IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version April 24, 2025.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

On Usage of Artificial Intelligence for Predicting Neonatal Diseases, Conditions and Mortality: A Bibliometric Review

FLÁVIO LEANDRO DE MORAIS[1], Raysa Carla Leal Da Silva[2], Anna Beatriz Silva[1], Estefani Pontes Simão[1], Maria Eduarda Ferro de Mello[1], Stephany Paula da Silva Canejo[3], Katia Maria Mendes[4], Waldemar Brandão Neto[3,4], Jackson Raniel Florencio da Silva[2], Maicon Herverton Lino Ferreira da Silva Barros[1](Member, IEEE), Patricia Takako Endo[1,2]

[1]Programa de Pós-Graduação em Engenharia da Computação (PPGEC), Universidade de Pernambuco (UPE), Recife, Pernambuco, Brazil (e-mails: flavio.leandromorais@upe.br; anna.beatrizs@upe.br; eps@ecomp.poli.br; mefm@ecomp.poli.br; mhlfsb@ecomp.poli.br; patricia.endo@upe.br)
[2]Universidade de Pernambuco (UPE), Caruaru, Pernambuco, Brazil (e-mail: raysa.silva@upe.br; jackson.florencio@upe.br)
[3]Programa Associado de Pós-Graduação em Enfermagem (PAPGenf), Universidade de Pernambuco, Recife, Pernambuco, Brazil (e-mail: stephany.canejo@upe.br; waldemar.neto@upe.br)
[4]Centro Universitário Integrado de Saúde Amaury de Medeiros (CISAM), Universidade de Pernambuco (UPE), Recife, Pernambuco, Brazil (e-mail: katia.mendes@upe.br)

Corresponding author: Patricia Takako Endo (e-mail: patricia.endo@upe.br).

ABSTRACT **Purpose:** Care and attention during the neonatal period are crucial to preventing negative outcomes. The literature presents artificial intelligence models as promising tools to assist healthcare professionals in disease prediction and support clinical decision-making. **Methods:** This study conducts a bibliometric review of the use of artificial intelligence models in predicting neonatal diseases, conditions and mortality. The review analyzed publications from 2014 to 2024. A total of 629 studies were selected after applying selection criteria. Subsequently, analyses of collaboration networks, keyword co-occurrence, citations and cluster analysis were performed. **Results:** The results show that the United States, China and the United Kingdom lead scientific production and international collaborations. 12 neonatal diseases were identified, with emphasis on ‘‘retinopathy of prematurity’’, ‘‘necrotizing enterocolitis’’ and ‘‘bronchopulmonary dysplasia’’; 7 clinical conditions, including ‘‘prematurity’’, ‘‘perinatal asphyxia’’ and ‘‘jaundice’’; and 5 neonatal outcomes, mainly ‘‘sepsis’’, ‘‘mortality’’ and ‘‘cerebral palsy.’’ Cluster analysis revealed that studies predominantly use clinical, laboratory, genetic and imaging data, with Logistic Regression, Random Forest and Convolutional. **Conclusion:** The study has growing interest in applying artificial intelligence to neonatal care. The models are increasingly used with clinical, laboratory, genetic and imaging data, enabling earlier and more accurate diagnoses. However, the study also underscores important ethical considerations, such as data quality, algorithmic transparency and equitable access to these technologies, particularly in underrepresented regions, with scientific production uneven and limited participation from low- and middle-income countries.

INDEX TERMS artificial intelligence, Bibliometric review, neonatal healthcare

I. INTRODUCTION

THE neonatal period encompasses the time from birth to the 28 days of life [1]. During this time, the newborn is in the phase of adapting to the extrauterine environment and is exposed to numerous physiological and environmental risks that may compromise their health and survival. According to the United Nations Children's Fund (UNICEF), in 2022, approximately 2.3 million deaths occurred in the neonatal period worldwide [2]. Of these, 75% were recorded in the first week of life (early neonatal death), with approximately 1 million neonates dying within the first 24 hours after birth. The remaining 25% occurred between 7 and 28 days of life (late death) [2]. The majority of negative neonatal outcomes occur in poor or developing countries, where avoidable factors such as limited access to healthcare, poor sanitation, socioeconomic vulnerability and other adverse factors [3] [4].

According to the World Health Organization (WHO) [5], factors that influence the vulnerability of neonatal health include biological conditions, such as prematurity and low birth weight, as well as issues related to the environment and access to adequate medical care. Furthermore, serious problems such as sepsis, necrotizing enterocolitis and respiratory diseases (bronchopulmonary dysplasia, pulmonary hypertension) are among the main causes of neonatal morbidity and mortality; other significant contributors include retinopathy of prematurity, hemorrhages, congenital malformations and various other health conditions [6] [7] [8].

The Sustainable Development Goal (SDG) 3, promoted by the United Nations (UN), aims to ensure healthy lives and promote well-being for all. For instance, it emphasizes maternal and child healthcare access in underprivileged regions. One of its goals is to eliminate avoidable neonatal and child deaths under the age of five in all countries by 2030 [9]. This goal reflects the urgency of addressing neonatal health challenges, particularly in low and middle-income countries, where neonatal mortality is more prevalent due to a lack of resources and public policies focused on promoting maternal health, equitable access to healthcare services and reducing socioeconomic disparities.

The current literature has highlighted artificial intelligence models as effective tools to address challenges related to improving maternal and neonatal health, characterized by their low costs associated with implementation and maintenance, making them accessible even in resource-limited settings [10]. Henry et al. [11] and Collin et al. [12] describe artificial intelligence models as promising instruments to assist healthcare professionals in predicting conditions, diseases and mortality, contributing to a more efficient and accurate approach, given that their outcomes provide healthcare professionals with support for risk prediction, early diagnosis and evidence-based clinical decision-making.

Although the current literature has explored the application of artificial intelligence in healthcare, there are still no specific bibliometric reviews on the use of these models in predicting neonatal conditions, diseases and mortality. The closest study to this topic was done by Lestari et al. [13], which focused on the usage of artificial intelligence in predicting obstetric complications affecting neonatal health, offering insights into early interventions during pregnancy. However, it does not directly address neonatal outcomes. Zhou et al. [14] conducted a bibliometric analysis of scientific production related to bronchopulmonary dysplasia in preterm neonates. The authors highlighted trends and risk factors associated with this condition, providing valuable data for clinical strategies to prevent and manage neonatal disease and morbidity. And Zhang et al. [15] investigated research trends on retinopathy of prematurity, including the usage of artificial intelligence. Their work emphasizes advancements in diagnostic precision and treatment approaches for this condition, though it primarily focuses on these aspects rather than predicting neonatal outcomes.

Otherwise, our study conducts a bibliometric review to investigate the use of artificial intelligence in predicting neonatal diseases, conditions and mortality. The review intends to identify key studies, influential authors and prominent collaboration networks, besides mapping research trends in the field. The most cited approaches, the most frequently used artificial intelligence models and the areas of greatest prominence in their application to neonatal health predictions will be analyzed. We also identify patterns and gaps in the existing literature that may provide a comprehensive view of the current state of artificial intelligence applications in neonatal healthcare and can be used as a start point for new research projects.

II. RELATED WORK

Lestari et al. [13] (2024) conducted a bibliometric analysis (2004-2023) on the application of intelligent techniques, such as machine learning, artificial intelligence and neural networks, in predicting obstetric complications affecting maternal and neonatal health. The authors analyzed 419 English studies from Web of Science using R language and VOSviewer software. The analysis revealed a trend of increasing publications starting from 2015. Analyzing co-author ships among countries, there was a strong presence of authors from China and the United States. Factor analysis demonstrated that "preeclampsia," "hypertensive disorders," and "epidemiology" were terms closely related to the concept of "prediction." Additionally, the growing focus on conditions like preeclampsia and gestational diabetes mellitus underscores the need for predictive tools to improve maternal and neonatal outcomes.

The study by Zhou et al. [14] (2021) presented a bibliometric analysis of bronchopulmonary dysplasia in preterm neonates, using the Web of Science database and the Cite Space software to identify trends in research output, collaborations and citation networks. A total of 610 articles published between 2008 and 2020 were analyzed, allowing the identification of research areas, influential authors, high-impact journals and institutions with the highest scientific output in the field of bronchopulmonary dysplasia. The objective is to understand the evolution of research, identify knowledge gaps and support the formation of new hypotheses and research outcomes on bronchopulmonary dysplasia in preterm infants, contributing to advancements in understanding and treating the disease. The results indicated that the United States leads research in the field, both in the number of publications and in institutional collaborations. The analysis also revealed that the main topics studied include risk factors such as prolonged ventilation, infection, inflammatory responses and pulmonary hypertension, as well as emerging therapeutic approaches, such as the use of low-dose corticosteroids, caffeine and mesenchymal stem cells.

