

Penda Health – OpenAI Studie

Zusammenfassung der Ergebnisse

Klinische Evaluation eines KI-Copiloten in der Primärversorgung in Kenia

März – April 2025

Bericht erstellt: August 2025

Autor: Zusammenfassung basierend auf Originalpublikation und Sekundärquellen

Zusammenfassung der Penda Health – OpenAI Studie

Executive Summary

Die Studie von Penda Health in Zusammenarbeit mit OpenAI stellt eine der ersten realen Anwendungen großer Sprachmodelle (LLMs) in der Primärversorgung dar. Zwischen März und April 2025 wurden in 15 Kliniken in Nairobi insgesamt 39.849 Patient:innenbesuche ausgewertet. Die KI „AI Consult“ wurde als Hintergrundsystem integriert und markierte potenzielle Fehler in der Dokumentation mittels Farbcodes (grün, gelb, rot).

Die wichtigsten Ergebnisse: • Reduktion diagnostischer Fehler um ca. 16 % • Reduktion therapeutischer Fehler um ca. 13–14 % • Deutlicher Rückgang kritischer Fälle („rot“) von ~40 % auf ~20 % • 100 % der befragten Ärzt:innen bewerteten das Tool als hilfreich • Keine Sicherheitsprobleme wurden festgestellt

Die Ergebnisse zeigen, dass KI-basierte klinische Copiloten einen wichtigen Beitrag zur Patientensicherheit leisten können. Die Studie liefert frühe, aber aussagekräftige Evidenz für die Integration solcher Systeme in den Alltag medizinischer Versorgung.

1. Einleitung & Hintergrund

Fehler in der Primärversorgung – sowohl bei der Diagnose als auch bei der Therapie – sind ein global bekanntes Problem. Studien zeigen, dass insbesondere in ressourcenlimitierten Gesundheitssystemen Fehlerraten hoch sind und die Versorgungsqualität beeinträchtigen. Künstliche Intelligenz (KI), insbesondere große Sprachmodelle (LLMs), bieten neue Möglichkeiten, Ärzt:innen in Echtzeit zu unterstützen.

Penda Health betreibt ein Netzwerk von Kliniken in Kenia mit Fokus auf qualitativ hochwertige, erschwingliche Primärversorgung. In Partnerschaft mit OpenAI wurde ein KI-System entwickelt, das als „Sicherheitsnetz“ fungiert und Ärzt:innen bei der Vermeidung von Fehlern unterstützt.

2. Methoden & Implementierung

Die Studie wurde vom 1. März bis 17. April 2025 durchgeführt. Insgesamt wurden 39.849 Patient:innenbesuche erfasst, davon 20.859 mit KI-Unterstützung (Interventionsgruppe) und 18.990 ohne (Kontrollgruppe).

Das System „AI Consult“ war nahtlos in das bestehende elektronische Patientendokumentationssystem (EHR) integriert. Es lief im Hintergrund, ohne den Arzt-Patient-Dialog zu stören, und signalisierte potenzielle Fehler über Farbcodes. Die Einführung wurde durch gezielte Schulungen, Peer-Support und kontinuierliches Monitoring begleitet.

3. Ergebnisse

Die Analyse zeigte eine signifikante Verbesserung der Behandlungsqualität:

- Diagnostische Fehler wurden um 16 % reduziert
- Behandlungsfehler sanken um 13–14 %
- Kritische „Rot“-Markierungen reduzierten sich von ~40 % auf ~20 %
- Befragte Ärzt:innen (n = ca. 67 % Rücklaufquote) bewerteten das Tool einstimmig als hilfreich
- Es wurden keine Sicherheitsprobleme festgestellt

Diese Ergebnisse unterstreichen die Wirksamkeit der KI als Echtzeit-Unterstützung im klinischen Alltag.

4. Diskussion

Die Ergebnisse legen nahe, dass KI-basierte Systeme nicht nur in der Forschung, sondern auch in realen klinischen Settings erfolgreich eingesetzt werden können. Besonders relevant ist der Rückgang kritischer Fehlerfälle, der unmittelbare Auswirkungen auf die Patientensicherheit hat.

Limitationen bestehen in der Durchführung in nur einem Land und über einen begrenzten Zeitraum. Zudem bleibt offen, wie sich die Langzeitnutzung auf ärztliche Entscheidungsprozesse auswirkt. Dennoch liefert die Studie wertvolle Hinweise für die Skalierbarkeit solcher Systeme.

5. Ethische & regulatorische Aspekte

Das Projekt erhielt Genehmigungen von AMREF Ethics Review, dem Gesundheitsministerium Kenias, dem Nairobi County Health Department sowie der nationalen Forschungskommission. Es wurde als Qualitätsverbesserungsprojekt eingestuft, wodurch keine gesonderte Patienteneinwilligung erforderlich war. Patient:innen hatten jedoch das Recht, die Teilnahme

abzulehnen.

6. Schlussfolgerungen & Ausblick

Die Studie liefert erste robuste Evidenz dafür, dass KI-Copiloten die Patientensicherheit signifikant erhöhen können. Mit zunehmender Reife und internationaler Validierung könnten solche Systeme ein zentraler Bestandteil der medizinischen Versorgung werden. Weitere Studien sollten sich auf Langzeiteffekte, internationale Vergleichbarkeit und ökonomische Aspekte konzentrieren.

Literaturverzeichnis

OpenAI & Penda Health. (2025). Clinical impact of AI Consult in primary care in Kenya. Preprint.
https://cdn.openai.com/pdf/a794887b-5a77-4207-bb62-e52c900463f1/penda_paper.pdf

OpenAI. (2025, Juli 22). Pioneering an AI clinical copilot with Penda Health.
<https://telegra.ph/Pioneering-an-AI-clinical-copilot-with-Penda-Health-07-22>

Northwestern University. (2025, August 11). OpenAI + Penda Health study overview.
<https://sites.northwestern.edu/pccmfellowship/2025/08/11/openaipenda-study>

World Health Organization. (2018). Technical series on safer primary care. Geneva: WHO.

Topol, E. (2019). Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again.
New York: Basic Books.