MACHINE LEARNING TOOLS IN CHRONIC DISEASE MANAGEMENT: SCOPING REVIEW

Ferramentas de machine learning na gestão da doença crónica: scoping review

Herramientas de aprendizaje automático en la gestión de enfermedades crónicas: revisión del alcance

Igor Pinto*, Marta Sá**, Ana Alves***, Maria Silva****, Ana Carvalho******, Cátia Moreira******

ABSTRACT

Background: the implementation of technologies based on Artificial Intelligence (AI) in the health sector, in particular machine learning (ML), has had a significant transformational effect. Their use improves disease prediction, classification and diagnosis, benefiting both patients and healthcare professionals. Objective: to map ML tools for chronic disease management, with relevance to nursing care for people with chronic diseases. Methodology: scoping review based on the recommendations of the Joanna Briggs Institute. The MEDLINE Complete via PUBMED, CINAHL Complete via EBSCO, SCOPUS, OpenGrey, RCAAP and DART-Europe databases were used with no time limit. Results: seven articles were included and 9 AI tools associated with chronic disease management were identified, namely chronic kidney disease, chronic obstructive pulmonary disease, hepatitis C, heart failure and chronic venous insufficiency. Conclusion: the tools identified have the potential to contribute to improving nursing care, particularly in identifying risk factors associated with chronic diseases, detecting exacerbations early, continuously monitoring and evaluating the effectiveness of treatment and supporting clinical decision-making.

Keywords: machine learning; chronic disease; nursing

RESUMO

Enquadramento: a implementação de tecnologias baseadas em Inteligência Artificial (IA) na área da saúde, nomeadamente o machine learning (ML), tem causado um efeito transformacional significativo. A sua utilização melhora a previsão de doenças, classificação e diagnóstico, beneficiando os utentes e os profissionais de saúde. Objetivo: mapear as ferramentas de ML para a gestão de doenças crónicas, com relevância para os cuidados de enfermagem à pessoa com doença crónica. Metodologia: scoping review com base nas recomendações do Instituto Joanna Briggs. A pesquisa foi efetuada nas bases de dados MEDLINE Complete via PUBMED, CINAHL Complete via EBSCO, SCOPUS, OpenGrey, RCAAP e DART-Europe, sem limite de tempo. Resultados: foram incluídos sete artigos e identificadas 9 ferramentas de ML associadas à gestão de doenças crónicas nomeadamente doença renal crónica, doença pulmonar obstrutiva crónica, hepatite C, insuficiência cardíaca e insuficiência venosa crónica. Conclusão: as ferramentas identificadas têm potencial de contribuir para a melhoria dos cuidados de enfermagem, nomeadamente na identificação de factores de risco associados a doenças crónicas, detectar precocemente exacerbações, monitorizar e avaliar continuamente a eficácia do tratamento e apoiar a tomada de decisões clínicas.

Palavras-chave: aprendizado de máquina; doença crónica; enfermagem

RESUMEN

Marco contextual: la implantación de tecnologías basadas en Inteligencia Artificial (IA) en el sector sanitario, en particular el aprendizaje automático (AM), ha tenido un importante efecto transformador. Su uso mejora la predicción, clasificación y diagnóstico de enfermedades, beneficiando tanto a los usuarios como a los profesionales sanitarios. Objetivo: mapear las herramientas de ML para la gestión de enfermedades crónicas, con relevancia para los cuidados de enfermería a personas con enfermedades crónicas. Metodología: revisión de alcance basada en las recomendaciones del Instituto Joanna Briggs. Se utilizaron las bases de datos MEDLINE Complete via PUBMED, CINAHL Complete via EBSCO, SCOPUS, OpenGrey, RCAAP y DART-Europe, sin límite de tiempo. Resultados: se incluyeron siete artículos y se identificaron 9 herramientas de IA asociadas a la gestión de enfermedades crónicas, a saber, la enfermedad renal crónica, la enfermedad pulmonar obstructiva crónica, la hepatitis C, la insuficiencia cardíaca y la insuficiencia venosa crónica. Conclusión: las herramientas identificadas tienen potencial para contribuir a la mejora de los cuidados de enfermería, especialmente en la identificación de factores de riesgo asociados a enfermedades crónicas, la detección precoz de exacerbaciones, el seguimiento continuo y la evaluación de la eficacia del tratamiento y el apoyo a la toma de decisiones clínicas.