The study by Zhang et al. [15] presented a bibliometric analysis of retinopathy of prematurity between 2003 and 2022, using the Web of Science Core Collection database and the CiteSpace software to identify research trends, collaborations and citation networks. A total of 4,957 publications were analyzed, revealing that the United States leads in the number of publications, with Harvard University as the most productive institution. The analysis highlighted seven main thematic clusters, including genetic polymorphism, neurodevel-

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

opmental outcomes, threshold retinopathy, oxygen-induced retinopathy, low birth weight, prematurity diagnosis and artificial intelligence. The analysis of emerging keywords revealed that terms such as "ranibizumab," "Deep Learning," and "artificial intelligence" remained prominent until 2022, indicating research trends in the field.

Another relevant and current topic regarding clinical complications is Neonatal Abstinence Syndrome (NAS). This condition is characterized by a set of signs and symptoms exhibited by neonates due to substance use or abuse during pregnancy. Zyoud et al. [16] (2022) examined global research trends on NAS, using the Scopus database and VOSviewer software to map research trends, map collaboration networks, citation impact and research priorities. The study analyzed 1,738 articles published between 1958 and 2019, with the majority being original articles (74.5%), followed by systematic reviews (15.4%) and letters to the editor (2.76%). The results showed that the United States leads publications on the topic, followed by Canada and the United Kingdom. The analysis revealed that the main research focuses include the clinical outcomes of NAS, the effectiveness of pharmacological treatment and evidence-based interventions for neonatal management. The study highlighted the need for further research on the long-term impacts of NAS and the development of standardized treatment protocols, aiming to improve neonatal outcomes and reduce associated complications.

Maulana et al. [17] (2024) analyzed research trends (2014-2023) on the use of machine learning in obstetric care, based on 829 studies from the Web of Science. It was observed that China and the United States had the highest number of author collaborations, with Iran and Saudi Arabia also standing out. The highest number of collaborations occurred between countries in North America, Europe and Asia. Through keyword co-occurrence analysis, recent literature emphasizes cost analyses and economic evaluations of interventions in obstetric care. Additionally, there is a notable focus on machine learning applications in fetal monitoring. The analysis of trend topics highlights an increased focus on 'preterm birth,' 'systematic review,' and 'meta-analysis.' These findings indicate a growing emphasis on using predictive analyses and risk assessments to optimize obstetric care practices.

Our study differs from the current literature because (a) we conduct a comprehensive bibliometric analysis on the use of artificial intelligence in predicting diseases, conditions and neonatal mortality, utilizing techniques to map global trends, collaboration networks, word analysis and research impact in the field; (b) we perform a temporal analysis to examine the evolution of terms over time, aiding in the identification of changes and emerging trends; (c) we apply cluster analysis to identify thematic cores and their interrelations, revealing connections between subtopics and enabling a deeper understanding of the evolution of the literature on the application of models in neonatal health; (d) we systematize and organize dispersed knowledge, providing a comprehensive and in-depth overview of the main research lines and their impacts on the use of artificial intelligence models, both in improving neonatal health quality and in identifying research gaps in the field.

III. MATERIALS AND METHODS

To conduct this bibliometric review, an integrated methodology was used, as shown in Figure 1, with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [18], applied in specific stages to ensure a standardized systematization of the selection and analysis of studies.

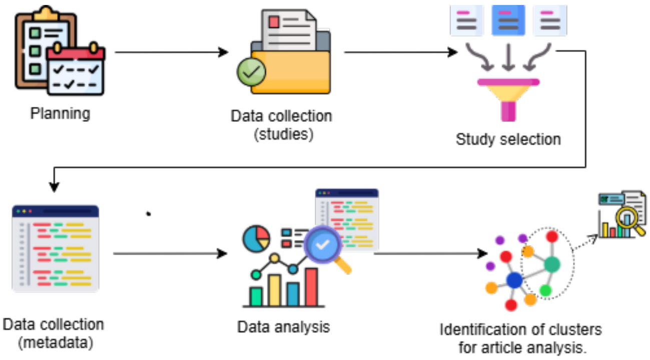


FIGURE 1. Research methodology.

A. PLANNINGS

1) Research questions

The initial planning phase involves formulating research questions relevant to the context of the proposed bibliometric review and outlining all the processes necessary for its execution. The main purpose of this work is to answer the following Research Questions (RQ):

- RQ1: What are the more influential studies in the prediction of neonatal diseases, conditions and mortality?
- RQ2: What are the strongest collaboration networks in the study for the prediction of neonatal diseases, conditions and mortality research?
- RQ3: What are the more influential institutions in the prediction of neonatal diseases, conditions and mortality?
- RQ4: What are the main conceptual themes and how have they evolved in the prediction of neonatal diseases, conditions and mortality?
- RQ5: What types of neonatal diseases, conditions and mortality are currently being studied in research using artificial intelligence?
- RQ6: What are the characteristics of the datasets used by studies using artificial intelligence models in predicting diseases, conditions and neonatal mortality?
- RQ7: What are the main attributes most frequently considered by studies using artificial intelligence models in predicting neonatal diseases, conditions and mortality?
- RQ8: What artificial intelligence models are used to predict neonatal diseases, conditions and mortality?

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

- RQ9: How is the performance of artificial intelligence models in predicting neonatal diseases, conditions and mortality evaluated?

2) Engines and Search String

The search engines were selected based on their scientific recognition and association with the area of information technology and health. Regarding the amount of search engines used, it's believed that they were adequate for the needs of the study. The search engines used in this study were:

- Association for Computing Machinery Digital Library (ACM Digital Library) [19]
- Institute of Electrical and Electronics Engineers (IEEE) [20]
- Institute for Scientific Information Web of Science (ISI Web of Science) [21]
- PubMed National Library of Medicine (PubMed) [22]
- Science Direct [23]
- Scopus [24]

The selection of these databases considered their relevance to artificial intelligence and neonatology, such as IEEE, which includes journals such as IEEE Transactions on Medical Imaging and IEEE Journal of Biomedical and Health Informatics, which are references in studies in the area. The ACM Digital Library, in addition to its contribution to computing, indexes journals like ACM Transactions on Computing for Healthcare, which addresses artificial intelligence in medical diagnostics. The ISI Web of Science includes journals like the Journal of Biomedical Informatics and Artificial Intelligence in Medicine, which are essential for studies on artificial intelligence in healthcare. PubMed was chosen for indexing essential biomedical journals like The Journal of Pediatrics and Pediatric Research. ScienceDirect includes renowned journals like Artificial Intelligence in Medicine, while Scopus, one of the largest scientific databases, indexes journals like Computer Methods and Programs in Biomedicine.

Right away, the search string was defined, which was established from keywords suggested and validated by experts and researchers in the areas of artificial intelligence and data science applied to health; as well as specialist nurses and researchers in neonatology and also in care in the Neonatal Intensive Care Unit (NICU). Subsequently, synonyms of the keywords that these experts validated were identified and then the keywords were combined and the sensitivity test was performed with their respective synonyms. Finally, the collection of primary studies was performed through searches in almost all databases, using the following search string:

("neonatal" OR "newborn" OR "preterm infants" OR "neonate" OR "perinatal") AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("prediction" OR "prognosis" OR "anticipation" OR "classification" OR "monitoring")

Except for the Science Direct source, which has specific restrictions on the use of Boolean connectors in its search queries. In particular, the platform allows the use of a maximum of eight boolean connectors per search field. This limitation required a strategic approach to the construction of searches, in which the search terms were adjusted to:

("neonatal" OR "newborn" OR "neonate" OR "perinatal") AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("prediction" OR "classification")

3) Selection Criteria

The selection criteria identify the primary studies that provide evidence related to the objective of this research and are divided into inclusion and exclusion criteria, as detailed below:

- Inclusion Criteria
  - Primary studies that have been published in conferences or peer-reviewed journals.
  - Primary studies are presented and available in full article format (but not books, posters, tutorials or editorials).
  - Primary studies that answer one or more of our research questions.
  - Primary studies published within the period defined for the research (from 2014 to 2024).
- Exclusion Criteria
  - Studies that do not directly contribute to our research questions.
  - Studies classified as secondary or tertiary, such as systematic reviews, meta-analyses or case studies.
  - Studies that are not written in English.
  - Duplicate studies.

B. DATA COLLECTION (STUDIES)

The collection considers only complete studies, considering the period from January 2014 to May 2024, an interval that guarantees a balanced analysis with the historical depth, updated and recent scientific and technological advances of artificial intelligence models and also neonatal health. Table 1 presents the results of the search for each database.

