

Accuracy of an Artificial Intelligence Chatbot's Interpretation of Clinical Ophthalmic Images.

臨床眼科画像における人工知能チャットボットの解釈の正確性

Mihalache A, Huang RS, Popovic MM, Patil NS, Pandya BU, Shor R, Pereira A, Kwok JM, Yan P, Wong DT, Kertes PJ, Muni RH.

JAMA Ophthalmol. 2024 Apr 1;142(4):321-326. doi: 10.1001/jamaophthalmol.2024.0017.

PMID: 38421670

すでに AI チャットボットとして広まっている OpenAI 社の ChatGPT を使ったことがある方も多いと思います。私もメールの日本語の校正や英文の和訳などに利用しており、その優秀さを実感しています。一方で、医学的な質問に対しては間違った情報をそれらしく回答することがあるため、使い方には注意が必要だと感じています。

本報告では、トロント大学眼科が Web で公開している医学教育トレーニングプラットフォーム「OCTcases」の教材を用い、ChatGPT に質問した際の正答率について研究した結果を示しています。OCTcases は、トロント大学眼科が各症例についての OCT、眼底などの画像データを提示し、4 つの選択肢から回答を選ばせる形式で、分野としては網膜、神経眼科、ぶどう膜炎、緑内障、眼腫瘍学、および小児眼科の症例が対象となっています。

結果ですが、全体の正答率は 70% でした。そのうち画像を伴わない質問の正答率は 82%、画像を伴う質問の正答率は 65% とのことです。画像認識に関してはまだ改善の余地があるようです。しかし、画像単独でのディープラーニングシステムを用いた AI 診断では 90% 以上の感度を示す報告が多数あり、今後はそのようなシステムとの連携により正答率が向上することが期待されます。

ただし、現時点では AI チャットボットは人間の専門知識を代替するものではなく、その出力に過度な信頼を寄せるべきではないことも強調する必要があると考えます。

(担当者：北海道大学 南場研一)