

## 生成AIによる救急外来受診の推奨に関する妥当性研究—生成AIの回答 に対する専門家と非医療従事者の解釈の差が明らかに—

日本救急医学会「救急医療における先端テクノロジー活用特別委員会」による報告

日本救急医学会 救急医療における先端テクノロジー活用特別委員会は、*Acute Medicine & Surgery*誌に研究論文“Medical Validity and Layperson Interpretation of Emergency Visit Recommendations by the GPT model: A Cross-Sectional Study”を公表しました。本研究では、大規模言語モデル(LLM)を用いて急病患者的の救急外来受診の必要性に関する推奨を生成し、それを救急医療の専門家が評価を行い、非医療従事者が解釈をしました。その結果、医学知識を持った専門家はLLMが救急外来受診の判断を高精度でアドバイスしていると評価した一方で、非医療従事者は同じ回答を見ても専門家と異なる解釈をする傾向があることが明らかになりました。これにより、AIが生成する医療アドバイスに関する理解や解釈において、専門家と非医療従事者の間に大きな隔たりがあることが示唆されました(図1)。

### 背景

日本では年間620万人以上が救急搬送され、そのうち約半数が軽症例とされています。適切な受診を促す取り組みが進められているものの、救急外来の混雑と医療リソースの不足は依然として社会課題となっています。

近年登場したLLMは自然言語での質疑応答が可能であり、医療に関する助言への応用が期待されています。しかしLLMが回答した救急外来受診要否の判断の正確性や、回答内容に対する利用者(非医療従事者)の理解度については検証されていませんでした。

### 研究概要

本研究では、総務省消防庁の『緊急度判定プロトコルVer.3救急受診ガイド(家庭自己判断)』をもとに、合計466のシナリオ(緊急度が高い314件、緊急度が低い152件)をLLM(GPT-3.5)に「救急受診の必要性」を尋ねました。その回答を、救急科専門医を含む7名の医療従事者が医学的観点から評価したうえで、非医療従事者157名がどのように受け取るかをアンケート方式で検証しました。

その結果、専門家の評価では、LLMの回答は緊急度の高いケースの96.5%において、合理的な理由を示した上で「救急受診が必要」と適切に判断されていると評価されました。また、軽症例についても88.8%のケースで「救急受診は不要」と適切に判断されていました。

一方、一般の人々(非医療従事者)の評価では、同じLLMの回答を見ても、緊急度の高いケースでは43.0%しか「救急受診が必要」と解釈されず、軽症例についても32.2%しか「救急受診は不要」と正しく理解されないという結果でした。

### 考察

LLMの回答は、医学知識を持った専門家は、救急外来受診の必要性を高精度で正しく判定していると判断した一方、非医療従事者にはその意図を正しく解釈できないケースが多く見られました。このことから、AIが提供する情報を活用する際には、医療従事者によるサポートや平易な言葉への置き換えなど、誤解を防ぐ工夫が求められることが示唆されました。特に緊急性の判断を要する場面では、AIのみに頼るのではなく、専門家との連携や追加の確認が重要です。本研究は今後、大規模言語モデルを救急医療に安全かつ有効に導入するための基礎データとなることが期待されます。

## 研究の制限

本研究は、GPT-3.5のモデルを使用した研究であり、最新のモデルでは異なる結果となる可能性はある。

## 論文情報

Chie Tanaka, Takahiro Kinoshita, Yohei Okada, Kasumi Satoh, Yosuke Homma, Kensuke Suzuki, Shoji Yokobori, Jun Oda, Yasuhiro Otomo, Takashi Tagami.

Medical Validity and Layperson Interpretation of Emergency Visit Recommendations by the GPT model: A Cross-Sectional Study Short Title: Emergency Visit Guidance by GPT model.

*Acute Med Surg.* 2025; 10:e913.

URL <http://dx.doi.org/10.1002/ams2.70042>

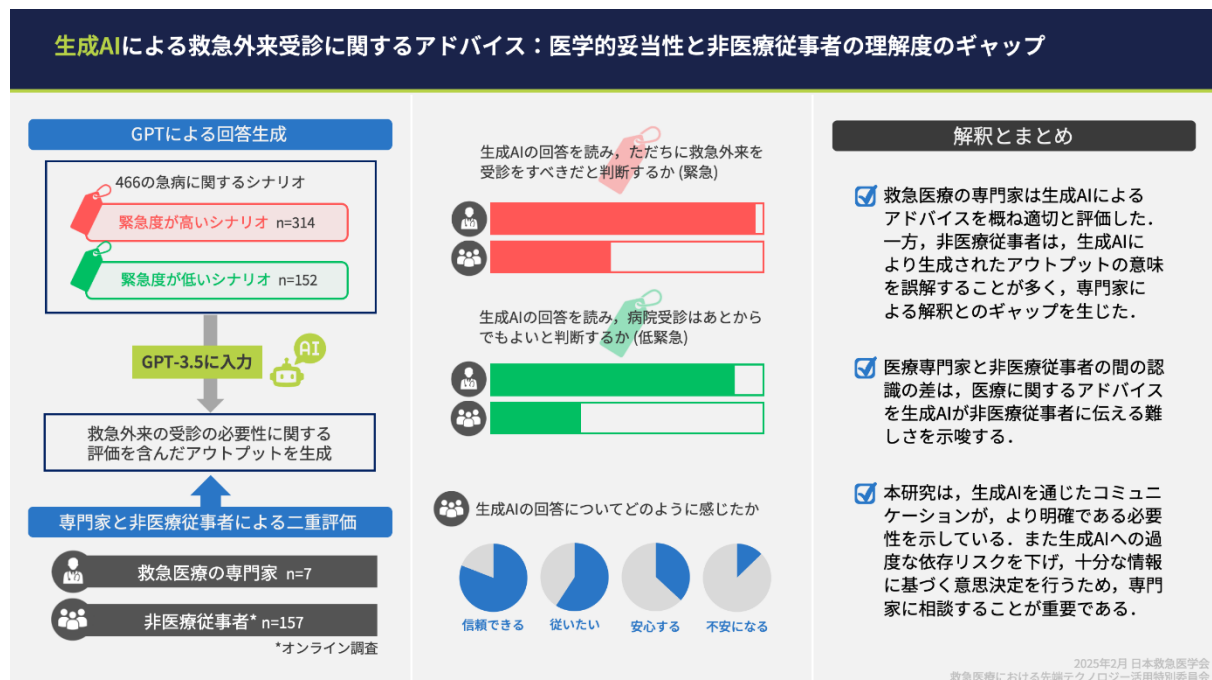


図1. 研究論文“Medical Validity and Layperson Interpretation of Emergency Visit Recommendations by the GPT model: A Cross-Sectional Study”の図解要約(和訳)