

Review

The HERMES Framework for Experimental Reproducibility

Eugene Oga , Jamea N. E. Yolandia  and Herbert Che Mughe 

Department of Chemistry, University of North Dakota, 151 Cornell Street, Grand Forks, ND 58202, USA; jamea.yolandia@und.edu (J.N.E.Y.); herbert.chemughe@und.edu (H.C.M.)

* Correspondence: eugene.oga@und.edu

Abstract

Experimental reproducibility remains a persistent challenge across scientific disciplines despite substantial advances in analytical instrumentation, quality assurance systems, statistical methodologies, and reporting standards. Although reproducibility is frequently examined through the lenses of experimental design, data analysis, and methodological transparency, hidden operational factors embedded within routine laboratory activities often remain underrecognized sources of variability. These factors can introduce cumulative and interacting effects that compromise experimental consistency even when formal protocols are followed. This review examines the influence of operational variability arising from human practices, environmental conditions, reagents and consumables, instrumentation, digital data management systems, and sample-related factors across multidisciplinary laboratory settings. Current mitigation approaches, including standard operating procedures, quality-control programs, electronic laboratory notebooks, and laboratory information management systems, are critically evaluated with emphasis on their strengths and limitations in controlling operational variability. To support a more systematic understanding of reproducibility, the HERMES framework (Human, Environmental, Reagent, Machine, Electronic, and Sample factors) is proposed as a systems-based model for identifying, classifying, and managing hidden sources of experimental variation. By integrating operational influences that are often considered independently, this framework highlights reproducibility as an emergent property of interconnected laboratory systems rather than solely a methodological or statistical outcome. Recognizing and controlling these hidden operational factors may improve experimental reliability, strengthen data quality, and enhance confidence in scientific findings across research disciplines. HERMES is presented as a conceptual framework that integrates principles of operational reproducibility across laboratory disciplines. Future empirical studies are needed to validate its implementation, evaluate its effectiveness, and develop quantitative operational metrics for routine laboratory practice.

Keywords: reproducibility; operational variability; laboratory reproducibility; analytical uncertainty; quality assurance; workflow stability; metrology



Academic Editor: Saverio Mannino

Received: 23 June 2026

Revised: 27 July 2026

Accepted: 3 August 2026

Published: 5 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction and Conceptual Framework

Experimental reproducibility remains a major challenge across modern scientific research despite advances in instrumentation, quality assurance, and standardized methodologies [1–4]. While widely discussed in the biomedical and psychological sciences, reproducibility issues are increasingly affecting analytical chemistry, materials characterization, and process engineering disciplines, in which methodological robustness and measurement precision are traditionally assumed to be well controlled [5–9]. Failed validation

studies, interlaboratory inconsistencies, and hidden workflow divergence highlight persistent challenges to the reliability and transferability of findings generated within these instrument-intensive environments [3,4].

Conventional discussions of reproducibility emphasize experimental design, statistical analysis, publication bias, and selective reporting [1,2,4], often treating experiments as static procedures. In reality, experiments occur within evolving laboratory conditions where undocumented operational factors influence analytical outcomes [3,5,10,11]. Nominally identical methodologies may therefore generate divergent results because of variations in sample handling, instrument state, environmental conditions, operator behavior, or procedural adaptations. Rather than causing catastrophic failure, these factors often introduce subtle systematic deviations that remain within acceptable tolerances while progressively affecting data reliability and comparability [3,5,9].

Experimental outcomes are governed by interacting physical, procedural, environmental, instrumental, computational, and human factors operating beyond formal methodological descriptions [3,9]. This hidden operational context, including environmental fluctuations, instrument maintenance, sample history, operator technique, and software metadata, is often incompletely documented [5,10,12]. As a result, reproducing a published method does not necessarily reproduce the original experiment. Minor variations in sample equilibration, contamination, reagent storage, or instrument control may alter analytical outcomes without violating formal protocol instructions [2,6,13–15]. Similarly, instruments may satisfy calibration requirements while exhibiting subtle drift, and laboratories may remain compliant despite operational differences that affect reproducibility [16,17].

Current scientific reporting practices rarely capture this operational context [11]. Methods sections are often constrained by journal formatting limitations, assumptions of reader familiarity, and prioritization of primary methodological parameters over routine operational details [3,4]. Consequently, two laboratories may follow the same written procedure while unknowingly conducting operationally distinct experiments.

In this review, operational reproducibility is defined as the consistency of experimental outcomes achieved through effective control, documentation, and management of operational conditions beyond formally documented experimental protocols. Unlike methodological reproducibility, which focuses primarily on reproducing documented procedures, operational reproducibility emphasizes the systematic management of interacting human, environmental, reagent, machine, electronic, and sample-related factors within dynamic laboratory environments.

Existing laboratory quality frameworks include ISO/IEC 17025 [18], Good Laboratory Practice (GLP) [19], Laboratory Quality Management Systems (LQMS) [20], FAIR principles [12], standard operating procedures (SOPs), laboratory information management systems (LIMS), and electronic laboratory notebooks (ELNs) have substantially improved documentation, traceability, validation, and data stewardship. However, these systems primarily address individual quality-management functions rather than the interactions among human, environmental, reagent, machine, electronic, and sample factors that collectively influence reproducibility. Consequently, laboratories may achieve procedural compliance while hidden operational differences continue to affect experimental outcomes.

Concerns regarding reproducibility extend well beyond analytical chemistry and metrology. Large-scale reproducibility initiatives in psychology, biomedical sciences, and molecular biology, including the Open Science Collaboration reproducibility project [21], the ARRIVE guidelines for animal research [22], and the MIQE guidelines for quantitative PCR [23], have highlighted the importance of transparent reporting, standardized methodologies, and comprehensive documentation in improving research reliability. While these initiatives primarily address discipline-specific methodological and reporting practices,

HERMES complements these initiatives by emphasizing interacting operational factors influencing reproducibility across laboratory environments.

This narrative critical review addresses these challenges by developing an integrated operational perspective on experimental reproducibility. We synthesize current evidence on hidden workflow divergence, examine how operational instability propagates through laboratory workflows, and evaluate the limitations of existing mitigation strategies. Based on this synthesis, we introduce the HERMES (Human, Environmental, Reagent, Machine, Electronic, and Sample factors) conceptual framework to support systematic identification, documentation, and management of operational sources of experimental variability across multidisciplinary laboratory environments. The framework is intended as a conceptual foundation to guide future implementation, quantitative assessment, and empirical validation across different laboratory settings.

2. Conceptual Foundations and Operational Metrology

An accurate discussion of experimental reproducibility requires clear distinction among closely related but conceptually different measurement and operational terms. Terms such as accuracy, precision, repeatability, and reproducibility are often used interchangeably despite representing distinct dimensions of measurement quality [19,24–26]. Such ambiguity may obscure the origins of experimental variability and complicate the interpretation of reproducibility failures across laboratories, instruments, and workflows.

These concepts provide the metrological foundation for distinguishing routine analytical variability from hidden operational instability. Table 1 summarizes the principal metrological and operational terms discussed throughout this review, together with their primary definition sources and operational relevance.

Table 1. Core Definitions in Experimental Reproducibility and Operational Metrology.

Term/References	Operational Definition (Adapted from Cited Sources)	Operational Relevance
Accuracy [24–27]	How close a result is to the true value	Indicates result correctness
Precision [25,27]	Degree of agreement among repeated measurements	Reflects measurement consistency
Repeatability [19,24,25]	Consistency of results under the same conditions	Evaluates local analytical stability
Reproducibility [2,3]	Consistency of results under changed conditions	Assesses interlaboratory transferability
Robustness [28,29]	The ability of a method to tolerate small variations in operating conditions	Evaluates method stability under variation
Analytical Bias [4,7,24]	Systematic deviation from the true value	Indicates persistent measurement distortion
Measurement Uncertainty [25–27]	Estimated range within which the true value may lie	Quantifies measurement confidence
Traceability [19,24–26]	Ability to link measurements to recognized standards	Supports calibration comparability
Procedural Drift [30–32]	Gradual deviation from established procedures over time	Introducing hidden workflow instability
Metadata Integrity [19,31,33,34]	Completeness and traceability of experimental information	Supports reproducibility and transparency

Repeatability evaluates measurement consistency under tightly controlled conditions involving the same laboratory, instrument, operator, and a short time interval [19,24,25]. Reproducibility extends this concept by assessing whether comparable results can be obtained across different laboratories, instruments, operators, and time scales [3,19]. Methods demonstrating excellent repeatability within one laboratory may nevertheless produce substantially different results elsewhere because of undocumented operational variability [25,26,35].

Operational reproducibility expands this perspective by considering whether the conditions surrounding experiments remain functionally equivalent across workflows. Accordingly, the “same experiment” is defined not only by methodological instructions but also by the stability of the surrounding operational environment. Consequently, two laboratories may follow identical procedures while conducting operationally distinct experiments because of differences in environmental conditions, instrument maintenance history, sample storage practices, or undocumented procedural adaptations [36].

Unlike classical metrological concepts, this perspective does not replace repeatability, reproducibility, or robustness. Rather, it extends these concepts by incorporating operational conditions that are often treated as background assumptions rather than measurable determinants of analytical performance. It therefore provides a complementary, integrated perspective that emphasizes coordinated stability among human, environmental, reagent, machine, electronic, and sample-related factors needed to preserve functionally comparable experimental conditions across laboratories.

