Review Title

The Use of Machine Learning in Occupational Risk Communication for Healthcare Workers – Protocol for scoping review

Authors

Gabriela Laudares Albuquerque de Oliveira[1] Clarice Alves Bonow[2] Itiberê de Oliveira Castellano Rodrigues[3] Amanda Xavier Geraldo[4]

ABSTRACT

Introduction: With the development of technology, the use of machine learning (ML), a branch of computer science that aims to transform computers into decision-making agents through algorithms, has grown exponentially. This protocol arises from the need to explore the best practices for applying ML in the communication and management of occupational risks for healthcare workers.

Methods and analysis: This scoping review protocol [1] details a search to be conducted in the academic databases Public Medical Literature Analysis and Retrieval System Online (PUBMED), through the Virtual Health Library (BVS): Medical Literature Analysis and Retrieval System (MEDLINE), Latin American and Caribbean Literature in Health Sciences (LILACS), Wanfang Data Co., Ltd. (WPRIM), Nursing Database (BDENF) and Scientific Electronic Library Online (SciELO), SCOPUS, Web of Science, IEEE Xplore Digital Library and Excerpta Medica Database (EMBASE). This scopingreview protocol outlines the objectives, methods, and timeline for a review that will explore and map the existing scientific evidence and knowledge on the use of machine learning (ML) in risk communication for healthcare workers. The protocol will follow the PRISMA-ScR [2] and JBI guidelines [3] and recommendations for protocol scoping reviews. The guiding question of the review is: How will Machine Learning be used in risk communication for healthcare workers? The search will use PCC (Population, Concept, Context) terms and the specific descriptors defined by each database. The narrative synthesis will describe the main themes and findings of the review.

The results of the scoping review will be disseminated through publication in an international peer-reviewed scientific journal.

Ethics and dissemination: Ethical approval is not required; data will rely on published articles. Findings will be published open access in an international peer-reviewed journal.

Strengths and limitations of this study

Strengths:

The study allows for comprehensive mapping of existing evidence on ML in occupational risk communication.

The methodology follows PRISMA-ScR and JBI guidelines, ensuring transparency and replicability.

The research employs a broad search strategy across multiple databases to capture relevant studies.

Limitations:

The accuracy of ML models is dependent on the quality of the data used.

The implementation of ML in healthcare requires careful evaluation of ethical, legal, and privacy issues.**Registration details OSF Registries:** The protocol for this review was registered in the Open Science Framework under DOI 10.17605/OSF.IO/92SK4 (available at https://osf.io/92SK4).

Keywords: Machine Learning, risk communication, occupational health and safety, healthcare workers.

INTRODUCTION

The International Labour Organization [4] The International Labour Organization estimates that approximately 2.3 million workers die annually due to work-related accidents or illnesses, which corresponds to more than 6,000 deaths per day. In Brazil, according to the Occupational Health and Safety Observatory, over 26,000 deaths and 6,804,060 Work Accident Notifications (CAT) were recorded between 2012 and 2022, with an average of approximately 618,000 accidents reported annually, while underreporting is estimated at nearly 19%. The observatory highlights priority areas, with Nursing Technicians being the occupation most affected by recorded Work Accidents (more than 313,000 between 2012 and 2022), followed by Nurses and Nursing Assistants, with 8,673 and 5,472 cases, respectively. These figures underscore the severity of this global issue. [5].

Therefore, the shortage of nursing professionals is recognized as a global health emergency [6], and unsatisfactory working conditions are directly linked to absenteeism and turnover [7 8 9].

Healthcare professionals face various occupational hazards, ranging from exposure to biological, chemical, physical, ergonomic substances to psychosocial risks such as shift work

and workplace violence [7]. In this context, risk communication plays a crucial role in informing, educating, and engaging workers about hazards and available protective measures [8 10].

The integration of artificial intelligence (AI) and Machine Learning in risk communication can yield significant benefits by enabling the analysis of large volumes of data to identify patterns and predict potential risks [9]. These technologies have the potential to enhance the safety of healthcare professionals, although their implementation requires ethical, legal, and privacy considerations[12 13 14].