TABLE 1. Amount of studies by the search engine.

Source	Amount
ACM Digital Library	440
IEEE	386
ISI Web of Science	1,037
PubMed	998
Science Direct	2,555
Scopus	1894
TOTAL	7,310

C. STUDY SELECTION

In this phase, we identify and remove 2,133 duplicate studies. Eight reviewers were then responsible for identifying eligible papers independently. In cases of uncertainty regarding the acceptability of a study, it was submitted to a joint evaluation by all reviewers and considered accepted if it received a majority of favorable votes (at least five votes). Initially, the titles, keywords and abstracts of the studies were analyzed, simultaneously applying the inclusion and exclusion criteria.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

Finally, 635 primary studies were selected. Figure 2 presents the PRISMA flowchart used to summarize the identified studies that were accepted or excluded due to duplication or selection criteria.

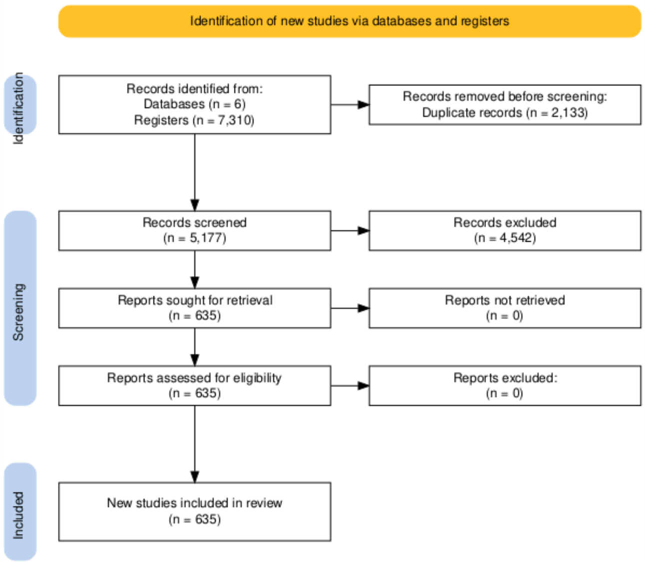


FIGURE 2. **Research methodology.**

D. DATA COLLECTION (METADATA)

In this phase, we collected bibliometric information (metadata) from the 635 selected studies. Initially, we collected 491 metadata directly from the Web of Science database. However, 138 studies whose metadata were not available in this database were obtained through the Scopus database. During the process of collecting metadata in both databases, it was not possible to obtain metadata from six studies. Despite the attempt to find this metadata by searching Google Scholar, it was also not possible to obtain the metadata due to the incompleteness of the information available on this platform. Thus, the final metadata database resulted from 629 primary studies. This consolidated set of metadata was structured into a final database, covering information such as title, author, affiliation, keywords, citations, abstract and other relevant information, which was the database for bibliometric analysis.

E. DATA ANALYSIS

The bibliometric analysis of the metadata was conducted systematically and comprehensively, encompassing everything from data preprocessing to the generation of thematic networks and maps [25]. The analysis began with a descriptive stage, seeking to identify publication patterns, including annual growth, document types, average citations per document and average age of publications. The main publication outlets were explored, highlighting the journals with the highest number of studies and impact measured by citations. The characterization also included information on the most productive authors, institutions and countries, both in individual publications and in international collaborations.

Co-word analyses of keywords [26] [27] and terms extracted from the article abstracts were conducted, with the aim of revealing the main trends and topics of interest in the use of artificial intelligence models in the prediction of neonatal diseases, conditions and mortality. To this end, the data underwent pre-processing, which included the standardization of synonymous terms and the removal of irrelevant words, ensuring greater consistency and relevance in subsequent analyses. We created co-word networks to explore the associations between terms, with minimum frequency criteria for inclusion. From this, we also created co-word networks for keywords and abstracts, using minimum frequency criteria for inclusion, which ensured the identification of significant terms. These networks allowed the visualization of thematic clusters, in which different colors represented related groupings between terms, while the size of the nodes and the thickness of the edges indicate the frequency and strength of the connections of the use of each keyword, respectively.

Citation analysis [28] was performed to identify the most influential studies, calculating metrics such as total number of citations, average annual citations and most cited journal. Pioneering studies and prominent journals were identified, providing a view of the scientific impact and relevance of contributions in the field. The analyses were performed with tools such as Bibliometrix, VOSviewer and additional libraries in R and Python. These libraries allowed the generation of network graphs and thematic maps that facilitated the interpretation of patterns and interrelationships.

F. IDENTIFICATION OF CLUSTER FOR ANALYSIS PAPERS

After analyzing the metadata, we carried the keyword co-word network node selection process out to further explore the use of artificial intelligence in predicting conditions, diseases and neonatal mortality. The cluster was selected and within that cluster one of the nodes was selected to identify the articles responsible for constructing that node. These studies were extracted from the initial set of the bibliometric review and organized for full reading, with a focus on extracting specific information about the methods, data, artificial intelligence models, preprocessing techniques, selected attributes and other relevant information within the clinical and technological contexts addressed by the researchers. To standardize information collection, we created a structured form in Google Forms [29]. The form included questions directed to methodological aspects (such as algorithms used, sample size and variables analyzed), results presented and limitations discussed in the studies.

Thereafter, we collect relevant information from the studies individually, with the data recorded directly on the form to ensure consistency and minimize the risk of omissions. This resulted in data sets that covered essential information to ensure quality in the analysis and discussion of the use of machine learning models in the prediction of neonatal diseases, conditions and mortality. This data set was analyzed to identify sample sizes, balance techniques, problems and proposed solutions regarding missing data, attribute selection, use of statistical methods, artificial intelligence models and several other pieces of information that are being addressed in multidisciplinary research on artificial intelligence and neonatology.

IV. RESULTS AND DISCUSSIONS

A. DATA ANALYSIS OF THE BIBLIOMETRIC INDICATORS

1) Descriptive analysis

Descriptive analysis seeks to explore the general characteristics of the bibliometric data collected between 2014 and 2024, allowing the identification of publication patterns, temporal trends and the distribution of documents and scientific collaborations in the use of models of artificial intelligence in the prediction of neonatal health outcomes. Table 2 presents an overview of the 629 studies included in this bibliometric review.

TABLE 2. Bibliometric summary.

Categories	Values
KEY INFORMATION ABOUT DATA	
Timespan	2014 - 2024
Sources (Journals, Books, conferences, etc.)	375
Total Documents	629
Annual Growth Rate (%)	19.44
Document Average Age	2.98
Average citations per document	11.04
Average citations per year per document	2.32
References	20,161
DOCUMENT TYPES	
Article (journal publication)	468
Article; early access	4
Article; proceedings paper	3
Article; retracted publication	1
Book chapter	5
Conference paper	96
Editorial material	1
Letter	1
Proceedings Paper	50
DOCUMENT CONTENTS	
Keywords plus (ID)	2,181
Author's keywords (DE)	1,740
AUTHORS	
Total of authors	3,570
Author appearances	4,798
Authors of single-authored documents	10
Documents per author	0.176
Co-Authors per document	7.63
International co-author ships (%)	23.69

The metadata characteristics indicate an increase in production science, with an annual growth rate of 19.44% from 2014 to 2024. Of the 629 documents analyzed, 468 are journal studies, representing the majority of production, followed by conference papers (96). The study identified 3,570 authors, highlighting the ample academic collaboration and an average of 7.36 co-authors per document. This level of co-authored indicates a collaboration between multidisciplinary teams, fundamentals to explore the complexity of neonatal health challenges [30] [23]. The analysis highlights the average of citations per document is 11.04, indicating strong interest from the scientific community in the neonate in the literature. With a total of 20,161 references cited across 629 studies, this reveals the interconnectedness within the field.

Figure 3 illustrates the annual scientific production from 2014 to 2024. It clearly shows a growth trend, where the amount of publications gradually increased, going from 11 publications in 2014 to 152 in 2023, suggesting an interest in the topic that started to be more exploited over these years.

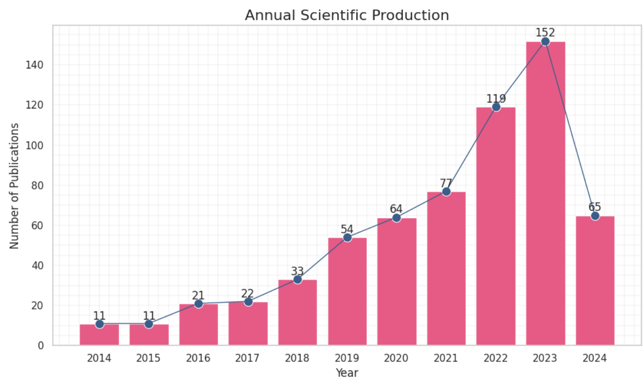


FIGURE 3. Annual distribution of research studies published in the period 2014 to 2024.

