AI-Driven Simulation of FDA DCOA Review to Inform Evidence Planning and Regulatory Strategy

MSR24

Rajdeep Kaur, PhD[1], Sarah Donelson, MA[2], Matthew Dixon, PharmD, PhD[3], Barinder Singh, RPh[1], Shubhram Pandey, MSc[1], Nicola Waddell, HNC[4], Siguroli Teitsson, BSc, MSc[5].
[1]Pharmacoevidence, Mohali, India, [2]Bristol Myers Squibb, Novato, CA, USA, [3]Bristol Myers Squibb, Princeton, NJ, USA, [4]Pharmacoevidence, London, United Kingdom, [5]Bristol Myers Squibb, London, United Kingdom.

Value of AI Simulation

- The U.S. Food and Drug Administration (FDA) requires robust, evidence-based Clinical Outcome Assessments (COAs) to evaluate patient-focused evidence and demonstrate treatment benefit
- Regulatory reviews are rigorous and dynamic, requiring sponsors to proactively identify and address potential evidence gaps to ensure comprehensive submissions
- Advances in artificial intelligence (AI), particularly Retrieval-Augmented Generation (RAG) and multi-agent architectures, provide an opportunity to simulate regulatory reviews
- By anticipating FDA reviewer considerations, AI simulations have the potential to provide manufacturers with early feedback, identify evidence gaps, and strengthen submission readiness before regulatory agency engagement

Objective

- The goal of this study was to develop an AI-powered, agent-driven platform using RAG to support early evaluation of COA strategies and identify evidence gaps in alignment with FDA regulatory guidance

Methods

Simulation Framework

- The COA simulation tool was designed using a multi-agent architecture integrated with a RAG framework to simulate the interaction between a sponsor and the FDA during COA evaluations
- The tool was deployed on a secure cloud-based infrastructure leveraging Python, microservices, databases, and AI frameworks integrated with large language models
- This architecture allowed agents to collaborate dynamically in a controlled and context-driven environment

Data migration and RAG pipeline

- The sponsor can upload documents such as the briefing book, FDA guidance, and related materials (PDF, PPT, DOC, and text formats) via a user interface
- The uploaded documents were processed through a RAG framework, ensuring AI reviewers access context-specific, traceable source information during evaluations
- Data files were securely stored in an S3 bucket, providing traceability, easy access for subsequent processing, and enabling users to view or retrieve files when needed as shown in Figure 1
- The standardized data were then ingested into the RAG pipeline, where content was chunked, indexed, and stored in a vector database

Agent Goal Setting

- Goal setting provided the scope and purpose of the review, defining what each AI reviewer agent (DCOA Lead, Statistician, Medical Officer, etc.) should evaluate in relation to their role
- The goals were grounded in relevant FDA guidance documents (e.g., FDA's 2009 Guidance for Industry, PFDD (Patient Focused Drug Development) Guidances 1-3, Draft PFDD Guidance 4, DDT Qualification Guidance (2014), and Substantial Evidence Guidance (2019))
- This ensured that the AI agents' responses modeled the structure and communication style characteristic of regulatory reviews

Setup of Reviewer Agents

- Different AI reviewer agents were assigned roles aligned with FDA reviewer responsibilities. For example, the DCOA Lead framed the sponsor's question, coordinated input from other agents, and delivered the final regulatory-style conclusion
- SMEs validated these role assignments and optimized the agents to ensure accuracy and relevance
- The AI reviewer team comprised one Sponsor agent and five FDA Reviewer agents; Division COA (DCOA) Lead, PFSS Team Leader, Statistician, Medical Officer, and Clinical Team Leader (as shown in Table 1). Each reflecting specific FDA reviewer roles and responsibilities

Table 1: AI Agent Roles and Responsibilities in PRO Review Simulation

AI agent name	AI agent role
DCOA	Leads the overall regulatory review, frames the sponsor question, coordinates feedback across review disciplines, and delivers the final regulatory conclusion consistent with FDA guidance expectations.
PFSS Team Leader	Provides statistical expertise on patient-reported outcomes and patient-focused drug development, ensuring valid and meaningful data capture
Statistician	Reviews statistical design, analysis methods, and data integrity to ensure methodological robustness and compliance in clinical trials
Medical Officer	Safeguards patient safety through risk assessment, monitoring oversight, and evaluation of safety protocols in clinical development
Clinical	Assesses pharmacokinetic, pharmacodynamic, and dosing data to ensure scientific rigor and regulatory compliance in IND submissions

