Advancing Electronic Clinical Quality Measure (eCQM) Interoperability: Model Context Protocol (MCP)-Orchestrated CQL-to-OMOP Translation

Star Liu[a], Robert B Barrett[a], Kyle Zollo-Venecek[b], Benjamin Riesser[c], Benjamin Martin[a]
[a]Biomedical Informatics and Data Science, Johns Hopkins University, Baltimore, MD, USA
[b]CTSI, Tufts University School of Medicine, Boston, MA
[c]Improving Health Outcomes, American Medical Association, Greenville, SC

Background

Electronic clinical quality measures (eCQMs) are measures published by the CMS to evaluate the quality of care (i.e., patient engagement, patient safety, population health, etc.)[1]. eCQMs are expressed using Clinical Quality Language (CQL) and value sets from Value Set Authority Center (VSAC), which include codes and terms that define a set of data in the EHR. Despite the CMS outlined recommendations to improve the support around reporting eCQMs, challenges remain[2-6]. Notably, a lack of standardized coding limits eCQM implementation and the ability to highlight opportunities to improve quality of care[3,4]. Furthermore, the lack of technical and process support for developing and implementing eCQMs among smaller institutions poses a threat to improving the quality of care and reimbursement at those sites[6].

The Observational Health Data Sciences and Informatics (OHDSI) community is uniquely positioned to help institutions meet CMS's call for reporting eCQM to improve the quality of care. OHDSI has been a leader in advancing health information interoperability by establishing the Observational Medical Outcomes Partnership (OMOP) common data model (CDM), which leverages standardized vocabularies. Additionally, OHDSI has developed a suite of analytical software and models to facilitate standards development, data characterization, and reproducible research on risk prediction, safety surveillance, treatment effectiveness, and quality improvement. However, OHDSI does not currently support quality measure calculations or reporting.

VSAC value sets are laborious and difficult to maintain manually[2,7]. Critical steps such as terminology and concept mappings are still done manually, driven by human judgment. However, manual processes are not scalable and tractable when concept sets get large, as in the case of complex cohort definitions and their dependency in eCQMs. Large language models (LLMs) have demonstrated similar efficacy to traditionally manual processes and present an opportunity for automated eCQM translation and reporting[8,9]. Moreover, introducing LLMs to facilitate this process requires a dedicated, scalable, and easy-to-maintain environment.

We published a Model Context Protocol (MCP) server, **Mercurius MCP** (https://github.com/StarLiu1/mercurius-mcp), a modular and low-overhead infrastructure that enhances LLMs with additional resources and tools to streamline the translation of eCQMs into executable SQL queries defined using OMOP concepts. Providing foundational support for eCQM reporting further incentivizes OMOP adoption beyond its capabilities for research functions.

Methods

Model Context Protocol (MCP) is a client-server architecture first proposed by Anthropic, aiming to standardize the augmentation of LLMs with additional contexts and resources, such as local databases and custom tools (Figure 1)[10]. Without additional contexts and tools, LLMs are next-word prediction algorithms trained using historical data and can be limited in their response relevance and accuracy[11]. MCP is the broader approach that supports the development of advanced AI and agentic systems, encompassing RAG systems, a technique used to provide additional information before generating outputs.

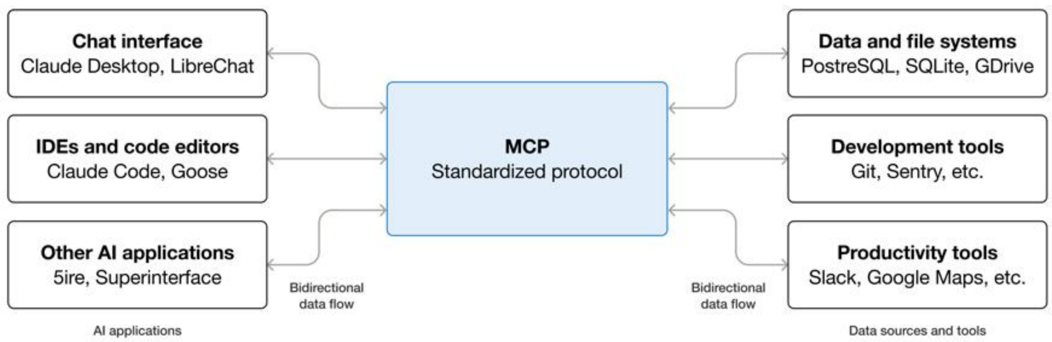

Figure 1. The role of MCP in connecting data and tools to the client applications.

