

The Application of Machine Learning Algorithms Integrated with Electronic Health Record for the Prediction of Chronic Diseases in Primary Care

A Aplicação de Algoritmos de Machine Learning Integrada ao Prontuário Eletrônico na Predição de Doenças Crônicas na Atenção Primária

Aplicación de Algoritmos de Aprendizaje Automático Integrados con la Historia Clínica Electrónica para Predecir Enfermedades Crónicas en Atención Primaria

RESUMO

OBJETIVO: identificar a aplicação de Inteligência Artificial integrado ao Prontuário Eletrônico do Paciente na atenção primária, para predição de doenças crônicas. **MÉTODO:** Foram utilizadas as bases de dados PubMed e Biblioteca Virtual em Saúde. Os critérios de inclusão utilizados foram: Artigos publicados entre 2019 e 2024, e estudos relacionados à predição de doenças em contextos de atenção primária com a utilização de Inteligência Artificial. **RESULTADOS:** Os estudos demonstraram melhorias no diagnóstico precoce, na identificação de fatores de risco e na redução de hospitalizações, embora desafios como a interoperabilidade limitada e a dependência de dados de qualidade persistam. **CONCLUSÃO:** Com base nas análises das potencialidades e limitações da integração da Inteligência Artificial com Prontuário Eletrônico do Paciente na gestão de doenças crônicas, recomenda-se que futuros esforços se concentrem na padronização de dados e na melhoria da infraestrutura técnica para otimizar o impacto na atenção primária.

DESCRIPTORES: Inteligência Artificial; Prontuário Eletrônico; Atenção Primária; Predição de Doenças.

ABSTRACT

OBJECTIVE: To identify the integration of artificial Intelligence into electronic patient records for the predicting chronic diseases in primary care settings. **METHOD:** The PubMed and Virtual Health Library databases were used. The inclusion criteria comprised articles published between 2019 and 2024 that investigated disease-prediction in primary models in primary contexts using Artificial Intelligence techniques. **RESULTS:** The reviewed studies demonstrated in early diagnosis, accurate identification of risk factors, and reductions in hospitalization rates. Persistent challenges included limited interoperability and dependence on high-quality data. **CONCLUSION:** Based on analyses of the potential and limitations of integrating artificial intelligence into electronic patient records for chronic disease management, we recommended that future efforts focus on comprehensive data standardization and on strengthening technical infrastructure to optimize impact in primary care.

KEYWORDS: Artificial Intelligence; Electronic Health Record; Primary Care; Disease Prediction.

RESUMEN

OBJETIVO: Identificar la aplicación de la Inteligencia Artificial integrada en la Historia Clínica Electrónica en atención primaria para la predicción de enfermedades crónicas. **MÉTODO:** Se utilizaron las bases de datos PubMed y Biblioteca Virtual de Salud. Los criterios de inclusión utilizados fueron: Artículos publicados entre 2019 y 2024, y estudios relacionados con la predicción de enfermedades en contextos de atención primaria utilizando Inteligencia Artificial. **RESULTADOS:** Los estudios mostraron mejoras en el diagnóstico precoz, la identificación de factores de riesgo y la reducción de hospitalizaciones, aunque persisten retos como la interoperabilidad limitada y la dependencia de datos de calidad. **CONCLUSIÓN:** A partir de los análisis del potencial y las limitaciones de la integración de la Inteligencia Artificial con los Registros Electrónicos de Pacientes en la gestión de enfermedades crónicas, se recomienda que los esfuerzos futuros se centren en estandarizar los datos y mejorar la infraestructura técnica para optimizar el impacto en la atención primaria.

DESCRIPTORES: Inteligência Artificial; Historia Clínica Electrónica; Atención Primaria; Predicción de Enfermedades.

