

ARTIGO ORIGINAL

CONFIABILIDADE E INFORMAÇÃO EM SAÚDE: AVALIAÇÃO DA PRECISÃO DAS RESPOSTAS DOS CHATBOTS DE IA NO CONTEXTO DO DIABETES

RELIABILITY AND HEALTH INFORMATION: EVALUATION OF THE ACCURACY OF AI CHATBOT RESPONSES IN THE CONTEXT OF DIABETES

Sabrina de Castro Rockenback¹
Maria Eduarda Vieira da Silva³
Claudia Giuliano Bica⁶

Giovana Diniz de Oliveira Bonetti²
Luis Fernando Marcelino Braga⁴ Tierre Aguiar Gonçalves⁵

¹Graduanda em Biomedicina. Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Rio Grande do Sul, Brasil. Área de atuação: Literacia digital.

²Biomédica, mestre e doutoranda em Patologia. Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Rio Grande do Sul, Brasil. Área de atuação: Literacia digital.

³Graduanda em Psicologia. Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Rio Grande do Sul, Brasil. Área de atuação: Literacia digital.

⁴Informático biomédico e mestrando em Ciências da Saúde. Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Rio Grande do Sul, Brasil. Área de atuação: Literacia digital.

⁵Médico e mestre em Patologia. Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA), Porto Alegre, Rio Grande do Sul, Brasil. Área de atuação: Literacia digital.

⁶Bióloga e doutora em Patologia. Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA). Porto Alegre, Rio Grande do Sul, Brasil. Área de atuação: Literacia digital.

Resumo

O uso crescente de chatbots baseados em modelos de linguagem de larga escala (LLM) como ferramenta de busca por informações em saúde representa uma mudança no acesso à informação. Isso é ainda mais relevante para usuários idosos, grupo com maior vulnerabilidade ao impacto de informações imprecisas. Este estudo transversal qualitativo analisa a precisão das respostas obtidas por seis plataformas de IA (Gemini 3 Flash, DeepSeek-V3, Meta IA Llama 4, ChatGPT-4.1 mini, Copilot GPT-5 e Perplexity GPT-4o mini) para a pergunta "Algum remédio cura o diabetes?", formulada a partir da coleta de dúvidas de uma amostra de 123 idosos participantes de iniciativas sem fins lucrativos de letramento digital em dois estados brasileiros. As respostas foram comparadas com as Diretrizes da Sociedade Brasileira de Diabetes por duas pesquisadoras, com desempate por um terceiro avaliador. Entre as seis ferramentas analisadas, somente a Meta IA e a Copilot forneceram respostas imprecisas. Os resultados evidenciam o risco informacional associado ao uso irrestrito de chatbots como fontes de informações em saúde, sinalizando a necessidade de avaliação contínua do seu uso.

PALAVRAS-CHAVE

Inteligência artificial. Precisão. Confiabilidade. Informação em saúde. Diabetes. Usuários idosos.

Abstract

The growing use of chatbots based on large language models (LLM) as a tool for searching health information represents a shift in how information is accessed, which is particularly relevant for elderly users - a group more vulnerable to the impact of inaccurate information. This cross-sectional, mixed-methods study analyzes the accuracy of responses provided by six AI platforms (Gemini 3 Flash, DeepSeek-V3, Meta AI Llama 4, ChatGPT-4.1 mini, Copilot GPT-5, and Perplexity GPT-4o mini) to the question "Is there any medicine that cures diabetes?", formulated from questions collected from a sample of 123 elderly adults participating in nonprofit digital literacy initiatives in two Brazilian states. The responses were compared against the Brazilian Diabetes Society Guidelines by two researchers, with a third researcher serving as tiebreaker in cases of disagreement. Among the six platforms analyzed, only Meta AI and Copilot provided inaccurate responses. The

findings highlight the informational risk associated with the unrestricted use of chatbots as sources of health information, signaling the need for continuous evaluation of their use.

KEYWORDS

Artificial intelligence. Accuracy. Reliability. Diabetes. Health information. Elderly users.

1 Introdução

Conforme o avanço do aprimoramento e da popularidade dos chatbots de inteligência artificial baseados em modelos de linguagem de larga escala (Large Language Model), observa-se uma expansão significativa de seu uso como ferramenta de busca por informações em saúde (Bean et al., 2026). Essa transformação, fundamentada no acesso rápido e personalizado garantido ao usuário por essas ferramentas, implica também uma alteração na forma de acesso e interação com os dados. Esse fenômeno é particularmente relevante no contexto da inclusão digital de pessoas idosas, grupo naturalmente mais inclinado a dificuldades de adaptação tecnológica (Rocha et al., 2025), que podem se beneficiar da acessibilidade e da facilitação de obtenção de informações proporcionadas por essas ferramentas.

