PHARMASUG CHINA 2026 - PAPER DS223-2026

AI IN REGULATORY SUBMISSIONS: HEALTH AUTHORITY EXPECTATIONS AND EMERGING OPPORTUNITIES

He Liu, Johnson & Johnson

ABSTRACT

Every major regulatory agency—FDA, EMA, PMDA, and NMPA—has now taken a public position on AI in drug development. This paper examines two landmark regulatory developments—FDA's January 2025 draft guidance on AI use in drug development and the joint FDA-EMA ten guiding principles published in January 2026—and their implications for AI use in regulatory submissions. The FDA draft guidance establishes a seven-step credibility framework for AI that produces information or data to support regulatory decision-making, and states that it does not address AI used for operational efficiencies that do not impact patient safety, drug quality, or the reliability of results. The joint FDA-EMA ten guiding principles apply to all AI use in drug development, with validation and oversight proportionate to risk.

INTRODUCTION

The use of artificial intelligence in the drug product life cycle has increased significantly. In an eighteen-month span, every major regulatory agency has taken a public position on AI in drug development:

- **September 2024**: EMA published a Reflection Paper on AI in the lifecycle of medicines [2]
- **January 2025**: FDA released its first draft guidance on AI for regulatory decision-making [1]
- **October 2025**: PMDA published an AI Action Plan for pharmaceutical regulation [5]
- **January 2026**: FDA and EMA jointly released ten guiding principles for good AI practice [3]
- **April 2026**: NMPA released Implementation Opinions on "AI + Drug Regulation" [4]

This convergence of regulatory guidance and industry adoption signals a pivotal moment. AI is no longer experimental in drug development—it is increasingly becoming embedded in routine workflows. As adoption accelerates, the industry needs clarity from health authorities on what is expected when AI is used across the drug product lifecycle. For anyone involved in regulatory submissions, the critical question is now practical: **which of these regulatory expectations apply to AI use in the submissions process, and what needs to be done differently?**

This paper addresses that question by examining the regulatory landscape, identifying what falls within and outside the scope of these guidances, discussing boundary cases where the answer is not yet clear, and presenting industry implementations that demonstrate effective governance approaches.

FDA DRAFT GUIDANCE: THE CREDIBILITY FRAMEWORK AND ITS SCOPE

What the Guidance Covers

In January 2025, FDA released draft guidance titled "Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products" [1]. For detailed analyses of this guidance, see [8]. The scope is defined precisely:

"This guidance discusses the use of AI models in the drug product life cycle, where the specific use of the AI model is to produce information or data to support regulatory decision-making regarding safety, effectiveness, or quality for drugs." [1]

Examples of AI use within scope include: (1) AI that predicts patient safety risk to determine the appropriate level of care—where the model output directly influences patient management decisions in a clinical trial; and (2) AI that analyzes digitized biopsy images to quantify disease activity as a clinical endpoint—where the AI-derived measurement directly constitutes efficacy evidence in the regulatory submission. In each case, the AI output directly

produces information or data to support regulatory decision-making regarding safety, effectiveness, or quality.

For AI applications within this scope, FDA recommends a seven-step credibility assessment framework:

Step	Activity	Purpose
1	Define the question of interest	What decision is the AI model supporting?
2	Define the Context of Use (COU)	What is the specific role and scope of the model?
3	Assess model risk	Based on Model Influence and Decision Consequence
4	Develop credibility assessment plan	Plan commensurate with model risk
5	Execute the plan	Perform validation activities
6	Document results	Credibility assessment report
7	Determine adequacy	Is the model fit for the COU?

The rigor of assessment activities should be commensurate with model risk [1]. High-risk models—where the AI has high influence on the decision and the decision has significant consequence for patient safety, drug quality, or the reliability of results—require detailed model architecture documentation, training data characterization, and stringent performance acceptance criteria.