Previous studies have shown that the application of Machine Learning (ML) algorithms, such as neural networks, not only increases the accuracy of risk prediction, but also facilitates the customization of communication strategies to meet the needs of different audiences. For instance, a study conducted with teachers found that it is possible to effectively predict cases of absenteeism due to morbidity, yielding promising results. The algorithm with the best predictive performance, according to standard ML evaluation metrics, was artificial neural networks, achieving an area under the curve of 0.79, with an accuracy of 71.52%, sensitivity of 72.86%, and specificity of 70.52% [15].

In Brazil, although there are no specific AI strategies for risk communication in the healthcare sector yet, it is crucial for this to be an area of interest to ensure the safety and well-being of workers and to anticipate safer practices in the face of imminent risks[15].

A preliminary search was conducted in The Cochrane Library, Open Science Framework, Database of Abstracts of Reviews of Effects (DARE), and the International Prospective Register of Systematic Reviews (PROSPERO), and no reviews or studies similar to this research were found.

This scoping review aims to explore the potential of AI in risk communication for healthcare professionals by analyzing relevant studies and applications.

The objective of this work is to understand how AI, particularly Machine Learning, can contribute to informing and protecting healthcare professionals regarding various risks, including biological, chemical, ergonomic, psychosocial, physical, among others. Through this approach, the aim is to promote a safer and healthier work environment for these professionals.

This scoping review aims to explore the potential of AI in risk communication for healthcare professionals by analyzing relevant studies and applications

METHODS AND ANALYSIS

The choice of a scoping review to address the topic "Use of Machine Learning in risk communication for healthcare workers" is justified by the need to explore and map the available scientific evidence and existing knowledge in this specific area. This protocol and review follow the guidelines and recommendations [1] of the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist [2] and JBI (Joanna Briggs Institute) [3]. The scoping review will meet the premises of

this research, as it does not intend to assess the quality of available evidence but rather to obtain a representation of the findings.

In the specific case of the scoping review on the use of Machine Learning in occupational risk communication for healthcare workers, this methodology will allow for a comprehensive and systematic search of different types of information sources, such as scientific articles, technical reports, and government guidelines. By adhering to the guidelines and recommendations of Prisma [2] or JBI [3], the scoping review will ensure transparency, rigor, and replicability in all stages of the process, from formulating the research question to analyzing and synthesizing the findings. This will help provide a comprehensive and up-to-date overview of the use of Machine Learning in occupational risk communication for healthcare workers, contributing to informed decision-making and the development of effective interventions in this area.

The review process [1] will follow the steps recommended by PRISMA-ScR [2], including Introduction, Review Question, Inclusion Criteria (Participants, Concept, Context), Types of Sources, Methods, Search Strategy, Study/Source of Evidence Selection, Data Extraction and Data Analysis and Presentation.

Review question

The research question was developed through alignment with the core elements/PCC framework, P = Population/ Participants, C = Concept, C = Context.

The PCC methodology is a widely used approach in scoping reviews to outline the characteristics of the population of interest, the relevant concepts or phenomena, and the context in which they occur, as recommended in the JBI Manual for Scoping Reviews [3]. In this study, we applied the PCC methodology to identify key elements related to the communication of occupational risks for workers exposed to hazardous work environments.

P - Population: Healthcare workers exposed to occupational risks.

C - Concept: Machine Learning is the conceptual technique to be explored in this scoping review.

C - Context: Occupational risk communication. The aim is to examine how effective communication of occupational risks can contribute to accident and work-related illness prevention, as well as to identify barriers and challenges encountered in this process.

With this mnemonic combination, the following guiding question was defined:

How has the use of Machine Learning been applied in risk communication for healthcare workers?

This question provides a clear direction for investigating the application of artificial intelligence in risk communication for healthcare workers and evaluating its outcome.

Before conducting this study, searches were conducted on The Cochane Library, Open Science Framework, Database of Abstracts of Reviews of Effects (DARE), and the International Prospective Register of Systematic Reviews (PROSPERO) websites to locate similar reviews or studies and prevent study duplication. No similar studies were found and this scoping review protocol was registered in the Open Science Framework under a protocol: DOI 10.17605/OSF.IO/92SK4 (available at https://osf.io/92SK4).

Google Scholar searches [13 14 16 17 18] related to the topic were conducted using the PCC keywords to identify the most recurrent descriptors and keywords. These descriptors and keywords were then listed in a spreadsheet according to Table 1 to initiate searches in PubMed.

