Paper ET09

Next Generation QC Workflow Optimization: Current Practices and Future Directions from the Emerging Trends and Technologies Working Group

Matthew Finnemeyer, Vertex Pharmaceuticals, USA
Caroline Phares, Siemens, USA

ABSTRACT

The QC Process Efficiency Project, part of the PHUSE Emerging Trends and Technologies Working Group, aims to improve Quality Control (QC) process efficiency while ensuring regulatory compliance. Building on previous PHUSE working groups and papers that outlined best practices for QC, this project focuses on not just sharing current practices but also advocating for innovative approaches enabled by new technologies. The first phase of research addresses three distinct topics: global regulatory considerations, similarities and differences in QC practices including risk-based approaches, and recent advancements in technologies and methodologies that are shaping the future of QC. By exploring these areas, the group seeks to pave the way for more efficient, compliant, and technologically driven QC practices. This paper discusses the findings of this research, the challenges identified, and the opportunities for future advancements in QC processes within Data/Statistical Programming.

INTRODUCTION

The QC Process Efficiency Project was launched in April 2025, aiming to collectively inspect existing QC practices prevalent in the industry and communicate best practices of working to stakeholders, given the evolving usage of Quality by Design (QbD), risk-based approaches, Agile, opensource and AI. In this initial phase of the project, the group was divided into subgroups in order to most effectively investigate three distinct topics impacting QC: global regulations, current industry practices, and emerging technologies. Additionally, as a starting point, the group decided to narrow the scope of focus to Clinical Trial Analysis & Reporting. This paper provides a summary of the activities and findings resulting from this initial phase of the project.

REGULATORY CONSIDERATIONS

For the initial phase of the project, the Regulatory subgroup of the QC Efficiency Project set out to inspect the regulations that are specifically applicable to clinical trial analysis & reporting and how they are being interpreted across the industry. The Regulatory subgroup also reviewed the draft guidance *"Considerations for the Use of Artificial Intelligence To Support Regulatory Decision-Making for Drug and Biological Products,"* released by FDA in Jan 2025. This guidance gives much-needed structure to the process of validating AI models in a regulatory setting. It provides a structured, risk-based framework that ensures AI systems used in regulatory decision-making for drugs and biological products are credible, reliable, and fit for their intended purpose. By requiring sponsors to clearly define an AI model's context of use, assess its risk, and develop a detailed credibility assessment plan—including documenting data sources, evaluating bias, ensuring transparency, and addressing performance drift—the guidance formalizes what level of validation is required for AI models in a regulatory setting. This ensures that AI-generated outputs influencing safety, effectiveness, or quality determinations undergo rigorous evaluation, thereby strengthening trust and consistency in how AI is applied across biologics and drug submissions. In the second phase of the project, this group will investigate whether there will be any regulatory challenges to using AI to make QC more efficient.

APPLICABLE GUIDELINES

The Regulatory subgroup researched guidelines from the following sources: ICH, FDA, MHRA, EMA, and PMDA. Appendix A contains specific excerpts and summaries of the relevant guidelines and regulations. Upon reviewing the various guidelines, the team decided that validation of software and computerized systems would be out of scope for this phase of the project. These are also referenced in Appendix A. Although the MHRA Good Clinical Practice (GCP) Guide does suggest specific techniques that can be used, including double-programming and peer review, most guidelines and regulations define principles and requirements rather than prescribing specifically how to perform validation of statistical programming. These principles and requirements were summarized into the following points.

- Programs and systems must be accurate, reliable, and consistent with intended performance (such as alignment with the Statistical Analysis Plan).
- Data integrity is key, and validation must ensure all data handling and derivations are correct.

- Validation should follow a risk-based approach, focusing on intended use.
- There must be traceability between input data, code executed, and the results produced.
- Results must be reproducible.
- Coding standards must be defined and followed (e.g. Good Programming Practices).
- There must be accurate and thorough documentation as evidence of validation.
- Thorough programming log review must be performed.

THE COMPLEXITY OF INTERPRETATION

In reviewing the guidelines and regulations, the Regulatory subgroup noted that there is complexity in interpretation. This is primarily because statistical programs blur the lines between software and statistical analysis. Programs for a study are usually a mix of general-purpose macros and tools, template code, code copied from previous studies, and study-specific code. Analysis programs are generally bespoke for a study. Even when based on standard macros or templates, analysis programs are usually modified to match the protocol, data collection, and statistical analysis plan (SAP) for a study. A set of programs is unique to a study because it is intended to analyze a set of data unique to the study. Programs evolve throughout the study lifecycle as data collection progresses and analysis deliverables are reviewed. The goal of validation is to confirm that the analyses are correct for the data in the study, not to validate the programs for general use.