RECEIVED: 09/18/2025 APPROVED: 10/06/2025

How to cite this article: Silva AJBD, Pisa IT, Novoa CG, Pinto VC. The Application of Machine Learning Algorithms Integrated with Electronic Health Record for the Prediction of Chronic Diseases in Primary Care. *Saúde Coletiva* (Edição Brasileira) [Internet]. 2025 [acesso ano mês dia];16(101):17422-17433. Disponível em: DOI: 10.36489/saudecoletiva.2025v16i101p17422-17433

**Andresa Jardim Bueno Damásio da Silva**

Family Health Specialist - UNIFESP 2016 Health Informatics Specialist - UNIFESP 2025. Health Services Manager - OS Monte Azul
ORCID: <https://orcid.org/0009-0009-2592-5855>

**Ivan Torres Pisa**

Professor, UNIFESP. Lecturer in health informatics, UNIFESP, 2013
ORCID: <https://orcid.org/0000-0002-5106-3904>

**Cláudia Galindo Novoa**

Associate Professor UNIFESP

**Vladimir Camelo Pinto**

PhD candidate in Information Systems - EACH-USP 2025 Master of Science (Health Informatics) - UNIFESP 2008 Bachelor of Computer Science - UNICSUL 2004 EBT Professor at the Federal Institute of São Paulo (IFSP)

INTRODUCTION

Primary health care is an essential pillar of health systems, responsible for managing chronic conditions and promoting health¹. With the increasing prevalence of chronic diseases such as heart failure and diabetes mellitus, technologies that aid in the early diagnosis and management of these conditions are becoming increasingly relevant². The integration of Artificial Intelligence (AI) into Electronic Patient Records (EPR) represents a promising solution to these challenges, promoting the analysis of large volumes of data and the identification of complex patterns³.

AI has the ability to identify patterns and predict outcomes based on large volumes of data, bringing significant advances to clinical practice. The integration of AI-based systems into EPRs reduces hospitalizations due to exacerbations of Chronic Obstructive Pulmonary Disease (COPD) and can improve the identification of risk factors². For example, evidence has shown a 74% increase in the identification of heart failure with the use of AI-associated systems¹.

Despite the advantages offered by new technologies in healthcare, their

implementation faces significant challenges such as limited interoperability, dependence on complete and high-quality data, lack of standardization in collection processes, and problems with information storage³.

Analysis of the benefits and limitations of technological innovations is essential, especially with regard to technological infrastructure, professional training, and adaptation to regional demands⁴. However, it is essential to understand how different technologies, such as EHRs, can be improved and applied to enhance early diagnosis and management of chronic diseases in primary care, exploring ways to optimize their integration into health systems⁵.

This study aims to identify the state of the art in the association of AI integrated with PEP, highlighting the types of AI used, its use in the detection of chronic diseases, and its main advantages and disadvantages when applied to the health field.

METHOD

This study conducts a literature review based on an approach to secondary sources, i.e., the theoretical contributions of various authors on specific topics⁶.

Inclusion criteria

For this study, the following inclusion criteria were considered: studies published between 2019 and 2024, articles that used quantitative and qualitative approaches, full articles with free access, which address the theme of chronic disease prediction and the application of Machine Learning (ML) algorithms integrated with EHRs in the context of primary care. In addition, studies published in English, Portuguese, or Spanish were included.

Articles that presented reviews or theoretical models and studies without integration with EHR or focusing on other types of pathologies different from those addressed in this study were excluded.

Databases consulted

Searches were conducted in the National Library of Medicine and the National Institutes of Health (PubMed) and Virtual Health Library (VHL) databases.

Search strategy

The following descriptors were used for the search strategy: “chronic disease,” “predict,” “primary care,” “electronic record,” “artificial intelligence,” and “machine learning,” in-

Literature Review

Silva AJBD, Pisa IT, Novoa CG, Pinto VC

The Application of Machine Learning Algorithms Integrated with Electronic Health Record for the Prediction of Chronic Diseases in Primary Care

cluding the use of Boolean operators AND and OR to combine descriptors.

