

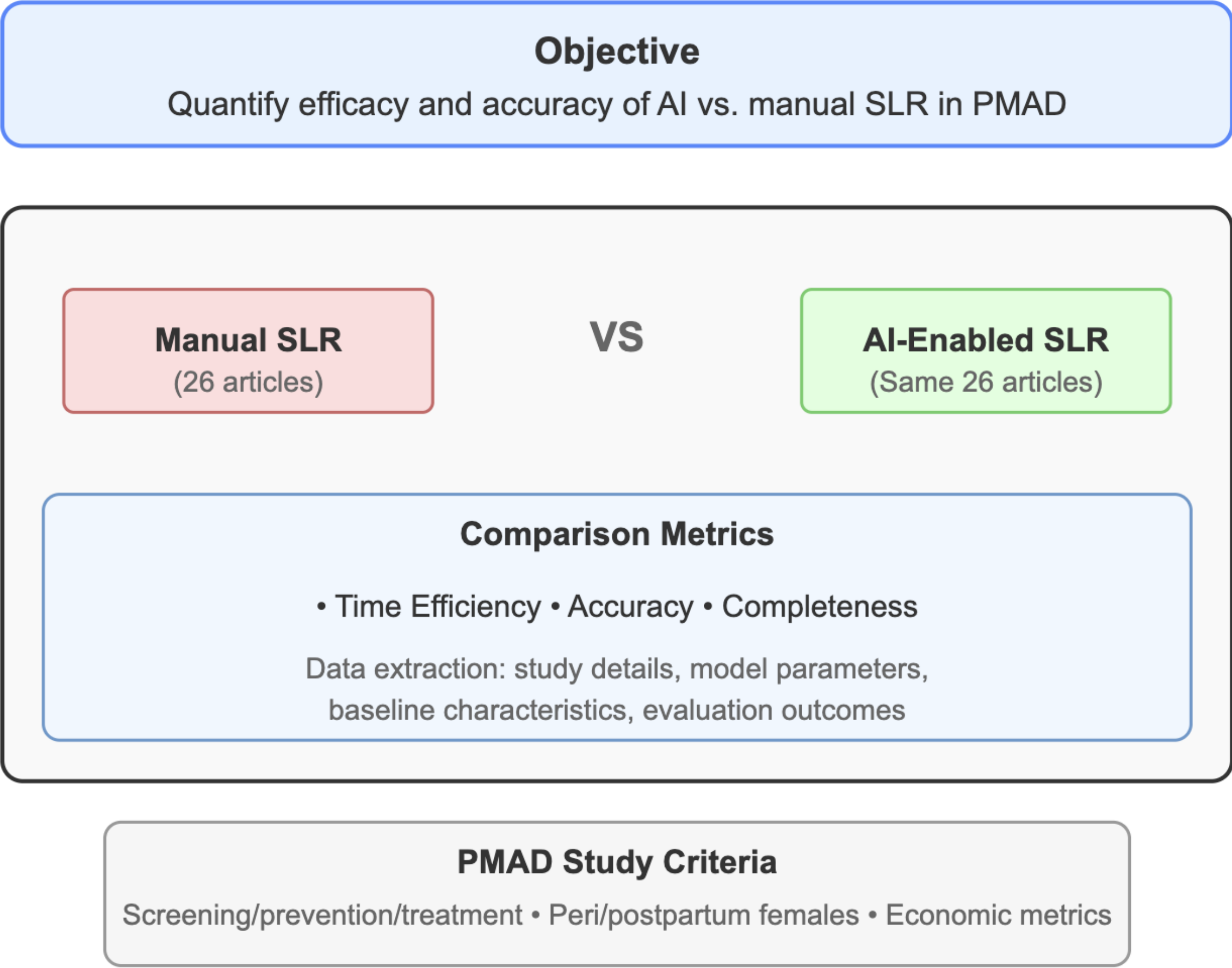
The Quantification of Efficacy and Accuracy for Screening and Data Extraction in Perinatal Mood and Anxiety Disorders Studies: A Comparative Analysis of Traditional vs. GenAI-Powered Approaches