ICH and other guidance distinguish between software/system validation and validation of statistical analysis, although there is some overlap. For example, ICH Harmonised Guideline for Good Clinical Practice E6(R3) has separate sections for Computerized System Validation and for Statistical Programming and Data Analysis.

Software/system validation uses techniques such as unit testing, integration testing, regression testing, along with software development life cycle. Guidance emphasizes adherence to ICH statistical principles and data integrity (e.g. providing traceability of data transformations and derivations). Techniques include independent double-programming, checks of output against raw data or data listings, and review of code/programs. General-purpose macros can be validated using software validation recommendations. Double-programming is often used for analysis programs for several reasons: it is considered a better method than code review; analysis programs tend to be bespoke for a study; even when based on standard templates or macros, the analysis programs are usually modified to match the protocol, data collection, and SAP for a study; programs evolve over course of a study as data comes in; it has not been considered worth the effort to validate study-specific programs using general software validation processes; and double-programming provides traceability and reproducibility, which are two regulatory requirements that can be easily evidenced.

FUNDAMENTAL ISSUES OF INDEPENDENT DOUBLE-PROGRAMMING

In terms of efficiency and efficacy, there are several issues with double-programming: First, it is resource intensive, as it requires two programmers independently creating the same deliverable. Second, it is difficult to maintain complete “independence”, as the QC programmer typically verifies their work against production results and, more rarely, code written by the Production programmer; at worst, the QC process can become an exercise in reverse-engineering the same results as the production programmer, resulting in duplication of the same errors. Third, both the Production and QC programmers develop their programs from the same resources (SAP, specifications, source data); when data misrepresentations are not programming errors, but rather misunderstandings of the SAP, specs, or source data, they can be hard to detect with double-programming and potentially overlooked.

RISK-BASED APPROACHES TO VALIDATING CLINICAL OUTPUTS

Although advocacy for a risk-based approach to validation has existed for at least a decade, the perceived high-risk impact of programming errors appears to trigger risk-aversion in the majority of programming groups; they default to the “most thorough” method. When producing a body of validated outputs for clinical submission, a thoughtful approach must balance the effort (e.g. human-hours, programmatic load, timelines, etc.) necessary to achieve the requisite degree of quality. As noted by Randall and Coar (Randall, 2020), the relationship between effort and quality” is “not linear”, and consequently, a maximalist approach to validation might fail to achieve significantly greater gains in quality while also reducing the available resources for concurrent or future work.

Although several publications in the past decade have recommended that the FQC process should not be used as the primary validation approach for all clinical outputs (e.g. SDTM/ADaM datasets, tables, figures, listings (TFLs), etc.) (Randall, 2020), an internal survey, by our Working Group, of programmers in pharmaceutical, biotechnology, and contract research organizations (CROs) suggests that FQC is consistently used in the generation of all CDISC-compliant datasets; divergent approaches to TFL validation were identified, ranging from independent review to triple-coded programmatic review, depending on the deliverable scope/context. Thus, considering both the established understanding that FQC is both over-utilized and potentially inefficient in achieving “quality” clinical outputs, we want

to reinforce the importance of a “risk-based” approach to validation and provide additional considerations for implementation.

A risk-based approach to decision-making orients around the likelihood of a risk and the severity of the impact (Marrer, 2019); in the context of clinical output validation, likelihood of risk describes the possibility that a programming decision fails to correctly assign, derive, or represent clinical data, while severity describes the impact of this incorrect representation on the accuracy of safety and/or efficacy reporting. Prior publications have articulated a broad perspective on the likelihood of risk and severity of impact (Marrer, PHUSE Best Practices), as below:

Degree of Risk	High	Medium	Low
Description	New or multiple-use programs (e.g. unvalidated standard macros), endpoint analysis, complex data manipulation, unique programs and complex algorithms.	Reusable document templates or programming code with study specific change, already validated multiple-use/standard programs and less complex derivations.	Programs that are neither high nor medium are categorized as low. Examples include operational programs and other components copied over from one reporting effort to another with minimal changes and repetitive programs.

Table 1 Categorizing risk along one dimension

While we concur with a central tenet of this prior work, that novel outcomes, new compounds, and novel or complex study designs raise the likelihood of a programming issue, we recommend taking a more granular approach, evaluating SDTM & ADaM datasets and TFL outputs discreetly for likelihood and severity of risk.