Minor operational deviations may interact cumulatively and nonlinearly. Variations in humidity, reagent aging, detector contamination, sample equilibration, or integration settings may individually appear insignificant while collectively influencing analytical outcomes beyond acceptable reproducibility limits [26,27]. This perspective reframes reproducibility as preservation of functional experimental equivalence rather than replication of methodology alone.

Because these operational factors interact across laboratory domains, their collective influence cannot be inferred from individual variables alone. Reproducibility, therefore, emerges from interactions among operational components rather than any single experimental factor, providing the conceptual basis for the HERMES framework developed in the following sections.

Metrology provides the theoretical foundation for evaluating measurement reliability and uncertainty propagation [19,24–27]. A key principle is measurement traceability, which links measurements to recognized reference standards through documented calibration chains [19,24,26]. However, traceability alone does not eliminate operational instability. Analytical systems may remain calibrated while variability arises from environmental fluctuations, procedural drift, maintenance inconsistencies, or incomplete metadata capture [24,30–32,37].

Operational uncertainty extends beyond conventional calibration models and instrument specifications. As laboratory workflows become increasingly automated and data-intensive, reproducibility must also account for interacting operational influences that extend beyond instrument performance alone.

Figure 1 summarizes the distinction between repeatability and reproducibility across increasing experimental variability. Repeatability reflects agreement under the same laboratory conditions, whereas reproducibility evaluates agreement when factors such as laboratories, operators, instruments, or time points differ.

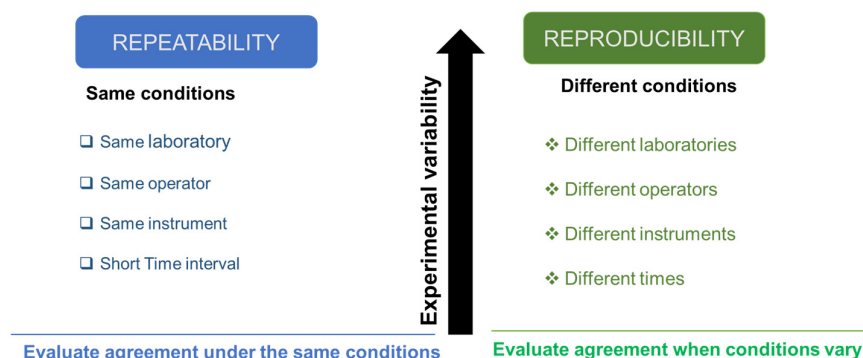


Figure 1. Simplified conceptual distinction between repeatability and reproducibility across increasing experimental variability. Repeatability reflects agreement under the same experimental conditions, whereas reproducibility evaluates agreement when laboratories, operators, instruments, or time points differ.

3. Taxonomy of Hidden Operational Error Domains

Hidden sources of laboratory variability rarely arise from isolated causes. Instead, it emerges through interactions among sample conditions, reagents, instrumentation, environmental influences, procedural practices, and digital workflows. Because these factors rarely trigger instrument failure or obvious protocol noncompliance, they often remain undetected by conventional quality-control systems while progressively degrading data reliability [2,3,37]. Figure 2 summarizes the major domains of hidden operational variability and their contributions to systematic bias, reduced precision, and compromised reproducibility.

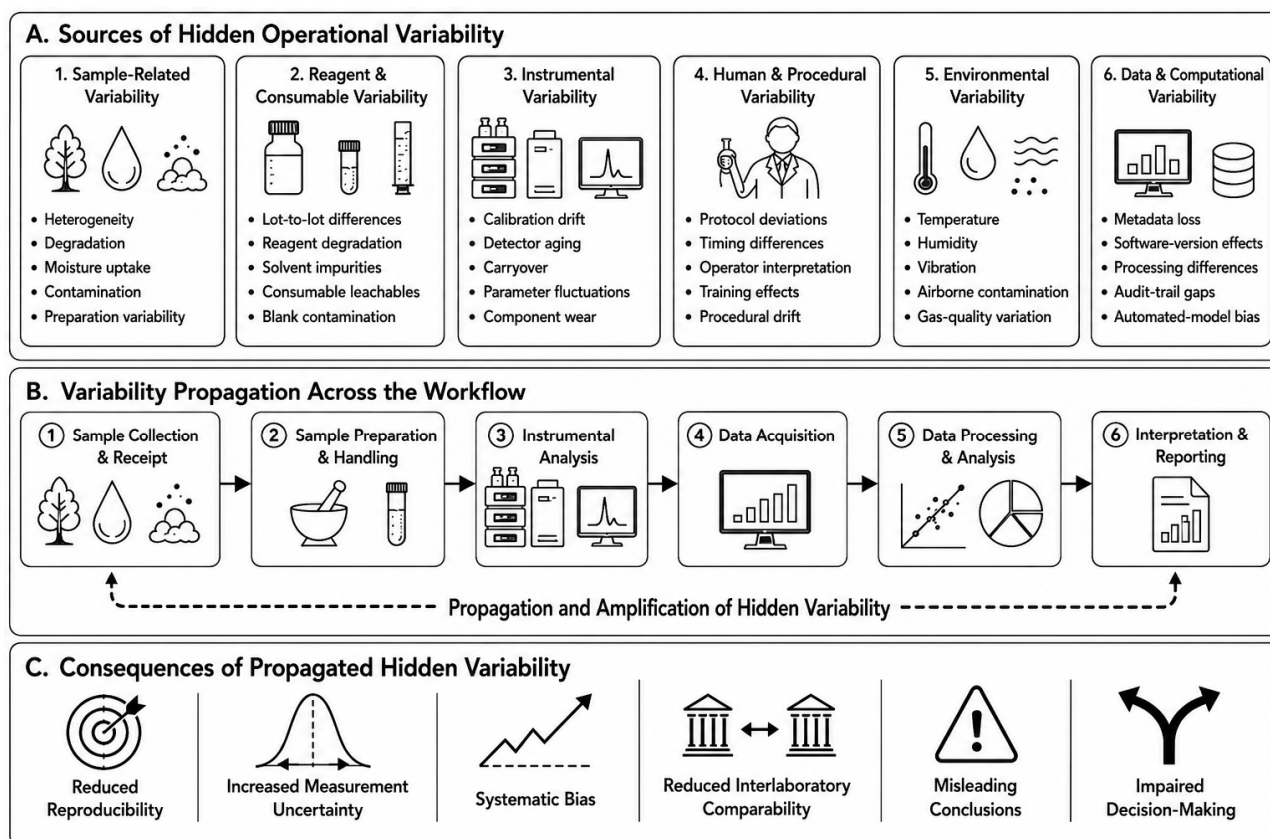


Figure 2. Taxonomy of hidden laboratory variability and its propagation through laboratory workflow. (A) Major sources of hidden operational variability. (B) Propagation pathways through analytical workflows. (C) Key consequences for analytical quality, comparability, and reproducibility.

The HERMES domains are conceptual categories rather than independent sources of error. Variability originating in one domain may influence others through interconnected operational pathways. The value of the taxonomy therefore lies in identifying both individual sources of variability and their interactions within laboratory workflows.

3.1. Sample-Related Errors

Sample-related variability is a fundamental source of experimental inconsistency because analytical measurements depend on the physical and chemical state of the analyte matrix [2,35,37]. Even with stable instrumentation, differences in sample preparation history, storage conditions, and handling practices may substantially influence reproducibility. Sample heterogeneity, including uneven analyte distribution, phase separation, and particle-size variability, can generate inconsistent responses between subsamples, particularly in biological and environmental matrices [3,37]. In addition, storage-related hydrolysis, oxidation, volatilization, photodegradation, microbial activity, and freeze–thaw cycles may progressively alter sample composition before analysis [37–39].

Physical interactions further contribute to hidden variability. Analyte adsorption onto container surfaces, tubing, or filtration materials may significantly affect trace-level measurements, while moisture uptake in hygroscopic materials can alter sample mass, extraction efficiency, and detector response [2,5,37].

Operational factors, including cross-contamination, incomplete dissolution, variable extraction efficiency, and operator-dependent preparation techniques, introduce additional sources of variability [27–31]. Because sample integrity depends on multiple operational conditions, sample-related variability frequently influences other HERMES domains.

3.2. Reagent and Consumable Errors

Although often treated as standardized inputs, reagents and consumables may introduce analytical variability through differences in purity, stability, and manufacturing lots [25,40–43]. Lot-to-lot variation can influence reaction kinetics, chromatographic behavior, derivatization performance, and analytical selectivity [41,44,45]. Similarly, trace solvent impurities, dissolved gases, peroxide formation, and degradation products may contribute to baseline instability, ghost peaks, analyte suppression, and other analytical artifacts [46–50]. Degradation of analytical standards may further introduce systematic calibration bias [25,51,52].

Consumable materials, including tubing, septa, filters, and pipette tips, may release plasticizers, silicone residues, fluorinated compounds, or other leachable species into analytical workflows [37,53–55]. Such contamination is particularly problematic in ultra-trace analyses, PFAS investigations, and mass spectrometry applications, where low-level background contributions may significantly influence analytical outcomes.

Measurable blank contamination originating from solvents, carrier gases, consumables, or laboratory surfaces further demonstrates that reagent quality cannot be assumed solely from supplier specifications [41,42,53,56–58]. Routine quality-control verification and contamination monitoring are therefore essential.

