

UC San Diego

Family Medicine Research Symposium 2026

Title

Microsoft Copilot as a Thematic Categorization Tool in Large-Scale Clinical Datasets: A Comparative Evaluation Against Independent Human Reviewers

Permalink

<https://escholarship.org/uc/item/1cw8c439>

Authors

Ho, Jenny

Huffman, Jared

Publication Date

2026-08-12

Peer reviewed

Background: Osteopathic manipulative treatment (OMT) is a noninvasive intervention with demonstrated inpatient benefits including reductions in analgesic use and hospital length of stay. UC San Diego Health is among the first academic health systems to establish a formal inpatient Osteopathic Manipulative Medicine (OMM) consult service. As large clinical datasets grow, efficient tools such as large language models (LLMs) are increasingly needed.

Objective: To evaluate Microsoft Copilot as a thematic categorization tool applied to a large clinical dataset and quantify agreement with independent human reviewers.

Methods: This study analyzed 5,531 free-text entries placed by referring services for OMM consultation at UC San Diego Health (May 2019–April 2025). Copilot was prompted using a standardized methodology to categorize each entry. Two osteopathic physicians independently categorized entries, then reconciled assignments referencing Copilot's output. Agreement was quantified using Cohen's kappa and percent agreement.

Results: Across 20 categories, overall percent agreement was 97%, with a Cohen's kappa of 0.79. Discrepancies were most common in entries with misspellings, ambiguous terminology, or insufficient context.

Conclusion: These findings support the utility of LLMs as scalable tools in clinical research. Future directions include standardizing order sets with predefined thematic selections to reduce free-text variability and improve downstream analysis.