RISK ASSESSMENT FOR TABLES, FIGURES, AND LISTINGS

When generating TFLs, the likelihood of a programming issue is dependent on the following qualities:

1. Complexity of the source data.
2. Complexity of the underlying output dataset structure.
3. Complexity of the programming logic for key variables/columns.

- **Low Likelihood:** safety listings or AE-/CM-related tables, with relatively simple source and output data structures and minimal derivations
- **Medium Likelihood:** threshold analysis tables (e.g. VS, EG, LB) or summary of analysis tables (e.g. MMRM), with relatively simple source data but increased complexity in the presentation of results
- **High Likelihood:** efficacy-related figures or threshold shift tables, with complex derivations and/or complex presentation of information

When generating TFLs, the severity of a programming issue is primarily dependent on the intended purpose of the output.

- **Low Severity:** Tables & listings with limited role in regulatory decisions
- **Medium Severity:** Safety tables, figures, and/or listings
- **High Severity:** Efficacy tables, figures, and/or listings

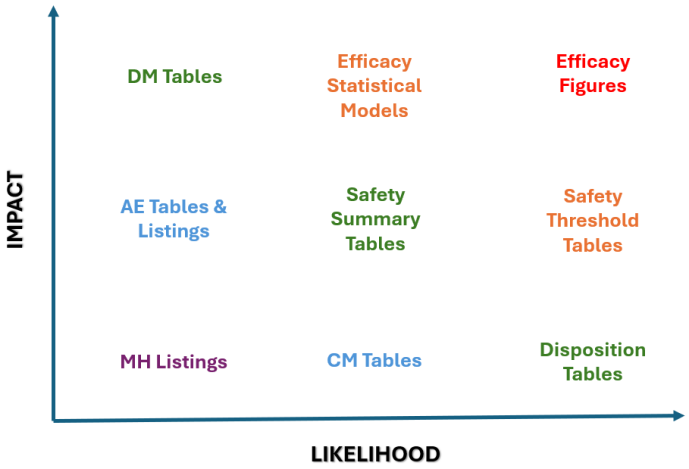

Figure 1 Examples of Likelihood and Impact for Tables, Figures and Listings

RISK ASSESSMENT FOR SDTM & ADAM DATASETS

When generating datasets, the likelihood of a programming issue is dependent on the following qualities; as an addendum, SDTM Datasets are highly structured and regulated, and can be treated as lower likelihood:

1. Complexity of the source data.
2. Complexity of the created dataset structure (e.g. derived records).
3. Complexity of the programming logic (e.g. derived and/or assigned variables).

- **Low Likelihood:** SDTM.PE or ADAM.ADAE, with relatively simple source and output data structures and minimal derivations
- **Medium Likelihood:** ADAM.ADEG or ADAM.ADVS, with relatively simple source data structures but more complex derivations (DTYPE = “AVERAGE” or baseline records)
- **High Likelihood:** ADAM.ADSL or ADAM.ADTTE, with complex derivations and data structure and/or multiple, integrated data sources

When generating datasets, the severity of a programming issue is primarily dependent on their downstream datasets and “terminal” outputs.

- **Low Severity:** SDTM.MH, with a limited role in ADaM-level datasets and safety- & efficacy-related outputs
- **Medium Severity:** ADAM.ADEG or ADAM.ADVS, with a direct role in safety-related outputs
- **High Severity:** ADAM.ADSL, with a direct impact on all safety- and efficacy-related outputs