What the Guidance Explicitly Excludes

Equally important is what falls **outside** scope. The guidance states:

"This draft guidance does not address the use of AI models: (1) in drug discovery, and (2) to streamline workflows and operational efficiencies (such as resource allocation or drafting or writing a regulatory submission) that do not impact patient safety, drug quality, or the reliability of results from a nonclinical or clinical study." [1]

This exclusion is particularly relevant for teams using AI in the submissions process. It means:

AI Use Case	Excluded When
AI-assisted SDTM/ADaM/TFL programming	AI assists execution under human direction
Specification & metadata generation (e.g., Define.xml, TFL specs)	Human reviews and approves
Submission document drafting (e.g., ADRG)	Human finalizes and takes accountability

The key principle: **AI assists programming and submission documentation under human oversight—it does not independently generate or alter regulatory evidence.**

However, falling outside the scope of this guidance does not mean these AI uses are unregulated or unimportant. It means the risk level is lower, and FDA chose not to address them in this particular guidance. The joint FDA-EMA ten guiding principles provide relevant guidance for all AI use in drug development, including operational efficiency use cases. Organizations should keep those principles in mind even when their AI use is outside the FDA draft guidance scope.

It is also important to note that this FDA guidance remains in draft form and does not establish legally enforceable responsibilities. As stated in the guidance: "guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations" [1]. However, sponsors should be prepared to provide credibility evidence if challenged, and early engagement with FDA is encouraged for uncertain cases.

Three Scenarios: Where Is the Line?

Consider three scenarios that illustrate the boundary:

Scenario A: AI assists ADaM programming; a human reviews the generated code, runs it, and validates the output. → **Out of scope.** AI assists execution, not decision-making—an operational efficiency that does not alter what is produced or who is accountable.

Scenario B: AI generates a synthetic control arm for a single-arm trial regulatory submission. → **In scope.** The AI output directly constitutes regulatory evidence.

Scenario C: AI interprets TFL specifications, generates SAS/R programs, triggers execution, self-reviews log errors, auto-fixes code, and validates output—with human involvement only at the QC stage. → **Gray area.**

Scenario C: It Depends on How Humans Are in the Loop

The classification of Scenario C depends on **how humans are involved in the production process**:

- If mandatory human checkpoints are built into the production process—AI can only proceed to the next step after human confirmation—this is **closer to out-of-scope**. The human retains decision authority throughout; AI remains an execution assistant.
- If AI operates end-to-end autonomously, and the human enters only at the QC stage—this is **closer to in-scope**. In this workflow, AI is not merely executing tasks but making decisions: it judges whether code is correct, self-corrects errors, and validates its own output. QC was designed as a second check, not the only check.

For workflows that fall in this gray area, early engagement with FDA is encouraged to align on expectations.

FDA-EMA TEN GUIDING PRINCIPLES

In January 2026, FDA and EMA jointly released ten principles for good AI practice in drug development [3]. Unlike the FDA draft guidance which provides a procedural framework, these principles articulate core values for responsible AI use.

Principles Overview

The following summary draws on the original joint statement [3] and Intuition Labs' analysis [9]:

#	Principle	Core Requirement
1	Human-Centric by Design	AI serves humans; qualified experts must review outputs
2	Risk-Based Approach	Validation proportionate to context of use and model risk
3	Adherence to Standards	Must adhere to relevant standards, including GxP and cybersecurity requirements
4	Clear Context of Use	Define role and scope precisely
5	Multidisciplinary Expertise	Cross-functional collaboration throughout lifecycle
6	Data Governance	Traceability, documentation, privacy protection
7	Model Design & Development	Software best practices, data fit for use, transparency
8	Risk-Based Performance Assessment	Evaluate complete system including human-AI interaction
9	Lifecycle Management	Monitor for drift, periodic re-evaluation
10	Clear Essential Information	Plain language disclosure of performance and limitations

These principles are not binding regulations but broad commitments ("good AI practice") that stakeholders should interpret within existing regulatory frameworks [3]. They apply across the entire drug product lifecycle—from discovery through post-marketing surveillance.

