RESEARCH PAPER

AI-CAPA: A Diagnostic Framework for Investigating Algorithmic Failures in GMP Pharmaceutical Manufacturing

Hemanth Kumar D[1*], Pramod N[2], Sneha H M[3]
[1*]Griffith College Dublin, Ireland.
Email: hemanthd9353@gmail.com
[2]Department of Pharmaceutical Chemistry, Harsha College of Pharmacy, Kambayyanapalya, Jakkanahalli post, Thymagondlu road, Nelamangala taluk, Bangalore - 562123, Karnataka, India
Email: pramodnayanapalli@gmail.com
[3]Department of Pharmacology, R R College of Pharmacy, Chikkabanavara, R R Layout, Medarahalli road, Bangalore - 560090, Karnataka, India

Received: 15th Jul, 2026 | Revised: 25th Jul, 2026 | Accepted: 5th Aug, 2026 | Available Online: 12th Aug, 2026

ABSTRACT

Background: The integration of artificial intelligence (AI) into pharmaceutical Good Manufacturing Practice (GMP) environments is increasing, with applications in inspection, process optimisation, and quality monitoring. Existing Corrective and Preventive Action (CAPA) frameworks, as defined within ICH Q9 and Q10, were developed for human and equipment-related failures and do not explicitly address AI-specific failure characteristics.

Objective: This paper presents AI-CAPA, a conceptual diagnostic framework designed to support post-failure investigation of AI-related quality events within GMP environments. The framework introduces Algorithmic Failure Mode Analysis (AFMA), a six-domain classification system, and a Domain Severity Index (DSI) to support structured investigation and prioritisation.

Methods: The framework was developed through structured literature synthesis of published AI failure cases, regulatory guidance, and established GMP quality management principles. The AFMA taxonomy was constructed through iterative conceptual analysis. No systematic review methodology was applied. Empirical validation is identified as a future requirement.

Results: AI-CAPA defines a five-phase investigative structure aligned with existing CAPA processes, incorporating AFMA as a diagnostic layer within Phase 2. A six-domain taxonomy and DSI-based prioritisation method are presented with proposed calibration guidance. A worked case study illustrates application in a pharmaceutical manufacturing context. An Algorithmic Failure Report (AFR) is proposed as a supplementary CAPA record.

Conclusion: AI-CAPA provides a structured, GMP-compatible approach to investigating AI-related quality events without modifying existing CAPA procedural architecture. The framework is intended as a conceptual extension requiring prospective validation and regulatory engagement.

Keywords: artificial intelligence; CAPA; GMP; ICH Q9; quality risk management; pharmaceutical manufacturing; algorithmic failure; diagnostic framework

How to cite this article: Hemanth Kumar D1* Pramod N2 Sneha H M3. AI-CAPA: A Diagnostic Framework for Investigating Algorithmic Failures in GMP Pharmaceutical Manufacturing. Int J Drug Deliv Technol. 2026;16(76s):2128-2137. DOI: 10.25258/ijddt.16.76s.208

.1. Introduction

The application of artificial intelligence (AI) in pharmaceutical manufacturing is expanding across multiple GMP-regulated functions, including visual inspection, process optimisation, environmental monitoring, and predictive quality control[1,2]. Regulatory authorities have acknowledged this trend. Recent guidance from the U.S. Food and Drug Administration highlights the need for robust lifecycle management of AI systems in drug manufacturing [1], while the European Medicines Agency has identified AI-related quality governance as an emerging area requiring further development [2].

Current GMP quality systems rely on Corrective and Preventive Action (CAPA) frameworks embedded within ICH Q9 (Quality Risk Management) and ICH Q10 (Pharmaceutical Quality System) [3,4]. These frameworks are well established for investigating deviations arising from human error or equipment malfunction. However, they do not explicitly address characteristics specific to AI systems, such as model drift, dependence on training data, or performance degradation without a discrete triggering event [5,6].

As a result, when AI systems contribute to quality events — such as misclassification in inspection systems or inaccurate process predictions — existing CAPA methodologies may lack the structure required to investigate root causes systematically. This reflects a gap between evolving manufacturing technologies and established quality governance tools.