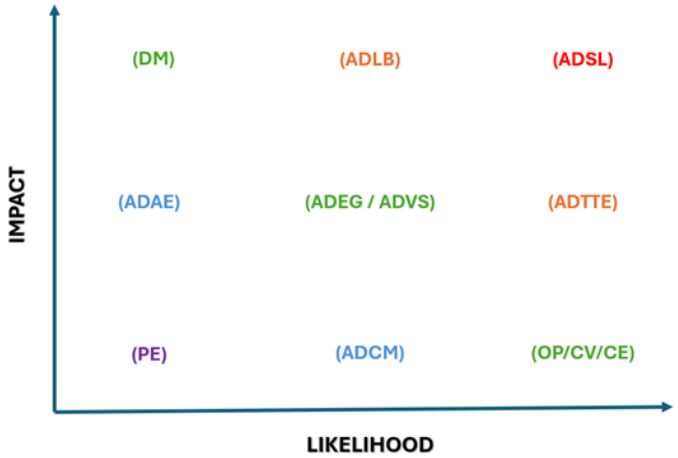

Figure 2 Examples of Likelihood and Impact for SDTM and ADaM Datasets

Following the assessment of risk, the responsible parties (e.g. study lead programmers, in consult with statisticians) should assign validation "levels" to each dataset and TFLs. Based on current industry practices, we recommend that FQC (full-quality control) should be prioritized for outputs in the top-right quadrant for Likelihood and Severity, with reduced degrees of validation correlated with reduced likelihood and severity; we encourage industry members to execute pilot programs for risk-based approaches to validation and remain attentive to new technological developments that might offer additional alternatives.

EXPLORATION OF EMERGING TRENDS

While the review of regulatory landscape and compilation of double-programming practices across industry kickstarted phase 1 of this project, an understanding of emerging tools and systems required further engagement with the community. After internal discussions, the team established areas of research as: Design & Process of QC, and the Implementation of QC tools.

This outline was further supported through the preliminary results of a recent public survey this group carried out via LinkedIn, in which 25% of responders expressed a need for additional investments in tools and technologies to enhance the QC process. 20% responders mentioned mismatching processes and systems, alongside siloed vendor-sponsor ecosystems aggravate the challenge of optimizing QC workflows.

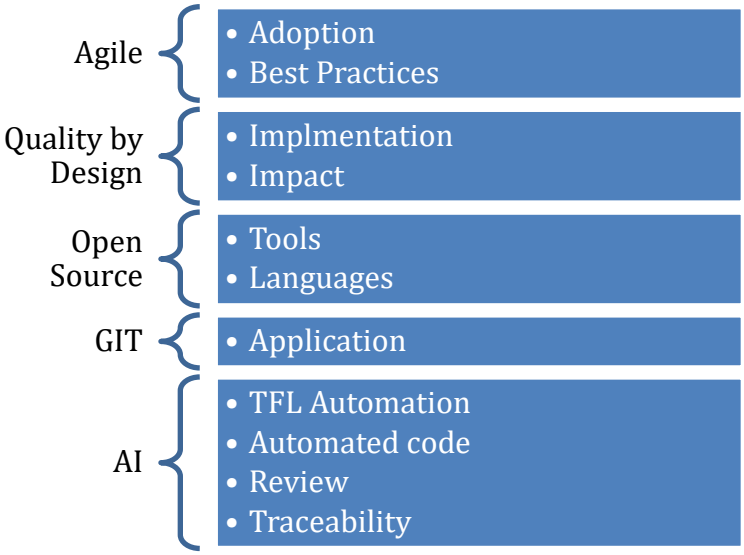

Figure 3 Outline of Research Areas

This examination was planned as a multistep process, involving a systemic review of published literature, collecting insights though public surveys and engagements with industry leaders, and workshops designed to create outline & structure for the final white paper.

LITERATURE REVIEW

RATIONALE

Performing QC in programming is a cornerstone in the Pharma industry and in 2015 PHUSE released the white paper "Best Practices for Quality Control & Validation". During recent years new/improved methodologies and processes emerged that may affect the way the QC is performed. This paper provides a revision of the 2015 best practices and will be integrated with a review of new methodologies/processes uptake. These include Artificial Intelligence Methodologies, Open-Source, GIT, Agile and Quality by design.

OBJECTIVES

According to Munn2018 we classified the review type as Methodology and the research question formulated based on "Types of Studies, Types of Data, Types of Methods, Outcomes (SDMO) "

The research question is "What is the effect of Artificial Intelligence Methodologies, Open-Source, GIT, Agile and Quality by design in terms of Statistical Programming Quality Control as reported in published reports?". The literature review will be performed according to the PRISMA2020 guidelines.

LITERATURE REVIEW METHODS

ELIGIBILITY CRITERIA

Inclusions:

Published reports and papers were included if

- Disseminated between 01 January 2023 and 31 December 2025
- In English Language

Exclusion:

- Unpublished reports as well as conference abstracts were not included

Note:

- 01 January 2023 has been set as starting timepoint as ChatGPT has been released in November 2022

INFORMATION SOURCES

Table 2 reports the repositories searched. Preliminary search was conducted in 2025, with further work planned in 2026.

