Agentic AI–Enabled Integration of a Hybrid System of Predictive Models for Accelerated Direct-Compression Drug Product Development

Theo Tait[*1] Pooria Soltani[*1] Mohammad Salehian[1] Daniel Markl[1]

[*]Equal contribution [1]CMAC, University of Strathclyde, Glasgow, United Kingdom

Correspondence to: Daniel Markl daniel.markl@strath.ac.uk

1. Introduction

Direct compression is widely used in pharmaceutical manufacturing due to its simplicity, scalability, and reduced processing burden. However, predicting tablet performance under direct compression remains challenging because final product properties arise from complex, non-linear interactions between material characteristics, blend attributes, and compaction conditions. Small changes in formulation composition, excipient grade, or particle morphology can lead to disproportionate effects on flowability, compressibility, and mechanical strength, making formulation development highly sensitive and difficult to generalise.

Addressing this sensitivity requires development workflows that systematically integrate mechanistic understanding, data-driven prediction, and process context within a unified digital framework. In this work, Quality by Digital Design (QbDD) principles [1,2] are adopted by embedding predictive models directly into formulation-development workflows, enabling structured exploration of formulation and process space while remaining accessible beyond formulation specialists. Existing modelling tools are often siloed or expert-driven, limiting their reuse, scalability, and value across cross-functional Chemistry, Monitoring and Control (CMC) teams.

To meet this need, *Digital Formulator* was developed as a hybrid predictive modelling system that combines mechanistic, domain-informed models with machine-learning approaches to predict critical attributes relevant to direct-compression formulation development at material, blend, and tablet levels. These include particle size and shape, flowability metrics, tapped density, feeder mass-flow performance , and compaction-derived porosity and tensile strength [3,4].

This study presents a Proof-of-Concept for operationalising QbDD within an agentic AI environment using a Model Context Protocol (MCP)-enabled architecture (Figure 1). By enabling AI agents to autonomously select, invoke, and sequence modular models within end-to-end CMC workflows, the system supports digitally driven formulation design, reducing experimental burden and broadening access to formulation insight.

2. Methodology

The hybrid modelling suite was engineered into a modular inferencing architecture compatible with AI-driven orchestration. Each predictive model was restructured to use a common set of inputs and outputs based on material properties, formulation composition, and compaction conditions. This enabled mechanistic models – such as those describing changes in particle populations during compression – to be combined with machine-learning models trained on high-resolution blend and tableting datasets.

To ensure reliable interaction between models, standardised feature definitions, consistent metadata handling, and robust, deterministic error-handling procedures were applied across the system. A dedicated computational integration layer was developed to convert externally supplied physicochemical information into the numerical formats required by each model. This allows models to be executed automatically, either sequentially or in parallel, without manual configuration.

All components were packaged within a containerised runtime environment to isolate software dependencies and ensure reproducible behaviour when models are invoked by AI agents. The full system was deployed within an agentic AI framework, where large-language-model–based agents can autonomously select, execute, and combine predictions from multiple models as part of multi-step formulation-design workflows. To assess the agentic framework, the proposed framework's predictions were benchmarked against experiments and domain knowledge, and orchestration quality was quantified via task-success rate, human-intervention rate, and end-to-end latency relative to non-agent baselines.