Table 1 - Searches in PubMed - Pelotas, RS, Brazil, 2024.			
How has the use of Machine Learning been applied in risk communication for healthcare workers?			
		Descriptors in DeCS/MeSH for PubMed search	
P	POPULATION	Healthcare Workers	"Health Personnel" OR "Personnel, Health" OR "Health Care Providers" OR "Health Care Provider" OR "Provider, Health Care" OR "Providers, Health Care" OR "Healthcare Providers" OR "Healthcare Provider" OR "Provider, Healthcare" OR "Healthcare Workers" OR "Healthcare Worker" OR "Providers, Healthcare" OR "Professional, Health Care" OR "Health Care Professional" OR "Health Care Professionals" OR nurses OR doctors OR "medical professionals" OR "healthcare team"
			AND
C	CONCEPT	Machine learning	"Machine Learning" OR "artificial intelligence" OR "Decision support systems OR expert system" OR "Cognitive system" OR "Learning, Machine" OR Algorithms OR "Deep Learning" OR "Neural Networks" OR "Natural Language Processing" OR "Cognitive Computing" OR "Data Mining" OR "Adversarial machine learning" OR Boosting OR "Dimensionality reduction" OR "Ensemble learning" OR "Federated learning" OR "Hyperparameter optimization" OR "Multi-armed bandit problem" OR "Random forests" OR "Reinforcement learning" OR "Relevance vector machines" OR "Representation learning" OR "Robot learning" OR "Large Language Models" OR "Computational and artificial intelligence" OR "Computational intelligence" OR "Feedforward neural networks" OR "Generative adversarial networks" OR "Minimax techniques" OR "Natural languages" OR "Robot learning" OR "Support vector machines" OR "Affective computing" OR "Autonomous robots" OR "Cognitive systems" OR "Decision support systems" OR "Intelligent systems" OR "Knowledge based systems" OR "Learning systems" OR "Prediction methods" OR "Virtual artifact"
			AND
C	CONTEXT	Risk Communication	"Risk m+A1:E7anagement" OR "Occupational risks" OR "Chemical risks" OR "Ergonomic risks" OR "Psychosocial risks" OR "Physical risks" OR "Biological risks" OR "Protective measures" OR "Risks mitigation" OR "Risks communication" OR "Risks perception" OR "Occupational risk" OR "Chemical risk" OR "Ergonomic risk" OR "Psychosocial risk" OR "Physical risk" OR "Biological risk" OR "Protective measure" OR "Risk mitigation" OR "Risk communication" OR "Risk perception" OR "occupational safety and health" OR "Workers health" OR "Employee Health" OR "Workplace safety" OR "Worker Safety" OR "Safety at work" OR "Professional Risk" OR "Professional Risks" OR "Hazard

			prediction" OR "work environment" OR "Communication system operations and management" OR "Occupational health" OR "Occupational safety"

Types of sources and inclusion and exclusion criteria

Inclusion and exclusion criteria will follow the dimensions outlined by the elements of the acronym PCC, with the Population (P) being healthcare workers, nurses, physicians, thus studies dealing with other populations (such as patients) will be excluded. In the Concept (C) element, all studies presenting the use of Artificial Intelligence, especially Machine Learning, will be included, while studies presenting other tools in their concept will be excluded. For Context (C), studies presenting communication and strategies for occupational risk management will be considered, while studies focused on other communication contexts (such as clinical decision support, patient safety, among others) will be excluded.

The criterion defined for selecting the databases used PUBMED, MEDLINE, LILACS, WPRIM, BDENF, SciELO, SCOPUS, Web of Science, EMBASE, IEEE Xplore, will be the availability to search for articles through search engines with support for descriptors and boolean operators, as they are updated databases. Studies published in all languages will be selected. After this step, the snowballing technique will be used to review the references of all included articles to identify other studies that may also meet the selection criteria.

The general inclusion criteria for articles will be: articles published in the last 5 years, as Machine Learning is constantly evolving and has seen increased application, especially in the healthcare field. The general exclusion criteria will be: grey literature (theses and dissertations, conference proceedings, reports, among others). Grey literature will not be considered in this work, as scientific articles were prioritized to ensure scientific equity is achieved in journal publications.

Search Strategy

The electronic search will be conducted between May and June/2024, using Health Science Descriptors (DECS) in English, or Medical Subject Headings (MESH), terms found in the IEEE Thesaurus and EMTREE Thesaurus (Embase). These terms will be used to search the following databases: Public Medical Literature Analysis and Retrieval System Online (PUBMED), through the Virtual Health Library (VHL): MEDLINE, Latin American and Caribbean Health Sciences Literature (LILACS), WPRIM, BDENF, and Scientific Electronic Library Online (SciELO), SCOPUS, Web of Science, IEEE Xplore, and EMBASE.