Since 2020, a more accelerated increase has been observed, with a jump from 77 publications in 2021 to 152 in 2023, possibly due to technological advancements and the large number of studies on the potential of using artificial intelligence to aid in solving health-related problems. In 2024, there is a decrease in the number of publications, totaling 65, due to the incompleteness of the data for 2024, as the data collection was up to May 2024.

2) RQ1: What are the more influential studies in the prediction of neonatal diseases, conditions and mortality?

Table 3 presents the 15 most cited studies, providing a comprehensive overview of the most impactful research in this area with key information: the Total of Citations (TC), the Average of Citations per Year (TCperYear) and the Normalization of Citations Totals (NCT).

The most cited study was that of Brown et al. [31] with 377 citations and an average of 53.86 citations per year, highlighting its impact on the prediction of Retinopathy of Prematurity and suggesting that the study has relevance and reference influence in the area. The studies by Gasparrini et al. [32], which examine the influence of early hospitalization and antibiotic use on the neonatal microbiome and Ansari et al. [33], focusing on the prediction of neonatal seizures, also stood out in the three quantitative metrics (TC, TCperYear, NCT), with TC values of 168 and 148, respectively and TCperYear values of 28.00 and 24.67 and NCT values of 7.45 and 6.56. The analysis of TCperYear also reveals that

TABLE 3. The 15 studies are the most cited.

Studies	Type of prediction	TC	TCperYear	NCT
Brown et al. (2018) [24]	Prediction of Retinopathy of Prematurity.	377	53.86	12.73
Gasparrini et al. (2019) [25]	Analyses of the impact of the early hospital and the use of antibiotics on the children's intestinal microbiome and its relationship with persistent metagenomic signatures.	168	28.00	7.45
Ansari et al. (2019) [26]	Prediction of neonatal seizures.	148	24.67	6.56
Mani et al. (2014) [27]	Early prediction of late-onset neonatal sepsis.	115	10.45	3.89
Antink et al. (2020) [28]	Body tracking and segmentation in neonatal videos for biometric monitoring.	100	20.00	5.44
Wang et al. (2018) [29]	Prediction of Retinopathy of Prematurity.	93	13.29	3.14
Boashash et al. (2016) [30]	Prediction of Electroencephalogram (EEG) abnormalities to identify neonatal seizures.	91	10.11	3.10
Masino et al. (2019) [31]	Early prediction of sepsis in NICU.	90	15.00	3.99
Smyser et al. (2016) [32]	Prediction of brain maturity in neonates.	90	10.00	3.06
Hu et al. (2018) [33]	Prediction of Retinopathy of Prematurity.	86	14.33	3.81
O'Shea et al. (2020) [34]	Prediction of neonatal seizures directly from multichannel EEG.	83	16.60	4.52
Zöllei et al. (2020) [35]	Prediction of brain maturity in neonates.	80	16.00	4.35
Villarroel et al. (2019) [36]	Non-invasive physiological monitoring of newborns in the NICU.	79	13.17	3.50
Ball et al. (2016) [37]	Characterization of brain functional connectivity in premature neonates.	76	8.44	2.59
Gephart et al. [38]	Prediction of necrotizing enterocolitis.	75	10.71	2.53

more recent studies, such as those by Brown et al. [31] and Gasparrini et al. [32], show higher citation rates, indicating ongoing academic interest and relevance. Older studies, such as those by Mani et al. [34] and Ball et al. [35], show a decline in citation rates, as evidenced by their much lower averages of 10.45 and 8.44 citations per year, respectively, reflecting a natural decrease in their academic influence over time.

The analyses of the 15 studies most cited evidence that Retinopathy of Prematurity is the disease most frequently investigated, representing 26.67 of the 15 studies most cited and a total of 655 citations, equivalent to 30.99% of the total citations. Neonatal seizures appear in second place, with 20% while neonatal sepsis and prediction of brain maturity are addressed in 13.33% of the studies each. Meanwhile, diseases such as necrotizing enterocolitis and imbalance in intestinal microbiome, investigated in only one of the 15 most cited studies (6.67%), show areas that are still little explored.

Table **??** presents the 15 top periodicals that have had the most influential studies in the area in the last 10 years, belonging to the 11 editorials. The Computers in Biology and Medicine is the journal that has the most published studies in this area, with 17 studies, representing 2.62% of the total. Others in emphasis are IEEE Access and NeuroImage, with 13 studies published each (2.01%). Between the 15 periodicals, NeuroImage led in the total of citations, with 381 citations, followed by Plos One (206) and Medical Image Analysis (186). The article most cited was published in NeuroImage (Smyser et al. (2016) [36]) and Plos One (Masino et al. (2019) [37]), both with 90 citations. The NeuroImage journal was also emphasized in the ranking by citations, contributing to an average of 29.3 citations by studied published, the larger among the top 15 journals.

The Computerized Medical Imaging and Graphics journal emphasizes the number of citations per article with an average of 21.0, followed by the Medical Image Analysis and Plos One, with an average of 20.7 and 20.6, respectively. It indicates a high scientific relevance of the studies published in these journals, even with fewer publications.

3) RQ2: What are the strongest collaboration networks in the study for the prediction of neonatal diseases, conditions and mortality research?

Our analysis identifies 70 countries that are leading contributors to scientific publications. Figure 4 illustrates the network of scientific collaborations among these countries. Each node (circle) represents a country and the size of each node reflects the volume of research output from that country. The edges (lines) connecting the nodes signify the collaborative ties between countries, with the thickness of the edges indicating the strength of these collaborations—the thicker the edge, the more intensive the collaborative production between the two countries. The color coding within the clusters denotes groups that engage in more frequent or stronger interactions within the analyzed network.

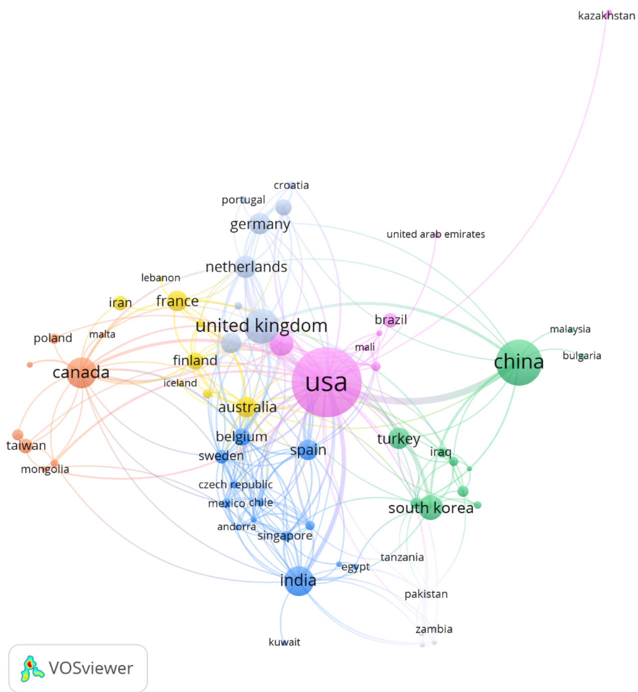


FIGURE 4. Collaboration network among countries.

The network comprises six clusters that represent distinct collaborative communities, each differentiated by color. The United States of America (USA) stands out as the central hub of collaboration, featuring the largest node in the network with 47 connections and 311 co-authorships, highlighting its leadership in scientific production and fostering international collaborations. China, the United Kingdom and India also occupy significant positions, with 16, 27 and 29 connections, respectively and substantial co-authorship figures (147, 132 and 121), thereby cementing their roles as key centers of scientific collaboration. Other nations, including Canada, the Netherlands, Germany, Australia and South Korea, make notable contributions to scientific production, though on a

TABLE 4. The top 15 most influent journals with their respective most cited article.