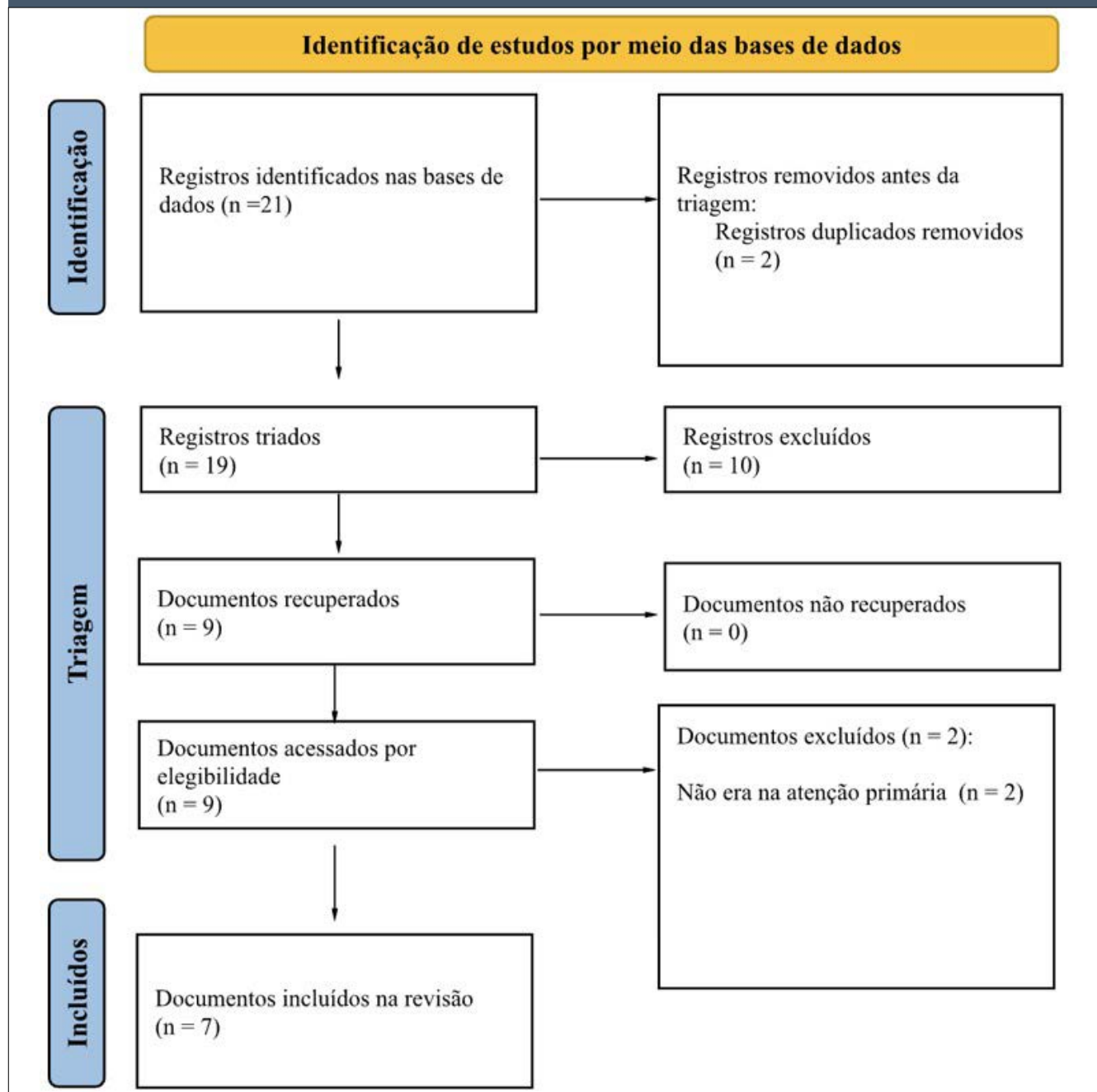
Screening process reduction steps

The search conducted in the da-

tabases resulted in the retrieval of 21 articles. After excluding duplicate studies, 19 articles were chosen for reading the title and abstract. After reading, 10 articles were excluded, and 9 of these were selected as eligible

for full-text reading. From this stage, seven studies were included in the review. The PRISMA flowchart was used to describe the phases of search, identification, selection of studies, and results, as shown in Figure 1.

Figure 1. Article selection process according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria⁷



Source: Adapted from Page⁷

RESULTS

The studies analyzed in this review show advances in the application of AI, integrated with PEPs in the context of primary care (Table 1).

Chronic diseases

The research highlights the use of artificial intelligence (AI) in the early detection of chronic conditions. The analysis of the seven studies demonstrated a diversity in the frequency of chronic conditions identified through AI. Among the diseases mentioned in the articles are: Diabetes Mellitus, Chronic Obstructive Pulmonary Disease (COPD), Heart Failure (HF), and Chronic Kidney Disease (CKD)^{1-3,8-11}.

It was found that Chronic Obstructive Pulmonary Disease (COPD) was addressed in four articles^{3,8,10,11}, as was Chronic Kidney Disease (CKD), present in an equal number of studies^{2,3,8,11}.