Kyeryoung Lee, PhD^{1*}, Hunki Paek, PhD^{1*}, Emma McNeill, MS², Brian Christman, MS², Sabrina Alam, MPH², Lizheng Shi, PhD³, Xiaoyan Wang, PhD^{1,3}
¹IMO Health, Rosemont, IL, USA, ²University of Mississippi Medical Center, Jackson, MS, USA, ³Tulane University, New Orleans, LA, USA
*Equal Contribution

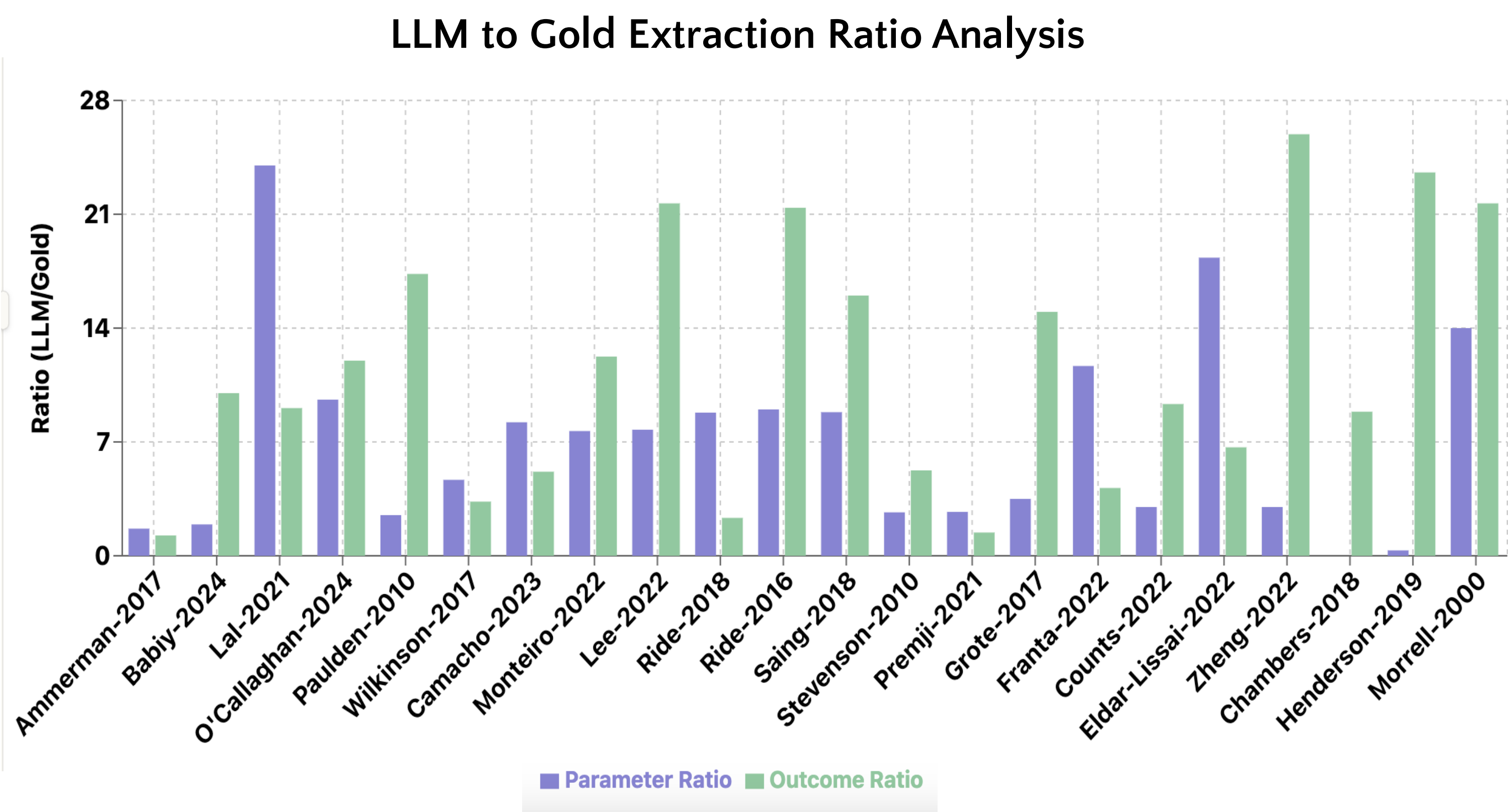


Introduction

AI-Enabled vs. Manual SLR: PMAD Case Study



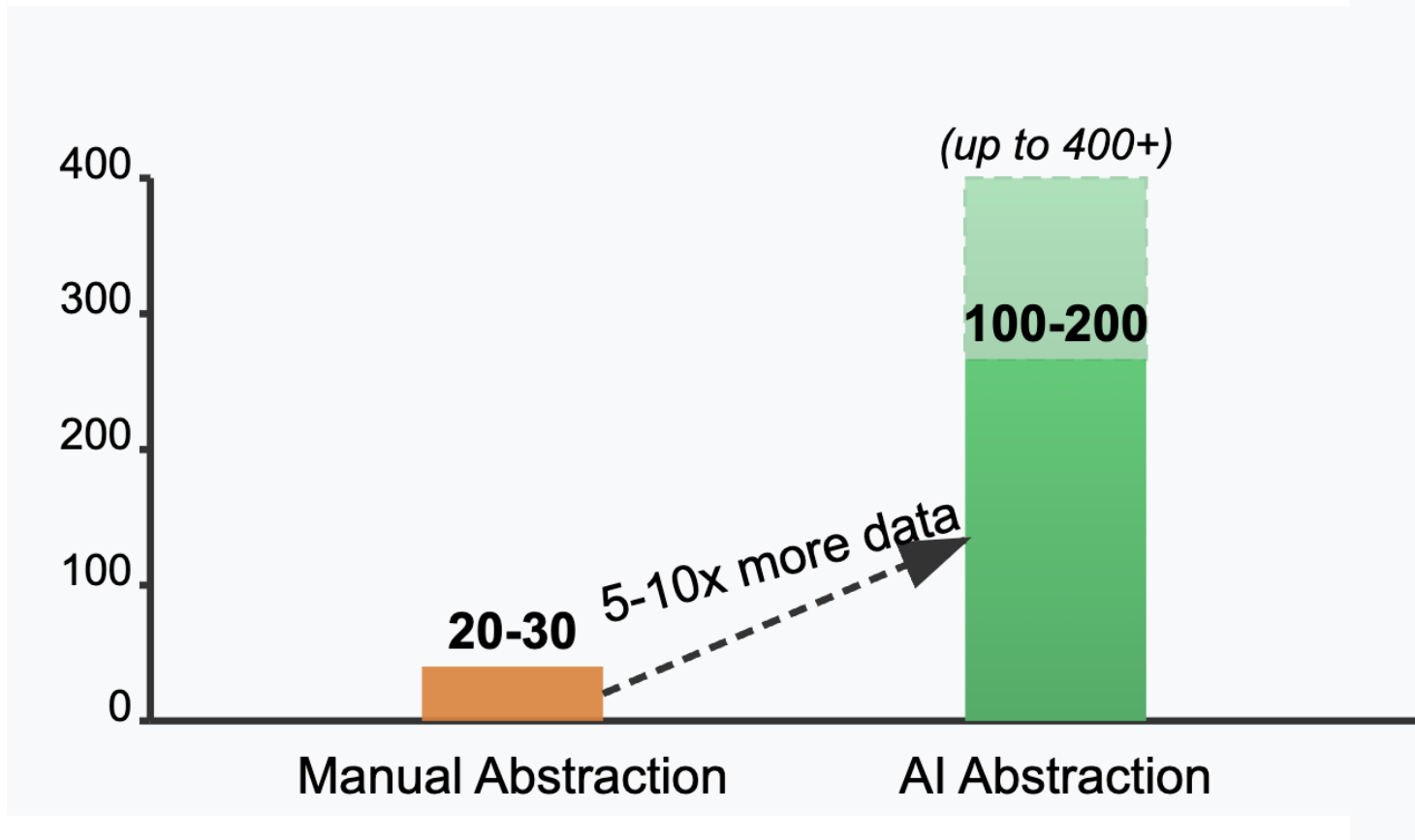
Results



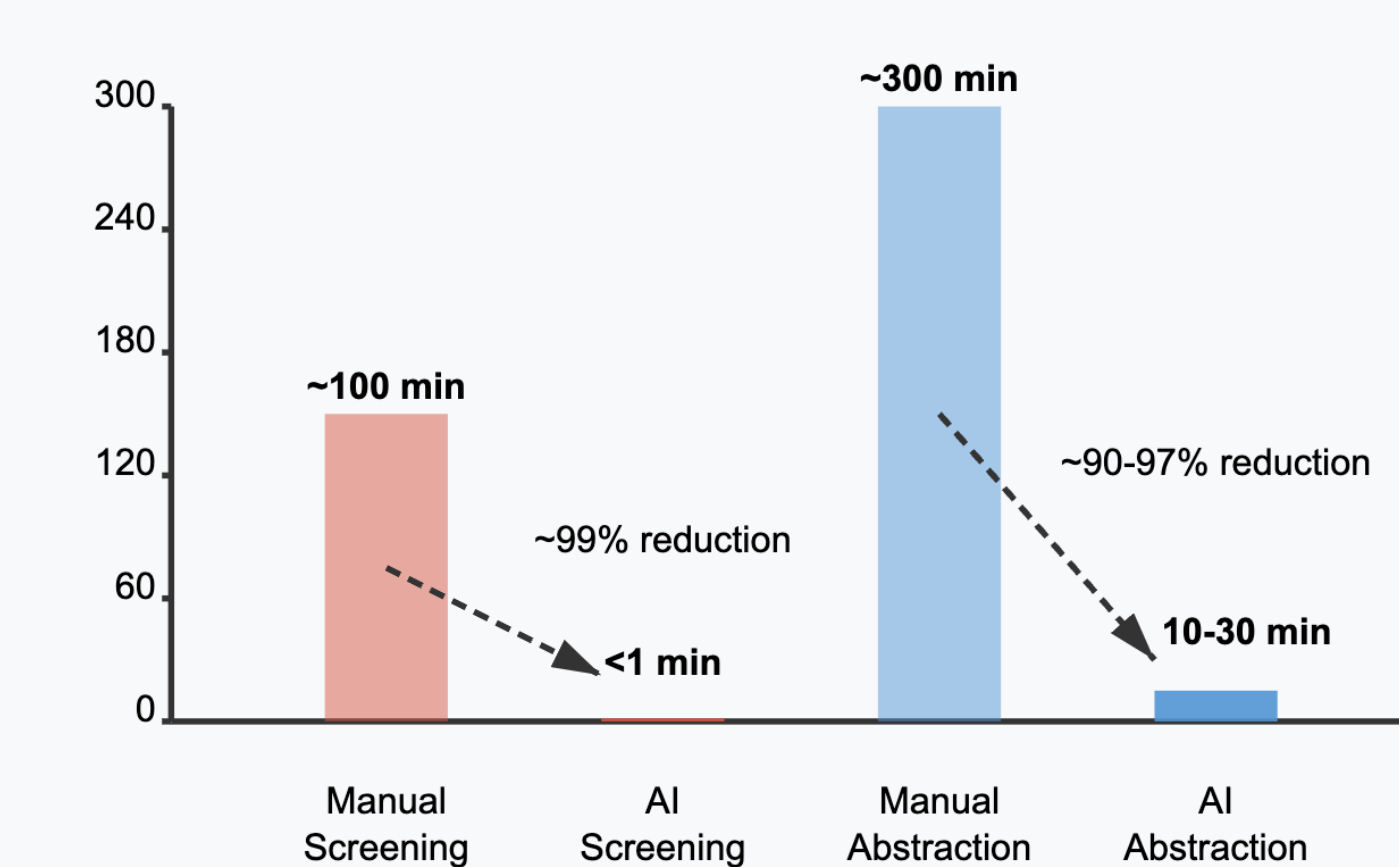
Key Observations:

- LLM extractions exceed Gold standards by 9.2x for parameters and 10.7x for outcomes
- In studies with zero Gold parameters, LLM still extracted significant numbers (e.g., Heslin-2022: 72, Petrou-2006: 76; not shown in this ratio graph)
- Most extreme parameter extraction: Morrell-2009 (7 Gold vs 456 LLM)
- Most extreme outcome extraction: Zheng-2022 (12 Gold vs 311 LLM)