This paper presents AI-CAPA, a conceptual framework designed to address this gap. The

framework introduces Algorithmic Failure Mode Analysis (AFMA), a structured diagnostic classification applied during CAPA investigation, and a Domain Severity Index (DSI) to support prioritisation of investigative effort and corrective actions. AI-CAPA is not proposed as a replacement for CAPA or ICH guidance, but as a targeted extension within the investigation phase of existing CAPA processes.

The framework is explicitly structured around three distinct layers: (1) prospective risk management (ICH Q9), unchanged; (2) post-failure diagnostic classification (AFMA), introduced in this work; and (3) CAPA procedural execution (ICH Q10), unchanged. This separation is intended to maintain alignment with established regulatory expectations while introducing AI-specific investigative capability. A worked example is provided to demonstrate application, followed by discussion of limitations and areas for future research.

2. Background and Rationale

2.1 AI Adoption and Quality Governance Gaps in GMP Manufacturing

Pharmaceutical manufacturers are deploying AI across core GMP processes, including computer vision for visual inspection, machine learning for batch failure prediction, and real-time monitoring of critical quality attributes [7,8]. The number of regulatory submissions incorporating AI components in manufacturing and development contexts has grown in recent years [1,2].

Regulatory guidance in this area is developing, but primarily at the level of documentation and lifecycle management requirements. The FDA's 2025 draft guidance addresses model credibility assessment, lifecycle maintenance planning, and documentation expectations [1]. It does not define post-failure investigative structures for GMP quality events. That absence is the operational gap AI-CAPA is designed to address.

2.2 Limitations of Traditional CAPA for AI-Related Quality Events

ICH Q9 defines quality risk management as a systematic process for assessment, control, communication, and review of risks to product quality [3]. The risk assessment tools endorsed by ICH Q9 — Failure Mode and Effects Analysis (FMEA), Hazard Analysis and Critical Control Points, and fault tree analysis — are prospective instruments applied before failure to identify and mitigate potential risk. They share the assumption that failures have identifiable causes traceable through a discrete triggering event.

AI systems may present a different investigative challenge. A machine learning model may degrade through distributional shift — where the statistical properties of production data diverge from training data over time — without producing a discrete triggering event [5,6]. The root cause of such a failure may reside in training data composition, in a validation study that did not adequately represent production conditions, or in a change control process that did not identify a supplier change as an AI system revalidation trigger. These causal structures are not addressed by standard CAPA root cause tools such as Fishbone analysis or 5-Why [9].

Published analysis of AI quality events in healthcare and pharmaceutical settings suggests that existing quality event management systems may lack specific investigative categories for this class of failure [10,11]. The AFMA taxonomy proposed in this framework is designed to address this by providing structured diagnostic domains applicable to AI system failures in GMP contexts.

2.3 Framework Positioning Relative to ICH Q9 and Q10

AI-CAPA is not a revision of ICH Q9 or Q10, nor does it claim regulatory equivalence with those documents. Its relationship to established GMP doctrine is one of targeted augmentation: it extends the investigative capability within CAPA's Phase 2 root cause analysis step for AI-related quality events, while leaving the ICH Q9 prospective risk framework and ICH Q10 CAPA procedural architecture unchanged.

This positioning reflects an important functional distinction. ICH Q9 FMEA is a prospective risk assessment tool applied before failure to prioritise controls. AFMA is a retrospective diagnostic classification tool applied after failure to structure investigation. These are different activities at different stages of the quality management lifecycle and are treated as such throughout this framework. Table 1 summarises the three-layer separation.

Layer	Tool	Stage	ICH basis
1 — Risk assessment	ICH Q9 FMEA / QRM	Pre-failure (prospective)	ICH Q9
2 — Failure investigation	AFMA diagnostic classification	Post-failure (retrospective)	Proposed extension of ICH Q10 CAPA Phase 2
3 — Action management	ICH Q10 CAPA	Post-failure (procedural)	ICH Q10

Table 1. Three-layer separation of quality management activities in AI-CAPA. AFMA operates in Layer 2 only. It does not replace or replicate ICH Q9 prospective risk assessment.