Palabras clave: aprendizaje automático; enfermedad crónica, enfermería
INTRODUCTION
The concept of Artificial Intelligence (AI) emerged in the 1950s with the simplified theory of human intelligence displayed through machinery (Helm, et al., 2020). Currently, it is recognized as one of the prominent computer sciences in the technological development of computerized systems, which simulate human capabilities. These intelligent computing systems are capable of performing tasks without direct interference from humans. In the field of health, this science explores ways to simulate human reasoning and the human decision-making process, using mathematical algorithms capable of recognizing problems, analysing data, indicating tasks, and supporting the decision-making of healthcare professionals (Lobo, 2018).

Computerized systems have existed for several decades; however, what is currently observed is the increased speed of processing and storage of computer information, enabling the analysis of large volumes of data in nanoseconds. In healthcare, AI analyses databases of mortality rates, birth rates, hospitalizations, disease reporting, prevalence, and incidence of diseases, making it an asset in anticipating pandemic outbreaks and applying preventive interventions by anticipating problems and diseases (Lobo, 2018).

The implementation of AI-based technologies, specifically in the field of healthcare, has had a particularly significant transformative effect on advanced nursing practice (Raymond, Castonguay, Doyon, & Paré, 2022). The use of AI in the realm of advanced nursing is relatively unexplored, with its relevance inferred from its application in disease prediction, classification, and diagnosis. This not only benefits patients but also healthcare professionals (Jimma, 2023).

AI has emerged in various healthcare contexts, especially in clinical databases for decision-making: machine learning (ML), algorithms of Clinical Decision Support Systems (CDSS) with ML, predictive analytics models, big data analysis, computer vision, natural language processing, and enhanced robotics (Raymond, Castonguay, Doyon, & Paré, 2022). ML is a field of artificial intelligence that applies algorithms to raw, absolute data to acquire knowledge in an automated manner. ML is a subdomain of AI, which can be translated as Machine Learning. This domain enables computers to evolve in learning and information accumulation without being explicitly programmed, through pattern recognition (Santos & Campos, 2022).

The techniques of ML are divided into supervised learning and unsupervised learning. In supervised learning, the training set comprises pairs of desired input and output, and the goal is to learn a mapping between input and output spaces. When the desired output is not part of the training set, and the output may return uncertain responses, it is termed unsupervised learning (Simeone, 2018).

The management of chronic disease represents a significant portion of the economic resources allocated to healthcare. The diagnosis of a chronic disease translates into the need for prolonged treatment for patients (Battineni, Sagaro, Chinatalapudi, & Amenta, 2020). According to data from the Pan American Health Organization (PAHO, 2020), seven of the top ten causes of death globally are non-communicable chronic diseases, linked to modifiable risk factors and behaviours. This reality emphasizes the crucial
importance of prevention, aligning with the Sustainable Development Goals (SDGs).
The availability of services to prevent, diagnose, and treat diseases is crucial for reducing deaths and disability, with the need for greater investment in these capabilities (PAHO, 2020). In Portugal, according to the Organization for Economic Cooperation and Development (2021), more than four in ten adults suffer from a chronic disease. Among Portuguese individuals aged sixteen or older, 41% reported suffering from at least one chronic disease (OCDE, 2021).

As healthcare professionals, nurses play a crucial role not only in delivering but also in improving healthcare for the population. However, knowledge about the nature, extent, and consequences of their involvement and experience with AI tools in healthcare is unknown (Raymond, Castonguay, Doyon, & Paré, 2022). Thus, the research question arises: What are the Machine Learning tools available in the literature to manage chronic diseases, with relevance to nursing care?

