

## ChatGPT, BingAI に医薬品情報を質問したデータの考察

塚口 真穂登† 五十嵐 洋平† 三澤 昭彦† 外山 聡†

新潟大学医歯学総合病院 薬剤部†

## 1 はじめに

筆頭筆者は、大学病院の薬剤部医薬品情報管理室に勤務し、医薬品に関する質問に回答することを主とした業務を行っている。我々は、医薬品情報に関する質問を AI スピーカーである Google Home に対する正答率を第 81<sup>[1]</sup>・83 回情報処理学会全国大会<sup>[2]</sup>にて発表した。今回、文章構造がより複雑になっても正しく解析し、正答を得られる新しい技術である ChatGPT-3.5, Bing AI 回ら正答が得られるようになれば、医薬品情報管理業務の軽減に繋がると考えた。今回、当院で実際にあった医薬品に関する問い合わせに対する ChatGPT-3.5, Bing AI からの回答について、正答率を調べたので報告する。

## 2 関連研究

医師<sup>[3]</sup>, 看護師<sup>[4]</sup>, 薬剤師<sup>[5]</sup>の国家試験, 医薬品情報(ハイデルベルグ大学病院薬剤部(ドイツ))<sup>[6]</sup>の ChatGPT での正答率を調べた報告がある。

## 3 実験手法

ChatGPT-3.5, Bing AI に対し, 新潟大学医歯学総合病院薬剤部医薬品情報管理室に記録された医薬品に関する問い合わせ(以下, 医薬品情報記録)(n=585)を質問し, 回答の正誤, BingAI では正答の際の参考情報を記録した。

## 4 結果と考察

ChatGPT-3.5 の正答率は 31.6%(185/585), BingAI では 61.2%(355/585)であった。

ChatGPT-3.5 に医薬品情報を質問した際, 医薬品銘柄名を入力すると, 医薬品銘柄名に対応する一般名を回答するが, 14.7%(86/585)において, 一般名を誤答した。

図 1 で Google Home における医薬品情報の正答率(2018 年 8 月, 2020 年 4 月), ChatGPT での医師・看護師・薬剤師国家試験の正答率, ChatGPT での医薬品情報の正答率(ハイデルベルグ大学・新潟大学), BingAI での医薬品情報正答率

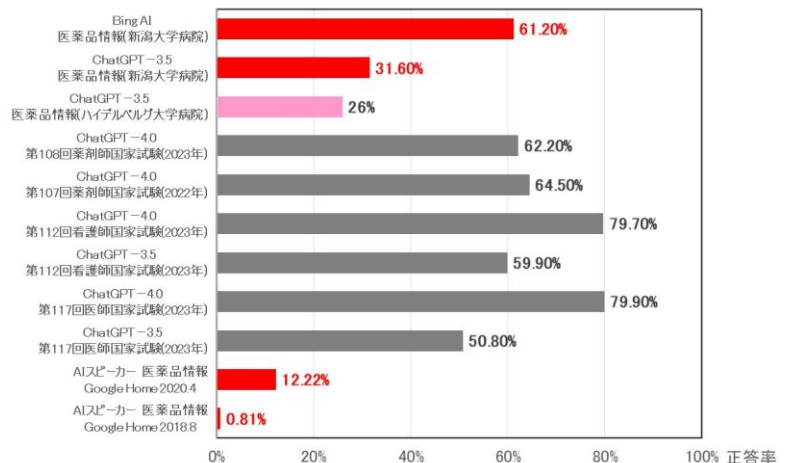


図 1 Google Home に医薬品情報, ChatGpt BingAI に医師・看護師・薬剤師国家試験, 医薬品情報を質問した時の正答率<sup>[1-6]</sup>

(新潟大学病院)を比較した。

2018 年当時, 新潟大学病院の医薬品情報記録では, 「**医薬品名〇〇〇の代替薬は**」のように, (代替薬の網羅的なデータは存在しないので)複数の医薬品の情報を組み合わせる必要がある場合, 正答率が 0.81%と低かったが, 2022 年に生成 AI の登場によって, 品詞の組み合わせが複雑になっても, 正答率が ChatGPT-3.5 では 31.6%, BingAI では 61.2%と向上したことが判明した。

ChatGPT では, 医師<sup>[3]</sup>, 看護師<sup>[4]</sup>, 薬剤師<sup>[5]</sup>の国家試験, 医薬品情報(ハイデルベルグ大学病院薬剤部(ドイツ))<sup>[6]</sup>の正答率を調べた報告がある。

医師・看護師国家試験では ChatGPT-3.5 での正答率は約 60%, ChatGPT-4.0 では約 80%であったが, 薬剤師国家試験では 4.0 でも 60%に留まっていた。薬剤師国家試験では, 臨床現場に近い臨床薬学分野の正答率が下がることが報告されている。さらに, より薬剤師現場に近い医薬品情報の正答率は日本でもドイツでも 30%の正答率に留まっている。

ChatGPT-3.5 における医薬品情報の正答率の低さの原因に, 医薬品名を示すものとして, 医薬品銘柄名(いわゆる商品名)と, 医薬品の名称を指す一般名を正しく対応付けできない場合があることが判明した(誤答率 14.7%(86/585))。ChatGPT に対して医薬品情報を質問する際に対して回答を得る際に, 医薬品銘柄名⇔一般名が正しく変換されないことに大きな課題がある。

Consideration of data on question of drug information on ChatGPT and Bing AI.

