WHAT HAPPENS WHEN AI MAKES A GMP MISTAKE?

Navigating Regulatory Realities, Accountability, and Governance in Pharmaceutical Quality Assurance

Artificial Intelligence (AI) is rapidly transitioning from a theoretical concept to an operational powerhouse within the pharmaceutical industry. From clinical trials (Morino, n.d.) to automating pharmacovigilance (Sardella, n.d.), the promise of AI is undeniable. But a critical question looms over this digital revolution in Quality Assurance (QA): **What happens when the AI is wrong?**

When a Large Language Model (LLM) hallucinates in a corporate email, the result is an embarrassing retraction. However, when an AI hallucinates in a Current Good Manufacturing Practice (cGMP) environment—inventing a non-existent regulatory citation, incorrectly disposing of a batch, or misinterpreting a critical process parameter—the consequences can compromise patient safety and trigger severe regulatory enforcement.

The Anatomy of an AI Error in a GMP Environment

Generative AI models do not "think." They predict the most logical next word or data point based on vast datasets. This reliance on statistical probability over factual deterministic reasoning naturally leads to *AI hallucinations*—instances where the system confidently generates plausible, structurally perfect, but factually incorrect information (Ramcharran, n.d.).

In a pharmaceutical QA environment, these AI missteps can manifest in highly dangerous ways:

- **Fabricated Root Causes:** An AI analyzing a temperature excursion might confidently invent a mechanical failure narrative based on generalized historical data, completely missing the actual localized sensor calibration error.
- **False Citations:** An AI might reference an invented ICH guideline to justify a deviation closure, circumventing required human oversight.
- **Incorrect Batch Disposition Recommendations:** An AI tool summarizing batch analytical data might wrongly recommend the release of an out-of-specification (OOS) product by failing to properly flag anomalies.

Hypothetical Case Study: The "Ghost" CAPA

Imagine a mid-sized biologics manufacturer deploys an AI tool to triage deviations. A bioreactor experiences a minor pressure drop. The AI digests the deviation and proposes a CAPA: *Update the Standard Operating Procedure (SOP) to bypass the secondary pressure valve during sterilization, citing a "historical precedent" in a previous batch.*

The QA specialist, experiencing alert fatigue and over-relying on the system, approves the AI-drafted CAPA without verifying the precedent. The problem? The "historical batch" never existed, and the secondary valve is critical for maintaining sterility. The AI optimized for linguistic completion, not process validation. The result is an unvalidated change control that compromises the next three commercial batches, leading to a massive recall.

The Accountability Crisis: Who Owns the Mistake?

When an AI makes a GMP mistake, accountability cannot be outsourced to the algorithm. Regulatory agencies have established that the manufacturer, and specifically the Quality Unit (QU), retains ultimate responsibility. FDA warning letters have consistently highlighted failures of the QU to thoroughly investigate unexplained discrepancies, such as OOS laboratory results, without addressing potential manufacturing causes (McDowall, n.d.). Relying entirely on an unverified output from an automated or AI system is a direct violation of 21 CFR § 211.22(c).

The FDA-EMA guiding principles emphasize that an AI system flagging an anomalous response for human review is fundamentally different from an AI system automatically accepting or rejecting a bioanalytical run (Oxley-King, n.d.). The former acts as an assistant where the human retains decision-making responsibility, while the latter makes the model the ultimate decision-maker—a scenario that requires an exponentially higher standard of justification.

Compliance Ripple Effects

An unchecked AI mistake creates a cascading ripple effect across the Quality Management System (QMS):

- **Validation Limitations:** Traditional software validation is deterministic: Input A always equals Output B. Conversely, probabilistic models may behave differently over time, requiring robust human review because their outputs vary depending on retraining and data drift (Oxley-King, n.d.).

- **Supplier Oversight Challenges:** Treat AI vendors as critical suppliers. Failing to verify the underlying data architectures of AI providers mirrors the critical failure of trusting raw material Certificates of Analysis (CoAs) without conducting identity tests or vendor audits.

> *"While AI systems can extend QA processes by analyzing larger volumes of data and highlighting patterns, they do not replace applied logic. The final assessment remains with the responsible scientist." (Oxley-King, n.d.)*

Building a Resilient AI Governance Framework

To balance operational opportunities with the stringent realities of cGMP, QA professionals must build robust governance models aligned with frameworks like the EU AI Act and the National Institute of Standards and Technology (NIST) guidelines, which categorize AI applications by risk tier (Ramcharran, n.d.).

1. Mandate a Human-AI Hybrid Model

AI should function as a "drafter" or "screener," never a final "decider." By adopting a hybrid model, AI systems can handle initial triage, technical consistency checks, and pattern recognition, while human experts retain the final scientific and compliance judgment (Dore, n.d.).

2. Define a Strict Context of Use (COU)

Do not give an AI system blanket access or authority. Define exact boundaries upfront. For example, specify that the AI is approved to summarize historical batch records but is strictly prohibited from authorizing product release (Oxley-King, n.d.).

3. Implement Risk-Based Validation

Oversight should be proportionate to the potential risk associated with the AI application. High-impact systems require rigorous, continuous testing against representative datasets to preempt and identify algorithmic degradation over time (Ramcharran, n.d.; Oxley-King, n.d.).

The Bottom Line

Artificial Intelligence will fundamentally transform pharmaceutical manufacturing. However, a modern Quality System cannot trade compliance for speed. When AI makes a GMP mistake, it is not a software glitch—it is a regulatory violation. By enforcing strict human oversight and

embedding AI within mature governance frameworks, QA professionals can harness its power without compromising product safety.

References

Dore, M. P. (n.d.). Artificial Intelligence as a Safeguard for Clinical Scientific Integrity: A Human–AI Hybrid Model for Medical Peer Review. *MDPI*.

Cited by: 0

Fung, M. C. (n.d.). Beyond the hype: navigating the real-world applications of artificial intelligence (AI) in healthcare. *PMC*.

Cited by: 3

McDowall, R. D. (n.d.). Are You Invalidating Out-of-Specification (OOS) Results into Compliance? *Chromatography Online*.

Cited by: 3

Morino, E. (n.d.). AI in clinical trials: Current status, challenges, and future directions for emergency infectious disease clinical trials —Insights from the 2025 iCROWN Symposium. *PMC*.

Cited by: 0

Oxley-King, S. (n.d.). Full article: Regulated bioanalysis in the age of AI: interpreting the FDA-EMA guiding principles for laboratory practice. *Taylor & Francis*.

Cited by: 0

Ramcharran, D. (n.d.). Orchestrating generative AI in pharmacovigilance: predicting and preempting the unpredictable. *PMC*.

Cited by: 4

Sardella, M. (n.d.). The 7th European Pharmacovigilance Congress: speaker abstracts. *PMC*.

Cited by: 0