On the other hand, both Heart Failure

^{1,9,10} and Diabetes Mellitus were addressed in three studies^{1,2,9}.

Type of AI used

Regarding the types of AI used in the prediction of chronic diseases, of the seven articles, four used artificial neural networks^{1,9-11}.

Machine learning², support vector machines⁸, and decision trees³ were also used.

Advantages of using AI

The articles highlight several advantages in the application of AI for the management of chronic conditions, evidencing high diagnostic accuracy and early prediction of clinical conditions, respectively^{1,3}. Faster processing of complex data² allows for more efficient identification of risk factors¹⁰, contributing to better personalization of care¹¹ and more robust support for clinical decision-making⁸. In addition, these technological advances promote the

reduction of complications, reinforcing the potential of AI to transform clinical practices and optimize treatment outcomes⁹.

Disadvantages of using AI

The articles show that, despite advances, the application of AI in healthcare faces significant challenges.

Among the disadvantages are the dependence on high-quality data, which is essential for the effectiveness of algorithms¹, and the high volume of data required for training, which can complicate the management and analysis of information¹¹. In addition, limited interoperability between systems and the difficulty of technological integration hinder the consolidation of robust solutions^{3,8,10}, while high implementation costs impose restrictions on the widespread adoption of these technologies in clinical practice^{2,9}.

Table 1. Articles found in the PubMed (2019-2024) and BVS (2019-2024) databases on the integration of ML algorithms into EHRs in primary care for the prediction of chronic diseases						
Author, Year	Type of AI used	Advantage cited	Disadvantage cited	Real-time responses	Prediction made	Chronic diseases studied
Raat et al., 2022 (1)	Artificial Neural Networks	Diagnostic accuracy	Dependence on high-quality data	Yes	Yes	Heart failure, diabetes mellitus
Mosa et al., 2022 (3)	Decision Trees	Early prediction	Limited interoperability	No	Yes	COPD, chronic kidney disease
Zafari et al., 2022 (2)	Machine learning	Faster processing	High implementation costs	Yes	Yes	Diabetes mellitus, chronic kidney disease
Nadarajah et al., 2023 (10)	Artificial Neural Networks	Better identification of risk factors	Difficulty in system integration	Yes	Yes	Heart failure, COPD
Ställberg et al., 2021 (11)	Artificial Neural Networks	Personalization in care	High volume of data for training	Yes	Yes	COPD, chronic kidney disease
Butzin-Dozier et al., 2024 (8)	Support Vector Machines	Better clinical decision support	Limited interoperability	Yes	Yes	Chronic kidney disease, COPD
Hrabok et al., 2021 (9)	Artificial Neural Networks	Reduction of complications	High costs	Yes	Yes	Diabetes mellitus, heart failure

Source: Prepared by the authors, 2025

DISCUSSION

The analysis of the seven selected

articles revealed significant results that highlight both the advances and challenges associated with integrating machine learning (ML) algorithms

into EHRs in the context of primary care.

The seven studies reviewed highlight the potential of AI in the early

identification of conditions such as COPD, CKD, heart failure, and diabetes mellitus. Studies report that the use of AI in healthcare systems can enable the early detection of chronic diseases such as diabetes, hypertension, and cardiovascular conditions, potentially preventing disease progression¹².

The studies analyzed highlighted several advantages of using AI in the management of chronic diseases. Recent research indicates that AI can process large amounts of medical data, identifying complex patterns and subtle indicators of diseases that may go unnoticed by human analysis, thus improving the accuracy of diagnosis of diseases such as diabetes, cancer, and cardiovascular diseases^{12,13}. In addition, AI models, including ML and deep learning, outperform traditional diagnostic methods by providing more accurate and timely predictions^{13,14}.

Early detection through AI enables timely medical interventions, potentially preventing disease progression and improving patient outcomes, as well as reducing hospitalizations and long-term healthcare costs^{11,13,15}.

AI-driven predictive modeling facilitates personalized treatment plans, improving long-term health outcomes¹⁶. Cutting-edge AI and mobile health (mHealth) applications enable real-time data processing and instant decision-making, reducing latency and improving public health management¹⁷.