Mahoto Tsukaguchi† Youhei Ikarashi† Akihiko Misawa† Akira Toyama†

†Division of Pharmacy, Niigata University Medical & Dental Hospital

ChatGPT の回答については、参考資料となる根拠が示されない。また、回答が遠まわしな表現であり、簡潔で、わかりやすい回答となっていないことも課題である。具体例として、アザルフィジン EN 錠は粉砕可能か？と質問すると、BingAI では、「アザルフィジン EN 錠は、腸溶性フィルムコーティング錠であり、粉砕不可です。粉砕すると、腸溶コーティングが損なわれ、胃障害の発現リスクが高まる可能性があるため、医師の指示に従って適切に使用してください。」と簡潔に回答を行うが、ChatGPT では、「アザルフィジン EN 錠 (Azulfidine EN Tablets) を粉砕する前に、医師または薬剤師に相談することが非常に重要です。薬物の形態や処方によって、錠剤を粉砕して摂取することが許可される場合もあれば、許可されない場合もあります。

一般的に、アザルフィジン EN 錠は飲み込むために設計された錠剤であり、粉砕することなく丸ごと服用することが一般的です。錠剤を粉砕することで、薬物の効果、吸収率、および安全性に影響を及ぼす可能性があります。

したがって、アザルフィジン EN 錠を粉砕したい場合、必ず医師または薬剤師に相談し、指示を受けるべきです。医師は、特定の病状や患者のニーズに基づいて、適切な薬物の摂取方法を提案します。粉砕せずに錠剤を服用することが適切である場合もあれば、特定のケースで粉砕が許可される場合もあります。自己判断せず、医療プロバイダーの指示に従うことが大切です。」と長文になってしまう。現状の ChatGPT は、医薬品情報を担当する現場の薬剤師としては、非常に使いにくいものとなっている。

BingAI では、参考資料となる根拠が示されるが、主な参考資料を図2に示した。Wikipedia の割合が2019年<sup>[1]</sup>当時と比して減少し、様々な参考資料を基に回答されていることが窺えた。

Bing参考資料	ヒット数
MEDLAY	169
金久実研究室データベース(京都大学)	129
日経メディカル	119
医薬品医療機器総合機構	102
Bing検索	83
Jstage	53
Wikipedia	35
その他	28
EPARKくすりの窓口	24
大塚製薬工場	22
YG研究会	19
大分大学病院薬剤部	18
厚生労働省	18
MSDマニュアル	18
くすりの適正使用協議会	17

図2 BingAI の参考資料の上位15件

## 5 おわりに

現状、医薬品情報に対する正答率は、ChatGPT-3.5 で31.6%、BingAI では61.2%であり、十分高いとは言えない。今後、正答率を向上させるには、臨床で必要な医薬品情報の参考資料の充実、生成AI技術のさらなる向上が必要であろう。今回は、ChatGPT、BingAI ともにキーボード入力によって質問を記したが、今後、医療現場の環境でも音声入力で誤変換なく質問でき、生成AIから回答が得られるようになれば、実際の現場で使用しやすくなるだろう。

### 参考資料

- [1] 塚口真穂登, 伊藤晶子, 鈴木直人, 外山聡: AIスピーカーに医薬品情報を質問したデータの考察, 第81回情報処理学会全国大会
- [2] 塚口真穂登, 伊藤晶子, 鈴木直人, 外山聡: AIスピーカーに医薬品情報を2年質問し続けた回答の考察, 第83回情報処理学会全国大会
- [3] Soshi Takagi, Takashi Watari, Ayano Erabi, Kota Sakaguchi: Performance of GPT-3.5 and GPT-4 on the Japanese Medical Licensing Examination: Comparison Study, JMIR Med Educ Vol.9 C48002
- [4] Yudai Kaneda, Ryo Takahashi, Uiri Kaneda, Shiori Akashima, Haruna Okita, Sadaya Misaki, Akimi Yamashiro, Akihiko Ozaki, Tetsuya Tanimoto: Assessing the Performance of GPT-3.5 and GPT-4 on the 2023 Japanese Nursing Examination, Cureus 15(8):e42924. DOI10.7759
- [5] Yuki Kunitsu: The Potential of GPT-4 as a Support Tool for Pharmacists: Analytical Study Using the Japanese National Examination For Pharmacists, JMIR Med Educ Vol.9 C48452
- [6] Benedict Morath, Ute Chiriach, Elena Jaszkowski, Carolin Deiß, Hannah Nürnberg, Katrin Hörth, Torsten Hoppe-Tichy, Kim Green: Performance and risks of ChatGPT used in drug information: an exploratory real-world analysis, Eur J Hosp Pharm 2023;0:1-7