Principles in Practice: DPS2R as a Worked Example

To illustrate how these principles translate from abstract commitments into concrete design decisions, this paper examines the DPS2R system [7]—a supervised agentic AI that reads TLF shells and ADaM metadata to generate draft R code. The following analysis maps DPS2R's design against each principle:

Principle	How DPS2R Reflects It
#1 Human-centric by design	6 mandatory checkpoints—AI cannot proceed without human confirmation
#2 Risk-based approach	3 autonomy tiers: Supervised → Semi-autonomous → Fully autonomous
#3 Adherence to standards	Designed for deployment into a validated regulatory reporting environment

Principle	How DPS2R Reflects It
#4 Clear context of use	Precisely scoped: TLF shells + ADaM metadata → draft R code
#5 Multidisciplinary expertise	AI engineers built it; statistical programmers evaluated it; domain experts drive feedback
#6 Data governance & documentation	Full audit trail with username + timestamp for traceability
#7 Model design & development	Tested across multiple therapeutic areas and studies
#8 Risk-based performance assessment	Measured the human+AI system: programmers re-generated known-correct TLGs with AI; compared outputs for execution success, accuracy, and efficiency
#9 Life cycle management	Pre-deployment; monitoring for drift and periodic re-evaluation planned post-launch
#10 Clear essential information	Published paper discloses which AI was used, performance results, limitations, and scope

The ten principles are not abstract compliance hurdles—they are a design guide that should inform every stage of AI development, deployment, and use.

A Note on Principle 1

One commentary noted [9]:

"If your 'human oversight' is a person who cannot independently complete the task themselves, you have not met this principle—you have only added formalism."

This has direct implications for AI-assisted workflows in regulatory submissions: the person reviewing AI-generated output must be capable of producing that output independently. AI augments expertise; it does not replace it. Organizations that reassign review to staff who lack domain understanding are violating the spirit of Principle 1, regardless of formal process compliance.

OTHER AGENCY POSITIONS

EMA Reflection Paper

EMA's September 2024 Reflection Paper [2] established the European perspective: a risk-based approach, transparency documentation, and lifecycle oversight. This paper laid groundwork for the joint FDA-EMA principles that followed in January 2026.

NMPA: "AI + Drug Regulation"

China's NMPA released Implementation Opinions on "AI + Drug Regulation" (GYJZ [2026] No. 6) in April 2026 [4], establishing a phased roadmap:

- AI empowering the regulator across the full drug lifecycle
- By 2030: integrated AI + drug regulation system preliminarily established
- By 2035: smart drug safety governance fully formed

Unlike the FDA and EMA frameworks that directly address industry's AI use, NMPA's document focuses on empowering the regulator itself—building AI-assisted review, intelligent risk monitoring, and smart inspection capabilities. However, it also commits to accelerating the development of guiding principles for industry's standardized application of AI in pharmaceutical development.

PMDA AI Action Plan

Japan's PMDA published its AI Action Plan in October 2025 [5], focusing on AI to enhance internal reviewer-side operational capabilities.

WHAT INDUSTRY IS DOING

Finding 1: Specification Quality Drives AI Output Quality

A PharmaSUG 2026 study [6] evaluated AI-generated SDTM transformation code under three levels of prompt governance:

Governance Level	Content	Functional Correctness
A1	Prose description only	20%
A2	BDD behavioral specifications (Gherkin syntax)	84%
A3	BDD + project conventions + test scripts	90%

Markov [6] reported that "BDD specifications alone improve functional correctness from 20% to 84%, representing the single largest gain observed. Adding project conventions and test contracts further improves correctness to 90%…"

Markov's core conclusion: **"Specification quality, rather than model sophistication, is the primary determinant of safe and scalable AI-assisted development in clinical data pipelines."** [6]

This suggests that higher-quality structured inputs can meaningfully improve AI output quality. Teams may consider developing and maintaining shared standards, templates, and conventions as reusable inputs for AI-assisted development.

Finding 2: Supervised Agentic AI Achieves Significant Efficiency Gains

The DPS2R team [7] reported the following efficiency gains:

Output Type	DPS2R Time	Estimated Traditional Time	Efficiency Gain
Graphs	0.8 hours	3.0 hours	73% faster
Tables	0.9 hours	2.2 hours	59% faster
Listings	0.7 hours	1.3 hours	46% faster

Output quality improved with iteration. The current system incorporates six human-in-the-loop checkpoints, ensuring human oversight throughout the workflow [7].

An unexpected benefit: "Carefully inspecting the resulting draft R code and how it handles both the creation of the tabular layout and the associated statistical analyses effectively acts as hands-on upskilling" [7].

These results align with a broader industry direction: leveraging AI to automate routine work and accelerate delivery, while relying on human expertise to craft high-quality inputs and maintain substantive oversight at critical checkpoints.