3.3. Instrumental and Equipment Errors

Instrument functionality does not guarantee analytical equivalence. Properly functioning and calibrated systems may still exhibit hidden variability arising from component aging, contamination states, calibration drift, maintenance history, and environmental fluctuations [2,29,37].

Calibration drift may alter detector responses, wavelength alignments, thermal sensors, and flow controllers over time [2,29,32]. Similarly, detector aging, ion-source fouling, sensor fatigue, and column degradation may influence selectivity, peak shape, and signal

intensity [2,37,48,59]. Minor variations in carrier-gas flow, pressure regulation, and thermal equilibration may further affect analytical performance without triggering instrument error codes [32,37].

These effects extend across diverse analytical platforms. For example, acquisition parameters and sample preparation conditions may influence morphology and phase characterization in microscopy and diffraction techniques, while dispersion conditions and measurement settings can affect particle-size and surface-charge determinations in solution-based analyses [60–63]. Similarly, emerging nanozyme-based sensing and nanomaterial-enabled remediation systems illustrate that analytical performance depends not only on the intrinsic properties of engineered nanomaterials but also on sample composition, matrix chemistry, material-dispersion state, and the microscopy, spectroscopy, fluorescence, particle-size, and surface-charge characterization conditions used to evaluate these materials. Consequently, hidden operational variability can influence both material characterization and subsequent analytical performance, underscoring the need for standardized operational practices throughout the analytical workflow [64,65]. These examples demonstrate that analytical outputs reflect both intrinsic sample properties and the operational conditions under which measurements are acquired.

Reproducible reporting should therefore include relevant acquisition parameters, calibration status, preprocessing settings, and quality-control practices in addition to instrument specifications [60,61]. Comprehensive reporting of instrument condition, calibration status, acquisition parameters, and quality-control practices is therefore essential for reproducible analytical measurements.

3.4. Human and Procedural Errors

Human and procedural variability remain exceptionally difficult to standardize because laboratory workflows rely heavily on judgment, procedural interpretation, and operator-specific tacit knowledge [3,18,66]. Over time, laboratories may accumulate undocumented workflow deviations, a phenomenon referred to as procedural drift [18,37,66]. Operators frequently adapt workflows in response to instrument behavior or sample characteristics by modifying equilibration times, altering rinsing practices, adjusting injection sequences, or employing selective data integration without corresponding updates to written protocols [18,37,66].

Timing inconsistencies further contribute to hidden variability. Delays between sample preparation and analysis, variable extraction durations, and inconsistent equilibration periods may alter reaction progress, degradation pathways, and analyte recovery [3,37,66]. In addition, critical operational practices, including troubleshooting strategies and contamination-control techniques, are often transferred informally rather than formally documented [3,18,66]. Although such tacit knowledge may support local workflow performance, reliance on undocumented expertise limits method transferability and interlaboratory reproducibility. Consistent documentation and adherence to standardized procedures are therefore essential for improving method transferability and interlaboratory reproducibility.

3.5. Environmental Errors

Many analytical methods assume stable ambient conditions; consequently, environmental fluctuations frequently bypass conventional quality-control thresholds while progressively degrading reproducibility [2,37]. Variations in temperature and relative humidity may alter sample mass through moisture uptake, influence reaction kinetics, accelerate solvent evaporation, and promote electrostatic charge accumulation [2,5,32,37]. These

effects are particularly important for gravimetric analyses, hygroscopic materials, and surface-sensitive measurements.

Mechanical vibrations may disrupt microbalances, microscopy platforms, and optical systems, while ambient dust, aerosols, and volatile organic compounds can contribute to background interference, particularly in trace-level workflows [2,5,32,37]. Similarly, fluctuations in carrier-gas or plasma-gas quality, including variations in moisture, oxygen, hydrocarbon impurities, and pressure regulation, may compromise baseline stability, ionization efficiency, and detector performance in chromatographic and mass spectrometric analyses [2,32,37]. Environmental conditions therefore act as cross-cutting modifiers capable of influencing nearly every HERMES domain.

3.6. Electronic and Digital Errors

Digital processing and laboratory automation should be regarded as active components of experimental workflows rather than neutral post-acquisition tools. Although automated instrumentation, robotic preparation systems, electronic laboratory notebooks (ELNs), and laboratory information management systems (LIMS) reduce manual handling errors, they do not eliminate reproducibility challenges. Instead, they redistribute variability toward software-dependent, algorithmic, and data-integration domains [12,67–70].

Incomplete metadata capture is a major source of instability in digital reproducibility. Missing instrument configurations, software versions, preprocessing parameters, and processing histories limit the reconstruction of experimental conditions and compromise reproducibility assessments [27,68,69]. This challenge is further compounded by metadata fragmentation across instrument software, databases, cloud storage platforms, statistical packages, and spreadsheets that store operational information inconsistently across disconnected systems [11,68–70].

Digital reproducibility also depends on maintaining the integrity, security, and long-term availability of electronic records. Cybersecurity incidents, unauthorized data modifications, database corruption, and failures in data backup or recovery may compromise reproducibility by disrupting data provenance, auditability, and access to experimental records. Although these risks are often addressed through institutional information technology and data-governance practices, they remain important operational considerations within increasingly digital laboratory environments [12,67–70].

Analytical outcomes may also be influenced by opaque preprocessing workflows and algorithmic drift. Minor software updates or user-defined settings can alter baseline correction routines, noise filtering, peak-detection thresholds, smoothing functions, and calibration models [12,67,69,71]. Because these processes frequently operate automatically and with limited user visibility, they may substantially affect quantitative outputs in chromatography, spectroscopy, imaging, and other computationally intensive analytical workflows [12,67,68,71–73]. Fragmented audit trails further complicate reproducibility by obscuring when data were reprocessed, excluded, recalculated, or otherwise modified during analysis [27,68,69].

The increasing adoption of artificial intelligence and machine-learning approaches in analytical chemistry workflows presents emerging considerations for reproducibility [69]. As these technologies become more widely integrated into data processing and decision support, maintaining transparency, auditability, explainability, and comprehensive metadata will be essential to ensure that AI-assisted analyses remain reproducible and interpretable.

Importantly, the apparent precision and standardization of digital workflows may create an illusion of stability, whereby computational outputs are assumed to be inherently reproducible despite hidden variability within software systems and data-processing pipelines. Improving digital reproducibility therefore requires metadata interoperabil-

ity, software transparency, algorithmic accountability, comprehensive auditability, and long-term digital traceability. Maintaining metadata integrity, software transparency, algorithmic accountability, and long-term digital traceability is therefore essential for preserving reproducibility.

4. Operational Error Propagation and Persistence

Operational variability rarely remains confined to its stage of origin. Hidden errors may propagate and accumulate throughout analytical workflows, producing downstream divergence [3,32,37]. Because many deviations remain within nominal acceptance tolerances, reproducibility degradation often emerges from the cumulative effects of multiple low-visibility perturbations rather than isolated failures. Accordingly, localized perturbations may be attenuated, propagated, or amplified as they move through analytical workflows, producing cumulative effects on reproducibility. Although standardized quality systems improve documentation and traceability, hidden operational variability may still propagate through laboratory workflows despite procedural compliance [2,3,11,18].

4.1. Workflow-Scale Error Cascades

Experimental workflows consist of sequential operational stages in which outputs from early processes become inputs for downstream analytical steps (Figure 3). Consequently, variability introduced during early stages may propagate and become amplified throughout the analytical lifecycle.

Minor moisture uptake during storage, for example, may alter extraction efficiency during preparation and subsequently influence detector response and final reporting conclusions [37,39,74]. Similarly, low-level contamination introduced during consumable handling may bypass initial quality-control checkpoints while later contributing to baseline instability or false-positive signals in high-sensitivity workflows [75].

Nonlinear amplification frequently occurs during data processing and interpretation. Instrumental drift producing negligible raw signal variation may become magnified through normalization, baseline correction, threshold filtering, and peak integration procedures [67,72]. Consequently, root-cause identification may be challenging because downstream failures often originate from interacting upstream instabilities rather than singular observable events.

Operational variability may also persist within archived datasets, calibration systems, protocols, and institutional practices long after the original disturbance has disappeared. Minor operational deviations may become normalized through routine workflows and incomplete documentation, allowing variability to propagate across experiments and personnel. Figure 3 summarizes these amplification and persistence mechanisms.

4.2. Interacting Error Systems

Hidden laboratory variability frequently develops through interacting operational components in which environmental conditions, sample states, instrumentation, procedural decisions, and computational workflows influence one another simultaneously [3,32,37].

Operational variables often exhibit coupled behavior. Thermal instability, for example, may simultaneously affect solvent evaporation, reaction kinetics, gas density, and detector performance [32,37]. In chromatography–mass spectrometry workflows, environmental fluctuations may further interact with carrier-gas stability, ionization efficiency, and automated data-processing routines [72,76,77].