3. The AI-CAPA Framework

AI-CAPA is a five-phase investigative framework that augments the investigation step within existing CAPA architecture. The five phases are: (1) Anomaly Detection and Triage; (2) Algorithmic Failure Mode Analysis; (3) Root Cause Determination; (4) Corrective and Preventive Action Planning; and (5) Effectiveness Verification and Reporting. Phases 1, 3, 4, and 5 map directly to standard CAPA steps. Phase 2 introduces AFMA as a new structured diagnostic activity between triage and root cause determination.

[Figure]

Figure 1. AI-CAPA five-phase investigative framework showing the integration of Algorithmic Failure Mode Analysis (AFMA) within the CAPA process.

3.1 Phase 1: Anomaly Detection and Triage

Phase 1 encompasses recognition, classification, and initial triage of an AI system quality event. In contrast to conventional CAPA initiation — which typically follows a discrete observed defect — AI anomalies may emerge from statistical process monitoring, retrospective batch record audit, or customer complaint analysis rather than direct operator observation [5]. This difference affects the time window between failure occurrence and detection, which is itself a data point captured in Phase 1.

Phase 1 outputs: AI system identifier, version, and GMP use case; severity classification (critical / major / minor / observation); scope definition covering affected batches and estimated period of AI system failure; and multi-disciplinary notification including the quality unit, data science team, and IT infrastructure personnel. The inclusion of data science and IT personnel from the outset reflects the technical complexity of AI failure investigation and is a structural difference from standard CAPA initiation.

3.2 Phase 2: Algorithmic Failure Mode Analysis (AFMA)

AFMA is a six-domain diagnostic classification tool applied post-failure to structure the investigation of an AI quality event. It is not a prospective risk assessment tool and is not equivalent to ICH Q9 FMEA. Its function is to provide investigators with a structured set of diagnostic domains — covering the principal system layers in which AI failures in GMP environments have been documented to originate — and to generate a prioritised investigation output that guides action planning in Phase 4.

The six AFMA domains are presented in Table 2. The domain structure is acknowledged as a systems decomposition of a GMP AI deployment rather than a pure failure mode taxonomy. This reflects a deliberate operational choice: AI failures in GMP contexts do not respect clean architectural boundaries. A training data gap (D1) frequently propagates through a validation process that did not detect it (D3) and surfaces as a regulatory documentation failure (D6) in the same incident. An investigative tool that separates these into isolated analytical layers would require investigators to coordinate multiple parallel analyses. The AFMA is designed for practical use by quality assurance teams, not as a theoretical taxonomy.

Domain	Description	Example indicators in GMP context
D1: Training data	Adequacy, representativeness, and provenance of data used to train the AI system	Class imbalance; excluded product or process sub-populations; undocumented data sources; outdated training sets no longer representative of current production
D2: Model drift	Performance degradation over time due to distributional shift between training and production data	Declining model confidence scores; increasing false-negative rate; covariate or concept drift detected in monitoring data
D3: Validation gap	Mismatch between conditions under which the AI system was validated and current production environment	Validation study did not represent current supplier material; change control did not trigger revalidation; inadequate stress testing coverage

Domain	Description	Example indicators in GMP context
D4: Human-AI interface	Failure arising from the interaction between operators and AI system outputs, including over-reliance and governance of override decisions	Automation bias; alert fatigue causing missed anomalies; undocumented or inadequately governed override decisions; insufficient operator training
D5: Integration	Failure at the interface between the AI system and GMP operational infrastructure	Data feed disruption between ERP and AI system; API failures; IT infrastructure change implemented without revalidation; data pipeline integrity failure
D6: Regulatory alignment	Failure to meet applicable regulatory expectations for AI system documentation, change control, and validation in GMP	Undocumented model updates; missing validation records; change control records absent AI impact assessment; non-compliant model lifecycle documentation

Table 2. AFMA six-domain diagnostic classification taxonomy. Domains reflect a systems decomposition of a GMP AI deployment. They are not mutually exclusive; multiple domains may be implicated in a single quality event.

The domain structure does not address all possible AI failure categories in GMP environments. Recognised gaps include cybersecurity and adversarial manipulation risks, third-party and vendor model risk (particularly relevant in cloud-hosted or SaaS AI deployments), and federated learning failure modes. These are identified as areas for taxonomy extension in future work.