KEY TAKEAWAYS

Based on the regulatory landscape and industry experience, five messages emerge for professionals involved in regulatory submissions:

1. Human Accountability Is Non-Negotiable

The person who reviews and submits AI-generated output bears full responsibility for its correctness. Whether code is written by a human or generated by AI, accountability remains with the sponsor. This is not a new requirement created for AI—it is the existing accountability model applied consistently.

2. The Reviewer Is Also Using AI

Health authorities are not merely setting rules—they are actively deploying AI to review submissions. FDA announced completion of its first AI-assisted scientific review pilot in May 2025 [10]. NMPA and PMDA are building AI-powered review capabilities.

This means the quality bar for submissions is **rising, not falling**. Organizations may benefit from proactively integrating AI into their own workflows.

3. Regulatory Expectations Are Risk-Based

The FDA draft guidance and joint FDA-EMA principles both adopt a risk-based approach: the level of validation and

oversight should be proportionate to the potential impact of AI on patient safety, drug quality, and the reliability of results. Using AI to draft an ADRG carries different risk than using AI to generate a synthetic control arm—and the governance requirements differ accordingly.

4. Review by Design

AI changes how we work. As AI handles more of the drafting work, the human role shifts toward review and decision-making. This shift must be intentional:

- Make sure human review stays **substantive, not ceremonial**
- Design workflows with explicit checkpoints where human judgment adds value
- Document what the human reviewer examined and confirmed
- Ensure reviewers have the expertise to identify errors AI might introduce

5. Think Critically

"Looks right" is not "is right." AI-generated output can appear logically sound and well-reasoned, which may lower the incentive to scrutinize further. Think critically at every stage: invest in structured, high-quality inputs that shape AI behavior, and evaluate the outputs rigorously rather than accepting them at face value. This is what makes human oversight substantive rather than ceremonial.

CONCLUSION

Health authorities' AI framework is still developing, but the direction is clear: governance should be proportionate to risk. As AI capabilities advance and organizations entrust AI with broader responsibilities, what is currently classified as operational efficiency could shift toward evidence generation—raising the associated risk level.

Organizations should monitor regulatory developments closely and engage with health authorities proactively—both to align on expectations and to contribute industry experience to the evolving framework.

REFERENCES

[1] FDA. *Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products — Draft Guidance.* January 2025. Federal Register Vol. 90, No. 4, FDA-2024-D-4689.

[2] EMA. *Reflection Paper on the Use of Artificial Intelligence in the Lifecycle of Medicines.* September 2024.

[3] FDA & EMA. *Guiding Principles of Good AI Practice in Drug Development.* January 2026.

[4] NMPA. *Implementation Opinions on "Artificial Intelligence + Drug Regulation"* (GYJZ [2026] No. 6). April 2026.

[5] PMDA. *AI Action Plan for Pharmaceutical Regulation.* October 2025.

[6] Markov R. "A Specification-Driven Approach to Improve the Reliability of AI-Generated SDTM Transformation Programs." *PharmaSUG 2026, Paper AI-420.*

[7] Rothe M, Kalappurakal S, Liu X. "Next-Gen Statistical Programming: AI as Your Coding Co-Pilot in R." *PHUSE Connect US 2026, Paper ML08.*

[8] Intuition Labs. "FDA Draft Guidance on AI in Drug Development Explained." June 2026.

[9] Intuition Labs. "FDA-EMA Good AI Practice Guidelines in Drug Development Explained." June 2026.

[10] Podichetty JT, Bauer AM, Xu R, et al. "How AI Transforms Regulatory Submission: Current Clinical Implementation and Future Prospects." *Clinical and Translational Science.* 2025;18(12):e70434.

CONTACT INFORMATION

He Liu Johnson & Johnson Email: hliu167@its.jnj.com

DISCLAIMER

The views expressed in this paper are those of the author and do not necessarily reflect the official policy or position of Johnson & Johnson.

ACKNOWLEDGMENTS

The author thanks the PharmaSUG China 2026 Program Committee for the opportunity to present this work. The author used AI tools (Claude) to assist with drafting and editing this manuscript. All content was reviewed, verified, and finalized by the author.