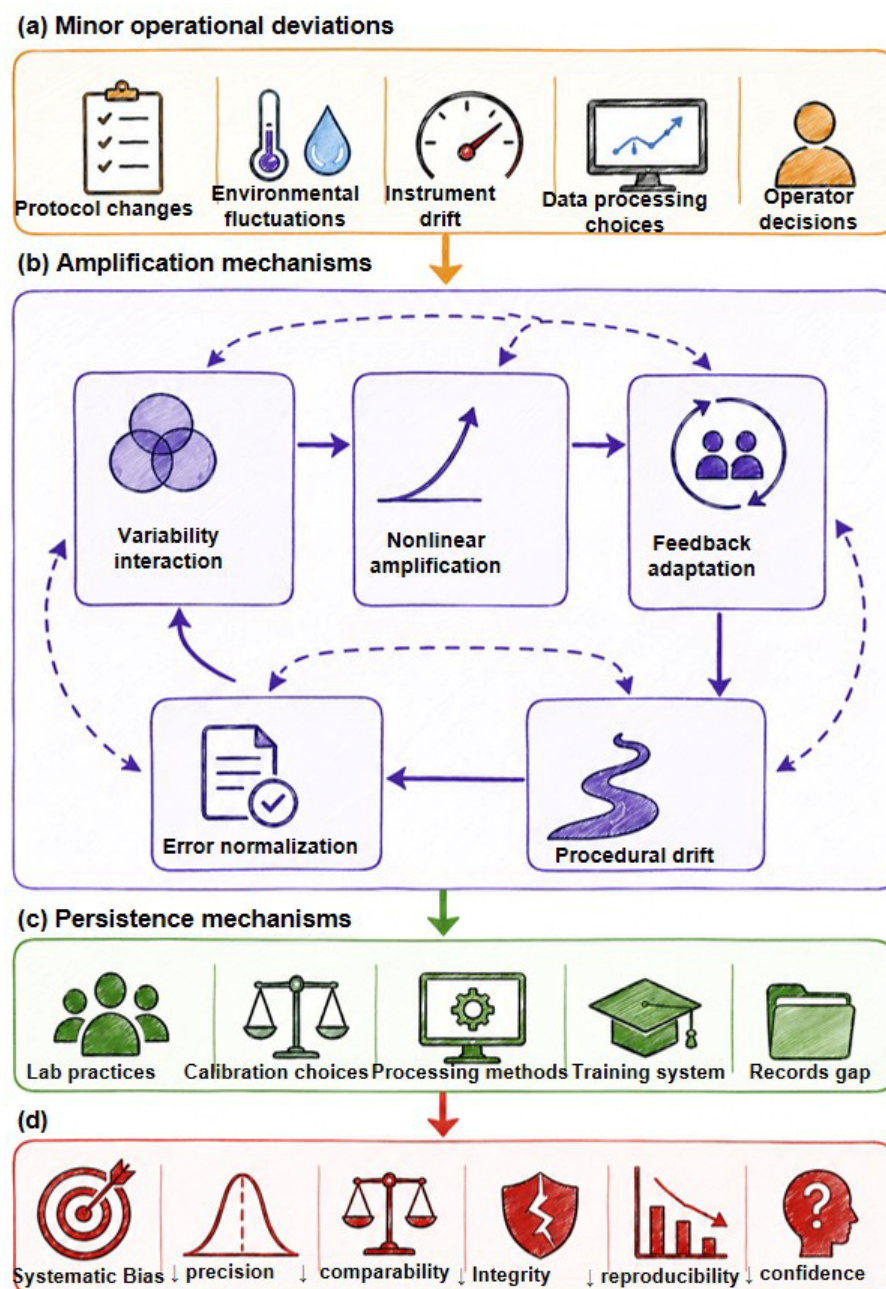


Figure 3. The Domain-based framework illustrates the origin, amplification, persistence, and consequences of hidden operational variability. (a) The orange domain identifies minor operational deviations arising from protocol changes, environmental fluctuations, instrument drift, data-processing choices, and operator decisions. (b) The purple domain shows how these deviations may be amplified through interacting variability, nonlinear responses, feedback adaptation, procedural drift, and error normalization. Solid arrows indicate the primary progression of operational variability through the framework, whereas dashed arrows represent feedback interactions and reinforcing pathways through which variability may propagate, recur, and become normalized across laboratory workflows. (c) The green domain highlights persistence mechanisms, including laboratory practices, calibration choices, processing methods, training systems, and documentation gaps. (d) The red domain presents the long-term consequences, including systematic bias, reduced precision, impaired comparability and integrity, diminished reproducibility, and lower confidence in analytical conclusions.

Interacting systems may also generate hidden covariance, whereby variables appearing independent during routine monitoring vary together under real laboratory conditions.

Seasonal humidity changes, for example, may simultaneously influence electrostatic behavior, powder-transfer efficiency, airborne contamination, and balance stability [5].

Laboratory personnel frequently respond to perceived instability through recalibration, parameter adjustment, selective reinjection, or procedural modification [18,37,66]. Although these adaptations may temporarily stabilize analytical outputs, they may also reinforce hidden covariance structures and contribute to long-term procedural drift. Consequently, identifying a single root cause may be insufficient because operational instability often reflects multiple interacting mechanisms rather than independent failures.

4.3. *The Persistence Paradox and the Compliance Artifact Trap*

A challenge in experimental reproducibility is that operational instabilities often persist within laboratory systems despite quality-control frameworks, validation protocols, and regulatory oversight [2,3,18]. This arises because hidden variability is frequently normalized, accommodated, or operationally absorbed rather than explicitly identified and eliminated.

Instability often survives through several structural mechanisms. Laboratories may implement localized troubleshooting modifications that restore apparent short-term functionality while obscuring deeper systems-level divergence [18]. Recurrent baseline drift may be accommodated through adjusted integration thresholds, unstable extraction efficiencies may be addressed through repeated sample reinjections, and inconsistent preprocessing practices may gradually evolve into laboratory-specific analytical conventions [18,38,59,72].

Fragmented metadata capture further reinforces persistence. Many workflows rely on incomplete instrument logs, inconsistent software-processing histories, and informal communication of critical operational practices [18,26,78]. When essential knowledge remains embedded within operator experience rather than formal infrastructure, retrospective reconstruction of environmental conditions, calibration states, and software modifications becomes increasingly difficult [68].

This review introduces the concept of a Compliance Artifact Trap, whereby quality systems become increasingly focused on demonstrating regulatory conformity rather than continuously evaluating operational stability. The Compliance Artifact Trap does not represent a failure of quality-management systems themselves but rather a limitation of relying predominantly on documentation-based indicators of performance. Frameworks such as ISO 17025, GLP, SOPs, LIMS, and ELNs remain indispensable for laboratory quality; however, when operational monitoring fails to evolve alongside laboratory processes, they may create a perception of reproducibility that exceeds the actual stability of experimental workflows [3,18,26]. Consequently, laboratories may maintain comprehensive SOPs, calibration records, and audit trails while simultaneously experiencing environmental fluctuations, instrument aging, contamination-state evolution, or software-version divergence [2,37]. Likewise, fixed calibration schedules and pass/fail acceptance criteria may overlook gradual changes in analytical performance occurring between scheduled evaluations [2,29].

Over time, these adaptations allow analytical systems to diverge from their originally validated operational states while preserving the appearance of regulatory compliance. Thus, reproducibility failures often arise not from isolated catastrophic events but from the gradual normalization of interacting, low-visibility operational instabilities distributed throughout experimental systems. Within the HERMES framework, recognizing these propagation and persistence mechanisms provides the basis for developing proactive operational monitoring strategies that complement existing quality-management systems by identifying emerging instability before reproducibility is measurably compromised.

5. Cross-Disciplinary Manifestations

Hidden operational variability manifests differently across scientific disciplines because laboratory workflows differ in sample complexity, instrumentation, workflow architectures, and experimental design. Despite these differences, many operational vulnerabilities are shared. Calibration drift, environmental sensitivity, metadata loss, sample instability, and procedural inconsistencies operate through discipline-specific pathways [3,32,37]. Table 2 summarizes representative operational vulnerabilities across major laboratory disciplines and illustrates how common operational mechanisms produce comparable challenges to measurement reliability and experimental reproducibility.

Table 2. Cross-Disciplinary Operational Vulnerability Matrix.

Laboratory Domain	Representative Hidden Operational Vulnerabilities	Potential Reproducibility Consequences
Analytical Chemistry Laboratories	Calibration drift, carryover, integration bias, matrix effects, and detector aging	Quantitative bias, peak distortion, and inconsistent analytical sensitivity
Biological and Biomedical Laboratories	Cell passage variability, incubation instability, freeze–thaw degradation, and operator-dependent handling	Altered cellular behavior, reduced biological comparability, variable assay response
Materials Science Laboratories	Humidity-sensitive materials, thermal history dependence, particle-size heterogeneity, surface contamination	Irreproducible material properties, altered degradation behavior, and inconsistent morphology
Environmental and Field Laboratories	Transport instability, preservation failure, field contamination, storage degradation	Sample alteration before analysis, false-positive contamination signals
Engineering and Process Laboratories	Scale-up instability, sensor drift, flow inconsistency, reactor nonuniformity	Process-performance divergence, unstable operational control, inconsistent scale translation

Analytical chemistry workflows are particularly vulnerable because they rely on trace-level quantitation and multistage computational processing [13,19,35]. In gas chromatography–mass spectrometry (GC–MS), residual analytes retained within injection ports, transfer lines, column surfaces, or ion sources may generate carryover effects that distort subsequent analyses, particularly during trace-level determinations [2,29,46]. Quantitative outcomes are also sensitive to preprocessing choices, including baseline correction routines, peak integration boundaries, and smoothing algorithms, which may produce different quantitative results despite identical raw datasets [67,78]. In liquid chromatography–mass spectrometry (LC–MS), matrix effects arising from coeluting compounds may cause ion suppression or enhancement, resulting in inaccurate quantification and reduced interlaboratory comparability [46,48].

