Annals of Emergency Medicine 2025

Ambient Artificial Intelligence Versus Human Scribes in the Emergency Department

Authors: Morey JD, Jones D, Walker L, Lindor R, Schupbach J, Mullan A, Heaton H

Presented by Thomas Hartka, MD, MS, MSDS

AI in Healthcare Elective January 20, 2026

Introduction

Background

Medical documentation have been in increasing and this creates significant challenges for physicians. Human medical scribes have emerged as one solution, reducing the documentation burden by transcribing patient encounters in real-time. However, scribe programs require substantial investment, extensive training infrastructure, and ongoing management resources.

Objective

The goal of this study was to assess if ambient AI technology match the documentation quality of human scribes.

Study Design & Evaluation Framework

01

Study Type

Prospective, observational, single-blind trial

02

Primary Outcome

Note quality assessed using the Physician Documentation Quality Instrument (PDQI-9) by two blinded reviewers

PDQI-9 Assessment Criteria

- **Up-to-date:** Current medical information
- **Accurate:** Error-free and true
- **Thorough:** Includes relevant details
- **Useful:** Aids decision-making
- **Organized:** Logical structure
- **Comprehensible:** Clear and understood
- **Succinct:** Concise, no excess
- **Synthesized:** Integrated narrative
- **Internally Consistent:** No contradictions

Results

710	284	426
Total Encounters	**Human Scribed**	**AI Scribed**
Patient visits analyzed across adult and pediatric populations	123 adult + 161 pediatric encounters	271 adult + 155 pediatric encounters

---

Primary Outcome:

PDQI-9 Quality Scores by Population

Adult Patients

AI Score: 38.22
Human Score: 40.59

Adjusted Rate Ratio: −2.37 (95% CI −5.43 to 0.70)

No statistically significant difference — AI performed comparably to human scribes in adult emergency encounters.

Pediatric Patients

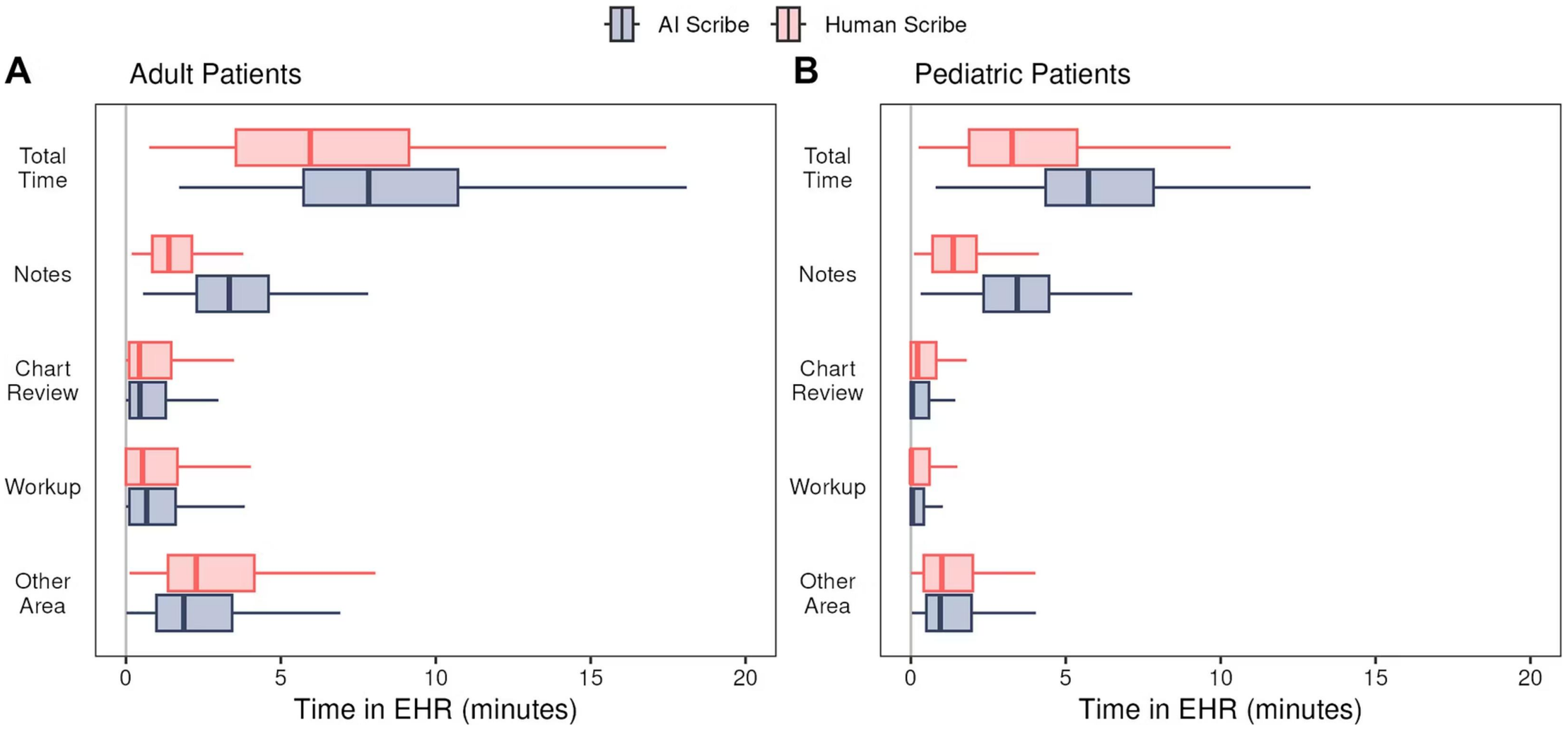
AI Score: 40.36
Human Score: 42.25

Adjusted Rate Ratio: −1.89 (95% CI −3.58 to −0.20)

Statistically significant difference — AI documentation quality was lower in pediatric encounters, suggesting room for improvement.

Secondary outcome

Time spent in electronic medical record



Conclusions

> "Conclusion: In comparison to human scribes, AI scribes were associated with more time spent in the electronic health record notes section, more physician note contribution, and similar to lower quality notes."

Thoughts

Cost Analysis Missing

The study didn't evaluate economic factors.

Learning Curve Effect

Physician proficiency with ambient AI tools will improve over time.

Rapid Technology Evolution

Ambient AI systems are improving rapidly through machine learning.