3.2.1 Domain Severity Index (DSI) for Investigation Prioritisation

To support prioritisation of investigative depth and corrective action sequencing across the six AFMA domains, AI-CAPA introduces a Domain Severity Index (DSI). The DSI is a structured scoring tool applied to each implicated domain during Phase 2. It is explicitly positioned as an investigation prioritisation overlay, not as a prospective risk assessment or a root cause indicator. Its function is to ensure that investigative effort and corrective action resources are allocated proportionately across identified domain findings.

The DSI adapts the multiplicative structure of FMEA Risk Priority Numbers, familiar to GMP quality professionals through ICH Q9, while redefining the parameters for post-failure investigative use. The three parameters and their scoring scales are defined in Table 3.

Score	Rating	Impact (I) Effect on product quality / patient safety	Recurrence likelihood (R) Probability of recurrence without intervention	Detectability (De) Inverse: difficulty of early detection
1	Negligible	No patient risk; cosmetic or administrative deviation	Very unlikely; no documented precedent in this system	Highly detectable via automated monitoring or SPC
2	Low	Minor quality impact; no batch failure risk	Unlikely; isolated single occurrence	Detectable via routine batch review or monitoring

Score	Rating	Impact (I) Effect on product quality / patient safety	Recurrence likelihood (R) Probability of recurrence without intervention	Detectability (De) Inverse: difficulty of early detection
3	Moderate	Batch failure potential; no direct patient harm	Possible; documented precedent in comparable AI systems	Detectable with targeted retrospective audit
4	High	Product recall risk; patient safety concern	Likely; systemic root cause identified	Difficult to detect without dedicated AI performance monitoring
5	Critical	Direct patient harm risk; GMP critical deviation	Very likely; recurring failure pattern in this or comparable system	Not detectable via current quality controls

Table 3. DSI parameter definitions and scoring scales. DSI = Impact (I) × Recurrence likelihood (R) × Detectability (De). Maximum DSI = 125. Note: these scales require calibration through cross-functional consensus before first use; see Section 3.2.2.

DSI calculation: **DSI = Impact (I) × Recurrence Likelihood (R) × Detectability (De)**. DSI scores range from 1 to 125. For investigation prioritisation: DSI ≥ 60 — primary cause requiring immediate corrective action and full root cause investigation; DSI 30–59 — contributing cause requiring planned corrective action within a defined timeframe; DSI < 30 — domain not implicated or flagged for preventive monitoring only. These thresholds are proposed as a starting point for organisational calibration and are not prescriptive.

3.2.2 DSI Calibration Protocol

The principal vulnerability of any multiplicative scoring system in a regulated environment is inter-rater variability: two quality teams scoring the same failure may assign different DSI values with no auditability trail. This is acknowledged as a structural limitation of the DSI and is addressed through a required calibration protocol.

Before first use of the DSI in any organisation, a calibration session is required involving a cross-functional team comprising at minimum: quality assurance, data science, IT infrastructure, and regulatory affairs representatives. The session establishes:

- Anchor examples at each score level for each parameter, specific to the organisation's AI systems and GMP context. For example, a Score 4 Impact in a sterile injectable environment may differ materially from Score 4 in an oral solid dosage operation.
- A documented rationale field accompanying each DSI score, recording the evidence basis for the assigned value. The score alone is not sufficient for audit purposes.
- A consensus rule for disputed scores: where cross-functional raters disagree by more than one point on any parameter, the higher score is applied until additional evidence resolves the dispute.

Calibration records are maintained as quality documents. Periodic review — recommended annually or following any significant new AI system deployment — ensures that scoring anchors remain fit for purpose as the organisation's AI portfolio evolves. This governance structure mirrors the cross-functional consensus mechanism used for FMEA scoring in ICH Q9 applications, adapted for post-failure investigative use.

3.3 Phase 3: Root Cause Determination

Root cause determination incorporates the AFMA domain findings and DSI scores to structure and prioritise investigation depth. Domains scoring DSI ≥ 60 are subject to full root cause analysis with documented evidence mapping. Domains scoring 30–59 receive targeted investigation proportionate to their contributing cause classification. Domains scoring below 30 are documented with a brief rationale for reduced investigative depth. This risk-proportionate approach to investigation resourcing is consistent with the principles of ICH Q9, applied retrospectively to post-failure investigation rather than prospectively to risk control.