Information source & URL	**Rationale**	**Access type**
PubMed https://pubmed.ncbi.nlm.nih.gov/	Popular repository used in Medical Research	Free
Google Scholar https://scholar.google.com/	Tool to search for scholarly literature from Google	Free
Copilot365	AI tool used to find white paper	Corporate subscription
Lexjansen https://www.lexjansen.com/papers/	Repository of PHUSE and PharmaSUG conference papers used to access full-text industry proceedings.	Free
PhUSE https://phuse.global/Communications/PHUSE_Archive	Source of recent conference proceedings related to statistical programming adavancements.	Free
PharmaSUG https://pharmasug.org/past-conference-proceedings/	Conference source focused on SAS and statistical programming in clinical research.	Free
Perplexity AI https://www.perplexity.ai/	Perplexity AI's Deep Research mode integrated automation by performing iterative web searches, reading sources, and prioritizing relevant 2024-2025 papers based on keyword strategies.	Paid Pro version
ChatGPT https://chatgpt.com/	ChatGPT, facilitates literature review by synthesizing and critically organizing information across diverse sources, leveraging deep semantic understanding to identify key themes, methodological linkages, and research gaps, while its unique strength lies in	Free

	contextual reasoning and coherent narrative construction that enables efficient structuring and interpretation of recent (2024–2025) scholarly work.	
ScienceDirect https://www.sciencedirect.com/	Peer-reviewed articles covering pharmaceutical drug development	Free
IJPS Journal https://www.ijpsjournal.com/	Pharmaceutical manufacturing and QC studies	Free

Table 2 Information Sources

SEARCH STRATEGY

The search focuses on the blocks reported in *Table 3*. Each block is related to a "and" condition i.e. "Block 1 and Block 2 and Block 3 and Block 4". Within blocks terms are related with an "or" condition.

Block	Terms
1	Artificial Intelligence or Machine Learning or AI or Statistical Programming or SAS or R Programming or Python or Data Analysis or Open-Source or GIT or Agile or Quality by design
2	Pharmaceutical Industry or Drug Development
3	QC or Quality Control
4	Dates between 01JAN2023 and 31DEC2025 included

Table 3 Building a Search Strategy

In line with PRESS 2015 Guideline Evidence-Based Checklist (PRESS2015), the search strategy will be adopted for each repository and system. Search String used for Google Scholar can be found in the Appendix as an example. For Generative AI tools, best practices for structured prompting were adopted e.g. COSTAR Prompt Engineering.

PEER REVIEW AND VALIDATION

Peer review of information sources and search strategy is currently in progress. Each reviewer will evaluate the search strategy using the PRESS 2015 Guideline Evidence-Based Checklist (PRESS2015). This activity will result in validation of above information.

OUTLOOK AND WHAT IS NEXT

The results of the search will be reviewed, and this literature review will provide insight on the state of the progress in tools, technologies and workflows as of 2025. The results of the review will be utilised to construct future community engagements and aid in the authoring of the final whitepaper.

COMMUNITY ENGAGEMENT PLAN

In the aim of having a broad overview on what is currently done within the industry, the working group created a survey. This public survey focused on the current state and asked for future vision of QC/validation workflows, tools and technologies. During the December 2025 PHUSE Webinar Wednesday, the survey was advertised. The recording of the session can be found on the PHUSE website in archive folder.

This survey was shared by PHUSE via LinkedIn, receiving 48 responses. Those responses are being used to create the content of workshops planned for PHUSE US and EU Connect 2026.

The working group will use the workshop at PHUSE US Connect 2026 to start discussions with industry peers, build new strategies, and refine the guidance on QC workflows. At PHUSE EU Connect 2026, another workshop will continue these discussions, and the outcomes of both workshops will be consolidated in the final whitepaper.

CONCLUSION

The PHUSE QC Workflow Optimization project's initial findings demonstrate that while regulatory requirements for quality control in clinical trial analysis are clear in principle, they provide flexibility in implementation. This creates an opportunity for the industry to move beyond resource-intensive independent double-programming toward more strategic, risk-based validation approaches.

Our analysis reveals three critical insights: First, global regulatory guidelines consistently emphasize accuracy, data integrity, and traceability without mandating specific validation methodologies. Second, independent double-programming, while widely used, has inherent limitations including resource intensity, potential for systematic bias, and inability to detect specification-level errors. Third, a risk-based framework that categorizes outputs by likelihood and severity of programming issues enables more efficient resource allocation while maintaining regulatory compliance.