Knowledge Base creation and execution of agents

- A dedicated knowledge base was created for each simulation, guided by the review goals and the assigned roles of the agents
- Relevant information chunks from the RAG pipeline were pre-selected to match each agent's responsibilities and context
- This process ensured that reviewer agents relied on the most relevant evidence when generating their evaluations

Agent Discussion Flow

- The DCOA Agent initiated the discussion and invited all reviewer agents to contribute based on their roles as shown in Figure 2
- Agents could interact, view, and exchange perspectives on the sponsor's regulatory question, simulating a collaborative FDA-style dialogue
- The DCOA Agent then synthesized the discussion into a regulatory-style summary highlighting gaps, data insights, and recommendations

Figure 1: Workflow for Data Pre-processing and Retrieval in AI-Driven COA Simulation

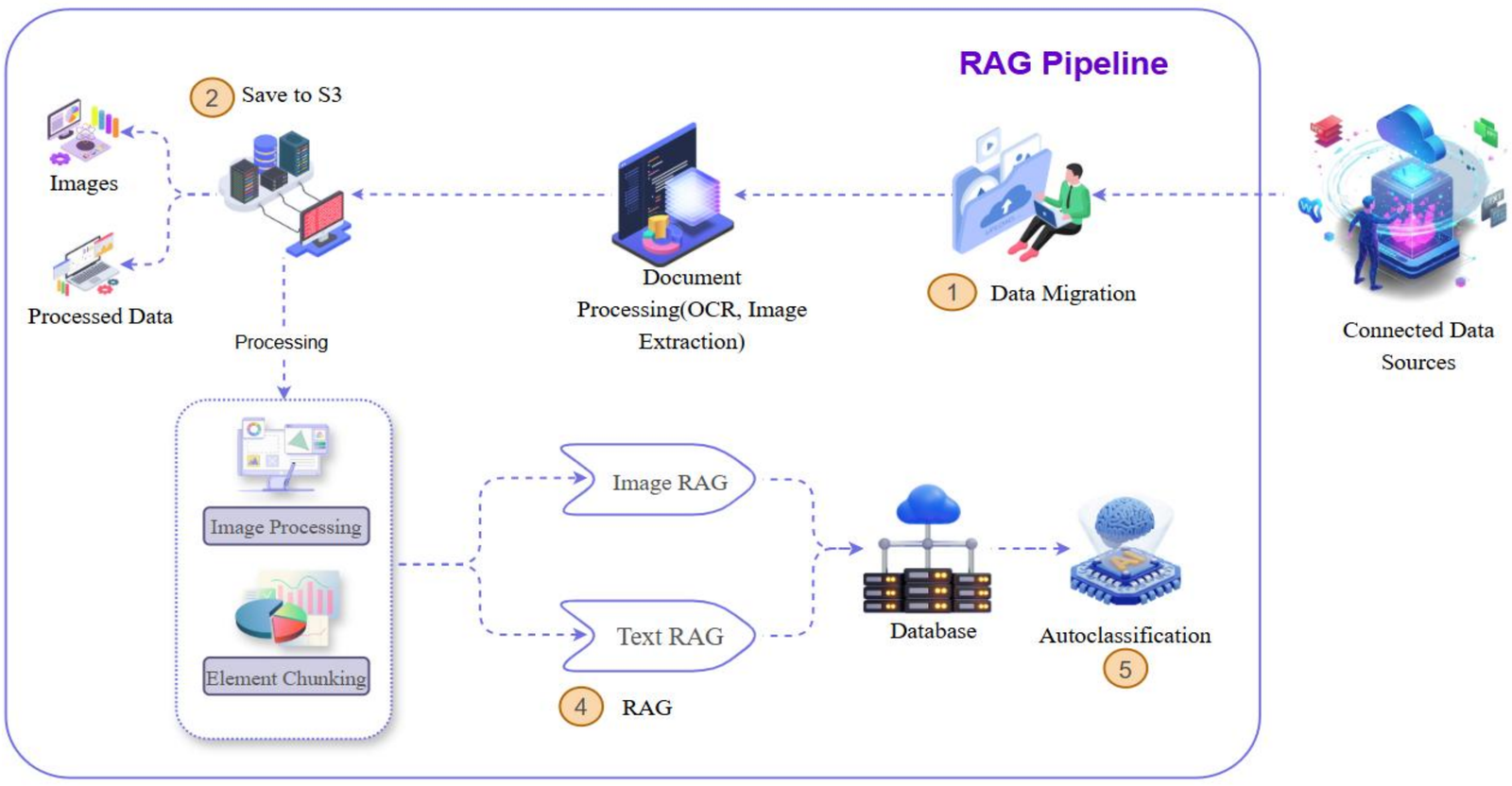
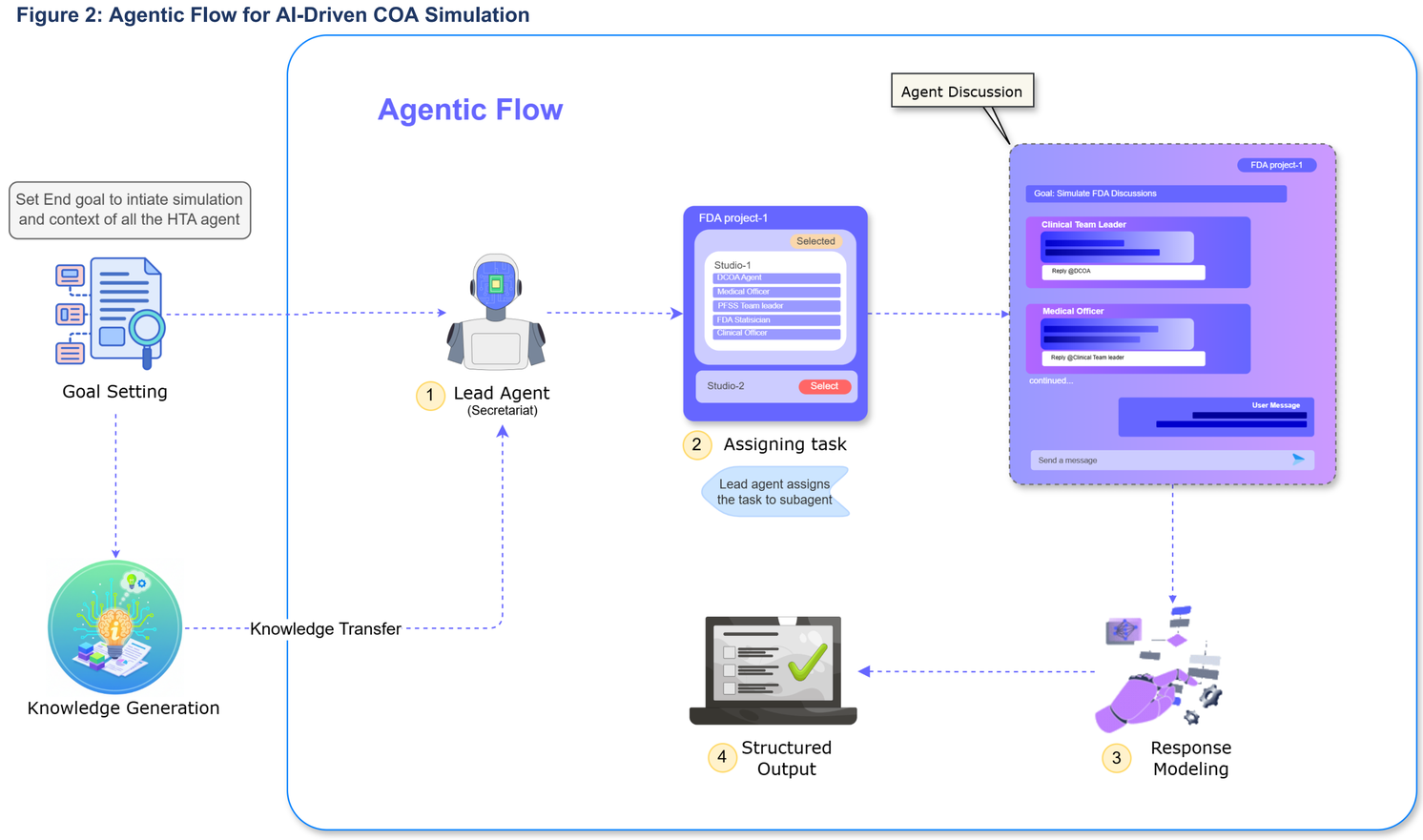