Given the conceptual considerations presented on the topic and the value that AI can have in the future of nursing care, the objective is to map the Machine Learning tools available in the literature for the management of chronic diseases, with relevance to nursing care for individuals with chronic illnesses, across all healthcare contexts.

**METHODOLOGICAL PROCEDURES OF REVIEW**

A Scoping Review was conducted, following the methodological strategy of the Joanna Briggs Institute for Scoping Reviews (Peters et al., 2022) and the guidelines of the PRISMA-ScR model ( Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Scoping Reviews) were adhered to (Tricco et al., 2018).

The selection of this methodology arises from the need to comprehensively explore and understand Machine Learning tools applied in the management of chronic diseases, with a particular focus on nursing care. This approach is ideal for mapping the extent, variety, and nature of existing research in an emerging and rapidly evolving field, such as AI in healthcare.

A scoping review protocol, which is not published, was developed.

The PCC mnemonic (Population, Concept, and Context) was employed to define eligibility criteria (Munn et al., 2018). The population refers to individuals with chronic diseases, the concept was the use of Machine Learning tools, and the context focused on nursing care in any healthcare setting.

Database search was conducted using the DeCS/MeSH® terms, Machine Learning, Chronic Disease, Nursing, operationalized by the Boolean operator “AND”. The terms were searched in the title and abstract.

The research strategy was conducted in three stages: (1) Initial search in MEDLINE (via PubMed) and CINAHL Complete (via EBSCO) to identify articles on the topic and analyse the words contained in the titles and abstracts of these articles, as well as any indexing terms used (Table 1); (2) Search in the databases of interest with the keywords identified in the first stage; (3) Analysis of the bibliographic reference list of all included articles for data extraction to find additional studies.
Table 1
Example of a search strategy in the MEDLINE database (via PubMed) with the identification of results by descriptor and possible combinations

<table>
<thead>
<tr>
<th>SEARCH STRATEGY</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 “Machine Learning”</td>
<td>109536</td>
</tr>
<tr>
<td>#2 “Chronic Disease”</td>
<td>334776</td>
</tr>
<tr>
<td>#3 Nursing</td>
<td>903339</td>
</tr>
<tr>
<td>#4 Nurs*</td>
<td>1154632</td>
</tr>
<tr>
<td>#5 “Machine Learning” AND “Chronic Disease” AND Nurs*</td>
<td>19</td>
</tr>
<tr>
<td>#5 &quot;Machine Learning&quot;[Title/Abstract] AND &quot;Chronic Disease&quot;[Title/Abstract] AND Nurs*[Title/Abstract]</td>
<td>3</td>
</tr>
</tbody>
</table>

The inclusion criteria encompassed primary and secondary studies, both qualitative and quantitative, including experimental and quasi-experimental studies, randomized and non-randomized studies, as well as observational, analytical articles, and cohort studies. Conference abstracts were excluded, and published studies were considered without restrictions on the publication period.

The research was conducted on 17 February 2023, in the MEDLINE Complete database via PUBMED, CINAHL Complete via EBSCO, and SCOPUS. Grey literature sources included OpenGrey, the Scientific Open Access Repository of Portugal (RCAAP) [Repositório Científico de Acesso Aberto de Portugal], and DART-Europe.

The bibliographic software ZOTERO 5.0.94 (Corporation for Digital Scholarship and Roy Rosenzweig Centre for & History and New Media, 2021), was used to collect and manage the search and remove duplicates. Subsequently, titles and abstracts were analysed by two independent reviewers to determine if they met eligibility criteria. Discrepancies were resolved by a third reviewer (Peters et al., 2022).

A table was used for data extraction, designed according to the study’s objectives and eligibility criteria was established. This process included the analysis of fundamental elements of the articles, such as the study type, objectives, methodology used, type of Machine Learning employed, sample characteristics, and results or performance indicators.