Os mecanismos dos chatbots amplamente utilizados atualmente operam a partir de extensos bancos de dados virtuais e utilizam técnicas avançadas de aprendizado profundo (deep learning), baseadas em arquiteturas complexas de rede neural do tipo transformer. Diferentemente dos mecanismos de busca tradicionais, que se limitam a encontrar correspondências por palavras-chave, os LLMs são capazes de lidar com as nuances da linguagem humana, não apenas identificando seus padrões, mas também replicando-os com diferentes graus de autonomia (Goodfellow et al., 2016). No entanto, a especialização em processamento de linguagem natural não garante, por si só, a prevenção da disseminação de informações cientificamente não comprovadas (Klingbeil et al., 2024), especialmente relevante no campo da informação em saúde.

Paralelamente, o diabetes mellitus, condição crônica caracterizada pela hiperglicemia persistente (Bandeira; Mancini; Graf, 2015) e sem possibilidade de cura com os recursos terapêuticos atualmente disponíveis (Rodacki et al., 2024; Silva Júnior et al., 2023), apresenta prevalência de 31,4% na população com idade superior a 65 anos, segundo o relatório estatístico do Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico - (VIGITEL) (Brasil, 2025). Em território brasileiro, estima-se que mais de 16 milhões de pessoas convivam com a doença, conforme estatísticas do Atlas Global do Diabetes 2025 da International Diabetes Federation, número que representa cerca de 8,2% da população (IBGE, 2023).

Considerando a elevada prevalência da doença na população idosa brasileira, sua maior vulnerabilidade em relação ao avanço tecnológico e o uso ascendente de chatbots de inteligência artificial para obtenção de informações em saúde, torna-se imprescindível investigar a validade e qualidade informacional do conteúdo produzido por essas ferramentas. Assim, o presente estudo tem como objetivo analisar a precisão das respostas fornecidas por diferentes chatbots de LLM a respeito de uma dúvida em saúde sobre diabetes formulada pelo público idoso.

2 Método

Trata-se de um estudo transversal quali-quantitativo com ênfase na avaliação de informações em saúde, a partir da seleção de dúvidas feitas por participantes de programas públicos de letramento digital de dois estados brasileiros (RS e MS), possuidores ou não da doença. A amostragem foi feita por conveniência. Os participantes deveriam ter pelo menos 55 anos, concordar em participar do estudo de forma voluntária e estar devidamente matriculados em cursos de letramento digital.

Coleta de dados: Os dados primários foram coletados a partir de formulários eletrônicos do Google Forms aplicados de forma dirigida, com um membro da equipe de pesquisa auxiliando no preenchimento. Além de questões para caracterização da amostra, o questionário incluía perguntas de campo aberto a respeito das principais dúvidas dos participantes sobre diabetes, hipertensão, câncer e saúde em geral.

Análise com uso de chatbots: Foram selecionadas as plataformas de IA Gemini 3 Flash, DeepSeek-V3, Meta IA Llama 4, ChatGPT-4.1 mini, Copilot GPT-5 e Perplexity GPT-4o mini. A pergunta foi aplicada a todas no mesmo dia, em outubro de 2025, utilizando suas versões gratuitas. As respostas foram avaliadas quanto à sua precisão.

Análise de precisão: As respostas foram comparadas com as Diretrizes da Sociedade Brasileira de Diabetes (SBD) - incluindo os capítulos referentes ao diagnóstico do diabetes mellitus (Rodacki et al., 2024), à insulino terapia no diabetes mellitus tipo 1 (Silva Júnior et al., 2023) e ao rastreamento no diabetes mellitus tipo 1 (Rodacki et al., 2025), sendo descritas como “precisas”, quando em conformidade com as diretrizes, ou “imprecisas” quando divergentes. Para esta análise, duas pesquisadoras (GB e SR) compararam as respostas com as diretrizes; um terceiro pesquisador (TR) auxiliou no desempate no caso de divergências quanto à classificação.

Análise de dados: Os dados coletados foram exportados para planilhas eletrônicas do Google Sheets, organizados pela equipe, agrupados pelos temas centrais das dúvidas e quantificados, com cálculo de frequência e percentis.

Considerações éticas: Essa pesquisa integra um projeto já aprovado no Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Ciências da Saúde de Porto Alegre, sob parecer número 78445224.8.0000.5345.

3 Resultados

Foram entrevistados 123 idosos participantes do Rio Grande do Sul (RS) e do Mato Grosso do Sul (MS). A amostra foi composta majoritariamente por mulheres (82,1%), acima dos 60 anos (86,2%), com escolaridade até o ensino fundamental (44,7%) e aposentadas (79,7%). Entre os participantes, 23,6% possuíam diagnóstico de diabetes. As dúvidas sobre diabetes corresponderam ao maior volume de ocorrências entre os temas investigados (72 ocorrências), e a categoria "medicação" concentrou o maior número de questões (22 ocorrências) dentro do tema diabetes, motivando o recorte temático do presente estudo.

