries in indelible ink, single-line corrections with a reason, signature and date, and witnessing of critical steps are the usual safeguards [1, 4, 10]. Second-person review of calculations catches transcription errors, which are far more common than falsification and equally damaging to the reliability of a result [4, 10].

Paper-based and hybrid practices bring other risks. Unofficial notebooks, loose sheets, written results transcribed from a screen, and pages recopied for neatness all break the principle that the first record should be preserved. Hybrid systems, in which an instrument produces electronic data but the official result is a printout signed by hand, are especially prone to inconsistency, because nobody can be sure that the paper reflects the file [11, 23, 24]. Entries made later from memory violate the requirement that data be contemporaneous, however accurate they may happen to be.

4. TECHNICAL AND PROCEDURAL CONTROLS

The control measures proposed in the literature fall into groups, set out with their sources in table 1. Technical controls include unique user accounts, role-based privileges, enabled and protected audit trails, secure time stamps, automatic backup and restricted data export. Procedural controls include second-person review of raw data and audit trails, defined rules for result invalidation and retesting, and periodic review of access lists [11, 23, 24]. Neither group is sufficient alone, as a well-configured system can be undermined by a procedure that tolerates workarounds. The distribution of entries across guidance, failure modes, technical controls and governance is shown in fig. 1.

Standalone instruments and spreadsheets are persistent weak points. Simple devices such as balances and pH meters may have no user management or audit trail at all, and the common workaround is a hybrid arrangement with a printed record and a controlled logbook [11, 24]. Spreadsheets used for calculations must be validated, locked against accidental change, and version-controlled, with the formulae protected from editing by analysts [24, 27]. Reviewers should check that the printout corresponds to the file, because an unlocked spreadsheet can be changed after printing. Where possible, replacing spreadsheets with validated functions in the laboratory information system removes an entire class of risk [23, 24].

Audit trail review has attracted particular attention because it is one of the few controls that can reveal what happened to a record after the fact. A reviewer is expected to look at the audit trail for entries relevant to a batch, including deletions, changes to processing methods, and repeated injections, before the result is approved [6, 11, 23]. Experienced practitioners point out that this is time-consuming when the system produces large undifferentiated logs, and that review should be risk-based, concentrating on critical events rather than every keystroke [6, 23]. Selecting instruments that filter the log by batch or sequence makes the task practical.

Outsourcing adds another layer. Many manufacturers send testing to contract laboratories, and regulators hold the contract giver responsible for the integrity of the data it receives [11, 14]. Quality agreements should therefore define data ownership, access to raw data and audit trails, retention periods and the right to audit, and the contract giver needs enough technical knowledge to review what it is sent [9, 11, 14]. Receiving only a certificate and a summary table, without the underlying electronic records,

leaves the sponsor unable to confirm that the testing was performed as reported, an arrangement that inspection guidance explicitly warns against [11, 12].

Table 1: Selected sources on data integrity in quality control laboratories and the points each addresses

Source / topic	Principal point	Ref.
FDA Q&A 2018	Defines data integrity, expects audit trails and individual user accounts	[1]
MHRA guidance 2018	Sets out ALCOA attributes, data life cycle and governance expectations	[2]
WHO guidance TRS 996	Good data and record management practices for all manufacturers	[3]
US v. Barr Laboratories 1993	Retesting without justification and averaging of results are unacceptable	[5]
FDA OOS guidance 2006	Phased investigation of out-of-specification results before any retest	[26]
Manual integration CDS	Changes to baselines need justification, reviewer check and audit trail	[6]
Unique accounts and audit trail Part 11	Secure computer-generated audit trails and user identification required	[16]
Instrument qualification USP	Staged model from design qualification to performance qualification	[27]
GAMP 5 risk-based validation	Scales validation effort to risk and software category	[7]
Psychological safety teams	Groups that feel safe to speak up report and learn from more errors	[28]

Computerised system validation underpins all this. The GAMP 5 framework offers a risk-based approach to categorising software and to scaling validation effort accordingly, and the PIC/S computerised systems guide and the FDA scope-and-application guidance on electronic records describe regulatory expectations for validation, access control and change management [7, 17, 19]. For laboratory instruments, the USP chapter on analytical instrument qualification describes a staged model of design, installation, operation and performance qualification [27]. Validating once is not enough; periodic review and change control keep a validated state from slipping. A rushed laboratory is more likely to skip these reviews, so staffing matters here as well.

Data governance as a concept ties the measures together. It means defining who owns each data set, how it moves between systems, how long it is kept and how it can be retrieved. Taking a lifecycle view, from creation through processing, review, reporting, storage and eventual destruction, helps identify gaps that a purely system-by-system approach misses [6, 23, 24]. Mapping data flows for a few key methods, such as an assay and a dissolution test, often shows unexpected manual steps and uncontrolled copies, and is a useful first exercise for a laboratory that is unsure where to start [6, 24].

