

AI in Healthcare: Enhancement, not Replacement

Daniel Pappalardo

Introduction

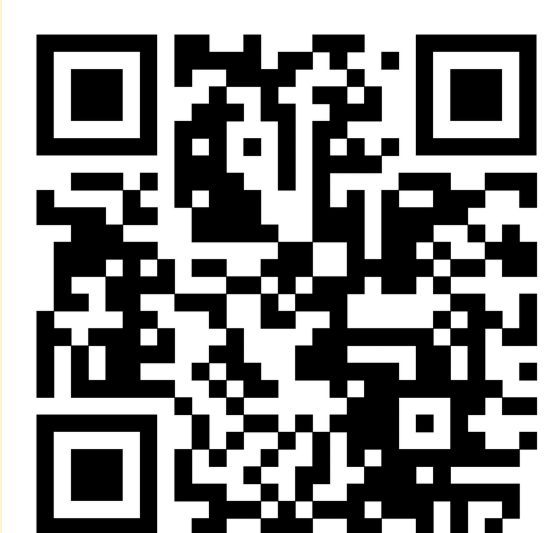
- With the introduction and growth of AI, we have seen it work its way from a sci-fi fantasy to the real thing.
- Self driving cars, AI call centers and autonomous factories are already existing, but what about healthcare?
- Research suggests that AI will not fully replace, but rather augment and enhance the roles of human physicians.

Comparison 1st specialty vs 2nd specialty	Difference in accuracy (%) (2nd specialty - 1st specialty)	Pooled estimate [95%CI]
General medicine vs Urology	38.1	[23.5, 52.6]
General medicine vs Dermatology	30.5	[16.2, 44.7]
General medicine vs Rheumatology	13.0	[-7.2, 33.2]
General medicine vs Orthopedic surgery	11.8	[-6.7, 30.3]
General medicine vs Gynecology	10.6	[-5.9, 27.2]
General medicine vs Psychiatry	3.5	[-16.3, 23.2]
General medicine vs Neurology	0.5	[-13.1, 14.1]
General medicine vs Radiology	-0.1	[-18.5, 18.3]
General medicine vs Endocrinology	-1.8	[-20.7, 17.1]
General medicine vs Ophthalmology	-2.1	[-27.0, 22.9]
General medicine vs Cardiology	-3.1	[-25.4, 19.2]
General medicine vs Pediatrics	-5.6	[-32.6, 21.4]
General medicine vs Plastic surgery	-10.1	[-25.5, 5.3]
General medicine vs Emergency medicine	-10.9	[-29.8, 8.0]
General medicine vs Otolaryngology	-13.6	[-31.4, 4.2]
General medicine vs Gastroenterology	-16.5	[-35.4, 2.3]

Table from Guo & Chen 2025

Methods & References

- Research methods included literature review of peer-reviewed healthcare AI-studies
- Sources: Google Scholar, PubMed, Microsoft Copilot
- Search Terms: AI in healthcare, AI in medicine, physicians and AI, future of AI in healthcare



Current Applications & Performance

- Shows great efficiency in administrative work (Spatharou et al 2020)
- Has been shown to match physicians in certain diagnostic tasks. (Takita et al. 2025)
- Performance declines in more complex or ambiguous cases (Hayat et al. 2025)
- Shown to be helpful as a decision support tool, not an independent agent.

Human Factors & Workforce Impact

- Patients rated human physicians higher than AI in trust and satisfaction (Guo & Chen, 2025)
- Providing explanations improved AI perception, but did not eliminate preference for humans.
- Trust in AI depends on transparency and consistency across cases (Steerling et al. 2023)
- Full AI adoption remains limited due to ethical and regulatory issues worldwide.

Limitations & Ethics

- AI systems can produce biased outputs if the data they are trained on is biased, this is known as algorithmic bias
- Many AI models lack explainability, making decisions difficult to interpret.
- Lack of transparency creates challenges for clinical accountability and decision justification.
- AI systems still require human guidance/oversight to ensure safe implementation. (Elechi et al. 2025)

Comparison between model and physician		
All models vs Overall physician	9.9	[-2.3, 22.0]
All models vs Overall Physician (Non-expert)	0.6	[-14.5, 15.7]
All models vs Overall Physician (Expert)	15.8	[4.4, 27.1]
Model (GPT-3.5) vs Overall physician	12.5	[-2.0, 27.0]
Model (GPT-3.5) vs Physician (Non-expert)	3.5	[-13.3, 20.3]
Model (GPT-3.5) vs Physician (Expert)	18.5	[4.7, 32.2]
Model (GPT-4) vs Overall physician	7.1	[-6.0, 20.2]
Model (GPT-4) vs Physician (Non-expert)	-2.1	[-19.2, 15.1]
Model (GPT-4) vs Physician (Expert)	12.9	[0.5, 25.3]
Model (Gemini 1.5 Pro) vs Overall physician	6.9	[-13.5, 27.3]
Model (Gemini 1.5 Pro) vs Physician (Non-expert)	-2.1	[-26.9, 22.8]
Model (Gemini 1.5 Pro) vs Physician (Expert)	12.9	[-6.5, 32.3]
Model (Claude 3 Sonnet) vs Overall physician	3.7	[-21.3, 28.7]
Model (Claude 3 Sonnet) vs Physician (Non-expert)	-5.5	[-34.0, 23.0]
Model (Claude 3 Sonnet) vs Physician (Expert)	9.4	[-13.8, 32.7]
Model (Claude 3 Opus) vs Overall physician	-4.0	[-24.6, 16.7]
Model (Claude 3 Opus) vs Physician (Non-expert)	-12.9	[-37.7, 11.8]
Model (Claude 3 Opus) vs Physician (Expert)	2.0	[-17.2, 21.2]

Table from Guo & Chen 2025

Conclusion

- AI has been shown to perform at physician level in controlled environments, but suffers drops in performance in more complex, real world cases.
- Patients and clinicians have shown greater trust in human providers as opposed to more accurate AI.
- Barriers to implementation and ethical concerns limit independent AI use in the healthcare setting.
- Evidence across multiple studies shows AI's ability to enhance physician performance, without replacing human clinical oversight.

Empowering Healthcare:

AI as a Partner, Not a Replacement



Graphic created by Microsoft Copilot