Tabela 1 | Ocorrência por tema

Tema	Número de ocorrências
Diabetes	72
Câncer	68
Hipertensão	65

Nota: Tabela de comparação sobre número de ocorrências por temas gerais. Fonte: Autoria própria.

Tabela 2 | Número de dúvidas sobre diabetes por tópico

Tema	Número de ocorrências
Medicação	22
Alimentação	18
Desenvolvimento	10
Tratamento	9
Cura	7
Sintomas	4
Outros	2

Nota: Tabela de comparação sobre número de dúvidas dos tópicos de diabetes. Fonte: Autoria própria.

A pergunta “Algum remédio cura o diabetes?”, selecionada pelo seu conteúdo e pela relevância da resposta para o público idoso por ofertar maior risco clínico em caso de imprecisão, foi submetida às seis plataformas de IA avaliadas (Gemini 3 Flash, DeepSeek-V3, Meta IA Llama 4, ChatGPT-4.1 mini, Copilot GPT-5 e Perplexity GPT-4o mini) traduzida para a língua inglesa, a fim de possibilitar uma resposta mais precisa pelos chatbots. As respostas obtidas foram comparadas com as Diretrizes da SBD (Rodacki et al., 2024; Silva Júnior et al., 2023; Rodacki et al., 2025) pelos pesquisadores para avaliar sua precisão.

Conforme demonstrado pela Tabela 3, quatro das seis plataformas - Gemini 3 Flash, DeepSeek-V3, ChatGPT-4.1 mini e Perplexity GPT-4o mini - forneceram respostas precisas para essa pergunta. As demais plataformas, Meta IA Llama 4 e Copilot GPT-5 tiveram o conteúdo de suas respostas avaliado como impreciso.

Tabela 3 | Classificação da precisão por comparação com as diretrizes da SBD

Chatbot IA	Precisão
Gemini 3 Flash	Precisa
DeepSeek-V3	Precisa
Meta IA Llama 4	Imprecisa
ChatGPT-4.1 mini	Precisa
Copilot GPT-5	Imprecisa
Perplexity	Precisa

Nota: Tabela de comparação da avaliação final dos pesquisadores sobre a precisão de cada resposta por chatbot. Fonte: Autoria própria.

Esses achados evidenciam que, embora a maioria das plataformas tenha demonstrado capacidade de atender os requisitos de compatibilidade com informações cientificamente comprovadas, ainda há possibilidade de inconsistências importantes. No caso da questão em pauta, que envolve a expectativa de um tratamento inexistente no estágio atual da medicina, essa variabilidade reforça a necessidade de cautela no uso dessas ferramentas como fonte de informação em saúde, especialmente em temas sensíveis.

A interpretação incorreta da resposta à pergunta “Algum remédio cura o diabetes?” pode impactar diretamente a compreensão e o manejo da doença pelo usuário idoso que poderia ter acesso fácil a esse conteúdo. Nesse sentido, é importante destacar que a plataforma Meta IA, que apresentou imprecisão, é amplamente utilizada no dia a dia por ser a ferramenta integrada a três dos aplicativos mais utilizados mundialmente, em ordem de maior número de usuários, respectivamente: WhatsApp Messenger, Facebook e Instagram (Kemp, 2025). Esse fato potencializa o alcance de informações incorretas a um público que já encontra barreiras no acesso ao meio digital, especialmente ao se tratar de uma doença de alto impacto para sua faixa etária.

4 Conclusões

As plataformas Gemini 3 Flash, DeepSeek-V3, ChatGPT-4.1 mini e Perplexity GPT-4o mini fornecem respostas precisas para a pergunta “Algum remédio cura o diabetes?”, quando comparadas às Diretrizes da Sociedade Brasileira de Diabetes (Rodacki et al., 2024; Silva Júnior et al., 2023; Rodacki et al., 2025). Meta IA Llama 4 e Copilot GPT-5 apresentam respostas imprecisas para essa questão. Há predominância de respostas precisas, mas as respostas imprecisas da Meta IA representam um risco informacional ampliado pela facilidade e extensão de seu acesso.

Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio ao desenvolvimento desta pesquisa. Agradecem também ao

Serviço Social da Indústria de Mato Grosso do Sul (SESI-MS) pela parceria e apoio às ações desenvolvidas, bem como ao Programa de Pós-Graduação em Patologia da Universidade Federal de Ciências da Saúde de Porto Alegre (UFCSPA) e à própria Universidade pelo suporte institucional que viabilizou a realização deste estudo.