Phase 3 outputs: a root cause statement referencing primary and contributing AFMA domains with

supporting DSI scores; an impact assessment covering affected batches, patient risk classification, and regulatory disclosure assessment; and a systemic versus isolated failure classification. Systemic classification requires mandatory preventive action planning in Phase 4.

3.4 Phase 4: Corrective and Preventive Action Planning

Phase 4 maps corrective and preventive actions to the AFMA domains identified as primary or contributing causes, with sequencing informed by DSI scores. Actions for high-DSI domains are initiated immediately; actions for contributing-cause domains are planned within defined timeframes. Table 4 presents the domain-mapped action framework.

Domain	Corrective actions	Preventive actions
D1: Training data	Data provenance audit; retraining with representative and documented dataset	Training data governance SOP; mandatory coverage and diversity requirements for new deployments
D2: Model drift	Model rollback or retraining; batch re-inspection for affected period	Statistical drift monitoring implementation; scheduled model performance review schedule
D3: Validation gap	Revalidation study under current production conditions	Change control SOP revised to include mandatory AI system revalidation trigger assessment
D4: Human-AI interface	Operator retraining; alert threshold and override protocol review	Human factors assessment programme; override audit and periodic training plan
D5: Integration	Integration audit; IT change control review for AI-affecting changes	AI system IT change control SOP; integration testing requirements in validation master plan
D6: Regulatory alignment	Documentation remediation; regulatory disclosure impact assessment	AI regulatory register; periodic compliance audit schedule aligned with product lifecycle

Table 4. Domain-mapped corrective and preventive action framework. Actions are sequenced by DSI score from Phase 2. Change control SOP revision (D3 preventive) is identified as the highest-priority systemic preventive action in most AI failure investigations.

3.5 Phase 5: Effectiveness Verification and Reporting

Phase 5 defines acceptance criteria against which Phase 4 actions will be assessed and specifies conditions for CAPA closure. Effectiveness verification in AI-CAPA differs from conventional CAPA closure in one material respect: completion of corrective actions is necessary but not sufficient. Closure requires evidence of sustained model performance recovery meeting pre-defined quantitative acceptance criteria over a defined monitoring period. The length and structure of the monitoring period are determined by the severity classification from Phase 1 and the DSI scores from Phase 2.

Phase 5 also generates the Algorithmic Failure Report (AFR). The AFR is proposed as a supplementary CAPA record — an extension of the existing deviation and CAPA documentation structure, not a new GMP document class. Its additional fields, beyond a standard CAPA record, are: AI system identifier and version at time of failure; AFMA domain findings with DSI scores; model performance data at time of failure, at corrective action completion, and at closure; training data provenance summary; and revalidation study reference. These fields can be implemented as supplementary sections within an existing CAPA record template in current quality management system software, without requiring new document type classification or additional software validation scope.

Feature	Traditional CAPA (ICH Q10)	AI-CAPA (proposed augmentation)
Framework stage	Post-failure investigation and action	Post-failure investigation and action (unchanged)

Feature	Traditional CAPA (ICH Q10)	AI-CAPA (proposed augmentation)
Root cause tools	5-Why; Fishbone; FMEA (retrospective use)	AFMA diagnostic classification + 5-Why / Fishbone
AI failure investigation	Not addressed	AFMA six-domain diagnostic taxonomy
Action prioritisation	Severity classification	DSI score from AFMA Phase 2
Change control link	Equipment / process change triggers	AI revalidation trigger embedded in change control SOP
Output document	CAPA record	AFR — supplementary CAPA record with AI-specific fields
ICH Q9 prospective QRM	Unchanged	Unchanged — AFMA is not a prospective risk tool

Table 5. Comparison of traditional CAPA and AI-CAPA. AI-CAPA augments Phase 2 investigation methodology only. ICH Q9 prospective quality risk management and ICH Q10 CAPA procedural architecture are preserved unchanged.

