

配信先：大阪科学・大学記者クラブ、文部科学記者会、科学記者会

2025年4月2日

大阪公立大学

生成 AI と医師の診断能力を比較 非専門医とは同等の精度であると判明

<ポイント>

- ◇医療分野における生成 AI の診断能力に関する研究論文 83 報についてメタ解析^{※1}を実施。
- ◇生成 AI と非専門医の診断能力を統計学的に比較すると、有意差がなかった。
- ◇医学教育や非専門医の診断支援などにおける生成 AI の活用に期待。

<概要>

近年、ChatGPT など生成 AI の活用が医療分野でも注目されており、これまでに生成 AI の診断能力に関する研究論文が多数発表されています。しかし、それぞれ評価基準が違うため、実際の医療現場で利用できる診断能力がどの程度あるのか、医師との比較においてどのような特徴があるのかなど、包括的な分析が必要でした。

大阪公立大学大学院医学研究科 放射線診断学・IVR 学の田北 大昂講師、人工知能学の植田 大樹准教授らの研究グループは、医療に関する生成 AI の診断能力について 2018 年 6 月から 2024 年 6 月までに発表された 83 報の研究論文を用いてメタ解析を実施しました。その結果、専門医は生成 AI よりも診断精度が 15.8% 高く、有意差がありました。しかし、生成 AI の平均診断精度は 52.1% で、医師全体と有意差がなく、非専門医との差も僅かでした。また、生成 AI の最新モデルについては、有意差はないものの非専門医と同等以上の診断精度を示す場合もありました。本研究により、生成 AI は専門医の完全な代替とはなりません。医学教育や非専門医の診断支援などでの活用が期待できます。

本研究成果は、2025 年 3 月 22 日に国際学術誌「npj Digital Medicine」にオンライン公開されました。

<掲載誌情報>

【発表雑誌】npj Digital Medicine

【論文名】A systematic review and meta-analysis of diagnostic performance comparison between generative AI and physicians

【著者】Hiroataka Takita, Daijiro Kabata, Shannon L Walston, Hiroyuki Tatekawa, Kenichi Saito, Yasushi Tsujimoto Yukio Miki, Daiju Ueda

【掲載 URL】<https://doi.org/10.1038/s41746-025-01543-z>

最近、世界中で急速に生成 AI が使われるようになり、医療現場にも AI の波が押し寄せています。便利ではあるものの、現状では生成 AI にはさまざまな問題がありますが、上手く使うことで医療の質の向上に貢献できると考えています。今後も生成 AI を含め人工知能の研究・開発を通じて医療の質の向上を目指します。



田北 大昂講師

<研究の背景>

近年、ChatGPTに代表される生成AIが医療分野でも注目を集めています。しかし、これらのAIが実際の医療現場でどの程度の診断能力を発揮するのか、また医師との比較においてどのような特徴があるのかについては、包括的な分析が行われていませんでした。そこで本研究では、生成AIの医療に関する診断能力について系統的レビュー^{※2}とメタ解析を実施しました。

<研究の内容>

本研究では、2018年6月から2024年6月までに発表された83報の研究論文を分析対象としました。その中で数多く評価されていたのはGPT-4（54件）とGPT-3.5（40件）で、その他にはGPT-4o、Claude 3、Gemini 1.5 pro、Llama 3 70Bなどの最新モデルも含まれていました。評価された医療分野は、一般内科が27件と最も多く、次いで放射線科が16件、眼科が11件、救急医療が8件など、多岐にわたりました。分析の結果、生成AIの平均診断精度は52.1%であり、医師全体と有意差がありませんでした（医師全体の方が9.9%高い）。特に非専門医と生成AIの診断精度の差は僅かで、両者の有意差はありませんでした（非専門医の方が0.6%高い）。一方で、専門医は生成AIよりも診断精度が15.8%高く、有意差がありました。ただし、GPT-4、GPT-4o、Claude 3、Gemini 1.5 pro、Llama 3 70Bなどの最新モデルは、非専門医と同等～それ以上の診断精度を示しました。

<期待される効果・今後の展開>

本研究により、生成AIは専門医の完全な代替とはなりません。医学教育での活用や非専門医の診断支援、医療資源の限られた地域での診断補助などでの活用が期待できます。今後は、より複雑な臨床シナリオでの評価や実際の医療記録を用いた性能評価、AIの判断根拠の透明性向上、多様な患者群での検証など、更なる研究が必要です。

<用語解説>

- ※1 メタ解析：Meta-analysis。複数の研究結果を統計的な手法を使って一つにまとめ、より確かな結論を導き出す分析方法。個々の小規模な研究では見えにくい傾向や差異を、より大きな視点から捉えることを可能にする。
- ※2 系統的レビュー：Systematic Review。特定のテーマについて、科学的な手法を用いて過去の研究論文を網羅的に集め、総合的に分析する研究手法。偏りのない結論を導き出すために、研究の選び方や分析方法を予め決めておく。

【研究内容に関する問い合わせ先】

大阪公立大学大学院医学研究科
田北 大昂（たきた ひろたか）
TEL：06-6645-3831
E-mail：z21227o@omu.ac.jp

【報道に関する問い合わせ先】

大阪公立大学 広報課
担当：谷
TEL：06-6967-1834
E-mail：koho-list@ml.omu.ac.jp