Training and supervision also count as controls. Analysts need to understand why a rule exists, not only what it says, and supervisors need the competence to review electronic data critically [9, 28]. Observational audits, in which a trained auditor watches work being done and compares it with records, are reported to find problems that document review does not. They are also unpopular, which is a hint of how much laboratory practice depends on goodwill [23, 28]. Courses and refreshers should include real examples of past failures. Refreshers should rely on concrete examples of what went wrong elsewhere.

5. CULTURE, GOVERNANCE AND REMEDIATION

Technical controls only work where people want them to. A great deal of the recent guidance therefore deals with quality culture. The EU GMP quality system chapter gives management responsibility for creating conditions in which problems are reported and addressed, and guidance on data governance asks senior managers to review data integrity metrics regularly [9, 23]. The aim is a laboratory where an analyst who makes a mistake or sees an unexpected result can say so without fear. Where mistakes bring blame and targets drive behaviour, concealment is a rational response. Without that, written policies will be followed on the

day of an audit and ignored for the rest of the year.

Management review needs data of its own. Metrics that have been suggested include the number of deviations raised per hundred analyses, the proportion of results invalidated, the time taken to close investigations, and the findings of periodic audit trail reviews [4, 9, 28]. None of these is a measure of integrity in itself, and each can be gamed once it becomes a target. Used sensibly, trends in them show where attention is needed; used crudely, they encourage under-reporting. The EU GMP quality system chapter asks for management review of process and product quality, and data integrity indicators fit naturally within it [9, 23].

Concepts borrowed from other fields help to explain the pattern. Reason's analysis of human error distinguishes slips and lapses from deliberate violations, and argues that most violations arise from system weaknesses such as poor procedures, time pressure and unworkable rules [29]. Edmondson's research on psychological safety in teams found that groups in which members felt safe to speak up reported more errors and learned more from them [28]. Applied to the laboratory, these ideas suggest that a rise in reported deviations after a culture programme may be a good sign and should not be read as decline.

Whistle-blowing and speaking up deserve explicit support. Several well-known data integrity cases came to light because an employee reported concerns to a regulator, and not through internal audit [1, 4]. A laboratory that wishes to hear about problems first needs a route by which analysts can raise them, for example a confidential line to quality assurance, and a visible record that earlier reports led to action and not to reprisals. Psychological safety is hard to build and easy to destroy, and a single episode in which a messenger is penalised can silence a team for years [28, 29].

When failures are found, remediation follows a recognisable path. Firms typically commission an independent review of historical data, assess the impact on released product, correct the underlying systems, retrain staff and appoint data integrity specialists [4, 23]. Corrective and preventive action under the quality system should address root causes, which are often management decisions on resourcing and priorities and not individual misconduct [4, 9]. Regulators have stated that they expect remediation to be thorough and verified, and not just a revised procedure [1, 2]. Records of such discussions are themselves evidence of an active quality system.

Smaller laboratories, including those in contract and academic settings, face the same expectations with fewer resources. The principles do not change, but implementation can be proportionate. Using a laboratory information system with built-in audit trails, limiting the number of people with administrator rights, and agreeing clear rules for invalidating results cost little compared with a failed inspection [6, 10, 11]. Guidance for good practice in quality control laboratories offers a baseline that can be adapted [10]. Smaller organisations can still adopt the core points without a large budget. Sharing templates with peers in other laboratories saves effort.

6. LIMITATIONS AND FUTURE DIRECTIONS

The evidence discussed here has limits. Most sources are regulatory guidance, textbooks and expert commentary, and not controlled studies. Descriptions of inspection findings are filtered through the regulator's choice of what to publish, and there is little information on laboratories that never receive a citation [4, 25]. No study that this review could identify compared rates of data integrity failure before and after the introduction of a particular control, so claims about effectiveness rest on plausibility more than on measurement. Evidence from laboratories outside the regulated sector is almost entirely missing.

Research that would help includes anonymous surveys of analysts about workarounds, audits of electronic records that compare actual practice with stated procedure, and comparisons of systems for how easily they support audit trail review. Analytical tools that detect unusual patterns in data, such as unlikely rounding or repeated values, could supplement manual review but

need validation themselves [6, 23]. Developments in artificial intelligence may offer assistance in the future, although the guidance on such tools is only beginning to be written. Whatever form research takes, it should be reported transparently, including cases in which a control did not help.

7. CONCLUSION

Data integrity in the quality control laboratory depends on a combination of well-configured systems, clear procedures and a working culture in which people report what really happened. The regulatory sources agree closely on the principles, from ALCOA to audit trail review and lifecycle governance, and the repeated enforcement examples show where laboratories fall short in practice: shared credentials, unreviewed integration changes, unofficial testing and weak handling of unexpected results [1, 2, 4, 5]. Management attention is the thread that links them.

Evidence on what works is thin. Guidance is plentiful, but trials or comparative studies of controls are not, so the strongest recommendation is also the plainest one: map the laboratory's data flows, close the obvious technical gaps, review audit trails on a risk basis and treat unexpected results as information instead of an inconvenience [6, 23, 29]. Quality leaders should also track metrics of reported deviations, since rising reports often mean improving openness and not worsening practice [28]. Small laboratories can start with the same steps at modest cost.