4. Worked Example: AI Visual Inspection Failure in Vial Production

This section illustrates AI-CAPA through a composite case scenario derived from documented failure patterns in the literature [10,11,12], constructed to demonstrate framework application across realistic GMP conditions.

4.1 Scenario

A sterile injectable manufacturer operates an AI-based visual inspection system (VisionCheck v2.3) for vial seal integrity assessment, deployed under a validated GMP configuration. Following a primary packaging supplier change — processed through change control as a supplier qualification event — the system begins releasing vials with seal defects falling outside the acceptable quality threshold for the new supplier's glass specification. Three batches are distributed before the deviation is identified through customer complaints. The change control record for the supplier qualification did not include an AI system impact assessment field.

4.2 AI-CAPA Application

Phase 1: Deviation classified critical. Scope: three released batches over six weeks. Multi-disciplinary notification completed: quality, data science, regulatory affairs. GMP documents reviewed: batch manufacturing records, supplier qualification change control record, CoA records for new supplier packaging.

Phase 2 — AFMA and DSI scoring: Six domains assessed. D3 (Validation gap) — Impact 4, Recurrence 5, Detectability 4, DSI 80. Primary cause: the change control process did not classify the supplier change as an AI system revalidation trigger. D1 (Training data) — Impact 4, Recurrence 4, Detectability 3, DSI 48. Contributing cause: training images did not include defect samples from the new supplier glass specification. D6 (Regulatory alignment) — Impact 3, Recurrence 4, Detectability 3, DSI 36. Contributing cause: change control SOP lacked an AI impact assessment field. D2, D4, D5 — all scored DSI ≤ 18; documented as not implicated with rationale.

Phase 3: Primary root cause: absence of a change-triggered AI revalidation protocol (D3, DSI 80). Contributing: insufficient training data coverage for new supplier specification (D1, DSI 48) and absent AI impact assessment in change control procedure (D6, DSI 36). Systemic classification assigned.

Phase 4: Immediate (D3 corrective): quarantine and re-inspection of three affected batches; suspension of VisionCheck v2.3 pending revalidation; revalidation study initiated. Planned 30 days (D1 corrective): model retrained with supplier-specific defect images. Planned 30 days (D6 corrective): change control SOP revised to include mandatory AI impact assessment field. Preventive: automated drift monitoring implemented; supplier change events linked to AI validation status flag in change control system (D3 systemic preventive — highest priority).

Phase 5: Effectiveness criteria: VisionCheck achieving ≥ 99.5% detection rate across 500 consecutive production vials under new supplier conditions over a 30-day monitoring period. AFR generated as supplementary CAPA record. Regulatory impact assessed; field alert report filed in accordance with applicable requirements.

4.3 DSI Scores Summary

Domain	I	R	De	DSI	Classification	Action priority
D1: Training data	4	4	3	48	Contributing cause	Planned corrective — 30 days
D2: Model drift	3	2	3	18	Not implicated	Preventive monitoring only
D3: Validation gap	4	5	4	80	Primary cause	Immediate corrective action
D4: Human-AI interface	2	2	3	12	Not implicated	No action required
D5: Integration	2	1	2	4	Not implicated	No action required
D6: Regulatory alignment	3	4	3	36	Contributing cause	Planned corrective — 30 days

Table 6. AFMA DSI scores for the vial inspection case study. D3 Validation gap (DSI 80) identified as primary cause; immediate corrective action required. DSI scores are illustrative; actual scores require cross-functional calibration per Section 3.2.2.

4.4 Comparative Outcome

Outcome dimension	Traditional CAPA	AI-CAPA
Root cause identified	Operator oversight or misclassification	Validation gap + training data scope (D3 DSI 80; D1 DSI 48)
Primary preventive action	Operator retraining	Change control SOP revision — AI revalidation trigger embedded systemically
Recurrence risk	High — systemic cause not addressed	Reduced — systemic change control gap closed
CAPA documentation	Standard CAPA record	AFR — supplementary CAPA record with DSI scores, model performance data, revalidation evidence

Table 7. Comparative outcomes of traditional CAPA versus AI-CAPA for the vial inspection case study.

