＋ 医知創造ラボ｜てんかん×AI ＃5

ChatGPTは、てんかんを診断できるのか？

生成AI・LLMの光と影【2026年版】

深掘り⑤　脳神経内科医が解説　所要 約10分

今日の問い TODAY'S POINT

AIに相談すれば、診断してくれるのか？

結論 教育・下書き・情報抽出には光。
自律診断と文献引用には影。専門家の監督下でのみ補助になります。

LLMは「次の語を予測する」モデル

大規模言語モデル（LLM）は、膨大な文章から「次に来そうな語」を確率で選ぶことで、自然な文を生成します。ChatGPTもこの仕組みです。

📋 ここが出発点

LLMは**事実を引くデータベースではありません。**「もっともらしさ」を生成する道具――この前提が、光と影の両方を説明します。

LIGHT — PATIENT EDUCATION

68.4%

ChatGPTがてんかんの一般質問に正確かつ包括的に回答（再現性82.3%）

Wu 2024 Epilepsy Behav（ILAEなどの378問・専門医評価）

公式ガイドと同等以上の説明も

韓国てんかん学会の患者ガイドに基づく57問で、GPT-4は40問で「十分な教育的価値」、完全な誤答はゼロでした。

> 📋 含意
>
> 患者さんへの**説明の下書きや、よくある質問への一次対応**には実用域。医療者の説明負担を軽くする余地があります。

光② BOARD EXAM

06

専門医試験に合格する知識

LLMはてんかん専門医の模擬試験で、合格相当の成績を示しました（JAMA Neurology）。知識の想起は得意分野です。

📋 ただし

試験に受かること＝臨床で診断できること、ではありません。知識と、目の前の患者への判断は別物です。

LIGHT — INFORMATION EXTRACTION

0.90 F1

診療documentから抗発作薬を抽出(てんかん型0.80・発作型0.76)

Fang 2025 Epilepsia(King's College・280レター・Llama)

光③ RESEARCH SUPPORT

08

カルテの山を読み解く力

ノイズの多いカルテから治療効果や症状を統合する模擬研究では、AIの解析が人間の解析と3%未満の差でした。

📋 得意なこと

「診断する」のではなく、**非構造化テキストから情報を抽出・整理する**——ここはLLMの明確な光です。

生の診断は、専門医に劣る

患者の語りから鑑別	正答率(balanced)
GPT-4(ゼロショット)	57%
GPT-4(1例提示)	64%(頭打ち)
経験ある神経内科医	71%

てんかん発作 vs 心因性発作の鑑別。専門医が全員正答した易しい例ではGPT-4も81%(Ford 2025)

「将来どうなる」には弱い

同じ研究でも、予後に関する質問で包括的だった回答は46.8%にとどまりました。

⚠ 著者の結論

「医療ガイダンスとしての直接利用は推奨しない。現時点の主用途は患者教育」。誤答を鵜呑みにすると危険な助言になりえます。

新しいモデルほど賢い？

✕ 最新・高度なモデルなら、必ず性能も上がる

◯ 小児診断では、高度な言語モデルが単純な古典モデルを上回らなかった（課題次第）

SHADOW — REFERENCE HALLUCINATION

20%

ChatGPT生成の参考文献で真正だったのは2割（実在62%・捏造31%）

Suppadungsuk 2023 J Clin Med（腎臓病領域・610文献を検証）

最新モデルでも引用は危うい

検証	引用・文献の問題
皮膚病理	引用の 27%が作話 、有用なのは24.5%
GPT-5レビュー	正確な引用は 約3分の1 のみ
事実そのもの	本文の事実精度は 比較的高い

「文章は正しそうでも、出典は捏造」——これがLLM最大の落とし穴（McCrary 2024・Ok 2026）

「もっともらしい」を作るからこそ

LLMは存在する文献を検索しているのではなく、「それらしい著者・誌名・年」を確率で生成しています。だから実在しない引用が、堂々と出てきます。

> 📋 だから対策は
>
> 引用は**必ず一次資料（PubMed等）で実在確認**。出典を生成させて鵜呑みにしない。これが必須の運用です。

専門知識で補強して使う

汎用ChatGPT単独ではなく、専門家の知識基盤（オントロジー）とEEGを組み合わせたカスタムGPTでは、発作の局在推定が93.8%まで向上しました。

📋 現実的な姿

LLMは**専門家の監督下で使う「下書き・たたき台」**。単独の判断者ではなく、人が検証する前提の補助ツールです。

LLM活用の現在地

光

教育と情報抽出

患者教育、診療documentからの情報抽出、説明文のたたき台づくりは実用域。

影

診断と文献引用

生の診断は専門医に劣り、参考文献はしばしば捏造。鵜呑みは危険。

原則

監督下で補助

人が必ず検証する前提で。専門知識で補強したカスタム運用が現実的。

＋ 医知創造ラボ

LLMは、賢い助手。
でも、判断するのは人。

本動画は教育目的の解説です。診療判断は最新の規制・GLをご確認ください。

🔔 チャンネル登録　👍 高評価

次回：「てんかん外科×AIと、薬剤反応予測」