

QUANDO OS MÉDICOS CONFIAM NAS MÁQUINAS: VIÉS DE AUTOMAÇÃO E SEU IMPACTO NA DECISÃO CLÍNICA E SAÚDE.

WHEN DOCTORS TRUST MACHINES: AUTOMATION BIAS AND ITS IMPACT ON CLINICAL DECISION-MAKING AND HEALTH.

João Pedro Cardoso Faccio e Yasmim Tres Rodrigues

1. Professor do curso de medicina da Faculdade afya Pato Branco.; 2 Yasmim Tres Rodrigues acadêmica do curso de medicina da Faculdade.
Yasmimtres8@gmail.com

Eixo Temático: 01 saúde

RESUMO

A utilização da inteligência artificial na área da saúde tem contribuído para o apoio ao diagnóstico e à tomada de decisões clínicas, ,entretanto,também podem favorecer o vício de indicações mediadas por tecnologias. Nesse sentido, o viés de automação representa um possível risco à segurança do paciente. O presente estudo teve como principal objetivo analisar a influência do viés de automação na tomada de decisões clínicas, com foco tanto diagnóstico como na segurança do paciente. Trata-se de uma revisão de literatura realizada na base de dados PubMed, utilizando os termos “automation bias”, “artificial intelligence” e “clinical decision-making”. Os estudos que foram analisados demonstraram que as recomendações de inteligência artificial podem até melhorar a precisão diagnóstica quando corretas, mas também podem favorecer erros quando incorretas. Além disso, fatores relacionados ao desempenho diagnóstico e às vivências profissionais na área da saúde podem influenciar a dependência dessas recomendações. Conclui-se então que a inteligência artificial deve atuar como ferramenta de apoio, mantendo-se a avaliação crítica e a intervenção do profissional de saúde..

Palavras-chave: Inteligência artificial,Tomadas de decisões clínicas e viés de automação.

ABSTRACT

The use of artificial intelligence in healthcare has helped support diagnosis and clinical decision-making, yet it can also foster reliance on technology-mediated recommendations. In this context, automation bias poses a real risk to patient’s safety. The primary objective of this study was to analyze the influence of automation bias on clinical decision-making, focusing on both diagnosis and patient safety. This literature review was conducted using the PubMed database, with the search terms "automation bias," "artificial intelligence," and "clinical decision-making." The studies analyzed demonstrated that artificial intelligence recommendations can improve diagnostic accuracy when correct, but can also lead to inappropriate decisions when incorrect. Furthermore, it was found that factors related to diagnostic performance and professional experience in healthcare also influence adherence to these recommendations. The study concludes that artificial intelligence should serve as a support tool, while preserving the healthcare professional's critical assessment and intervention.

Keywords: artificial intelligence, clinical decision-making, automation bias.

INTRODUCTION

In recent years, artificial intelligence (AI) has rapidly expanded its role in medical decision-making, aiding in early diagnosis, image analysis, treatment, and triage. However, this advancement has implications for both professionals and patients, as its growing use can influence clinical decision-making. In this context, automation bias emerges:a phenomenon directly related to the tendency to strongly follow the recommendations of automated systems.

Automation bias refers to the tendency to place excessive trust in and rely on automated technologies, a phenomenon studied across various healthcare fields. Clinical Decision Support Systems (CDSS) were developed to assist in clinical decision-making; although some studies report improved performance with their use, users may also fail to detect errors that the CDSS themselves might introduce ¹.

Automation bias thus poses a challenge to effective clinical decision-making and may compromise diagnostic accuracy. Previous studies have identified overconfidence as a factor associated with automation bias ¹, suggesting that a professional's individual profile can also influence reliance on AI-generated recommendations ².

However, the professional's profile is not the only relevant factor; the technology itself can alter decision-making processes, treatment choices, and diagnoses. Kücking et al ². found that correct AI recommendations improved decision-making, whereas incorrect ones led to negative effects ².

The influence of AI can therefore be beneficial or detrimental, depending on the accuracy of its recommendations.

Notable concerns include the potential loss of professional skills and impacts on clinical reasoning ³. As AI becomes increasingly integrated into healthcare, it is crucial to understand its limitations and the extent to which it can assist without introducing errors, given that an inappropriate decision influenced by these tools could compromise patient safety and health. The key question is no longer simply whether AI provides correct answers, but rather how it alters the reasoning of the professionals who use it.

This study, therefore, aims to critically analyze the available evidence on the influence of automation bias on clinical decision-making, focusing on its potential impacts on diagnose.

METHOD

This is a literature review conducted via a search of the PubMed database. The descriptors "automation bias," "artificial intelligence," and "clinical decision-making" were used, combined with the Boolean operator AND. Articles available in English regarding the use of AI in clinical decision-making and automation bias were selected; duplicate studies or those unrelated to the review's objective were excluded. Certain classic references were also considered in developing the topic.

RESULTS

An article by Goddard, Roudsari, and Wyatt identifies approximately 74 studies related to automation bias ¹. The authors noted that individual and cognitive factors, such as trust, can influence the tendency to follow a technological recommendation. In a more recent study, Kucking et al. evaluated 210 participants and analyzed diagnostic

performance as influenced by interaction with AI recommendations ². The results show that technology can aid decision-making but can also lead to errors when recommendations are accepted without critical analysis. In a review of 29 studies, Al-Anezi highlighted the presence of automation bias and its potential impact on diagnostic reasoning. Strategies to mitigate these risks include professional training and maintaining human involvement in decision-making. Collectively, these studies demonstrate that the benefits of AI depend not only on the technology itself but also on how the professional interprets and utilizes the recommendations.

CONCLUSION

This literature review indicates that automation bias poses a significant challenge to clinical decision-making. Artificial intelligence can contribute to diagnostic accuracy, but its excessive or inappropriate use tends to foster errors and compromise patient safety. Therefore, it is considered that these technologies should be employed as support tools, while preserving the professional's critical judgment and active participation in the decision-making process.

Understanding the factors that foster automation bias therefore appears to be a necessary condition for the safer use of artificial intelligence in medical practice. Balancing technological resources with professional expertise tends to promote safer clinical decisions, thereby contributing to patient protection. In this context, professional training and the development of skills for critically evaluating AI recommendations are essential to reduce inappropriate reliance on these systems. Thus, the integration of AI into healthcare should be accompanied by strategies that strengthen clinical reasoning and ensure that technological recommendations complement, rather than replace, professional judgment.

REFERENCES

1. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc.* 2012;19(1):121-127. doi:10.1136/amiajnl-2011-000089.

2. Kücking F, Hübner U, Przysucha M, Hannemann N, Kutza JO, Moelleken M, et al. Automation bias in AI-decision support: results from an empirical study. *Stud Health Technol Inform*. 2024;317:298-304. doi:10.3233/SHTI240871.

3. Al-Anezi FM. Generative artificial intelligence in healthcare: automation bias, deskilling, and cognitive implications: a systematic review. *J Healthc Leadersh*. 2026;18:590498. doi:10.2147/JHL.S590498.