5. Discussion

AI-CAPA addresses a specific operational gap: the absence of a structured investigative methodology for AI system failures within existing GMP CAPA systems. The AFMA six-domain taxonomy and DSI prioritisation method together provide quality assurance teams with a structured, documentable approach to AI failure investigation that is compatible with existing CAPA procedural architecture without modifying it.

The framework's most practically significant element is the linkage between change control and AI system revalidation. In the worked case study, the primary root cause — a supplier change that was not classified as an AI revalidation trigger — is a gap that is both realistic and systemic. It reflects the current state of change control procedures in GMP environments that have adopted AI systems without updating their change management infrastructure to account for AI-specific validation dependencies. The corrective action — revising the change control SOP to include a mandatory AI impact assessment field — is implementable today, within existing quality system infrastructure, without requiring new regulatory guidance or system investment. This is the element of AI-CAPA most likely to have immediate practical value for adopting organisations.

The diagnostic framing of AFMA — as distinct from prospective FMEA — is intentional and important.

ICH Q9 FMEA is applied before failure to prioritise risk controls. AFMA is applied after failure to structure investigation. These are different activities at different points in the quality management lifecycle, and conflating them would create ambiguity in regulatory audit situations. The decision to rename the scoring tool as a Domain Severity Index, rather than Risk Priority Number, reflects this distinction: DSI values are investigation prioritisation outputs, not risk assessments.

The AFMA taxonomy is acknowledged as a systems decomposition rather than a pure failure mode taxonomy. A reviewer might reasonably argue that its six domains span multiple abstraction levels — data, model behaviour, process validation, human factors, infrastructure, and regulatory compliance — and that a theoretically pure taxonomy would separate these into distinct analytical layers. The counter-argument, embedded in the framework design, is that operational utility in GMP investigation contexts depends on a single structured tool that covers the space of AI failure modes, even if that space does not map cleanly onto a single analytical layer. The empirical question — whether the six domains adequately cover the failures encountered in practice — is identified as the priority for prospective validation research.

The DSI calibration requirement (Section 3.2.2) directly addresses the inter-rater reliability problem inherent in any multiplicative scoring system. Without calibration, two quality teams could assign materially different DSI values to the same failure with no audit trail. The calibration protocol — cross-functional anchor examples, documented score rationale, consensus rules for disputed scores — provides the governance structure that makes DSI scores auditable and reproducible in regulated environments. This is a necessary condition for GMP adoption, not an optional enhancement.

6. Limitations and Future Research

The principal limitations of this work are as follows. First, the AFMA taxonomy and DSI thresholds have not been empirically validated against a prospective dataset of pharmaceutical AI quality events. All scores in the worked example are illustrative and calibration-dependent. Second, this paper presents a conceptual framework developed through structured literature synthesis, not a systematic review; inclusion and exclusion criteria were not applied and the methodology is not reproducible in the formal sense. Third, the six AFMA domains do not address all AI failure categories relevant to GMP environments. Recognised gaps include: cybersecurity and adversarial manipulation risks; third-party and vendor model risk, particularly relevant in cloud-hosted or SaaS AI deployments; ETL data pipeline integrity failures as distinct from training data quality; and federated learning failure modes. Fourth, the DSI threshold values (60, 30) are proposed on the basis of FMEA precedent and expert reasoning; empirical calibration against real-world event data is required before these thresholds can be considered validated. Fifth, the AFR supplementary record, while conceptually consistent with existing CAPA documentation, has no current regulatory basis and would require industry consultation and potential regulatory guidance to standardise.

Future research priorities are: (1) prospective pilot application of AI-CAPA in a pharmaceutical GMP environment, with DSI scores compared against expert consensus ratings to establish inter-rater reliability data; (2) retrospective application of the AFMA taxonomy to historical AI failure cases in published literature and regulatory databases to assess domain coverage; (3) engagement with ICH working groups and regulatory agencies to explore possible incorporation of AI failure investigation guidance into future revisions of Q9 and Q10; (4) taxonomy extension to incorporate the identified domain gaps, particularly third-party model risk and cybersecurity; and (5) development of a computer system validation framework for digital AI-CAPA implementations, consistent with EU GMP Annex 11 and FDA 21 CFR Part 11.