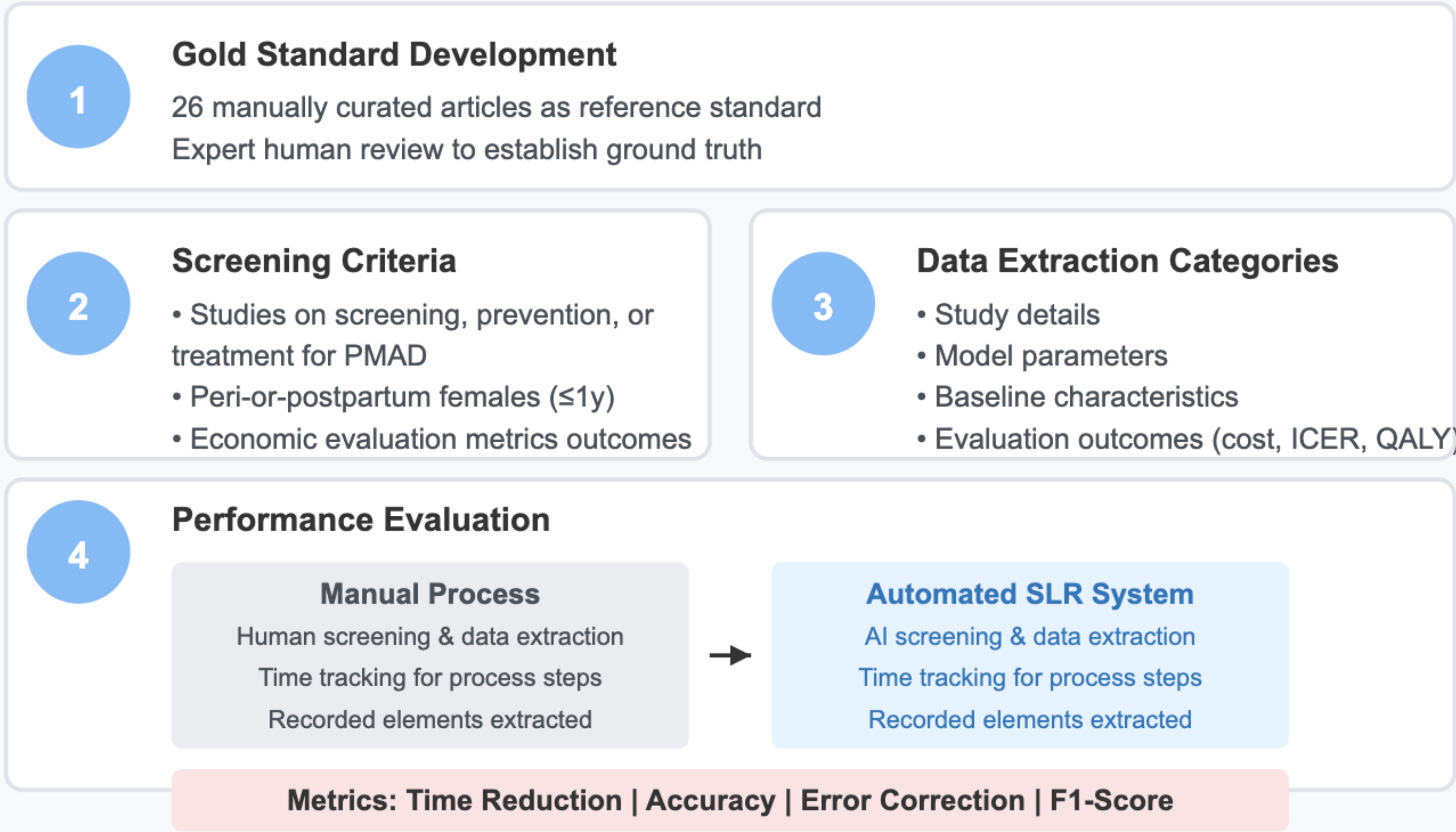
Data Elements Captured per Article



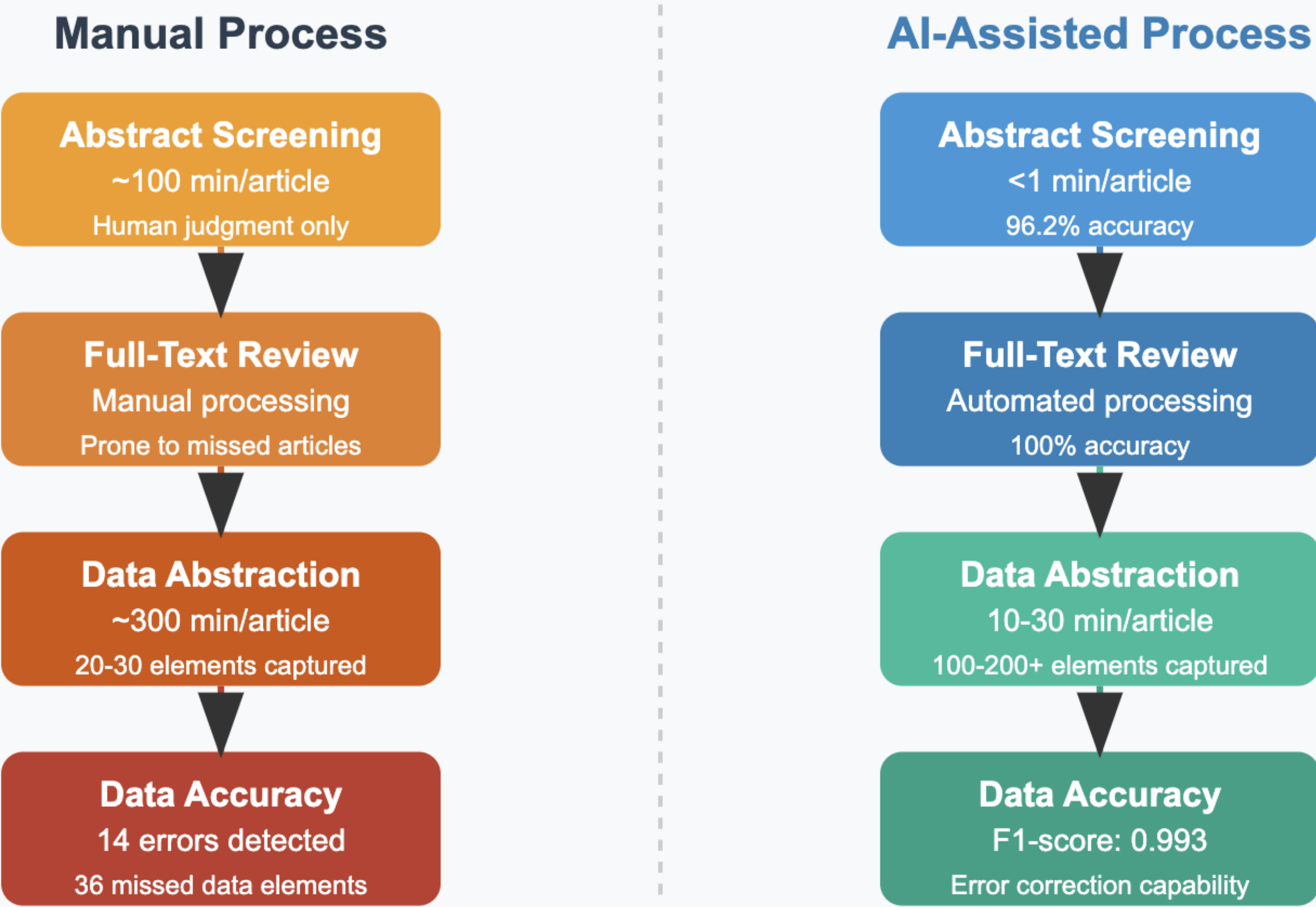
Time Comparison: Manual vs AI



Methodology



AI-Assisted vs. Manual Systematic Review Workflow

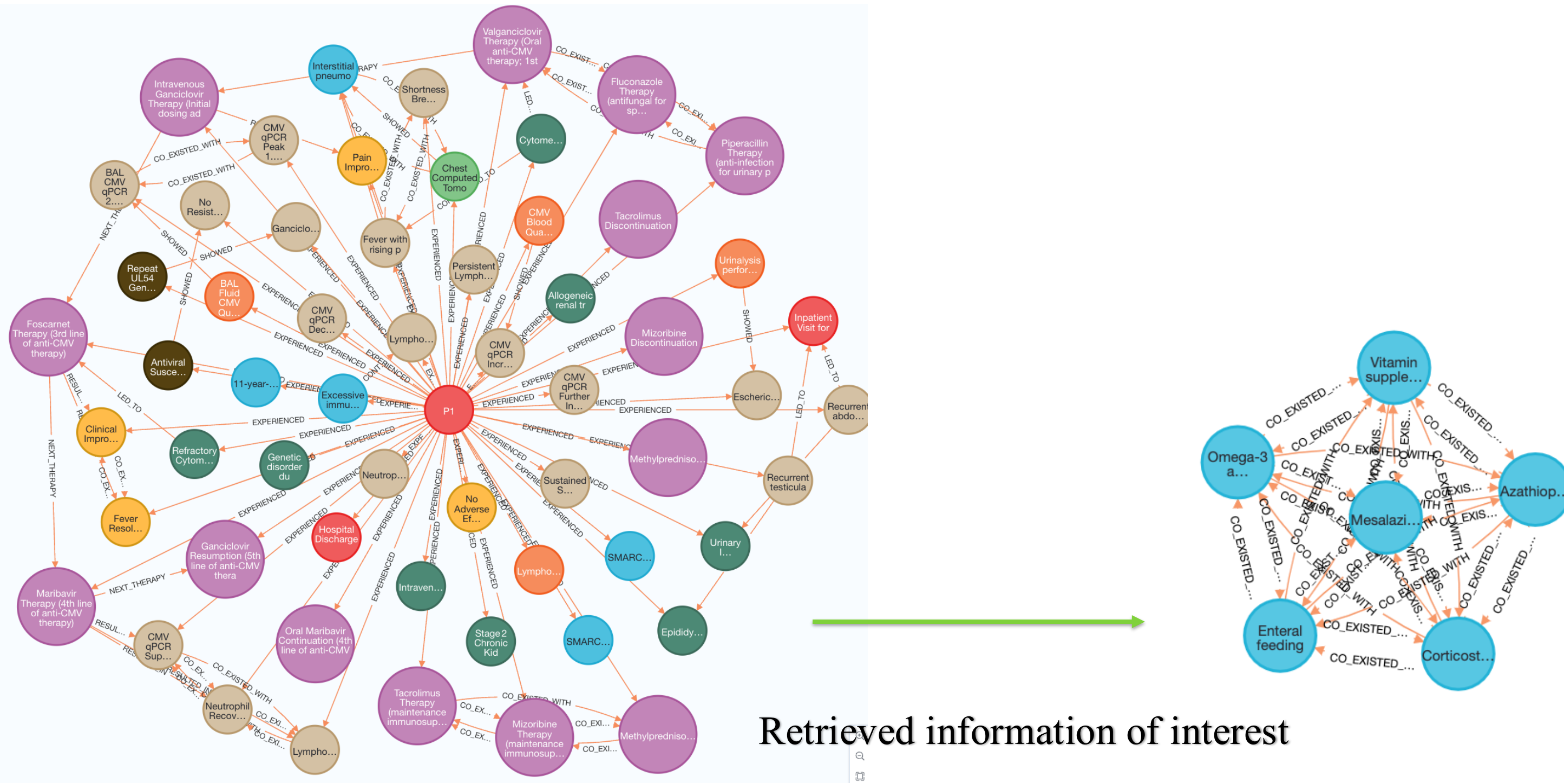


Conclusion

Our AI-system significantly reduced time and error while effectively enhancing accuracy and comprehensiveness in screening and data extraction from full-texts/tables. This approach holds promise to advance SLRs and health economics and outcomes research.

Plan

Construct a knowledge graph using the extracted data for the efficient retrieval and deep analysis



Contact Information

Kyeryoung Lee: klee@imoheal.com
Hunki Paek: hpae@imohealth.com
Xiaoyan Wang: xwang@imohealth.com

