

# Artificial Intelligence in Healthcare: Innovation and Impact in Medical Practice

Amelia Maria ISAC<sup>a</sup> and Corina VERNIC<sup>b</sup>

<sup>a</sup> “Victor Babeș” University of Medicine and Pharmacy Timișoara, Romania.

<sup>b</sup> Department of Medical Informatics and Biostatistics, “Victor Babeș” University of Medicine and Pharmacy Timișoara, Romania.

E-mails: isac.amelia@yahoo.com; cvernic@umft.ro

\* Author to whom correspondence should be addressed;

## Abstract

**Background:** Artificial Intelligence (AI) is transforming medical healthcare by improving diagnostics, optimizing treatment plans, and enhancing patient outcomes. However, the statistical robustness and generalizability of AI models in clinical settings remain critical research areas. This study evaluates the predictive accuracy, reliability, and clinical applicability of AI-based models in evidence-based decision-making. **Methods:** A systematic review and meta-analysis were conducted to assess AI applications in healthcare, focusing on predictive modeling, diagnostic accuracy, and treatment optimization. Studies were selected following PRISMA guidelines, ensuring the inclusion of peer-reviewed research with statistical validation. Data were extracted from two key studies published between 2016 and 2021: (1) “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs” (Gulshan et al., 2016) and (2) “Deep Learning for Alzheimer’s Disease: A Systematic Review” (Wang et al., 2021). These studies incorporated various AI methodologies—including deep learning, decision trees, and ensemble models—and applied both Bayesian and frequentist statistical approaches to evaluate performance. Sensitivity analyses were performed to assess the impact of dataset size, feature selection, data heterogeneity, and potential biases. **Results:** Preliminary findings indicate that AI-based models outperform traditional statistical methods in disease detection, with an average AUC improvement of 10-15%. Additional metrics such as sensitivity, specificity, and F1-score further support AI’s superior predictive capabilities. However, significant variability exists, particularly in smaller datasets ( $n < 500$ ) and when suboptimal feature selection strategies are used. The analysis highlights the necessity of robust validation techniques, such as cross-validation and external dataset testing, to mitigate overfitting and improve model reliability across diverse clinical settings. **Conclusions:** AI-driven models show promise in augmenting evidence-based clinical decision-making, but rigorous statistical validation and generalizability remain critical challenges. Future research should prioritize the integration of AI with traditional statistical methodologies to enhance interpretability, reliability, and real-world applicability. Furthermore, federated learning and domain adaptation techniques may help address data privacy concerns and improve model robustness across diverse populations. One limitation of this study is the variability in dataset sizes and the potential lack of standardization in feature selection across different AI models.

**Keywords:** Artificial Intelligence (AI) in Healthcare; Medical AI; Healthcare Statistics; Predictive Modeling; Medical Data Analysis; Sensitivity Analysis; AUC-ROC; Cross-Validation; Model Validation; Federated Learning.