Figure 2: Agentic Flow for AI-Driven COA Simulation



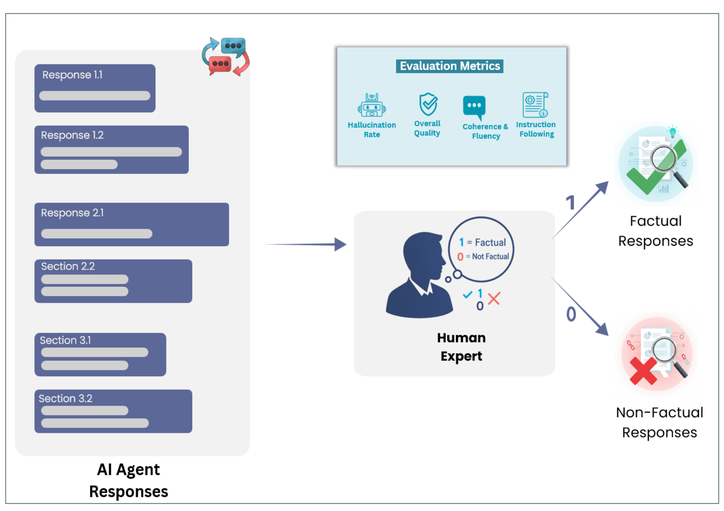
SME Evaluation

- SMEs reviewed these responses using predefined evaluation metrics (hallucination rate, coherence & fluency, instruction following, relevance, completeness and overall quality)
- For evaluation, the tool was applied to two therapeutic areas using regulatory questions addressing the conceptual model, content validity, and psychometric adequacy
- A binary scoring system (1 = factually accurate and aligned with guidance, 0 = not aligned) was applied to maintain objectivity, ensuring robust validation of AI outputs

Results

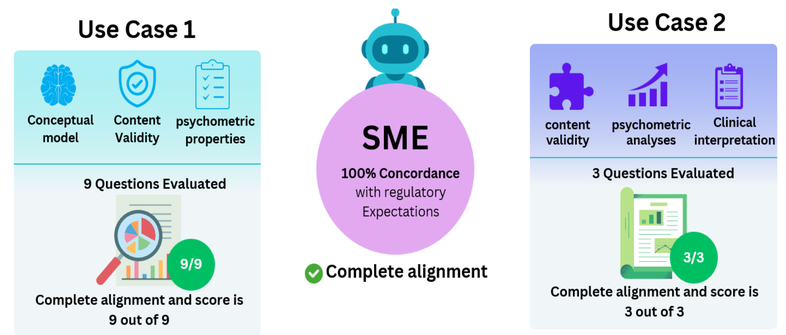
- The tool successfully simulated regulatory responses as determined by SMEs who assessed accuracy and alignment with regulatory guidance as shown in Figure 3
- In the first therapeutic area, SMEs evaluated nine regulatory questions covering key aspects of conceptual model and validity, and confirmed complete alignment, resulting in a score of 9 out of 9
- In the second therapeutic area, SMEs evaluated three regulatory questions that focused on content validity and approaches for defining clinically meaningful score differences, confirming complete alignment with a score of 3 out of 3 (Figure 4)

Figure 3: Evaluation and Validation of AI Agent Responses in COA Simulation



- These findings indicate strong concordance between the AI-generated outputs and regulatory expectations across both therapeutic areas
- SMEs confirmed that the tool-generated responses were accurate, complete, and aligned with regulatory expectations, demonstrating reliability and relevance across therapeutic areas
- SMEs noted that the tone of agent responses could be further refined to better reflect regulatory reviewer style

Figure 4: SME Validation Outcomes Across Use Cases



Conclusion

- ✓ This study demonstrated the feasibility of using a RAG-based, agent-driven simulation framework to support early evaluation of COA strategies
- ✓ By aligning agent evaluations with sponsor evidence and regulatory guidance, the system enables early detection of evidentiary gaps and improves submission readiness
- ✓ SME assessment confirmed strong concordance between AI-generated responses and regulatory expectations across the two evaluated therapeutic areas, while recommending further refinement of tone to better mirror the style of regulatory reviewers
- ✓ Future efforts will focus on extending validation across additional disease areas and global regulatory agencies, and advancing the proof-of-concept into a full-scale product with secure access, centralized management, and exportable outputs

Presented at ISPOR Europe 2025; November 09-12, 2025; Glasgow, Scotland, UK