RESULTS

A total of 16 articles were identified in the databases and one in grey literature repositories, as shown in the diagram of the article selection process (Figure 1). No duplicates were identified. After reading the titles, three articles were excluded, and one article was excluded after reading the abstract. Full reading included 13 articles, with six articles being excluded: one related to the population, two based on the concept, and three based on the context.”
Machine learning tools in chronic disease management: scoping review

Figure 1
Diagram of the article selection process

The seven articles included in the quantitative data synthesis were published between 2020 and 2022, with the majority in the United States of America (USA) (three) and China (two). Table 2 provides characteristics of the studies included regarding authors, year, country, study objectives, and methodology.

Table 2
Characterization of the studies included with respect to authors, year, country, study objectives, and methodology

<table>
<thead>
<tr>
<th>Study/Country</th>
<th>Objective</th>
<th>Methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Alzghoul et al., 2020) USA</td>
<td>To develop a predictive model to assist in identifying potential future patients with pulmonary embolism (PE) and assessing the impact of acute PE on hospital outcomes</td>
<td>Retrospective Cohort Study</td>
</tr>
<tr>
<td>(Chatterjee et al., 2020) Norway</td>
<td>To provide a review of various machine learning methods for identifying risk factors for diseases such as obesity, cardiovascular diseases (CVD), and type 2 diabetes</td>
<td>Systematic Review and Meta-Analysis</td>
</tr>
</tbody>
</table>
The results regarding the machine learning tools described in the literature and used to manage chronic diseases are mapped in Table 2. The most addressed chronic disease was Chronic Kidney Disease (CKD); however, the tools were also applied to other pathologies such as Chronic Obstructive Pulmonary Disease (COPD), Hepatitis C (HCV), Heart Failure (HF), and Chronic Venous Insufficiency (CVI).

For a clearer understanding and to address the research objective, the results were grouped into specific categories relevant to nursing practice. Thus, areas of nursing intervention focusing on health promotion, disease prevention, and care during illness were considered: (1) identification of risk factors for the development of chronic diseases; (2) identification of risk factors for complications of chronic disease; (3) continuous monitoring and treatment adequacy; (4) improvement in healthcare delivery; (5) support for clinical decision-making.

### Table 3
Machine learning tools and their relevance to nursing practice

<table>
<thead>
<tr>
<th>Study</th>
<th>Machine Learning</th>
<th>Sample/Chronic disease</th>
<th>Results/Indicator</th>
<th>Relevance to nursing practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Alzghoul et al., 2020) USA</td>
<td>Random forest ML based on clinical variables: a prediction model</td>
<td>758 people hospitalized with asthma exacerbation</td>
<td>Predicting asthma exacerbation with Pulmonary Embolism (PE) and without PE. Accuracy of 88% in classifying the state of Acute Pulmonary Embolism</td>
<td>(1)</td>
</tr>
<tr>
<td>(Chatterjee et al., 2020) Norway</td>
<td>Digital eCoaching – capturing data related to potential risk factors</td>
<td>Not applicable</td>
<td>Risk associated with the development of obesity/excess weight</td>
<td>(2)</td>
</tr>
<tr>
<td>(Du et al., 2021) China</td>
<td>Takagi-Sugeno-Kang Fuzzy System-Based Model (G-TSKFS) for predicting dialysis adequacy</td>
<td>250 people with CKD, undergoing dialysis for four hours</td>
<td>Method for predicting dialysis adequacy</td>
<td>(3) (4) (5)</td>
</tr>
<tr>
<td>(Nakayama et al., 2021) USA</td>
<td>Decision tree and random forest models (Care Records, Antiviral)</td>
<td>Individuals born between</td>
<td>Indicates the prioritization of patients with more advanced diseases, illness, and/or</td>
<td>(2) (3) (4) (5)</td>
</tr>
</tbody>
</table>
Machine learning tools in chronic disease management: scoping review