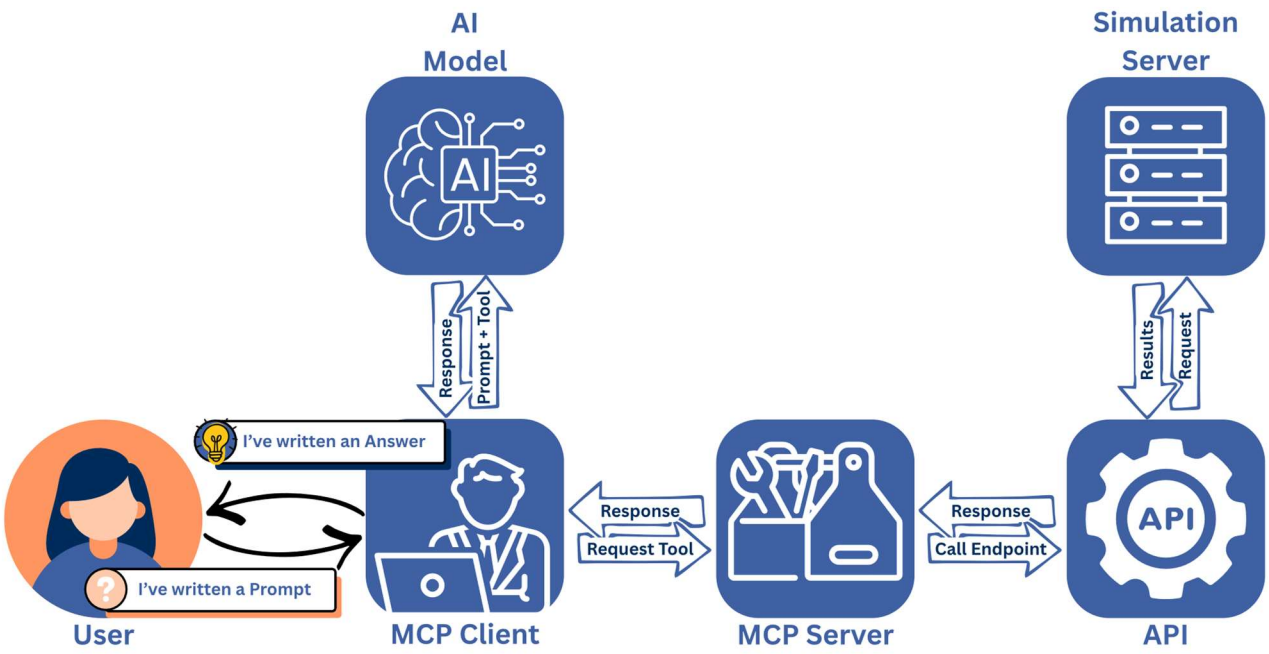

Figure 1: A simplified illustration showing the structure for a a Proof-of-Concept for operationalising Quality by Digital Design (QbDD) within an agentic AI environment using a Model Context Protocol (MCP)-enabled architecture.

3. Scientific Rationale and Implications

The integration of a hybrid predictive modelling system within an enterprise agentic AI environment demonstrates how advanced computational tools can accelerate direct-compression tablet development. Our *Digital Formulator* Proof-of-Concept shows that data-driven formulation exploration is not only feasible but can meaningfully reduce experimental workload, enhance process understanding, and shorten development timelines. While recent advances in pharmaceutical digitalisation have explored digital twins, hybrid models, and machine learning for process optimisation and prediction of tablet quality attributes, most approaches remain process-centric or focus on predicting individual properties such porosity, tensile strength, or dissolution [3, 4, 5]. Embedding hybrid models within an agentic AI framework to enable real-time, autonomous formulation exploration remains largely unexplored. By connecting hybrid predictive modelling directly to formulation optimisation workflows, our work bridges the gap between model-based prediction and operational autonomy, illustrating a practical route for industrial AI integration and laying the groundwork for AI-native drug-product design and broader digital CMC transformation.

Acknowledgments

Authors are thankful to the Scottish Funding Council (SFC) for financially supporting the Proof of Concept Project "AI-enabled CMC Data-Factory™"(project code: H26011).

References

[1] Mustoe, C.L., et al., Quality by digital design to accelerate sustainable medicines development. International Journal of Pharmaceutics, 2025: p. 125625.
[2] Houson, I., et al., Quality by digital design in action: a workflow for crystallisation and isolation. International Journal of Pharmaceutics, 2025: p. 126343.
[3] Salehian, M., et al., A hybrid system of mixture models for the prediction of particle size and shape, density, and flowability of pharmaceutical powder blends. International Journal of Pharmaceutics: X, 2024. 8(1): p. 100298.
[4] Abbas, F., et al., Accelerated Medicines Development using a Digital Formulator and a Self-Driving Tableting DataFactory. arXiv preprint arXiv:2503.17411, 2025.
[5] Tait, T., *et al.*, Empirical model variability: Developing a new global optimisation approach to populate compression and compaction mixture rules. International Journal of Pharmaceutics, 2024: p. 124475.