The Mercurius MCP server provides several key modules:

1. **CQL Parser**: Extracts VSAC value set identifiers from CQL query.
2. **VSAC API Connector**: Fetches concepts from the VSAC.
3. **OMOP Mapper**: Maps extracted VSAC concepts to the OMOP standardized vocabulary.
4. **OMOP Validator**: Validates mapped OMOP concepts for relevance and eligibility.
5. **SQL Generator**: Generates executable SQL queries against OMOP databases.

The extraction of value sets from a CQL query is done through an executable regular expression script to extract the VSAC value set identifiers. Concept details are fetched from VSAC via available APIs. VSAC concepts are then mapped to OMOP concepts through SQL queries that map source VSAC concepts to OMOP concepts through the *concept_relationship* table. The LLM proceeds to filter the list of mapped OMOP concepts for eligibility, given metadata from the value set and the intention of the measure. Finally, the LLM generates an executable SQL query based on the logical representation of a measure represented in CQL, incorporating the OMOP concepts retrieved from mapping processes.

An LLM can use the entire suite of modules to automate the pipeline from taking a structured CQL as input, to fetching the concepts in each Value Set, to mapping to the proper OMOP concept, and to generating a relevant and executable SQL query against an OMOP database.

Results

As shown in Figure 2, Claude Desktop was able to access the various tools offered through this MCP server.

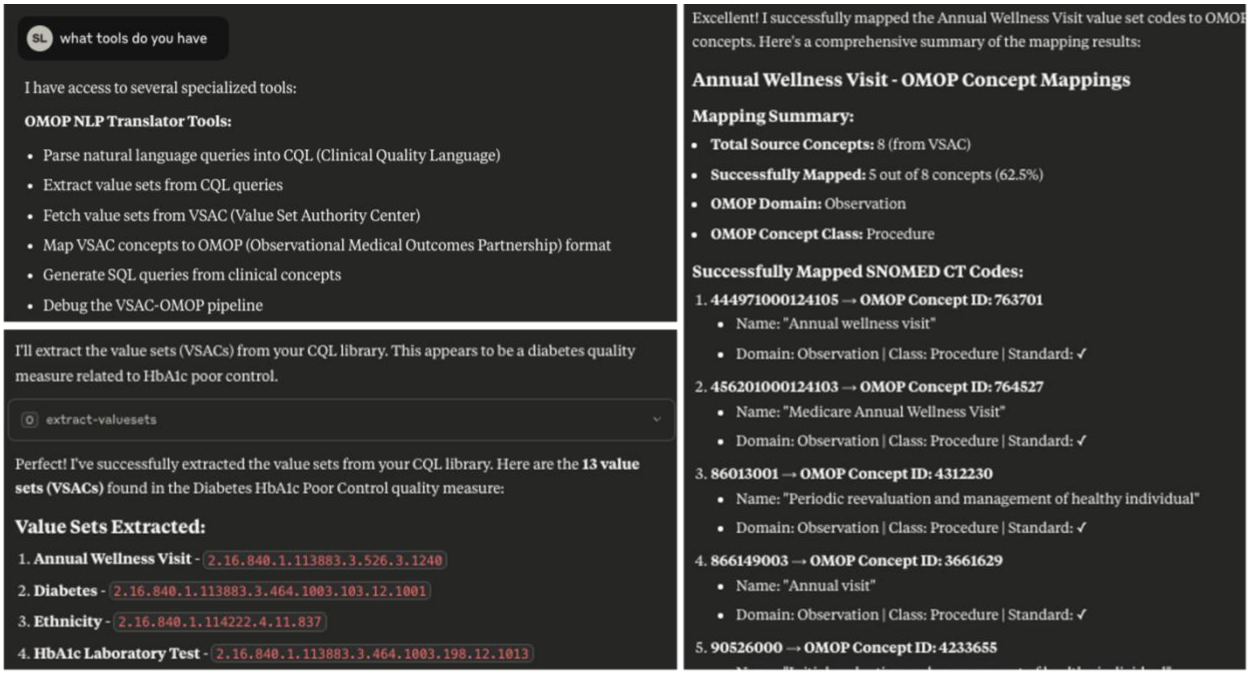

Figure 2. Custom tools (CQL parser, VSAC API call, mapping to OMOP) were made available to Claude Desktop. Claude successfully used the tools to extract value sets, fetch concepts from VSAC, and map Annual Wellness Visit-related concepts to OMOP standardized concepts.