<table>
<thead>
<tr>
<th>Study (Year)</th>
<th>Country</th>
<th>Methodology</th>
<th>Participants</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Chan et al., 2022) Singapore</td>
<td></td>
<td>Handheld 3-D infrared wound imaging device (WoundAide [WA] imaging system, Konica Minolta Inc, Tokyo, Japan)</td>
<td>52 people with CVI with chronic wounds</td>
<td>High intra- and inter-rater reliability achieved for the WA imaging systems. High reliability among raters for WA measurements compared to traditional measurement.</td>
</tr>
<tr>
<td>(Su et al., 2022) China</td>
<td></td>
<td>Six ML Models – Logistic Regression (LR), Lasso Regression, Random Forests (RF), Gradient Boosted Decision Tree (GBDT), Support Vector Machine (SVM), and Deep Neural Network (DNN) – Developed to Predict the Probability of Chronic Kidney Disease (CKD) Among the Elderly</td>
<td>925 seniors with CKD</td>
<td>The OML model can successfully capture the linear and non-linear relationships of risk factors for CKD in the elderly.</td>
</tr>
<tr>
<td>(Jørgensen et al., 2022) USA</td>
<td></td>
<td>Inferential Random Forests used for Multivariate Logistic Regression with the Top Three Chronic Diseases for each outcome adjusted for sociodemographic characteristics was conducted to quantify associations.</td>
<td>Data from 4,993 individuals (4,420 NH White and 573 NH Black) aged ≥ 65</td>
<td>Prediction of hospitalization and potentially avoidable hospitalization.</td>
</tr>
</tbody>
</table>

Caption: (1) Identification of risk factors for the development of chronic diseases; (2) Identification of risk factors for complications of chronic disease; (3) Continuous monitoring and adequacy of treatments; (4) Improvement in healthcare delivery; (5) Support for clinical decision-making

**DISCUSSION**

Data analysis of the findings confirms the importance of developing and using ML algorithms to create AI tools that assist in decision-making and the daily clinical practice of nursing care. Difficulties such as errors in diagnoses, inadequate treatments, wasted resources, inefficient workflows, among others, can be mitigated through the use of AI.

Through the use of ML models, it is possible to identify risk factors and symptoms that nurses may not have considered previously, thereby improving the accuracy of diagnosis and prognosis (Habehh & Gohel, 2021). In the study by Alzghoul et al. (2020), a prediction model for exacerbation of the disease was developed using the random forest technique, achieving an 88% accuracy.

The same technique was simultaneously used with logistic regression by Jørgensen et al. (2022) to predict hospitalizations and potentially avoid them.

Regarding the accuracy of the diagnosis, Su et al. (2022) demonstrated that a model composed of six ML algorithms—logistic regression, lasso regression, random forests, gradient boosted decision tree, support vector machine, and deep neural network—successfully establishes relationships between risk factors for Chronic Kidney Disease and the probability of diagnosis.
The use of ML models allows for a more precise assessment of patient data, identifying patterns and hidden risk factors that can contribute to a more accurate diagnosis and prognosis of chronic disease. This can lead to better disease management and an improvement in the patient’s quality of life (Habehh & Gohel, 2021).

Nurses play a prominent role in collecting and interpreting data used in ML models, emphasizing the need for training and education to make the most of available technologies (Habehh & Gohel, 2021). Personalization of care is crucial in managing chronic disease, as the needs of patients can vary widely. ML models enable a more personalized approach, considering the characteristics and medical history of each patient (Alzghoul et al., 2020; Chatterjee et al., 2020).

ML models can be used to analyse large amounts of data from individuals with chronic diseases, enabling healthcare professionals to develop more personalized treatment plans (Rajkomar et al., 2019).

The study by Du et al. (2021) presents a model based on the Takagi-Sugeno-Kang Fuzzy System (G-TSKFS) to predict dialysis adequacy. These models can be used to predict disease progression based on individual risk factors and help healthcare professionals understand which treatments work best for specific patients based on their medical histories and personal characteristics. This can lead to better treatment adherence and improved outcomes for patients (Rajkomar et al., 2019).