On the other hand, some studies have highlighted significant barriers to the implementation of ML in healthcare. The main challenges identified include dependence on high-quality data, limitations of interoperability between systems, and the high costs associated with training and applying predictive models and ethical issues^{10,18}. Hrabok⁹ also highlighted ethical challenges related to the collection and use of sensitive

data, especially in conditions such as epilepsy.

Although simplicity and customization are fundamental goals in the implementation of predictive tools, Rafiq identified challenges related to adapting strategies to the primary care environment. These barriers include the need to align tools with managers' preferences and the practical realities of integrated care centers. This study offers valuable insights into how co-design with physician managers can improve the adoption and effectiveness of patient-centered predictive tools⁵.

A comparison between the studies revealed that the advances observed still depend on well-structured contexts, such as the existence of robust technological infrastructure and standardized data. Compared to sectors such as finance and logistics, the use of ML in healthcare faces unique challenges, including the complexity of clinical data and the need to ensure information security and privacy^{18,19}.

Despite this, studies converge in indicating that ML has the potential to transform healthcare, particularly in primary care. However, for this impact to be widely achieved, it will be essential to invest in clear guidelines for data standardization, ethical implementation, and interoperability between systems. Thus, ML can truly contribute to more accurate diagnoses and more timely interventions, optimizing clinical outcomes and promoting greater efficiency in healthcare systems.

CONCLUSION

This study demonstrated that the integration of ML algorithms into EHRs represents a promising approach to improving early diagnosis and management of chronic diseases in primary care. The analysis revealed significant gains, such as increased diagnostic accuracy for heart failure

and the identification of risk factors for chronic kidney disease. These results highlight the potential of ML to transform clinical practice by offering more personalized and effective interventions.

However, significant barriers were identified, including dependence on high-quality data, limitations of interoperability between systems, and high implementation costs. These challenges reinforce the need for investments in technological infrastructure and professional training, as well as the development of standards to ensure the ethical and efficient application of these technologies.

The integration of AI and machine learning tools into EHRs has great potential to transform clinical practice in primary care. Collaborative approaches, such as those discussed by Rafiq⁵, highlight the importance of user-centered strategies centered on users to overcome technical and operational challenges. Future implementations may benefit from a focus on simplicity of design and customization of solutions, ensuring that technologies effectively meet the demands of professionals and patients.

This study achieved its objective by highlighting both the possibilities and challenges of using machine learning (ML) in electronic patient records (EPRs), as well as suggesting ways for its practical application. To advance in this area, it is essential that future research investigate the impact of these solutions in different contexts, considering not only clinical aspects but also economic and social dimensions. It is also increasingly urgent to develop solid guidelines to standardize and regulate these technologies, ensuring that ML truly contributes to the improvement of healthcare systems.