Biological reproducibility is challenged by the dynamic nature of living systems [3]. Serial passaging of commonly used cell lines, such as HeLa or HEK293 cells, may introduce genetic drift, phenotypic adaptation, and altered metabolic behavior over time, contributing to variability between laboratories [3]. Minor fluctuations in incubator temperature, humidity, or CO₂ regulation may further influence cellular growth kinetics and assay sensitivity [79]. In addition, repeated freeze–thaw cycles of serum, plasma, proteins, or nucleic acid extracts can induce degradation and compromise comparability across experiments [59].

Materials characterization workflows are highly sensitive to environmental exposure and processing history [32,37]. Hygroscopic materials, including metal oxides, porous polymers, and biomass-derived materials, may undergo moisture adsorption that alters mass, conductivity, and surface reactivity [37]. Similarly, minor differences in thermal history during annealing, calcination, or polymer curing may substantially influence crystallinity, degradation pathways, and mechanical performance [32,37]. Sample heterogeneity arising from particle-size variation, surface contamination, or inconsistent packing may further contribute to variability in spectroscopic, microscopic, and diffraction-based measurements [37].

Environmental measurements are particularly susceptible to variability introduced during sampling, transportation, and preservation [3,33]. For example, delays in processing water samples may alter dissolved oxygen concentrations and microbial composition, whereas volatilization losses can affect measurements of organic contaminants [80]. Inadequate cooling, inappropriate preservatives, or unsuitable container materials may further modify sample composition or introduce contamination artifacts, particularly in ultra-trace analyses such as PFAS determination [20,33].

Engineering workflows must account for scale dependence and interacting physical variables. Processes optimized at laboratory scale frequently encounter challenges during pilot-scale translation because heat transfer, mixing behavior, residence-time distributions, and fluid dynamics change nonlinearly with reactor volume [32]. In continuous manufacturing systems, sensor drift and spatial nonuniformities in flow distribution, turbulence, or pressure regulation may further influence reaction kinetics, heat dissipation, and final product selectivity [2,37].

Various scientific disciplines have developed field-specific initiatives to improve reproducibility. Psychology has undertaken large-scale replication studies to quantify reproducibility challenges [21], while biomedical sciences have established reporting standards such as the ARRIVE guidelines for animal research [22] and the MIQE guidelines for quantitative real-time PCR experiments [23]. Although these initiatives have substantially improved methodological transparency and reporting within their respective disciplines, they remain largely domain-specific.

In contrast, HERMES is not intended to replace discipline-specific reporting standards or regulatory frameworks. Rather, it complements these efforts by providing a systems-oriented operational framework that identifies common sources of hidden variability arising from interactions among human, environmental, reagent, machine, electronic, and sample factors across diverse laboratory settings.

These examples demonstrate that hidden operational variability is a common challenge across laboratory disciplines despite substantial differences in experimental methods. Although its manifestations differ according to scientific context, many underlying operational vulnerabilities are shared. This cross-disciplinary consistency supports the application of the HERMES framework as a conceptual structure for identifying and interpreting operational sources of variability across diverse laboratory environments.

6. The HERMES Operational Reproducibility Framework

The preceding sections demonstrate that reproducibility degradation rarely arises from isolated methodological failures but from interactions among personnel, environmental conditions, reagents, instrumentation, computational infrastructure, and evolving sample states throughout experimental workflows. Existing quality management systems often evaluate individual components while incompletely assessing their combined operational behavior [2,3,37]. To address this limitation, this review proposes the HERMES Operational Reproducibility Framework. This system-oriented model organizes six interconnected

domains: Human, Environmental, Reagent, Machine, Electronic, and Sample, to identify, interpret, and manage interacting reproducibility risks across the experimental lifecycle.

Importantly, HERMES is not intended to replace existing QA/QC or regulatory systems. Rather, it functions as a complementary operational framework that could enhance laboratory infrastructure by identifying interacting sources of variability that are often underrepresented in conventional quality management approaches. The framework is founded on the principle that reproducibility depends not only on repeating experimental procedures but also on preserving operational equivalence, that is, maintaining functionally comparable experimental conditions despite the inevitable variability present in laboratory environments. Consequently, HERMES emphasizes the identification, interpretation, and management of operational factors that meaningfully influence reproducibility rather than the elimination of variability itself. Existing systems such as GLP, ISO 17025, FDA guidance, SOPs, LIMS, and ELNs establish requirements for documentation, validation, calibration, and quality assurance, whereas HERMES provides a complementary systems-oriented perspective for monitoring interactions among operational factors that are not routinely evaluated within these quality-management frameworks.

6.1. Framework Philosophy and Architecture

HERMES shifts reproducibility assessment away from static procedural conformity toward dynamic operational-state management. Rather than asking only whether documented procedures were followed, the framework evaluates whether operational conditions remained functionally comparable throughout experimental execution. Consequently, reproducibility failures may occur despite procedural compliance because environmental fluctuations, procedural adaptations, contamination-state evolution, software-version divergence, and workflow drift alter analytical outcomes [18,33].

Accordingly, HERMES treats reproducibility as an emergent property of interconnected operational stability [81]. Variability within one domain may influence the stability or sensitivity of another, emphasizing operational visibility, contextual awareness, traceability, and early instability detection rather than complete elimination of variability.

The operational implementation component of HERMES is organized through the HERMES Control Matrix (Table 3), which evaluates each domain according to five operational questions: what should be monitored, standardized, documented, audited, and recognized as a source of hidden uncertainty.

The Control Matrix serves as a conceptual decision-support framework that enables laboratories to prioritize operational domains according to analytical technique, workflow complexity, regulatory requirements, available resources, and susceptibility to hidden variability, rather than functioning as a prescriptive checklist. Consequently, its indicators are intentionally expressed in technology-neutral terms to accommodate differences in laboratory disciplines, instrumentation, and organizational contexts. Quantitative performance indicators, monitoring frequencies, decision thresholds, and standardized assessment tools should be developed and empirically validated within specific laboratory settings before routine implementation.

Monitoring priorities will therefore vary according to analytical sensitivity, workflow complexity, instrumentation, and environmental exposure. Figure 4 illustrates the system-level architecture of the HERMES framework.

Table 3. The HERMES Operational Control Matrix.

Domain	Monitor	Standardize	Document	Audit	Hidden Uncertainty Sources
Human	Workflow execution and training consistency	Procedural interpretation and timing practices	Deviations and operator actions	Training records and procedural adaptation	Tacit knowledge and operator habits
Environmental	Temperature, humidity, and gas quality	Environmental-control conditions	Environmental logs	Monitoring system stability	Seasonal variability and airborne contamination
Reagent	Purity, lot consistency, and blank response	Reagent sourcing and storage	Lot numbers and expiration records	Contamination-control practices	Leachable, degradation, and background contamination
Machine	Calibration state and detector performance	Maintenance intervals and instrument settings	Calibration and maintenance history	Instrument-performance trends	Drift, aging, and hidden contamination
Electronic	Software versions and metadata continuity	Preprocessing workflows	Audit trails and computational settings	Data-integrity systems	Hidden algorithms and synchronization failures
Sample	Storage conditions and degradation indicators	Preparation procedures	Handling history and sample state	Chain-of-custody systems	Moisture uptake, adsorption, degradation, and heterogeneity

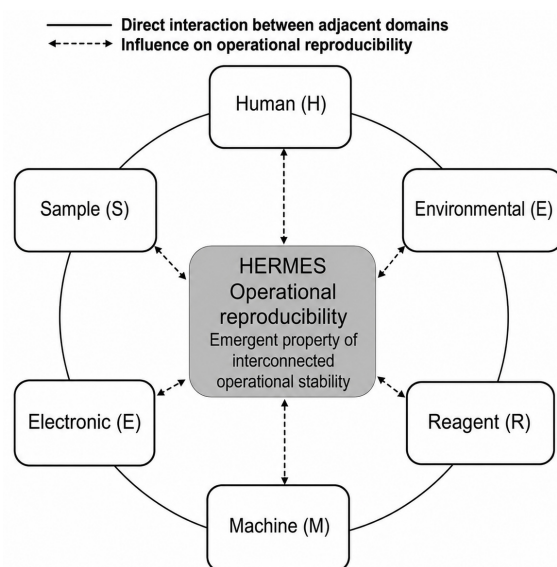


Figure 4. HERMES Operational Reproducibility Model. Conceptual systems model of six interacting operational domains within the HERMES framework: Human, Environmental, Reagent, Machine, Electronic, and Sample. Each domain is shown as a distinct component that contributes to overall operational reproducibility, while the central HERMES node represents reproducibility as an emergent property of interconnected operational stability. The model emphasizes that interactions among domains throughout experimental workflows may amplify hidden variability across laboratory systems.

6.2. Scalable Implementation Continuum

Implementation of operational reproducibility systems should not be interpreted as a binary transition between “noncompliant” and “fully reproducible” laboratories. Instead, reproducibility infrastructure exists along a continuum of operational maturity influenced by laboratory resources, workflow complexity, analytical sensitivity, computational integra-

tion, regulatory requirements, and institutional priorities. Foundational implementation emphasizes metadata integrity, contextual documentation, environmental awareness, contamination monitoring, calibration verification, preprocessing transparency, and personnel training [37,68].