As this project enters its second phase, the systematic literature review and planned workshops at PHUSE US and EU Connect 2026 will provide deeper insights into successful implementation of emerging technologies including AI, open-source tools, Git-based workflows, and Agile methodologies. We envision a collective addressing of how these innovations can be best implemented to meet regulatory standards while delivering meaningful efficiency gains, while not going overboard with additional levels of systemic checks that may impact the ability to implement at scale.

Where not practiced, we encourage organizations to evaluate their current QC practices against the risk-based framework presented here, considering where Full Quality Control provides genuine value versus where alternative validation methods may be more appropriate. The path forward requires not just technological adoption but a fundamental reconceptualization of quality—moving from effort-based metrics to outcome-based assurance that protects patient safety while enabling more efficient drug development.

This paper establishes the foundation for next-generation QC workflows. The PHUSE QC Process Efficiency Project invites continued industry collaboration to transform these findings into practical guidance that balances innovation, efficiency, and unwavering commitment to quality and compliance.

REFERENCES

1. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, McKenzie JE. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021 Mar 29;372:n160. doi: 10.1136/bmj.n160. PMID: 33781993; PMCID: PMC8005925.
2. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol 2018;18:5. doi:10.1186/s12874-017-0468-4
3. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. J Clin Epidemiol. 2016 Jul;75:40-6. doi: 10.1016/j.jclinepi.2016.01.021. Epub 2016 Mar 19. PMID: 27005575.
4. COSTAR Prompt Engineering https://portkey.ai/blog/what-is-costar-prompt-engineering/
5. International Conference on Harmonisation, Integrated addendum to ICH E6(R1): guideline for good clinical practice E6(R2), 2016 https://database.ich.org/sites/default/files/E6_R2_Addendum.pdf
6. International Conference on Harmonisation, Integrated addendum to ICH E6(R3): guideline for good clinical practice E6(R3 Annex 2), 2024 https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-e6r3-guideline-good-clinical-practice-annex-2-step-2b_en.pdf
7. International Conference on Harmonisation, Integrated addendum to ICH E6(R3): guideline for good clinical practice E6(R3), 2025. https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf
8. E9 Statistical Principles for Clinical Trials, Docket Number: FDA-1997-D-0508, Issued by Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e9-statistical-principles-clinical-trials
9. MHRA Good Clinical Practice Guide, Page 327, TSO Information and Publishing Solutions, 2012. https://www.tsoshop.co.uk/product/9780117081079/Medical/MHRA/Good-Clinical-Practice-Guide-Paperback?srsltid=AfmBOoqPN4Q0SdyUArLfsOdzefcWqnHPeR9HL3bJhMk-horYGOJzsCva
10. Revision of Technical Conformance Guide on Electronic Study Data Submissions, PMDA/CPE Notification No. 1345, PMDA/CRS Notification No. 21. April 8, 2024. https://www.pmda.go.jp/files/000267935.pdf
11. Patel, H., Pandya, C., Handling Health Regulatory Information Requests: Best Practices and Strategies, Paper SS-146, PharmaSUG 2025 https://pharmasug.org/proceedings/2025/SS/PharmaSUG-2025-SS-146.pdf

12. ICH Harmonised Tripartite Guideline, Statistical Principles for Clinical Trials, 5 February 1998, https://www.pmda.go.jp/files/000156905.pdf.
13. PMDA FAQs on Electronic Study Data Submission, Provisional Translation, March 2025. https://www.pmda.go.jp/files/000274654.pdf.
14. Bommichetty, G., Gurrala, H., e-Submissions key differences of FDA and PMDA, Paper EP-146m, PharmaSUG 2021, https://pharmasug.org/proceedings/2021/EP/PharmaSUG-2021-EP-146.pdf
15. Randall, A., Coar, W. Assessing Performance of Risk-based Testing, Paper 5112-2020, SAS Global Forum 2020. https://support.sas.com/resources/papers/proceedings20/5112-20202.pdf
16. Marrer, J. Risk-based Program Validation for Clinical Trial Analysis and Reporting. Paper SP06, PhUSE US Connect 2019. https://www.lexjansen.com/phuse-us/2019/sp/SP06.pdf
17. Marrer, J., Dalton, M., PHUSE Best Practices for Quality Control and Validation. Doc ID: WP-24, Version: v1.0, PHUSE Working Group: Standard Analyses & Code Sharing. https://phuse.s3.eu-central-1.amazonaws.com/Deliverables/Standard+Analyses+and+Code+Sharing/Best+Practices+for+Quality+Control+%26+Validation.pdf

ACKNOWLEDGMENTS

The primary authors acknowledge the contributions of the PHUSE QC Workflow Optimization project volunteers and thank the PHUSE Emerging Trends & Innovation Working Group leadership, the PHUSE Admin and Project Managers, the PHUSE US Connect Organizers and Stream Chairs, and the responders of the Public Survey for making this paper possible.