REFERENCES

1. Raat W, Smeets M, Henrard S, Aertgeerts B, Penders J, Droogne W, Mullens W, Janssens S, Vaes B. Machine learning optimization of an electronic health record audit for heart failure in primary care. *ESC Heart Fail*. 2021;9(1):39-47. Disponível em: <https://doi.org/10.1002/ehf2.13724>.
2. Zafari H, Langlois S, Zulkernine F, Kosowan L, Singer A. AI in predicting COPD in the Canadian population. *Biosystems* [Internet]. 2022; 211:104585. Disponível em: <https://doi.org/10.1016/j.biosystems.2021.104585>.
3. Mosa AI, Watts D, Tangri N. Impacting management of chronic kidney disease through primary care practice audits: a quality improvement study. *Can J Kidney Health Dis*. 2022; 9:205435812211448. Disponível em: <https://doi.org/10.1177/20543581221144840>.
4. Gomes SA, Montanini JF, Rocha Sobrinho HM. O uso da Inteligência Artificial na Medicina: os benefícios e desafios da parceria homem-tecnologia na saúde. *Rev Eletronica Acervo Saude*. 2024; 24(12):e18374. Disponível em: <https://doi.org/10.25248/reas.e18374.2024>.
5. Rafiq M, Mazzocato P, Guttman C, Spaak J, Savage C. Predictive analytics support for complex chronic medical conditions: an experience-based co-design study of physician managers' needs and preferences. *SSRN Electron J*. 2021. Disponível em: <https://doi.org/10.2139/ssrn.3976786>.
6. Sá-Silva JR, Almeida CD, Guindani JF. Pesquisa documental: pistas teóricas e metodológicas. *Rev Bras Hist Ciênc Soc*. 2009;1(1).
7. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *J Clin Epidemiol*. 2021;134:178-89. Disponível em: <https://doi.org/10.1016/j.jclinepi.2021.03.001>.
8. Butzin-Dozier Z, Ji Y, Li H, Coyle J, Shi J, Phillips RV, Mertens A, Pirracchio R, van der Laan MJ, Patel RC, Colford JM, Hubbard AE. Predicting long COVID in the national COVID cohort collaborative using super learner: cohort study (preprint). *JMIR Public Health Surveill*. 2023. Disponível em: <https://doi.org/10.2196/53322>.
9. Hrabok M, Engbers JD, Wiebe S, Sajobi TT, Subota A, Al-mohawes A, Federico P, Hanson A, Klein KM, Peedicail J, Pillay N, Singh S, Josephson CB. Primary care electronic medical records can be used to predict risk and identify potentially modifiable factors for early and late death in adult onset epilepsy. *Epilepsia*. 2020. Disponível em: <https://doi.org/10.1111/epi.16738>.
10. Nadarajah R, Wu J, Hogg D, Raveendra K, Nakao YM, Nakao K, Arbel R, Haim M, Zahger D, Parry J, Bates C, Cowan C, Gale CP. Prediction of short-term atrial fibrillation risk using primary care electronic health records. *Heart*. 2023 heartjnl—2022-322076. Disponível em: <https://doi.org/10.1136/heartjnl-2022-322076>.
11. Ställberg B, Lisspers K, Larsson K, Janson C, Müller M, Łuczko M, Kjølner Bjerregaard B, Bacher G, Holzhauer B, Goyal P, Johansson G. Predicting hospitalization due to COPD exacerbations in swedish primary care patients using machine learning – based on the ARCTIC study. *Int J Chronic Obstr Pulm Dis* [Internet]. 2021;16:677-88. Disponível em: <https://doi.org/10.2147/copd.s293099>.
12. Emeihe EV, Nwankwo EI, Ajegbile MD, Olaboye JA, Maha CC. The impact of artificial intelligence on early diagnosis of chronic diseases in rural areas. *Comput Sci Amp IT Res J*. 2024;5(8):1828-54. Disponível em: <https://doi.org/10.51594/csitrj.v5i8.1447>.
13. Freek A, Elgalb A, Gerges M. AI for Early Disease Detection: Developing models for early diagnosis of diseases like cancer or Alzheimer's using biomarkers and imaging data. *J AI Powered Med Innov (Int Online ISSN 3078 1930)*. 2024;3(1):1-24. Disponível em: <https://doi.org/10.60087/japmi.vol.03.issue.01.id.001>.
14. Wake AP, Pawar AP. Review paper on early detection of chronic kidney disease using predictive AI. *Int J Res Appl Sci Eng Technol*. 2024 ;12(11):2378-80. Disponível em: <https://doi.org/10.22214/ijraset.2024.65604>.
15. Moreno-Sánchez PA. Data-driven early diagnosis of Chronic Kidney Disease: development and evaluation of an explainable AI model. *IEEE Access*. 2023 [citado 23 jun 2025];1. Disponível em: <https://doi.org/10.1109/access.2023.3264270>.
16. Panguraj AR. Exploring how LLMs can analyse patient data, medical literature, and symptoms to assist in the early detection of chronic conditions like diabetes, cancer, or cardiovascular diseases. *Interantional J Sci Res Eng Manag*. 23 jan 2025 [citado 23 jun 2025];09(01):1-7. Disponível em: <https://doi.org/10.55041/ijrsrem22924>.
17. Bhatt P, Liu J, Gong Y, Wang J, Guo Y. Emerging artificial intelligence-empowered mobile health: a scoping review (preprint). *JMIR mHealth uHealth* [Internet]. 18 nov 2021 [citado 23 jun 2025]. Disponível em: <https://doi.org/10.2196/35053>.
18. Mirkin S, Albensi BC. Should artificial intelligence be used in conjunction with Neuroimaging in the diagnosis of Alzheimer's disease? *Front Aging Neurosci* [Internet]. 18 abr 2023 [citado 23 jun 2025];15. Disponível em: <https://doi.org/10.3389/fnagi.2023.1094233>.