The resources required to implement HERMES will vary according to laboratory size, workflow complexity, existing quality-management infrastructure, and regulatory requirements. Foundational implementation emphasizes improvements in documentation, metadata capture, personnel training, and operational awareness, whereas more advanced implementation may incorporate integrated environmental monitoring, predictive maintenance, interoperable digital infrastructure, and AI-assisted operational analytics. Accordingly, HERMES is designed as a scalable framework that laboratories may adopt progressively according to their operational needs and available resources.

Implementation of HERMES may also present practical challenges, including increased documentation requirements, personnel training, resistance to workflow changes, and additional demands on laboratory time and resources. These considerations are particularly relevant for laboratories with limited infrastructure or staffing. Accordingly, HERMES is intended as a flexible, risk-based framework that supports incremental adoption, allowing laboratories to prioritize operational domains according to their specific needs and resource constraints rather than implementing all components simultaneously. Future studies should evaluate the practical costs, benefits, and organizational impacts of HERMES implementation across diverse laboratory settings.

Intermediate implementation emphasizes integration of environmental monitoring, instrument-state tracking, audit systems, preprocessing traceability, and metadata management within coordinated workflow frameworks. Interlaboratory comparisons and operational trend analyses further support assessment of workflow transferability and long-term analytical consistency [68,78]. Advanced implementation incorporates predictive monitoring to identify emerging instability before substantial analytical divergence occurs. Integrating environmental conditions, instrument performance, preprocessing trends, contamination indicators, and metadata continuity may support early detection of instability. Predictive maintenance and transparent AI-assisted analytics may further strengthen reproducibility when computational traceability and auditability are maintained [78].

Importantly, advanced reproducibility systems should not be interpreted as attempts to eliminate all variability. Rather, the objective is to strengthen operational visibility, reduce normalization of hidden divergence, and preserve long-term operational equivalence. Future development of the HERMES framework should include empirical validation through controlled interlaboratory studies, quantitative operational indicators, and integration with existing laboratory quality-management systems. Such studies will enable refinement of monitoring priorities, evaluation of domain interactions, and assessment of the framework's effectiveness across diverse scientific disciplines.

6.3. Emerging Directions in Operational Reproducibility

Future reproducibility efforts will likely shift from isolated methodological validation toward integrated operational systems analysis. As laboratory workflows become increasingly computationally intensive, automated, multidisciplinary, and interconnected, reproducibility challenges will arise from interactions among physical, digital, environmental, and organizational systems. Within this context, the HERMES framework should be viewed as a foundation for future operational reproducibility research rather than a finalized implementation model. Its principal contribution is to provide a system-oriented structure that can guide the development, evaluation, and integration of future reproducibility technologies and operational monitoring strategies.

One important direction involves the development of reproducibility-aware metadata infrastructures capable of preserving operational context throughout analytical workflows [68]. Automated capture of environmental conditions, instrument configurations, preprocessing histories, and workflow interactions, supported by standardized metadata models and interoperable digital infrastructures, could improve reproducibility across laboratories by enabling consistent documentation and exchange of operational information.

Computational reproducibility will become increasingly important as laboratories increasingly adopt automated preprocessing, statistical modeling, and AI-assisted analytical tools. These technologies may have the potential to improve efficiency and consistency. However, their contribution to reproducibility will depend on transparent documentation of algorithms, software versions, preprocessing workflows, model evolution, auditability, explainability, and metadata continuity.

Another important transition will be a shift from episodic validation toward continuous assessment of operational states. Rather than relying solely on periodic calibration checks, validation exercises, or interlaboratory studies, future reproducibility frameworks will increasingly emphasize continuous monitoring of operational equivalence across evolving workflows.

Unlike conventional quality-control systems, which primarily evaluate predefined analytical performance at discrete intervals, continuous operational assessment emphasizes ongoing monitoring of operational conditions that influence reproducibility throughout the experimental workflow. Examples include environmental status, instrument performance, metadata continuity, contamination indicators, and workflow integrity. The development of standardized operational indicators, monitoring frequencies, and decision thresholds represents an important area for future empirical research.

Ultimately, progress in reproducibility science will depend on both technological advancement and laboratory cultures that prioritize operational transparency, systems awareness, and early recognition of emerging instability. Future research should focus on empirical validation of the HERMES framework through controlled interlaboratory studies, development of quantitative operational indicators, assessment of domain interactions, and evaluation across diverse scientific disciplines. Such efforts will determine the practical utility of HERMES while guiding its continued refinement and integration into laboratory quality-management systems.

7. Conclusions

Experimental reproducibility is often discussed in terms of methodology, statistical rigor, and experimental design. However, the evidence examined throughout this review demonstrates that reproducibility is also strongly influenced by operational factors embedded within routine laboratory workflows. Rather than representing isolated technical deficiencies, these operational factors interact across laboratory systems, allowing small perturbations to propagate, accumulate, and influence experimental outcomes over time.

A key observation is that reproducibility failure rarely results from a single catastrophic event. Instead, it frequently emerges through the gradual accumulation of low-visibility operational deviations, including environmental fluctuations, procedural adaptation, calibration drift, sample instability, metadata discontinuity, and computational inconsistency. Individually, these factors may appear insignificant, but collectively they can substantially affect experimental outcomes and data reliability. Such challenges are increasingly important in modern laboratory environments characterized by automation, digital integration, multidisciplinary collaboration, and complex analytical workflows.

This review further highlights the limitations of relying solely on traditional quality management measures. While standard operating procedures, quality assurance systems, calibration protocols, electronic laboratory records, and regulatory frameworks improve standardization and traceability, they may not fully capture the dynamic interactions that influence operational equivalence over time. Consequently, laboratories can remain procedurally compliant while gradually diverging from previously validated analytical conditions.

To support a more systematic approach to reproducibility management, this review presents the HERMES Operational Reproducibility Framework as a conceptual systems model for identifying and evaluating hidden sources of operational variability. By considering human, environmental, reagent, machine, electronic, and sample factors as interconnected components of laboratory systems, the framework encourages greater operational awareness, contextual traceability, and proactive recognition of instability throughout experimental workflows. Importantly, HERMES is intended to complement rather than replace existing quality-management systems. By providing a structured operational perspective spanning human, environmental, reagent, machine, electronic, and sample domains, the framework extends conventional reproducibility practices through greater emphasis on operational equivalence and systems-level awareness.

Ultimately, reproducibility should be viewed not only as a methodological objective but also as a systems-level operational property. Continued progress in reproducibility science will require integration of analytical rigor with environmental awareness, computational traceability, metadata integrity, workflow evaluation, and operational-state monitoring. Strengthening these interconnected dimensions may improve data reliability, enhance scientific confidence, and support more robust and reproducible research across disciplines. Future work should focus on empirical validation of the HERMES framework through controlled interlaboratory studies, development of quantitative operational indicators, and evaluation across diverse laboratory environments. Such efforts will help determine the framework's practical effectiveness while refining its application to evolving scientific workflows.

Author Contributions: Conceptualization, E.O. and J.N.E.Y.; methodology, E.O.; formal analysis, E.O. and H.C.M.; investigation, E.O., J.N.E.Y. and H.C.M.; resources, E.O. and J.N.E.Y.; data curation, E.O. and H.C.M.; writing—original draft preparation, E.O.; writing—review and editing, E.O., J.N.E.Y. and H.C.M.; visualization, J.N.E.Y. and H.C.M.; supervision, E.O. and H.C.M.; project administration, E.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new experimental data were generated or analyzed in this study. Data sharing is therefore not applicable to this review article.

Acknowledgments: During the preparation of this manuscript, ChatGPT (GPT-4, OpenAI, 2025) was used to assist with language refinement, grammar correction, structural organization, and editorial clarity. It was not used to generate scientific concepts, interpret findings, evaluate the literature, formulate conclusions, or make scientific decisions. All scientific content, analysis, interpretation, and final manuscript revisions were developed, verified, and approved by the authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Definition
AI	Artificial Intelligence
BIPM	Bureau international des poids et mesures
CLSI	Clinical and Laboratory Standards Institute
ELN	Electronic Laboratory Notebook
FAIR	Findable, Accessible, Interoperable, and Reusable
GLP	Good Laboratory Practice
HERMES	Human, Environmental, Reagent, Machine, Electronic, and Sample
HVAC	Heating, Ventilation, and Air Conditioning
ICH	International Council for Harmonization
ISO	International Organization for Standardization
IEC	International Electrotechnical Commission
LIMS	Laboratory Information Management System
QA	Quality Assurance
QC	Quality Control
SOP	Standard Operating Procedure
USP	United States Pharmacopeia
VIM	International Vocabulary of Metrology