7. Conclusion

AI-CAPA proposes a diagnostic extension of GMP CAPA practice for the structured investigation of AI system failures in pharmaceutical manufacturing. The framework is grounded in three explicitly separated layers: ICH Q9 prospective quality risk management, unchanged; AFMA diagnostic classification applied post-failure within CAPA Phase 2; and ICH Q10 CAPA procedural architecture, unchanged. This separation is the framework's core structural principle.

Of the framework's components, the change control to AI revalidation trigger linkage is the most immediately actionable for GMP organisations. It requires no new regulatory guidance, no new system infrastructure, and no significant resource investment — only a revision to existing change control procedure templates. Based on the worked case study and the published literature, it is also one of the most common systemic gaps enabling AI quality failures in practice.

The framework requires prospective validation before it can be considered empirically grounded. The conceptual contributions — the three-layer separation, the AFMA taxonomy, the DSI calibration protocol, and the AFR supplementary record — are offered as a basis for that validation work and for regulatory dialogue on the governance of AI in GMP pharmaceutical manufacturing.

8. Declarations

Author Contributions

H.K.D.: conceptualisation, framework development, taxonomy construction.
Pramod N: conceptualisation, framework development.
Sneha M: literature synthesis, manuscript preparation.

Competing Interests

The authors declare no competing interests.

Funding

No external funding was received for this work.

Data Availability

A proof-of-concept digital implementation of the AI-CAPA framework is available at: https://ai-capa-gmp.netlify.app. The implementation is provided for demonstration purposes only and is not a validated GMP computer system.

AI Use Declaration

AI-assisted writing tools were used during manuscript preparation for structural editing and writing refinement. All intellectual content, framework development, taxonomy construction, DSI design, and analytical judgements are the sole work of the authors. This declaration is made in accordance with emerging journal standards for AI-assisted research.

References

[1] U.S. Food and Drug Administration. Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products: Draft Guidance for Industry. FDA; January 2025. Docket No. FDA-2024-D-4689.

[2] European Medicines Agency. Reflection Paper on the Use of Artificial Intelligence (AI) in the Medicinal Product Lifecycle. EMA/CHMP/CVMP/QWP/166500/2023. EMA; 2023.

[3] International Council for Harmonisation. ICH Q9(R1): Quality Risk Management. ICH; 2023.

[4] International Council for Harmonisation. ICH Q10: Pharmaceutical Quality System. ICH; 2008 (reaffirmed 2023).

[5] Lu J, Liu A, Dong F, Gu F, Gama J, Zhang G. Learning under concept drift: a review. IEEE Trans Knowl Data Eng. 2019;31(12):2346-2363. doi:10.1109/TKDE.2018.2876857.

[6] Quionero-Candela J, Sugiyama M, Schwaighofer A, Lawrence N. Dataset Shift in Machine Learning. Cambridge: MIT Press; 2009.

[7] Rathore AS, Mishra S, Nikita S, Priyanka P. Bioprocess control: current progress and future perspectives. Life (Basel). 2021;11(6):557. doi:10.3390/life11060557.

[8] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg: NIST; 2023.

[9] Tague NR. The Quality Toolbox. 2nd ed. Milwaukee: ASQ Quality Press; 2005.

[10] Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc. 2012;19(1):121-127. doi:10.1136/amiajnl-2011-000089.

[11] Sendak MP, D'Arcy J, Kashyap S, Gao M, Nichols M, Corey K. A path for translation of machine learning products into healthcare delivery. EMJ Innov. 2020;4(1):1-8.

[12] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56. doi:10.1038/s41591-018-0300-7.

[13] Varshney KR, Alemzadeh H. On the safety of machine learning: cyber-physical systems, decision sciences, and data products. Big Data. 2017;5(3):246-255. doi:10.1089/big.2016.0051.

[14] World Health Organization. Ethics and Governance of Artificial Intelligence for Health. Geneva: WHO; 2021. Licence: CC BY-NC-SA 3.0 IGO.

[15] International Council for Harmonisation. ICH Q13: Continuous Manufacturing of Drug Substances and Drug Products. ICH; 2022.

[16] U.S. Food and Drug Administration. 21 CFR Part 211: Current Good Manufacturing Practice for Finished Pharmaceuticals. Washington: U.S. Government Publishing Office; 2023.