As per Table 1. After each descriptor/keyword/search strategy search, titles of the found articles will be read, looking for words related to the researched topic, following the PCC strategy. When the title and abstract are unclear, introduction and conclusion will be read.

Study Selection

The description and summarization of data will be conducted by two independent reviewers for reading titles, abstracts, and keywords to find studies that meet the inclusion criteria. In case of any disagreement, a third reviewer will be called upon to analyze and decide on the inclusion or exclusion of articles.

Zotero and Rayyan managers will be used to identify duplicate articles and organize eligible or non-eligible studies.

Data Extraction

For information mapping, data collection will be conducted using an instrument adapted from the JBI [2] form, developed by the authors in Microsoft Excel to record relevant study categories, as shown in Table 2.

Table 2: Data items to be collected

Table 2 Data items to be collected	
Category	**Data**
Bibliographic Information:	Author(s)
	Title of the study
	Year of publication
	Journal or source name
	Digital object identifier (DOI) or International standard book number (ISBN) (if applicable)
Study characteristics:	Study design
	Study type
	Setting (location)
	Sample size and population studied
	Aims and objectives
Data sources:	Data sources used
Population (P) Information:	Population studied
	Sample size
Concept (C) Information:	Machine Learning Techniques
	Inputs
	Outputs
	Type of problem
	Description of interventions or exposures
	Metrics for evaluating the model
	Algorithms performance

Context (C) Information	Type of risk
	Detailing the type of risk
Outcomes	Key findings and results
	Summary of main findings
	The best performing model
	Observed associations, relationships, or legal or ethical effects
	Discussion and conclusions
	Limitations of the study/approach as mentioned by authors
	Level of evidence

Planned dates

The study is planned to begin in May 2024 and end in December 2024.

Author affiliations

[1]Social Communicator. Master in Administration. PhD student in Nursing. Universidade Federal de Pelotas (UFPEL). Email: gabiabuquerque83@gmail.com. https://orcid.org/0000-0001-5053-9993

[2]Nurse. PhD in Nursing. Universidade Federal de Pelotas (UFPEL). Email: claricebonow@gmail.com. https://orcid.org/0000-0001-9918-9234

[3]Doctor of Law. Professor in the area of State Law at the Faculty of Law of the Universidade Federal de Pelotas (UFPEL). Email: itibere.rodrigues@hotmail.com. https://orcid.org/0000-0002-8665-4149

[4]Nurse. Master's student in Nursing. Universidade Federal de Pelotas (UFPEL). Email: amanda.xg@hotmail.com. https://orcid.org/0009-0003-4718-8434

Acknowledgements

This review is part of a PhD trajectory of Oliveira, GLA.

Author contributions

GLAO, AXG, CAB, and IOCR contributed to the development of this study protocol. CAB and GLAO proposed the study, which was then developed by GLAO and AXG with the help of all the authors. GLAO, AXG, CAB, and IOCR discussed the results of the exploratory search and prepared the supporting theoretical framework. CAB and IOCR provided specialized knowledge in Risk Communication and legal aspects of

Artificial Intelligence use, respectively. GLAO identified the search terms and established the search strategy together with CAB and the university library. The scoping review planning will follow the following strategy: GLAO and AXG will be responsible for data collection and will read the titles and abstracts of the articles found after each search for descriptors/keywords, looking for words related to the research topic, according to the PCC strategy. When the title and abstract are not clear, the introduction and conclusion or the entire study will be read if necessary, and if there is any disagreement, CAB will be called to analyze and decide whether or not to include the articles. After this step, the articles and the work will be divided among the authors for screening and secondary analysis. GLAO will prepare the first drafts of the manuscript. The final manuscript will be prepared through interactive refinement and group discussions between GLAO, AXG, CAB, and IOCR.

Funding

GLAB is supported by a doctoral scholarship [[141906/2023-5] from the National Council for Scientific and Technological Development (CNPq).

Competing interests

None declared.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication

Not applicable.

Ethics and dissemination

This review does not require ethical approval. Our dissemination strategy includes publishing in peer-reviewed journals, presenting at conferences, and sharing findings with relevant stakeholders.