References

- Goodman, S.N.; Fanelli, D.; Ioannidis, J.P.A. What Does Research Reproducibility Mean? *Sci. Transl. Med.* **2016**, *8*, 341ps12. [CrossRef]
- National Academies of Sciences, Engineering, and Medicine. *Reproducibility and Replicability in Science*; National Academies Press: Washington, DC, USA, 2019. [CrossRef] [PubMed]
- Baker, M. 1500 Scientists Lift the Lid on Reproducibility. *Nature* **2016**, *533*, 452–454. [CrossRef] [PubMed]
- Ioannidis, J.P.A. How to Make More Published Research True. *PLoS Med.* **2014**, *11*, e1001747. [CrossRef] [PubMed]
- Masharipov, S. *A Beginner's Guide to Uncertainty of Measurement: Measurement Good Practice Guide No. 11, Issue 2*; National Physical Laboratory: Teddington, UK, 1999. Available online: <https://www.npl.co.uk/resources/gpgs/beginners-guide-measurement-uncertainty-gpg11> (accessed on 15 June 2026).
- Plant, A.; Hanisch, R. Reproducibility in Science: A Metrology Perspective. *Harv. Data Sci. Rev.* **2020**, *2*, 1–26. [CrossRef]
- Han, R.; Walton, K.S.; Sholl, D.S. Does Chemical Engineering Research Have a Reproducibility Problem? *Annu. Rev. Chem. Biomol. Eng.* **2019**, *10*, 43–57. [CrossRef] [PubMed]
- Ferreira, B.D.; Olivares, I.R.B.; Carrilho, E.; Paccas, V.H.P. Is Everything Wrong in Analytical Chemistry? A Study on Reproducibility. *Accredit. Qual. Assur.* **2025**, *30*, 361–366. [CrossRef]
- Sharifi, S.; Reuel, N.; Kallmyer, N.; Sun, E.; Landry, M.P.; Mahmoudi, M. The Issue of Reliability and Repeatability of Analytical Measurement in Industrial and Academic Nanomedicine. *ACS Nano* **2023**, *17*, 4–11. [CrossRef] [PubMed]
- McNutt, M. Journals Unite for Reproducibility. *Science* **2014**, *346*, 679. [CrossRef] [PubMed]
- Oga, E.; Yolandia, J.N. Lab Notebook Commentary: Transforming the Laboratory Notebook into a Framework for Risk Management. *ACS Chem. Health Saf.* **2026**, 1–8. [CrossRef]
- Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef] [PubMed]
- Sweedler, J.V. Reproducibility and Replicability. *Anal. Chem.* **2019**, *91*, 7971–7972. [CrossRef] [PubMed]
- Ciriminna, R.; Angellotti, G.; Petri, G.L.; Pagliaro, M. Reproducibility in Chemistry Research. *Heliyon* **2024**, *10*, e33658. [CrossRef] [PubMed]
- Gundersen, O.E. The Fundamental Principles of Reproducibility. *Philos. Trans. R. Soc. A* **2021**, *379*, 20200210. [CrossRef] [PubMed]
- Roy, S.; Mandal, B.; Das, B.K. Quality Assurance in Laboratory Practices and Equipment Maintenance: Ensuring Precision, Reliability, and Compliance. In *Laboratory Techniques for Fish Disease Diagnosis*; Springer: Singapore, 2025; pp. 79–103. [CrossRef]
- Holloway, M.D. *Accuracy and Calibration Failures in Science and Engineering*; John Wiley & Sons: Hoboken, NJ, USA, 2025. [CrossRef]
- ISO/IEC 17025:2017; General Requirements for the Competence of Testing and Calibration Laboratories. ISO: Geneva, Switzerland, 2017. Available online: <https://www.iso.org/standard/66912.html> (accessed on 17 June 2026).

19. Organisation for Economic Co-operation and Development. *OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring No. 1: OECD Principles of Good Laboratory Practice*; OECD Publishing: Paris, France, 1998. [CrossRef]
20. Clinical and Laboratory Standards Institute. *Quality Management System: A Model for Laboratory Services*, 5th ed.; CLSI Guideline QMS01; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2019.
21. Open Science Collaboration. Estimating the Reproducibility of Psychological Science. *Science* **2015**, *349*, aac4716. [CrossRef] [PubMed]
22. Percie du Sert, N.; Hurst, V.; Ahluwalia, A.; Alam, S.; Avey, M.T.; Baker, M.; Browne, W.J.; Clark, A.; Cuthill, I.C.; Dirnagl, U.; et al. The ARRIVE Guidelines 2.0: Updated Guidelines for Reporting Animal Research. *PLoS Biol.* **2020**, *18*, e3000410. [CrossRef] [PubMed]
23. Bustin, S.A.; Benes, V.; Garson, J.A.; Hellemans, J.; Huggett, J.; Kubista, M.; Mueller, R.; Nolan, T.; Pfaffl, M.W.; Shipley, G.L.; et al. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* **2009**, *55*, 611–622. [CrossRef] [PubMed]
24. International Organization for Standardization. *ISO 5725: Accuracy (Trueness and Precision) of Measurement Methods and Results*; ISO: Geneva, Switzerland, 1994. Available online: <https://www.iso.org/standard/69418.html> (accessed on 17 June 2026).
25. Magnusson, B.; Örnemark, U. (Eds.) *Eurachem Guide: The Fitness for Purpose of Analytical Methods—A Laboratory Guide to Method Validation and Related Topics*, 2nd ed.; Eurachem: Teddington, UK, 2014.
26. Ramsey, M. Sampling as a Source of Measurement Uncertainty: Techniques for Quantification and Comparison with Analytical Sources. *J. Anal. At. Spectrom.* **1998**, *13*, 97–104. [CrossRef]
27. Barker, M.; Chue Hong, N.P.; Katz, D.S.; Lamprecht, A.-L.; Martinez-Ortiz, C.; Psomopoulos, F.; Harrow, J.; Castro, L.J.; Gruenpeter, M.; Martinez, P.A.; et al. Introducing the FAIR Principles for Research Software. *Sci. Data* **2022**, *9*, 622. [CrossRef] [PubMed]
28. Van Breemen, R.; Betz, J.; Evans, L.; Hall, A.; Jennens, M.; Kreider, R.; Kuszak, A.; Mudge, E.; Patel, P.; Rimmer, C.; et al. AOAC SMPR 2016.017: Standard Method Performance Requirements (SMPRs) for Quantitative Measurement of Vitamin B12 in Dietary Supplements and Ingredients. *J. AOAC Int.* **2017**, *100*, 294–296. [CrossRef] [PubMed]
29. Joint Committee for Guides in Metrology. *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM)*, 3rd ed.; Joint Committee for Guides in Metrology: Paris, France, 2012. [CrossRef]
30. Heckert, N.A.; Filliben, J.J. *NIST/SEMATECH e-Handbook of Statistical Methods*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2003. Available online: <https://www.itl.nist.gov/div898/handbook/> (accessed on 10 June 2026).
31. Berte, L.M. *Quality Management System: A Model for Laboratory Services*; Approved Guideline; Clinical and Laboratory Standards Institute: Wayne, PA, USA, 2011.
32. McLaughlin, E. *Data Integrity and Compliance With Drug CGMP: Questions and Answers—Guidance for Industry*; U.S. Food and Drug Administration: Silver Spring, MD, USA, 2018. Available online: <https://www.fda.gov/media/119267/download> (accessed on 19 June 2026).
33. Schiavone, C.; Portesi, C. PFAS: A Review of the State of the Art, from Legislation to Analytical Approaches and Toxicological Aspects for Assessing Contamination in Food and Environment and Related Risks. *Appl. Sci.* **2023**, *13*, 6696. [CrossRef]
34. Shinde, S.S.; Khulbe, P. Analytical Method Validation: A Comprehensive Review of Current Practices and Applications. *South East. Eur. J. Public Health* **2025**, *XXI* 2023, 244–254. [CrossRef]
35. Mitra, S. *Sample Preparation Techniques in Analytical Chemistry*; John Wiley & Sons: Hoboken, NJ, USA, 2004. [CrossRef]
36. Ullagaddi, P. Safeguarding Data Integrity in Pharmaceutical Manufacturing. *J. Adv. Med. Pharm. Sci.* **2024**, *26*, 64–75.
37. Taylor, B.N.; Kuyatt, C.E. *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*; NIST Technical Note 1297; National Institute of Standards and Technology: Gaithersburg, MD, USA, 1994. Available online: <https://www.nist.gov/pml/nist-technical-note-1297> (accessed on 16 June 2026).
38. Nayebi, B.; Khurana, P.; Pulicharla, R.; Karimpour, S.; Brar, S.K. Preservation, Storage, and Sample Preparation Methods for Freshwater Microplastics: A Comprehensive Review. *Environ. Sci. Adv.* **2023**, *2*, 1060–1081. [CrossRef]
39. Gika, H.G.; Theodoridis, G.A.; Wilson, I.D. Liquid Chromatography and Ultra-Performance Liquid Chromatography–Mass Spectrometry Fingerprinting of Human Urine: Sample Stability under Different Handling and Storage Conditions for Metabonomics Studies. *J. Chromatogr. A* **2008**, *1189*, 314–322. [CrossRef] [PubMed]
40. Ravisankar, P.; Gowthami, S.; Rao, G.D. A Review on Analytical Method Development. *Indian J. Res. Pharm. Biotechnol.* **2014**, *2*, 1183–1195.
41. Luo, Y.; Pehrsson, M.; Langholm, L.; Karsdal, M.; Bay-Jensen, A.-C.; Sun, S. Lot-to-Lot Variance in Immunoassays—Causes, Consequences, and Solutions. *Diagnostics* **2023**, *13*, 1835. [CrossRef] [PubMed]
42. Workman, J., Jr.; Lavine, B.; Chrisman, R.; Koch, M. Process Analytical Chemistry. *Anal. Chem.* **2011**, *83*, 4557–4578. [CrossRef] [PubMed]
43. Meier, P.C.; Zünd, R.E. *Statistical Methods in Analytical Chemistry*, 2nd ed.; John Wiley & Sons: Chichester, UK, 2005. [CrossRef]
44. Chandrul Kaushal, K.; Srivastava, B. A Process of Method Development: A Chromatographic Approach. *J. Chem. Pharm. Res.* **2010**, *2*, 519–545.