Periodicals	Total publications Quantity (%)	Total citations	Citation per document	Ranking (cited)	Publisher	Most cited article	Quantity of citations
Computers In Biology And Medicine	17 (2.62)	161	9.47	4	Pergamon-Elsevier Science Ltd	Delannoy et al. (2020) [39]	44
IEEE Access	13 (2.01)	152	11.7	5	Ieee-Inst Electrical Electronics Engineers Inc	Zhang et al. (2018) [40]	43
Neuroimage	13 (2.01)	381	29.3	1	Academic Press Inc Elsevier Science	Smyser et al. (2016) [41]	90
Scientific Reports	12 (1.86)	142	11.8	6	Nature Portfolio	Podda et al. (2018) [42]	47
Frontiers In Pediatrics	11 (1.70)	50	4.55	4	Frontiers Media Sa	Helguera-Repetto et al. (2020) [43]	27
Plos One	10 (1.55)	206	20.6	2	Public Library Science	Masino et al. (2019) [44]	90
Medical Image Analysis	9 (1.39)	186	20.7	3	Elsevier	Išgum et al. (2015) [45]	73
Pediatric Research	9 (1.39)	119	13.2	8	Springer Nature	Kim et al. (2020) [46]	36
Biomedical Signal Processing And Control	8 (1.24)	82	10.2	11	Elsevier Sci Ltd	Rosales-Pérez et al. (2015) [47]	36
Computer Methods And Programs In Biomedicine	8 (1.24)	87	10.9	9	Elsevier Ireland Ltd	Bolón-Canedo et al. (2015) [48]	31
Diagnostics	8 (1.24)	26	3.25	14	Mdpi	Khalilzad et al. (2022) [49]	10
IEEE Journal Of Biomedical And Health Informatics	7 (1.08)	84	12.0	10	Ieee-Inst Electrical Electronics Engineers Inc	Leon et al. (2017) [50]	35
Artificial Intelligence In Medicine	6 (0.93)	22	3.67	15	BMC pediatric	Leigh et al. (2022) [51]	10
BMC Pediatrics	6 (0.93)	33	5.5	13	Pergamon-Elsevier Science Ltd	Nisha et al. (2019) [52]	17
Computerized Medical Imaging And Graphics	6 (0.93)	126	21.0	7	Elsevier Sci Ltd	Galdi et al. (2020) [53]	36

smaller scale, with 21, 18, 11, 25 and 11 connections respectively and co-authorship counts below 100.

To enhance the examination of scientific output across countries, Figure 5 displays the 30 most productive countries, categorizing their publications into Single-Country Publications (SCP) and Multi-Country Publications (MCP). It also details the relative frequency of each country's contributions to the global total, expressed as percentages (%). Additionally, the figure includes the ratio of MCPs to the total number of studies published by each country, with the values indicated in parentheses (MCP-Ratio).

The USA leads with 22.6% of global publications but has an MCP-Ratio of 0.25, indicating that a majority of its publications are produced domestically without international collaborations. China follows, accounting for 16.9% of the total publications and exhibiting a similar trend with an MCP-Ratio of 0.17. India contributes 7% to global publications but shows even less internationalization, with an MCP-Ratio of 0.13. In contrast, European countries such as Germany (0.46), the United Kingdom (0.45), Ireland (0.44) and Belgium (0.83) demonstrate significant levels of international collaborations, underscoring their crucial roles in the global integration of scientific research. Other nations noted for high internationalization include Switzerland, Norway and Portugal, all displaying elevated MCP-Ratios. Meanwhile, Australia, with an MCP-Ratio of 0.50, maintains a balance between domestic production and international collaborations.

Regional analysis indicates a concentration of publications in global hubs such as North America, Western Europe and East Asia. North America, predominantly the USA and Canada, leads in the volume of studies, while European countries exhibit notable levels of internationalization. In Asia, South Korea, India and China show limited participation in international collaborations, with MCP-Ratios below 0.20. However, low- and middle-income countries, including Bangladesh (MCP-Ratio of 0.25) and Ethiopia (0.00), are underrepresented, displaying minimal involvement in scientific production and collaboration.

Countries such as the United States, Canada, the United Kingdom, Germany, France, Australia and China are central to global scientific output, driven by their favorable economic environments and strong research infrastructures. These nations are integral to the burgeoning field of artificial intelligence applied to disease prediction, neonatal conditions and mortality, enhancing the global exchange of knowledge and technological innovation. However, the underrepresentation of low- and middle-income countries, including Egypt, Pakistan, Tanzania, Zambia and Mali, underscores persistent inequalities in the global distribution of scientific resources, which remain concentrated in affluent, technologically advanced nations. While there are emerging collaborations in these lesser-represented countries, the prevailing dominance of established hubs in North America, Western Europe and certain parts of Asia highlights significant regional and socioeconomic disparities. These disparities reveal that while wealthier countries lead global scientific endeavors, nations with fewer resources encounter substantial barriers to equitable participation in scientific progress.

4) RQ3: What are the more influential institutions in the prediction of neonatal diseases, conditions and mortality?

Table 5 displays the top 15 institutions ranked by the number of studies published, accompanied by the total citations received and the average citations per article. This data facilitates the identification of not only the most productive institutions but also those that exert a significant impact in the research field, as evidenced by their recognition and citation rates.

TABLE 5. The 15 most productive institutions based on the number of studies published.

Institutions	Quantity of studies	Total citations	Average citations per article
Harvard University	64	2862	44.71
University of Pennsylvania	40	1412	35.30
Stanford University	37	1823	49.27
University of London	30	1723	57.43
Fudan University	29	1076	37.10
Massachusetts General Hospital	26	1131	43.50
University College Cork	26	1178	45.30
National Institutes Of Health (Nih) - Usa	25	940	37.60
University of Helsinki	25	1075	43.00
Children's Hospital Of Philadelphia	24	822	34.25
Pennsylvania Medicine	24	822	34.25
Johns Hopkins University	22	1384	62.90
Oregon Health \& Science University	22	910	41.36
Utrecht University Medical Center	22	899	40.86
University College London	21	1107	52.71

The results indicate that Harvard University tops the list with 64 studies published, amassing 2,862 citations and an average of 44.72 citations per article. This high citation count underscores the significant influence and relevance of the

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

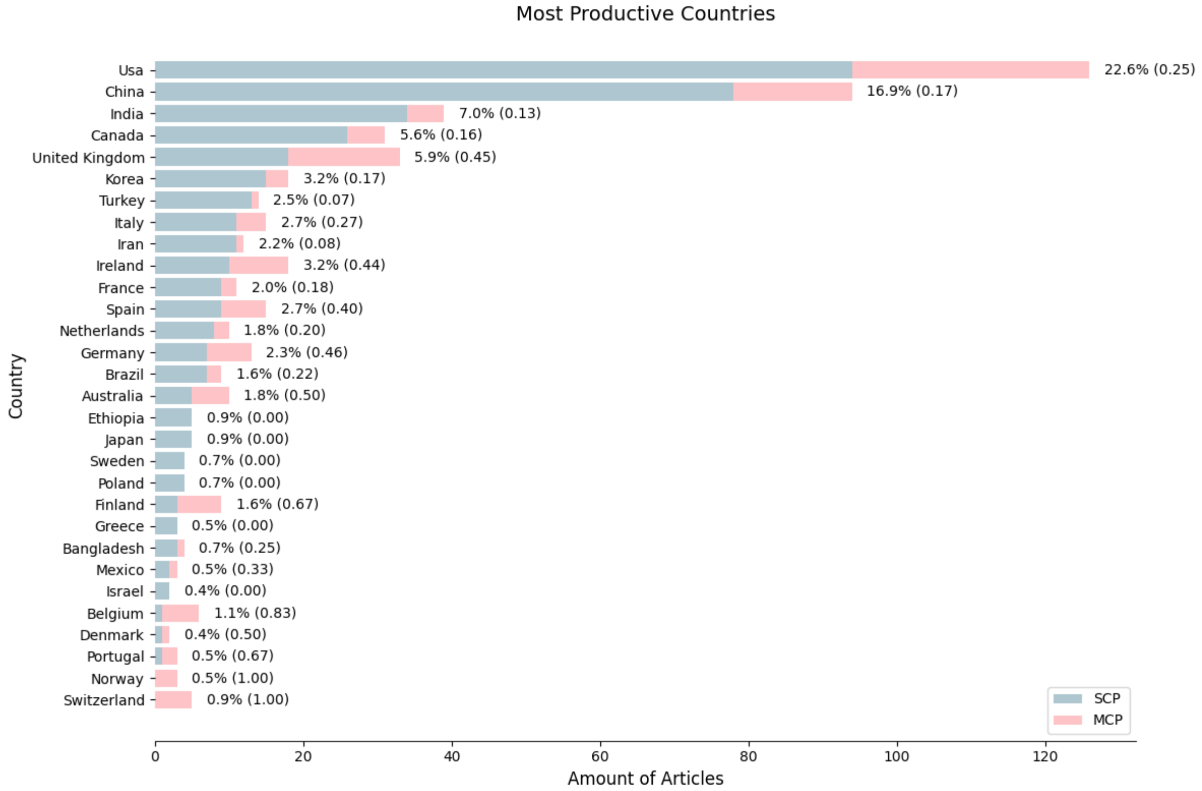


FIGURE 5. Proportion of exclusive and collaborative productions among the 30 most productive countries.

research conducted at this institution. The University of Pennsylvania ranks second, publishing 40 studies, which have garnered 1,412 citations and an average of 35.3 citations per article, reflecting robust productivity and impact. Stanford University also stands out, with 37 studies resulting in 1,823 citations and an average of 49.27 citations per article—the highest among the top three institutions.