Key contributors to this paper included Elsa Lozachmeur, Enesh Vashisht, Korak Datta, Maria Dalton, Matthew Tendler, Neha Srivastava, Oscar Bortolami, and Yash Karampuri.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Author Name: Matthew Finnemeyer

Company: Vertex Pharmaceuticals Inc.

Address: 50 Northern Ave., Boston, MA

Email: matthew_finnemeyer@vrtx.com

Secondary Author Name: Caroline Phares

Company: Siemens Digital Industry, Inc.

Address: Plano, TX

Email: caroline.phares@siemens.com

This paper has not been explicitly funded by any organization. Relevant IT infrastructure for collaboration has been hosted by PHUSE. Time and effort are attributed to volunteering/external engagement time by individual employers and can vary for individuals involved.

APPENDIX A – REGULATORY GUIDELINES

ICH

Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice E6(R2)	5.1 Quality Assurance and Quality Control 5.1.3 Quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly.
	5.5.3 When using electronic trial data handling and/or remote electronic trial data systems, the sponsor should: (a) Ensure and document that the electronic data processing system(s) conforms to the sponsor's established requirements for completeness, accuracy, reliability, and consistent intended performance (i.e., validation).
	ADDENDUM The sponsor should base their approach to validation of such systems on a risk assessment that takes into consideration the intended use of the system and the potential of the system to affect human subject protection and reliability of trial results.
	ADDENDUM (h) Ensure the integrity of the data including any data that describe the context, content, and structure. This is particularly important when making changes to the computerized systems, such as software upgrades or migration of data.
ICH Harmonised Guideline for Good Clinical Practice E6(R3)	3.11.3 Quality Control Quality control should be applied using a risk-based approach to each stage of the data handling to ensure that data are reliable and have been processed correctly. Within clinical trials, monitoring and data management processes are the main quality control activities. Where appropriate, quality control activities may also be applied to facilities outside of investigator sites (e.g., central image reading facilities).
ICH Harmonised Guideline for Good Clinical Practice E6(R3)	3.16.2 Statistical Programming and Data Analysis This section concerning documentation of operational aspects of clinical trial statistical activities should be read in conjunction with ICH E9 Statistical Principles for Clinical Trials and ICH E9(R1) Addendum on Estimands and Sensitivity Analysis in Clinical Trials to The Guideline on Statistical Principles for Clinical Trials, which provides detailed guidance on statistical principles for clinical development, trial design, conduct, analysis and reporting. (a) The sponsor should develop a statistical analysis plan that is consistent with the trial protocol and that details the approach to data analysis, unless the approach to data analysis is sufficiently described in the protocol. (b) The sponsor should ensure that appropriate and documented quality control of statistical programming and data analysis is implemented (e.g., for sample size calculations, analysis results for IDMC review, outputs for clinical trial report, statistical or centralised monitoring). (c) The sponsor should ensure the traceability of data transformations and derivations during data processing and analysis. (d) The sponsor should ensure that the criteria for inclusion or exclusion of trial participants from any analysis set is pre-defined (e.g., in the protocol or the statistical