45. Kundu, A.; Wahome, J.; Zhou, W. Impact of Lot-to-Lot Variability of Cation Exchange Chromatography Resin on Process Performance. *BioPharm Int.* **2008**, *21*, 48–56.
46. Bonfiglio, R.; King, R.C.; Olah, T.V.; Merkle, K. The Effects of Sample Preparation Methods on the Variability of the Electrospray Ionization Response for Model Drug Compounds. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1175–1185.
47. Hu, X.; Li, C.; Deng, J. The Origins, Identification, and Solutions for Ghost Peaks in Gas Chromatography. *Chromatographia* **2025**, *89*, 121–133. [\[CrossRef\]](#)
48. Krueve, A.; Rebane, R.; Kipper, K.; Oldekop, M.-L.; Evard, H.; Herodes, K.; Ravio, P.; Leito, I. Tutorial Review on Validation of Liquid Chromatography–Mass Spectrometry Methods: Part II. *Anal. Chim. Acta* **2015**, *870*, 8–28. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Knapp, G.; Schramel, P. Sources of Analyte Contamination and Loss during the Analytical Process. In *Comprehensive Analytical Chemistry*; Elsevier: Amsterdam, The Netherlands, 2003; Volume 41, pp. 23–45. [\[CrossRef\]](#)
50. Kumar, A.; Sharma, C. Recent Update of the Various Sources Originating Ghost Peaks in Gas Chromatography: A Review. *J. Chromatogr. A* **2022**, *1685*, 463625. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Hibbert, D.B. Systematic Errors in Analytical Measurement Results. *J. Chromatogr. A* **2007**, *1158*, 25–32. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Kane, J.S. Analytical Bias: The Neglected Component of Measurement Uncertainty. *Analyst* **1997**, *122*, 1283–1288. [\[CrossRef\]](#)
53. Jahn, M. Leachables and Extractables: From Regulatory Expectations to Laboratory Assessment. In *Challenges in Protein Product Development*; Springer: Cham, Switzerland, 2018; pp. 337–351. [\[CrossRef\]](#)
54. Casey, B.K.; Schimmelbusch, H.R.; Holmes, D.T.; Yi, R.; Middleberg, R.A. Specific Lab Techniques for LC–MS/MS Analysis. In *Practical Guide to Implementing Liquid Chromatography–Mass Spectrometry in the Clinical Laboratory*; Springer: Cham, Switzerland, 2025; pp. 101–124. [\[CrossRef\]](#)
55. Bello, W.; Pezzatti, J.; Rudaz, S.; Sadeghipour, F. Study of Leachable Compounds in Hospital Pharmacy-Compounded Prefilled Syringes, Infusion Bags and Vials. *J. Pharm. Sci.* **2024**, *113*, 3227–3237. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Botnick, E.L.; Suga, R.S.; White, K.T. *Laboratory Quality Assurance Manual*; AIHA Press: Fairfax, VA, USA, 2005.
57. World Health Organization. *Laboratory Quality Management System: Handbook*; World Health Organization: Geneva, Switzerland, 2011. Available online: <https://stacks.cdc.gov/view/cdc/12202> (accessed on 19 June 2026).
58. Greener, J.; Pearson, G.; Cakmak, M. Roll-to-Roll Manufacturing: Process Elements and Recent Advances. In *Roll-to-Roll Manufacturing*; Greener, J., Pearson, G., Cakmak, M., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2018; Chapter 1. [\[CrossRef\]](#)
59. Broadhurst, D.; Goodacre, R.; Reinke, S.N.; Kuligowski, J.; Wilson, I.D.; Lewis, M.R.; Dunn, W.B. Guidelines and Considerations for the Use of System Suitability and Quality Control Samples in Mass Spectrometry Assays Applied in Untargeted Clinical Metabolomic Studies. *Metabolomics* **2018**, *14*, 72. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Bhattacharjee, S. DLS and Zeta Potential—What They Are and What They Are Not? *J. Control. Release* **2016**, *235*, 337–351. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Berne, B.J.; Pecora, R. *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics*; Courier Corporation: Mineola, NY, USA, 2000.
62. Harris, D.C. *Quantitative Chemical Analysis*, 9th ed.; W. H. Freeman and Company: New York, NY, USA, 2019.
63. Skoog, D.A.; Holler, F.J.; Crouch, S.R. *Principles of Instrumental Analysis*, 6th ed.; Saunders College Publishing: Philadelphia, PA, USA, 1998.
64. Mughe, H.C.; Kola, H.A.; Ersan, M.S.; Darland, D.C.; Zhao, J.X. Nanozymes: Promising Frontiers in Detection and Nanoremediation of Environmental Pollutants. *Energy Environ. Sustain.* **2026**, *2*, 100070. [\[CrossRef\]](#)
65. Mughe, H.C.; Okosun, B.O.; Ozdogan, M.; Kola, H.A.; Oncel, N.; Darland, D.C.; Zhao, J.X. Iron and Nitrogen Co-Doped Graphene Quantum Dots-Based Nanozyme for the Detection of NADPH. *ACS Appl. Nano Mater.* **2026**, *9*, 14553–14565. [\[CrossRef\]](#)
66. White, G. Basics of Estimating Measurement Uncertainty. *Clin. Biochem. Rev.* **2008**, *29*, S53–S60. [\[PubMed Central\]](#)
67. Stodden, V.; Leisch, F.; Peng, R.D. (Eds.) *Implementing Reproducible Research*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2014; p. 448. [\[CrossRef\]](#)
68. Sandve, G.K.; Nekrutenko, A.; Taylor, J.; Hovig, E. Ten Simple Rules for Reproducible Computational Research. *PLoS Comput. Biol.* **2013**, *9*, e1003285. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Nam, Y.; Park, H.-D. Revolutionizing Laboratory Practices: Pioneering Trends in Total Laboratory Automation. *Ann. Lab. Med.* **2025**, *45*, 472–483. [\[CrossRef\]](#) [\[PubMed\]](#)
70. Bai, Y.; Mosbach, S.; Taylor, C.J.; Karan, D.; Lee, K.F.; Rihm, S.D.; Akroyd, J.; Lapkin, A.A.; Kraft, M. A Dynamic Knowledge Graph Approach to Distributed Self-Driving Laboratories. *Nat. Commun.* **2024**, *15*, 462. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Shen, X.; Yan, H.; Wang, C.; Gao, P.; Johnson, C.H.; Snyder, M.P. TidyMass: An Object-Oriented Reproducible Analysis Framework for LC–MS Data. *Nat. Commun.* **2022**, *13*, 4365. [\[CrossRef\]](#) [\[PubMed\]](#)
72. Panko, R.R. What We Know About Spreadsheet Errors. *J. Organ. End User Comput.* **1998**, *10*, 15–21. [\[CrossRef\]](#)
73. Alewijn, M.; Rasker, S.; van Doorn, D.; Blokland, M. Reducing Manual Targeted LC–MS/MS Peak Integration Using a Supervised Learning Peak Evaluation and Automated Review Tool. *Microchem. J.* **2024**, *207*, 111578. [\[CrossRef\]](#)

74. Weitzel, M.J. The Estimation and Use of Measurement Uncertainty for a Drug Substance Test Procedure Validated According to USP <1225>. *Accredit. Qual. Assur.* **2012**, *17*, 139–146. [[CrossRef](#)]
75. Groh, K.J.; Mückel, C. In Vitro Toxicity Testing of Food Contact Materials: State-of-the-Art and Future Challenges. *Compr. Rev. Food Sci. Food Saf.* **2017**, *16*, 1123–1150. [[CrossRef](#)] [[PubMed](#)]
76. Karpovich, V.; Haiduk, N.; Oga, E.; Bala, N.; Kubátová, A.; Kozliak, E.; Sulkes, M. PAH Induction upon Pyrolysis of Hydroxyl-Terminated Polybutadiene-Based Solid Rocket Fuels. *J. Phys. Chem. A* **2025**, *129*, 6356–6373. [[CrossRef](#)] [[PubMed](#)]
77. Oga, E.; Bala, N.; Fisher, S.; Kozliak, E.; Kubátová, A. Investigation of Window Silicone Sealant Weathering Using Evolved Gas Analysis and Pyrolysis Gas Chromatography with Mass Spectrometry. *Polymers* **2025**, *17*, 2884. [[CrossRef](#)] [[PubMed](#)]
78. Powell, S.G.; Baker, K.R.; Lawson, B. Errors in Operational Spreadsheets. In *End-User Computing: Concepts, Methodologies, Tools, and Applications*; IGI Global: Hershey, PA, USA, 2011; pp. 236–247. [[CrossRef](#)]
79. Labuza, T.P.; Fu, B. Growth Kinetics for Shelf-Life Prediction: Theory and Practice. *J. Ind. Microbiol.* **1993**, *12*, 309–323. [[CrossRef](#)]
80. Richardson, H. *Medical Laboratories—Requirements for Quality and Competence: An ISO Perspective*; Vox Sanguinis: Basel, Switzerland, 2002; Volume 83, pp. 333–335. [[CrossRef](#)] [[PubMed](#)]
81. Leveson, N.G. *Engineering a Safer World: Systems Thinking Applied to Safety*; MIT Press: Cambridge, MA, USA, 2016. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.