Other notable institutions include Johns Hopkins University, with 22 studies yielding 1,384 citations and an impressive average of 62.91 citations per article and the University of London, which published 21 studies accumulating 1,107 citations, with an average of 57.43 citations per article. These high average citations highlight the widespread recognition and influence of their research contributions in the field.

5) RQ4: what are the main conceptual themes and how have they evolved in the prediction of neonatal diseases, conditions and mortality?

a: Keyword terms co-word graph

The analysis of the keyword co-word network is designed to identify the strongest and most frequent connections among the keywords used in the studies. Each node in the network represents a keyword and the edges connecting these nodes signify the frequency with which these keywords co-word in the analyzed metadata. The thickness of the edges and the proximity of the nodes to one another reflect the strength of their associations. Additionally, the size of the nodes indicates the frequency of keyword usage, highlighting the relevance of particular topics within the research area. The colors of the nodes denote different communities or clusters of terms, based on their proximity or interrelationships within the dataset.

After standardizing synonymous keywords and excluding terms that did not contribute meaningful value to the analysis, a total of 283 keywords were retained. Out of these, 178 keywords met the minimum occurrence threshold of nine, a criterion established for constructing the keyword co-word network map based on the metadata of the selected studies. Figure 6 illustrates the visual mapping of the interrelationships among the keywords, with connections displayed only if the co-word frequency is at least three.

The keyword co-word network effectively highlights the interconnections between various research areas and topics of interest. The node colors and sizes emphasize the most frequently explored areas and their interconnections, revealing a strong focus on topics like machine learning in neonatal health, conditions related to prematurity and artificial intelligence techniques for image segmentation. The largest cluster, marked in yellow, includes terms such as "machine learning," "NICU," and "artificial intelligence". This grouping indicates a significant application of machine learning techniques within the context of neonatal intensive care, especially in predicting adverse outcomes for hospitalized neonates. The prominence of these terms underscores the increasing priority

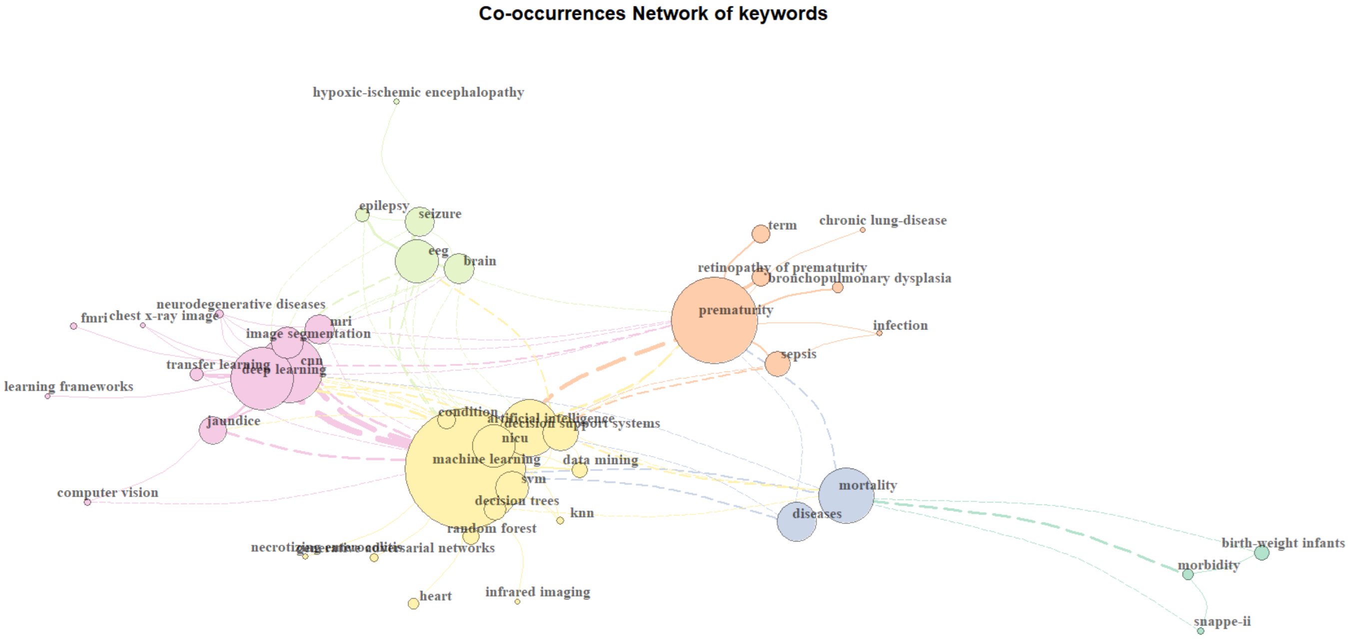


FIGURE 6. **Co-word network of the authors' keywords from the studies.**

placed on integrating technological tools to enhance survival rates and improve the quality of life for neonates.

The salmon-colored cluster includes keywords such as ''prematurity'', ''sepsis'', ''Retinopathy of Prematurity'' and ''bronchopulmonary dysplasia'', recommending an emphasis on the prediction and management of diseases and specific conditions associated with prematurity, which are common and high-risk for neonates. This grouping indicates a particular focus on diseases and conditions related to prematurity, which represent frequent and critical challenges in neonatal care. The strong interconnection of these terms suggests a collaborative effort in predicting and managing these conditions, highlighting the importance of using artificial intelligence to address high-risk and clinically impactful issues in neonatal health.

Other groupings identified in the analysis of the keyword co-word network reveal a diversity of topics that complement the understanding of neonatal diseases, conditions and mortality. The smaller clusters address respiratory complications, such as ''respiratory distress syndrome'' and ''persistent pulmonary hypertension'', emphasizing the critical management of these conditions in neonates, especially preterm infants. Metabolic and neurological issues, which can be predicted using imaging exams or photographs, such as ''jaundice'', ''epilepsy'' and ''seizures'', are conditions frequently diagnosed that require rapid interventions to prevent long-term complications. Notably, congenital anomalies, such as ''heart malformations'' and ''neural tube defects'', are important causes of neonatal morbidity and mortality.

b: Abstracts terms co-word graph

Rigorous criteria were applied to construct the co-word network of keywords extracted from the abstracts of the analyzed studies, including a minimum frequency of 5 occurrences and at least 3 mutual mentions to establish connections between terms. These parameters ensured the robustness and relevance of the represented relationships. Irrelevant terms were removed and synonymous keywords were grouped. Figure 7 illustrates the network generated from this analysis.

The keyword co-word network was organized into four clusters, identified by distinct colors, representing groups of common terms related to each other. The thematic clusters identified in the network reinforce the main research focus on artificial intelligence applied to predicting neonatal outcomes. The analysis highlights that central terms such as ''machine learning'' and ''artificial intelligence'' exhibit high frequency and strong connections with other concepts, emphasizing their fundamental role in the application of predictive methods in the neonatal context. These terms are interconnected with various models, such as ''Logistic Regression'', ''SVM'' and ''XGBoost'', suggesting the diversified use of predictive approaches.