	analysis plan). The rationale for exclusion for any participant (or particular data point) should be clearly described and documented. (e) Deviations from the planned statistical analysis or changes made to the data after the trial has been unblinded (where applicable) should be clearly documented and justified and should only occur in exceptional circumstances (e.g., data discrepancies that must be resolved for the reliability of the trial results). Such data changes should be authorised by the investigator and reflected in an audit trail. Post-unblinding data changes and deviations from the planned statistical analyses should be reported in the clinical trial report. (f) The sponsor should retain the statistical programming records that relate to the output contained or used in reports of the trial results, including quality control/validation activities performed. Outputs should be traceable to the statistical software programs, dated and time stamped, protected against any changes, and have access controls implemented to avoid inappropriate viewing of information that may introduce bias. 4.3.4 Validation (a) The responsible party is responsible for the validation status of the system throughout its life cycle. The approach to validation of computerised systems should be based on a risk assessment that considers the intended use of the system; the purpose and importance of the data/record that are collected/generated, maintained and retained in the system; and the potential of the system to affect the well-being, rights and safety of trial participants and the reliability of trial results. (b) Validation should demonstrate that the system conforms to the established requirements for completeness, accuracy and reliability and that its performance is consistent with its intended purpose. (c) Systems should be appropriately validated prior to use. Subsequent changes to the system should be validated based on risk and should consider both previously collected and new data in line with change control procedures. (d) Periodic review may be appropriate to ensure that computerised systems remain in a validated state throughout the life cycle of the system. (e) Both standard system functionality and protocol-specific configurations and customisations, including automated data entry checks and calculations, should be validated. Interfaces between systems should also be defined and validated. Different degrees of validation may be needed for bespoke systems, systems designed to be configured or systems where no alterations are needed. (f) Where relevant, validation procedures (until decommissioning) should cover the following: system design, system requirement, functionality testing, configuration, release, setup, installation and change control. ICH E6(R3) Guideline (g) The responsible party should ensure that the computerised systems are validated as fit for purpose for use in the trial, including those developed by other parties. They should ensure that validation documentation is maintained and retained. (h) Validation should generally include defining the requirements and specifications for the system and their testing, along with the associated documentation, to ensure the system is fit for purpose for use in the trial, especially for critical functionality, such as randomisation, dosing and dose titrations and reductions, and collection of endpoint data. (i) Unresolved issues, if any, should be justified and, where relevant, the risks identified from such issues should be addressed by mitigation strategies prior to and/or during the continued use of the system.

FDA

Type of Validation	Sources	Key Points
Validation of Statistical Analysis	Refers to ICH Guidance	
Validation of Software and Computerized Systems - ***Out of Scope of this paper***	General Principles of Software Validation; Final Guidance for Industry and FDA Staff (Jan. 2002) Note that Section 6 will be replaced by Computer Software Assurance for Production and Quality System Software Electronic Systems, Electronic Records, and Electronic Signatures in Clinical Investigations: Questions and Answers | FDA (October 2024) Guidance for Industry - Part 11, Electronic Records; Electronic Signatures — Scope and Application (August 2003) Guidance for Industry Computerized Systems Used in Clinical Investigations (May 2007)	General Principles for Software Validation written for medical devices but also applied to drug development software and systems System requirements include security safeguards, audit trails, date/time stamps, SOPs, controls for system changes, and training of personnel Defines principles and requirement for validation – does not prescribe how to meet requirement Computer systems must be validated to ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records. Should have coding standards Lists software verification methods: Testing (e.g. unit tests, integration tests), Static and dynamic analyses, code and document inspection, walkthroughs Accurate and thorough documentation is essential (e.g. specs, risk assessment, validation process and results)

EMA

Type of Validation	Sources	Key Points
Validation of Statistical Analysis	Refers to ICH Guidance	
Validation of computerised systems and electronic data in clinical trials - ***Out of Scope of this Paper***	https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-computerised-systems-and-electronic-data-clinical-trials_en.pdf (2023) https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/notice-sponsors-validation-and-qualification-computerised-systems-used-clinical-trials_en.pdf (2020)	EMA mandates specific validation requirements for clinical trial data, focusing on the integrity, accuracy, and reliability of data generated and computerized systems used in trials Primarily aligned and outlined in ICH E6(R2) **ALCOA Principles** Attributable, Legible, Contemporaneous, Original, and Accurate

APPENDIX B – LITERATURE REVIEW RESULTS

The below records are retrieved from Google Scholar on year 2025 with the below search query:

("artificial intelligence" OR "machine learning" OR "AI" OR "statistical programming" OR "SAS" OR "R programming" OR "Python" OR "data analysis" OR "Open-Source" OR "GIT" OR "Agile" OR "Quality by design") AND ("pharmaceutical industry" OR "drug development") AND ("QC" OR "Quality Control")

The sorting is “by relevance” and the articles with more than 30 citations (as of 21JAN2026) are selected on the first 10 Google Scholar pages

https://journals.sagepub.com/doi/full/10.1177/20420986251321704

https://www.mdpi.com/1424-8247/18/1/47

https://pmc.ncbi.nlm.nih.gov/articles/PMC11944607/

https://www.mdpi.com/1424-8247/18/6/788

https://link.springer.com/article/10.1186/s12943-025-02321-x

https://jurnal.delitekno.co.id/index.php/jcbd/article/view/634

https://link.springer.com/article/10.1208/s12248-025-01079-w

https://journals.sagepub.com/doi/full/10.1177/20420986251321704

Similar exercises will be conducted on each information source, as listed in *Table 2*, during the Literature Review period.