Conclusion

We built an open-source MCP server to support the extraction of concepts from CQL, execution of API calls to VSAC, the translation to OMOP concepts and validation, and the generation of executable OMOP-based SQL queries. This MCP-driven pipeline offers a scalable, accessible solution to eCQM development, implementation, and reporting challenges. This MCP architecture completes the chain from narrative specification to reporting, expanding the ecosystem to support both community-defined measures and local adaptations. Enabling MCP positions organizations to leverage real-time eCQM insights within their analytics frameworks without high overhead, vendor dependency, or specialized human resources. Furthermore, the flexible, modular design lays the groundwork for future enhancements, such as quality dashboards, regulatory reporting, and advanced AI-augmented quality platforms.

Efforts are underway to strengthen the LLM validation of concept mapping; We will also develop support for human-in-the-loop evaluation of LLM translation and review of concept mappings flagged for removal. We will expand the MCP server to support FHIR integration. The MCP server serves as a building block to support eCQM reporting and the broader health information interoperability efforts, with OMOP at the heart. This process lowers the barrier for organizations to develop and implement quality measures, driving improved quality of care and care delivery, as well as reimbursement for understaffed or economically challenged organizations.

Acknowledgments

This research was supported by the NLM, part of the NIH, under award number T15LM013979.

References

1. Electronic Clinical Quality Measures Basics | CMS2024; . Accessed Jun 26, 2025.
2. Zahn L, Ahmad H, Sittig D, Russo E, Koh B, Nimocks A, et al. The Fault in Our Sets: A Mixed Methods Analysis of Clinical Value Set Errors. 2025 March 1:2025.02.27.25323054.
3. Baumann Kreuziger L, Keenan M, Dykhoff H, Hall M, Campbell K, Cahill E, et al. Lack of Standardized Coding Limits Accuracy of Electronic Clinical Quality Measure for Pulmonary Embolism Diagnosis. Am J Med Qual 2025 Mar-Apr 01;40(2):38–43.

4. Miller AO, Chin AS, Carli AV, Sayegh G, Chee D, Buchalter DB, et al. Electronic Clinical Quality Measures for Prosthetic Joint Infection Diagnosis: Pitfalls and Potential. J Healthc Qual 2025 Feb-Apr 01;47(1):e0467.
5. Schreiber M, Krauss D, Blake B, Boone E, Almonte R. Balancing value and burden: the Centers for Medicare & Medicaid Services electronic Clinical Quality Measure (eCQM) Strategy Project. J Am Med Inform Assoc 2021 August 12;28(11):2475–2482.
6. Parchman ML, Anderson ML, Penfold RB, Kuo E, Dorr DA. The Ability of Practices to Report Clinical Quality Measures: More Evidence of the Size Paradox? J Am Board Fam Med 2020;33(4):620–625.
7. Fung KW, Xu J, Gold S. The Use of Inter-terminology Maps for the Creation and Maintenance of Value Sets. AMIA Annu Symp Proc 2019;2019:438–447.
8. Huh S. Comparative Analysis of ChatGPT-4 for Automated Mapping of Local Medical Terminologies to SNOMED CT. Stud Health Technol Inform 2025 -05-15;327:813–817.
9. Yoon D, Han C, Kim DW, Kim S, Bae S, Ryu JA, et al. Redefining Health Care Data Interoperability: Empirical Exploration of Large Language Models in Information Exchange. Journal of Medical Internet Research 2024 -05-31;26(1):e56614.
10. Model Context Protocol: Introduction. Accessed Jun 26, 2025.
11. Lewis P, Perez E, Piktus A, Petroni F, Karpukhin V, Goyal N, et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. 2021 April 12.