Data availability statement

No data are available.

REFERENCES

1. Gupta R, Gunjawate DR, Nguyen DD, Jin C, Madill C. Voice disorder recognition using machine learning: a scoping review protocol. BMJ open. 2024 Feb 1;14(2):e076998.
2. PRISMA for Scoping Reviews (PRISMA-ScR). Available online: https://www.prisma-statement.org/scoping. PRISMA-ScR (accessed on 03 February 2024).
3. Aromataris E, Munn Z (Editors). JBI Manual for Evidence Synthesis. JBI, 2020. Available: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-20-01 [Accessed 10 Feb 2024]
4. The International Labour Organization (ILO). World Statistic. Available: https://www.ilo.org/moscow/areas-of-work/occupational-safety-and-health/WCMS_249278/lang--en/index.htm [Accessed 03 Jan 2024].
5. Superior Labor Court (TST). Available: https://www.tst.jus.br/web/trabalhoseguro/observatorio-sst disponível em: https://smartlabbr.org/sst [Accessed 03 Jan 2024].
6. International Council of Nurses. Occupational health and safety for nurses. Geneva, 2017. Acessado em 20 jun. 2024. Online. Disponível em: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.icn.ch/sites/default/files/2023-04/PS_C_Occupational_health_safety_0.pdf
7. World Health Organization et al. Communicating risk in public health emergencies: a WHO guideline for emergency risk communication (ERC) policy and practice. World Health Organization, 2017.
8. U.S. EPA, 2007. Integrated risk information system (IRIS), record for arsenic. Available online: www.epa.gov/iris/search.htm. U.S. Environmental Protection Agency. [Accessed on January 2024].
9. American Nurses Foundation. COVID-19 impact assessment survey the second year. 2022. Acessado em 10 jun. 2024. Online. Disponível em: https://www.nursingworld.org/practice-policy/work-environment/health-safety/disaster-preparedness/coronavirus/what-you-need-to-know/covid-19-impact-assessment-survey---the-second-year/
10. U.S. EPA, 1991. Risk assessment guidance for Superfund. Volume I: Human health evaluation manual supplemental guidance. Standard default exposure factors. Interim final.
11. Lundgren, Regina; Mcmakin, Andrea. Risk communication: a handbook for communicating environmental, safety and health risks. Ohio: Battelle Press, 2000.
12. Touzet, Sandrine et al. Impact of a comprehensive prevention programme aimed at reducing incivility and verbal violence against healthcare workers in a French ophthalmic emergency department: an interrupted time-series study. BMJ open, v. 9, n. 9, 2019
13. Zhang, Xin et al. A prediction model of sleep disturbances among female nurses by using the BP-ANN. Journal of Nursing Management, v. 27, n. 6, p. 1123-1130, 2019
14. Rangel, Maria L. Comunicação no controle do risco à saúde e segurança na sociedade contemporânea: uma abordagem interdisciplinar. Ciência e Saúde Coletiva, Rio de Janeiro, v. 12, n. 5. set./out. 2007. Available:

<http://www.scielo.br/scielo.php?pid=S141381232007000500035&script=sci_abstract&tlng=pt> [Accessed 06 May 2024].

15. Fernandes, Fernando T and Filho, Alexandre DPC. "Prediction of teacher absenteeism in public schools with machine learning." Revista de Saúde Pública 55: 23, 2021.
16. Climent-Rodríguez JA, Navarro-Abal Y, García-Iglesias JJ, Vaca-Acosta RM, Ortega-Moreno M, Gómez-Salgado J. Violencia en el trabajo y compromiso laboral en los profesionales de enfermería en España: un estudio transversal [Violence at work and work engagement against nursing professionals in Spain: a cross-sectional study.]. Rev Esp Salud Publica. 2023 Jan 16;97:e202301003. Spanish. PMID: 36655388; PMCID: PMC10541725.
17. Shishavan, Hossein Hamidi et al. Continuous physiological signal measurement over 24-hour periods to assess the impact of work-related stress and workplace violence. Applied ergonomics, v. 108, p. 103937, 2023.
18. Souza, Fernanda Mattos et al. Classification and regression trees for predicting the risk of a negative test result for tuberculosis infection in Brazilian healthcare workers: a cross-sectional study. Revista Brasileira de Epidemiologia, v. 24, p. e210035, 2021.