Another relevant aspect is the presence of connections with specific diseases and clinical conditions, such as ''prematurity'', ''sepsis'', ''bronchopulmonary dysplasia'', ''necrotizing enterocolitis'' and ''jaundice'', as well as critical neonatal outcomes like ''mortality'' and ''morbidity''. This indicates that predictive models are widely applied to foresee common and severe complications in newborns. Mentions of ''NICU'' emphasize the focus of these tools on intensive care settings. The analysis also reveals a growing emphasis on advanced techniques, such as ''Deep Learning'' and its subfields, in-

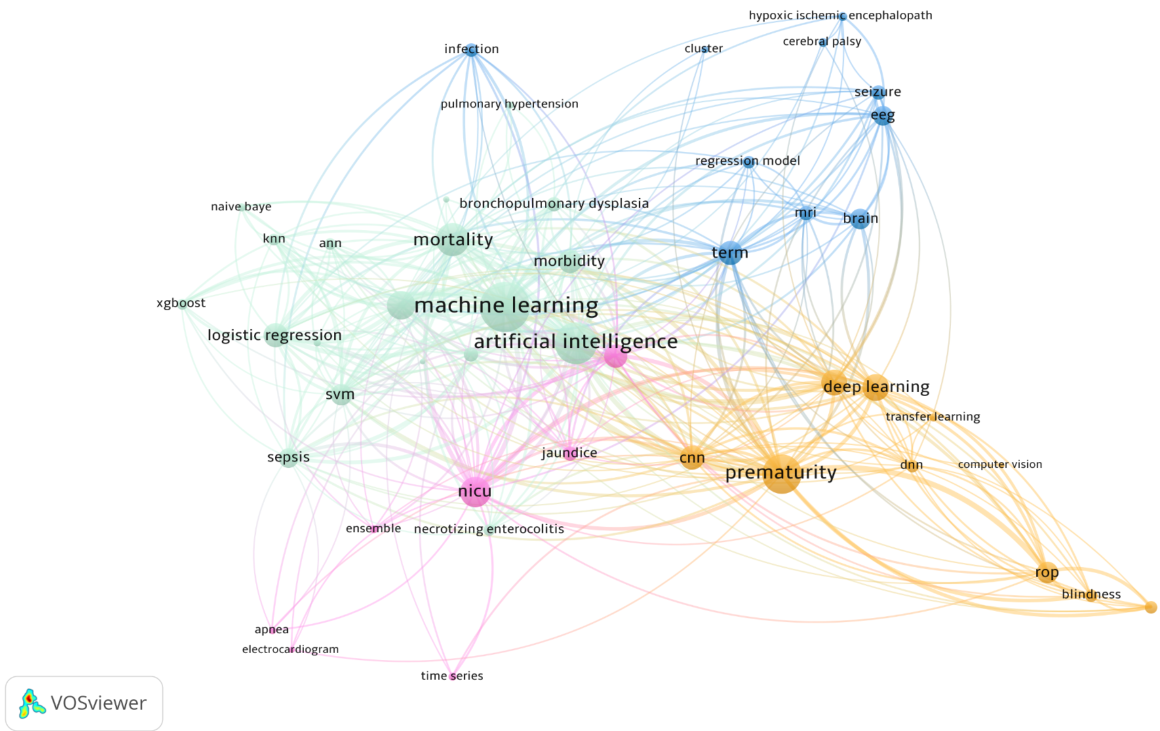


FIGURE 7. Co-word of keywords extracted from the abstracts.

cluding Convolutional Neural Networks (CNN)'', ''Deep Neural Networks (DNN)'' and ''Transfer Learning'', which are frequently associated with medical imaging analysis, such as ''Magnetic Resonance Imaging (MRI)'' and ''EEG''. This connection suggests the application of models in laboratory tests for detecting neurological conditions, such as ''seizure'' and ''hypoxic-ischemic encephalopathy'', highlighting the potential of artificial intelligence for early and more accurate diagnostics.

A bibliometric overlay map was also created using this keyword co-word network to understand how these research trends and focuses have evolved over time (Figure 8). This map offers a new perspective by linking the keywords extracted from the abstracts with the average year of publication of the documents in which they appear.

The overlay map uses a color scale ranging from blue (older years) to red/yellow (more recent years), allowing for the identification of the average period of prominence for each term. Terms such as ''machine learning'' and ''artificial intelligence'' display lighter orange tones, indicating that they have been widely used in recent publications, particularly since 2021. Similarly, keywords like ''Deep Learning'', ''CNN'' and ''Transfer Learning'' stand out as emerging and recent approaches applied to the neonatal context.

The terms ''bronchopulmonary dysplasia'', ''pulmonary hypertension'' and ''morbidity'' are associated with darker red tones, indicating that the study and management of these conditions using artificial intelligence have gained recent prominence, particularly in publications since 2022. Other diseases and clinical conditions, such as ''sepsis'', ''necrotizing enterocolitis'', ''apnea'' and ''cerebral palsy'', are also explored in recent research, but with greater prevalence in publications from 2019 to 2021.

c: Analysis frequencies of keyword

Figure 9 complements this analysis by presenting the frequencies of keywords extracted from the literature, providing a comprehensive view of the main areas of interest and trends in the use of artificial intelligence for predicting neonatal outcomes. It highlights the relevance of techniques in critical areas for predicting specific neonatal outcomes.

Figure 9 displays the 40 most frequently used keywords in the studies, with Machine Learning being the most commonly found keyword in the keyword co-word network (Figure 6), appearing in 404 studies. In the same group of terms related to the use of technologies, other prominent keywords include prematurity (278 occurrences), CNN (218 occurrences), Deep Learning (213 occurrences), artificial intelligence (183 occurrences) and Decision Support Systems (122 occurrences). Additionally, specific neonatal and medical terms such as mortality (168 occurrences), NICU (144 occurrences), EEG (140 occurrences) and diseases (123 occurrences) were also frequently mentioned, highlighting the relevance of these topics in studies on artificial intelligence

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

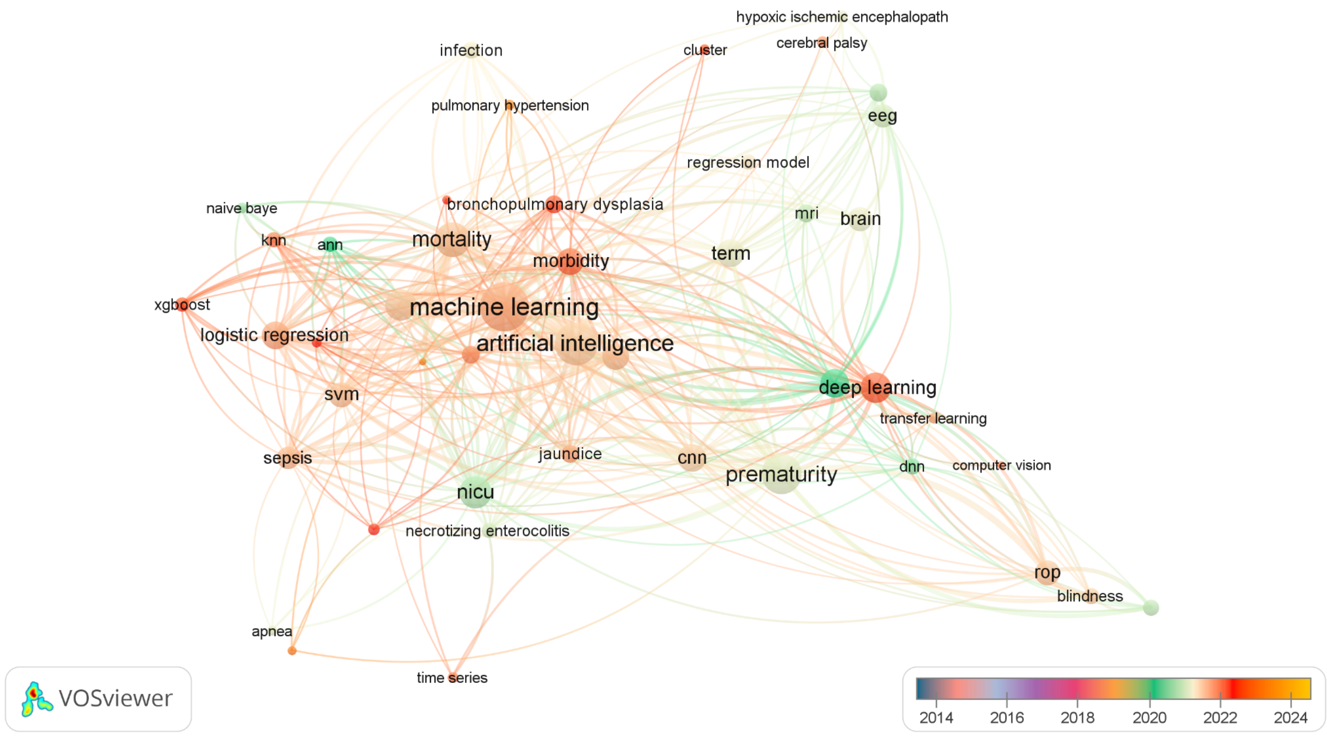


FIGURE 8. **Average annual publication year of the documents in which a keyword co-word.**

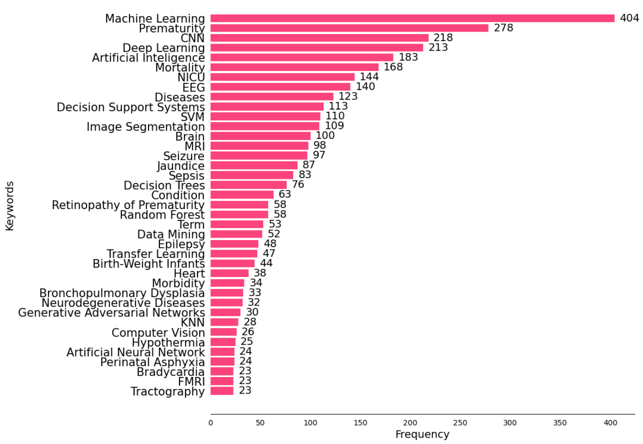


FIGURE 9. **Analysis of keyword frequencies of the studies.**

applied to neonatal health.

