

FastOMOP-Skills: Agentic Skills for Automated Real-World Evidence Generation Using LLMs

Farr, I.¹, Shenzhang, S.³, Moller-Grell, N.³, Dobson, R.^{3,4,5,6,7}, Chandrabalan, V.^{1,2}

1. Lancashire Teaching Hospitals Trust, Preston, UK
2. Lancaster University, Lancaster, UK
3. Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology and Neuroscience (IoPPN), King's College London, London, UK
4. Institute for Health Informatics, University College London, London, UK
5. NIHR Biomedical Research Centre, University College London Hospitals National Health Service Foundation Trust, London, UK
6. Health Data Research UK London, University College London, London, UK
7. NIHR Biomedical Research Centre, South London and Maudsley National Health Service Foundation Trust and King's College London, London, UK

Background

The translation of clinical research questions into actionable Real-World Evidence (RWE) is traditionally hindered by manual data engineering bottlenecks and vital governance constraints. While Large Language Models (LLMs) offer remarkable potential to accelerate analytics, their unconstrained application to sensitive healthcare databases introduces critical risks, including syntax hallucinations, non-reproducible code, and potential privacy breaches. To challenge these vulnerabilities and bottlenecks, we introduce a secure LLM framework within the FastOMOP [1] ecosystem, using a dedicated repository of discrete skills to safely automate RWE workflows using the OMOP Common Data Model.

Methods

We developed a library of 28 parameterized skills designed to reduce the latency between a user-defined research question and a high-quality initial report draft. These skills execute core epidemiological operations, such as identifying concept phenotypes, cohort definition, propensity score estimation, statistical tests, and data visualization. The system architecture enforces a formal separation between the underlying LLM, the operational harness, and the skill layer. The LLM serves strictly as a reasoning engine, while the harness controls execution, data access, and state management.

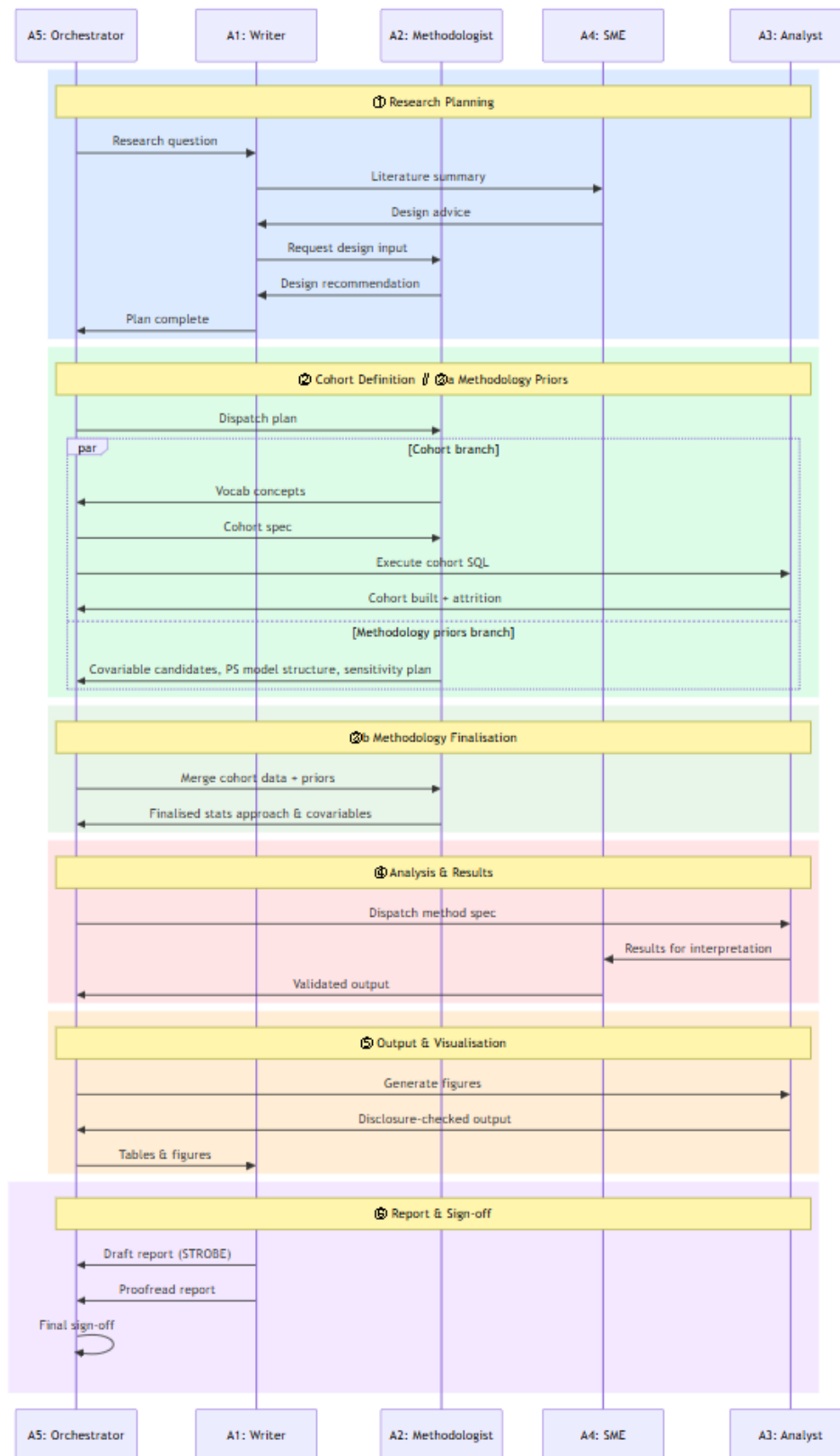
Skills operate as self-contained, composable units with dynamically isolated context. By providing only task-essential inputs and restricting access to the broader pipeline state, the harness strictly constrains the LLM's focus at each step. Above this foundational layer, we tested a pathway managed by five specialized roles:

Orchestrator, Writer, Methodologist, Analyst, and Subject Matter Expert. These personas handle strategic sequencing, delegation, and synthesis. Because each role is defined by its assigned skillset, the underlying code remains static, enabling flexible recombination to support alternative configurations or custom report requirements.

Table 1: Atomic skills assigned to agents for different stages of the research pathway

Stage	Key Skills	Agent Layer
Research Planning	literature-review, study-design-selector, sample-size-feasibility	Writer, SME
Cohort Definition	omop-vocabulary-lookup, concept-set-generation, phenotype-validation, drug-exposure-era, temporal-logic, cohort-definition, omop-cdm-query, data-quality-assessment, cohort-attrition	Methodologist, Analyst
Methodology Selection	covariable-selection, propensity-score, self-controlled-designs, negative-control-analysis, sensitivity-analysis, multi-site-analysis, stepped-wedge-design, randomised-controlled-trial, orchestration-pipeline	Methodologist
Analysis & Results	cohort-characterisation, incidence-prevalence, time-to-event-analysis, meta-analysis	Analyst, SME
Output Checking & Visualisation	output-disclosure-check, visualisation	Analyst
Summary & Discussion	rwe-report-generator, review-proofreading, interpretation-implications	Writer, SME

Figure 1: Agentic skills research pathway swimlane chart, condensed for illustration



Results

This framework establishes a reliable pipeline capable of converting natural language prompts into reproducible, governance-compliant workflows. Through dynamic orchestration, the system demonstrated that agents could safely execute complex steps via the skill repository. For example, the Analyst agent leveraged specific skills to generate high-resolution baseline tables while simultaneously applying automated disclosure controls. Furthermore, the architecture enabled robust iterative feedback loops, intermediate diagnostics produced by the Analyst were reviewed by the SME agent, triggering refinement cycles until pre-defined data quality thresholds were achieved.

Conclusion

This robust methodological framework enables agentic RWE generation. By anchoring LLM agents to a specific skill repository rather than allowing fully unconstrained code generation, complex epidemiological workflows can be automated rapidly without sacrificing integrity or safety. Ultimately, this architecture provides a scalable, hallucination-limiting foundation for integrating LLMs securely into sensitive electronic health record environments.

References

1. FastOMOP. (2026). *FastOMOP: LLM-enabled Real World Evidence and Real Time Intelligence using OMOP Common Data Model*. GitHub.
<https://github.com/fastomop>