Referências

- BANDEIRA, Francisco; MANCINI, Márcio C.; GRAF, Hans (eds.). **Endocrinologia e Diabetes**. 3. ed. Rio de Janeiro: MedBook Editora Científica Ltda., 2015. Parte VIII: Pâncreas endócrino e diabetes mellitus. pg. 759.
- BEAN, Alexander; PAYNE, Rebecca Elizabeth; PARSONS, Guy et al. **Reliability of LLMs as medical assistants for the general public: a randomized preregistered study**. *Nature Medicine*, [s.l.], v. 32, p. 609–615, 2026. DOI: <https://doi.org/10.1038/s41591-025-04074-y>
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde e Ambiente. **Vigitel Brasil 2006-2024: estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal**. Brasília: Ministério da Saúde, 2025. p. 142, tabela 51. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/svsa/vigitel/vigitel-brasil-2006-2024.pdf/view>. Acesso em: 22 abr. 2026.
- GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning**. [s.l.]: MIT Press, 2016. p. 443.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Censo demográfico 2022: resultados**. Rio de Janeiro: IBGE, 2023.
- INTERNATIONAL DIABETES FEDERATION (IDF). **IDF Diabetes Atlas**. 11. ed. Brussels: International Diabetes Federation, 2025. **Capítulo 3: The global picture of diabetes, tabela 3.4**. Disponível em: <https://diabetesatlas.org>. Acesso em: 5 jan. 2026.
- KEMP, Simon. **Digital 2025: Global Overview Report**. [s.l.]: DataReportal, 2025. Disponível em: <https://datareportal.com/reports/digital-2025-global-overview-report>. Acesso em: 28 abr. 2026.
- KLINGBEIL, Artur; GRÜTZNER, Cassandra; SCHRECK, Philipp. **Trust and reliance on AI — An experimental study on the extent and costs of overreliance on AI**. *Computers in Human Behavior*, [s.l.], v. 160, 108352, 2024. DOI: <https://doi.org/10.1016/j.chb.2024.108352>. Acesso em: 28 abr. 2026.
- ROCHA, Mayara Elisabeth Ferreira da; OLIVEIRA, Augusto César Alves de; TRINDADE FILHO, Euclides Mauricio. **Utilização da internet por idosos e sua contribuição no conhecimento sobre a COVID-19. Contribuciones a las Ciencias Sociales**, São José dos Pinhais, v. 18, n. 6, e18643, 2025. DOI: <https://doi.org/10.55905/revconv.18n.6-140>. Acesso em: 20 abr. 2026.
- RODACKI, Melanie; COBAS, Roberta A.; ZAJDENVERG, Lenita; SILVA JÚNIOR, Wellington Santana da et al. **Diagnóstico de diabetes mellitus**. Diretriz Oficial da Sociedade Brasileira de Diabetes, 2024. DOI: <https://doi.org/10.29327/5412848.2024-1>. Disponível em: <https://diretriz.diabetes.org.br/diagnostico-de-diabetes-mellitus/>. Acesso em: 20 abr. 2026.
- RODACKI, Melanie; GABBAY, Monica Andrade Lima; CALLIARI, Luis Eduardo; ZAJDENVERG; SILVA JÚNIOR, Wellington Santana da. *et al.* **Rastreamento no diabetes mellitus tipo 1 (DM1)**. Diretriz Oficial da Sociedade Brasileira de Diabetes, 2025. DOI: <https://doi.org/10.29327/5738823.2025-1>. Disponível em: <https://diretriz.diabetes.org.br/rastreamento-do-diabetes-mellitus-tipo-1-dm1/>. Acesso em: 20 abr. 2026.
- SILVA JÚNIOR, Wellington Santana da; GABBAY, Monica; LAMOUNIER, Rodrigo; BERTOLUCI, Marcello. **Insulinoterapia no diabetes mellitus tipo 1 (DM1)**. Diretriz Oficial da Sociedade Brasileira de Diabetes, 2023. DOI: <https://doi.org/10.29327/557753.2022-5>. Disponível em: <https://diretriz.diabetes.org.br/insulinoterapia-no-diabetes-mellitus-tipo-1-dm1/>. Acesso em: 20 abr. 2026.

Submissão: 04/08/2026

Aceite: 16/08/2026

Como citar o artigo:

ROCKENBACK, Sabrina de Castro et al. Confiabilidade e informação em saúde: avaliação da precisão das respostas dos chatbots de IA no contexto do diabetes. **Estudos Interdisciplinares sobre o Envelhecimento**, Porto Alegre, v. 31, e 2316-2171, 2026. DOI : 10.22456/2316-2171.157286