In the study conducted in the Southern U.S. (Nakayama et al., 2021), through decision tree and random forest models, it was possible to prioritize patients with Hepatitis C Virus (HCV), understand which ones were in a more advanced stage of the disease and/or complications, and thus define the next steps of treatment, such as transplantation.

With continuous monitoring, nurses can monitor patients’ health conditions in real-time and intervene quickly in case of changes in health indicators. That is what the study by Jørgensen et al. (2022) highlighted, as it managed to predict hospitalization and prevent potential hospitalizations in people with Chronic Kidney Disease (DRC), Heart Failure (HF), and Chronic Obstructive Pulmonary Disease (COPD).

The relevance of ML in continuous monitoring of the health of patients with chronic diseases allows for the early identification of signs of worsening conditions and timely interventions to prevent complications. This can be particularly important for nurses working in home care environments, where constant monitoring is essential for the safety and well-being of patients.

ML models can assist in clinical decision-making, especially in complex cases of chronic diseases, by providing nurses with relevant information about patients and their health conditions, enabling them to make more informed and accurate decisions. Additionally, ML models can be integrated into clinical decision support systems, providing evidence-based treatment recommendations and insights from previous experiences. This can help improve healthcare efficiency and reduce the risk of medical errors (Liao et al., 2018).

Through the use of a handheld 3-dimensional infrared wound imaging device, WoundAide [WA], wound measurements demonstrated high reliability among assessors, maintaining care uniformity and aiding in decision-making regarding the treatments to be applied (Chan et al., 2022).
With the assistance of ML models, nurses can enhance the quality of life for individuals with chronic diseases by providing personalized and effective care, helping them better manage their condition and avoiding unnecessary complications (Battineni et al., 2020). We can thus tailor treatments, such as dialysis techniques (Du et al., 2021), to each person’s home, prioritize treatments, and define new interventions (Nakayama et al., 2021), managing the scarce and finite human and material resources we know of, or even predict hospitalizations and potentially prevent them (Jørgensen et al., 2022), thereby improving the quality of life for patients.

Some limitations were identified in the review. The search was conducted in specific databases, which may have resulted in the exclusion of some ML tools available in the literature. Additionally, the literature review was limited to three languages, which may have excluded relevant studies in other languages.

It is important to emphasize that, although ML tools may hold promise for clinical practice, their implementation may require additional resources and training for healthcare professionals, aspects that were not addressed in this review.

CONCLUSION

The literature review identified several ML tools focused on chronic disease management that have the potential to contribute to the improvement of nursing care. The tools identified aim to identify risk factors for chronic diseases, early detection of chronic disease exacerbations, improvement in healthcare delivery, continuous monitoring, and evaluation of treatment effectiveness, as well as support for clinical decision-making.

With the advancement of ML tools, nurses need to stay up-to-date on the possibilities offered by these tools and how they can be applied in their clinical practice. The use of these tools can contribute to improving the accuracy of diagnosis and prognosis, personalizing care according to the patient’s needs, monitoring health in real-time, and making clinical decisions based on data. Nurses should be aware of potential challenges associated with the use of ML tools, particularly in terms of data privacy and security. It is crucial to ensure that their use complies with local regulations and safeguards the privacy of patient information.

For future research, it would be relevant to assess the effectiveness and acceptance of ML models by nursing professionals and individuals with chronic diseases. It would also be interesting to evaluate the impact of these tools on improving health outcomes for patients, comparing them with traditional care approaches. Furthermore, there is a need to explore how ML models can be integrated into nursing clinical practice and identify the key challenges and limitations associated with their large-scale implementation.

REFERENCES


Corporation for Digital Scholarship and Roy Rosenzweig Center for & History and New Media. (2021). Zotero | Your personal research assistant (Zotero 5.0.94). https://www.zotero.org/