8. ABBREVIATIONS

ALCOA-Attributable, legible, contemporaneous, original, accurate, CDS-Chromatography data system, EU GMP-European Union good manufacturing practice, FDA-United States Food and Drug Administration, GAMP-Good Automated Manufacturing Practice, ICH-International Council for Harmonisation, MHRA-Medicines and Healthcare products Regulatory Agency, OOS-Out of specification, USP-United States Pharmacopeia, WHO-World Health Organization

9. FUNDING

Nil

10. AUTHORS CONTRIBUTIONS

All the authors have contributed equally.

11. CONFLICT OF INTERESTS

Declared none

12. REFERENCES

1. US Food and Drug Administration. Data Integrity and Compliance With Drug CGMP: Questions and Answers. Guidance for Industry. Silver Spring: FDA; 2018.
2. Medicines and Healthcare products Regulatory Agency. 'GxP' Data Integrity Guidance and Definitions. London: MHRA; 2018.
3. World Health Organization. Guidance on good data and record management practices. WHO Technical Report Series No. 996, Annex 5. Geneva: WHO; 2016.
4. Rattan A. Data integrity: history, issues and remediation of issues. PDA J Pharm Sci Technol 2018;72:105–16.
5. United States v. Barr Laboratories, Inc. 812 F Supp 458 (D NJ 1993).
6. McDowall RD. Validation of Chromatography Data Systems: Ensuring Data Integrity, Meeting Business and Regulatory Requirements. 2nd ed. Cambridge: Royal Society of Chemistry; 2017.
7. International Society for Pharmaceutical Engineering. GAMP 5: A Risk-Based Approach to Compliant GxP Computerized Systems. Tampa: ISPE; 2008.
8. European Commission. Quality Risk Management, Annex 20 to the EU Guidelines for Good Manufacturing Practice. Brussels: European Commission; 2008.
9. European Commission. EudraLex Volume 4: EU Guidelines for Good Manufacturing Practice, Chapter 1: Pharmaceutical Quality System. Brussels: European Commission; 2013.
10. World Health Organization. Guidelines on good practices for pharmaceutical quality control laboratories. WHO Technical Report Series No. 957, Annex 1. Geneva: WHO; 2010.
11. Pharmaceutical Inspection Co-operation Scheme. Good Practices for Data Management and Integrity in Regulated GMP/GDP Environments. PI 041-1. Geneva: PIC/S; 2021.
12. World Health Organization. Guideline on data integrity. WHO Technical Report Series No. 1033, Annex 4. Geneva: WHO; 2021.

13. Woollen SW. Data quality and the origin of ALCOA. The Compass 2010;Summer.
14. European Medicines Agency. Questions and answers: good manufacturing practice, data integrity. London: EMA; 2016.
15. US Food and Drug Administration. Code of Federal Regulations Title 21, Part 211: Current good manufacturing practice for finished pharmaceuticals. Silver Spring: FDA; 2008.
16. US Food and Drug Administration. Code of Federal Regulations Title 21, Part 11: Electronic records; electronic signatures. Silver Spring: FDA; 1997.
17. US Food and Drug Administration. Guidance for Industry: Part 11, Electronic Records; Electronic Signatures. Scope and Application. Rockville: FDA; 2003.
18. European Commission. EudraLex Volume 4: Good Manufacturing Practice, Chapter 4 Documentation. Brussels: EC; 2011.
19. Pharmaceutical Inspection Convention and Pharmaceutical Inspection Co-operation Scheme. Good practices for computerised systems in regulated GxP environments PI 011-3. Geneva: PIC/S; 2007.
20. European Commission. EudraLex Volume 4: Good Manufacturing Practice, Chapter 6 Quality Control. Brussels: EC; 2014.
21. ICH Harmonised Tripartite Guideline. Good manufacturing practice guide for active pharmaceutical ingredients Q7. Geneva: ICH; 2000.
22. International Organization for Standardization. ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories. Geneva: ISO; 2017.
23. McDowall RD. Data Integrity and Data Governance: Practical Implementation in Regulated Laboratories. Cambridge: Royal Society of Chemistry; 2019.
24. International Society for Pharmaceutical Engineering. GAMP Guide: Records and Data Integrity. Tampa: ISPE; 2017.
25. Parenteral Drug Association. Technical Report No. 80: Data Integrity Management System for Pharmaceutical Laboratories. Bethesda: PDA; 2018.
26. US Food and Drug Administration. Guidance for Industry: Investigating Out-of-Specification (OOS) Test Results for Pharmaceutical Production. Rockville: FDA; 2006.
27. US Pharmacopeial Convention. General Chapter 1058 Analytical instrument qualification. In: USP 41–NF 36. Rockville: USP; 2018.
28. Edmondson A. Psychological safety and learning behavior in work teams. Adm Sci Q 1999;44:350–83.
29. Reason J. Human error: models and management. BMJ 2000;320:768–70.