6) RQ5: What types of neonatal diseases, conditions and mortality are currently being studied in research using artificial intelligence?

The analysis provides an overview of research on using artificial intelligence in the neonatal field, focusing on a wide range of diseases, conditions and mortality factors affecting this vulnerable population. The data analysis highlights a significant emphasis on conditions associated with prematurity, including "sepsis", "bronchopulmonary dysplasia" and "Retinopathy of Prematurity", due to their critical impact on neonatal health. Other widely studied topics include metabolic and neurological disorders, such as "jaundice", "seizures" and "epilepsy", often diagnosed with the aid of artificial intelligence models trained on tests like "EEG" and "MRI". Severe respiratory conditions, such as "respiratory distress syndrome" and "persistent pulmonary hypertension", are also explored, particularly in preterm neonates.

Another relevant aspect highlighted by analysis of the keyword co-word is the diversity of diseases, health conditions and negative neonatal outcomes analyzed in the studies. These terms were grouped in Table **??** to organize the main categories observed in the keyword analysis.

A total of 22 terms were identified in the co-word and frequency analyses, distributed across three main categories: diseases, which account for 8 terms (36.4%); health conditions, with 10 terms (45.5%); and neonatal outcomes, represented by 4 terms (18.1%). This analysis reflects the diversity of aspects covered in studies related to neonatal health.

B. IDENTIFICATION OF CLUSTERS FOR DOCUMENT ANALYSIS

Among the identified clusters, we selected the node related to bronchopulmonary dysplasia due to its central position in the prematurity cluster and high connectivity with other nodes in the network, indicating its relevance within the field

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

TABLE 6. Classification of keywords based on co-word networks and frequency analysis of words.

Diseases	Health Conditions	Outcomes
neurodegenerative diseases, Retinopathy of Prematurity, chronic lung disease, infection, necrotizing enterocolitis, bronchopulmonary dysplasia, epilepsy, pulmonary hypertension, hypoxic-ischemic, encephalopathy, chronic lung disease.	prematurity, birth-weight infants, perinatal asphyxia, bradycardia, hypothermia, jaundice seizure.	mortality, morbidity, cerebral palsy, blindness, sepsis.

of application of artificial intelligence in neonatology. This choice was also justified by the clinical importance of bronchopulmonary dysplasia as a frequent and high-risk condition in premature newborns, making it a priority topic for the development of predictive models based on artificial intelligence. A total of 34 studies were identified under the node of bronchopulmonary dysplasia that mentioned this keyword. However, only 19 of them consisted of research focused on predicting the condition, as presented in Table 7.

Of the 19 studies analyzed, 13 predicted bronchopulmonary dysplasia considering severity levels (mild, moderate, severe and Bronchopulmonary Dysplasia-Associated Pulmonary Hypertension (BPD-PH)). Six studies [40], [41], [44], [47]–[49] predicted bronchopulmonary dysplasia based on its presence or absence, regardless of severity. Kostekci et al. [40] and Khurshid et al. [44] predicted bronchopulmonary dysplasia based on gestational age and/or specific days of analysis, without a direct focus on predicting cases by severity.

1) RQ6: What are the main attributes most frequently considered by studies using artificial intelligence models in predicting neonatal diseases, conditions and mortality?

The analysis presents various attributes used for predicting bronchopulmonary dysplasia. These attributes can be categorized into clinical, demographic, laboratory, genetic, imaging and other data. Among the 19 studies analyzed, there was a predominance of clinical and demographic data, while laboratory, genetic and imaging data were explored less frequently. Clinical data were used in 15 studies [38]–[40], [42]–[45], [47], [49]–[51], [53]–[56], covering variables such as surfactant use, mechanical ventilation and conditions like sepsis and intraventricular hemorrhage, demonstrating their direct relevance in monitoring and intervening in newborns. Demographic data, such as gestational age, birth weight and sex, were used in 16 studies [38], [40], [42]–[45], [47], [49]–[56], highlighting their importance in characterizing populations and generating robust predictive models.

Laboratory exams, such as inflammatory markers and biochemical profiles, were present in five studies [40], [43], [50], [53], [54], contributing as complements to clinical variables for the identification of bronchopulmonary dysplasia. Genetic attributes appeared in three studies [43], [46], [48], focusing on the exploration of genetic predispositions and molecular characteristics associated with neonatal conditions, such as in the study by Moreira et al. [46], which investigated genes related to metabolism and immune response. Finally, imaging data were used in four studies [41], [52], [56], primarily chest X-rays and magnetic resonance imaging, to analyze structural and functional characteristics.

The use of genetic and imaging data illustrates methodological advancements in specific areas, contributing to more personalized approaches and advanced diagnostics. The integration of multiple data types, as observed in studies such as [41]–[43], [53], presents a distinct approach to predicting neonatal diseases and conditions. The study by Sun et al. [41] mentioned the use of clinical and demographic data, along with chest X-rays, but did not describe them. Additionally, the analysis reveals that clinical and demographic data are widely used in most studies, as shown in Figure 10.

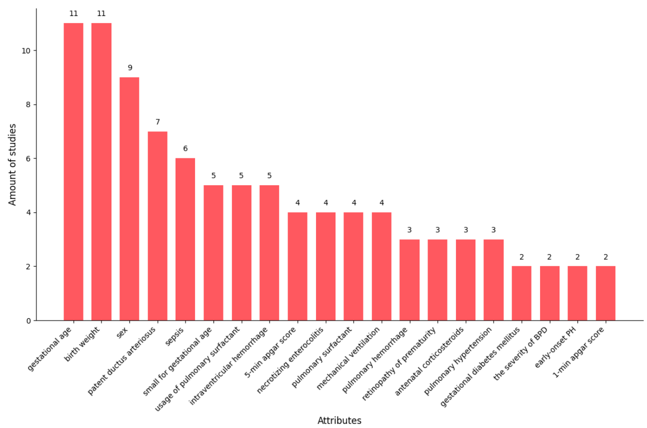


FIGURE 10. The 15 most used attributes for predicting bronchopulmonary dysplasia.

Figure 10 shows the 15 most commonly used attributes for predicting bronchopulmonary dysplasia, with emphasis on gestational age and birth weight, which are frequently cited and appear in 11 studies each [38], [40], [42]–[45], [47], [49]–[51], [55]. The sex of the neonate, present in 9 studies, is highlighted in [38], [40], [42]–[45], [49], [51], [55]. The persistence of the ductus arteriosus, mentioned in 7 studies [40], [43], [45], [47], [50], [51], [53], demonstrates its importance. Sepsis appears in 6 studies [40], [43], [44], [50], [53], [55], reinforcing its impact. The attribute small for gestational age, also present in 6 studies [38], [44], [45], [49], [51], [55], is recurrent in the literature. The use of various integrated attributes is essential for predictive modeling of bronchopulmonary dysplasia, allowing for an understanding of the multiple factors associated with clinical outcomes. The predominance of respiratory, neonatal and gestational variables reflects the need for rigorous and personalized monitoring of these aspects in neonates at risk of developing this condition.

TABLE 7. Summary of the data of studies of bronchopulmonary dysplasia.

Studies	Location	Number of records (used)	No bronchopulmonary dysplasia	Bronchopulmonary dysplasia level	Prediction by bronchopulmonary dysplasia level	Balancing	Number of attributes	Attribute selection	Missing data	Outliers
Wang et al. (2024) [38]	China	801 (761)	0	761 (mild = 312 (BPD = 306, BPD-PH = 6), moderate = 89 (BPD = 84, BPD-PH = 5), severe = 360 (BPD = 272, BPD-PH = 88))	Yes	Was used balanced (761) and imbalanced (original)	46	Yes	Yes	Not
Wang et al. (2023) [39]	China	761	0	761 (BPD (mild, moderate and severe) = 662, BPD-PH = 99)	Yes	Was used balanced (662) and imbalanced (original)	47	Yes	Yes	Not
Kostekci et al. (2023) [40]	Turkey	143 (124)	29	95 (moderate and severe (day1 = 95, day7 = 64, day14 = 43, day28 = 41))	No (done by day)	Not balancet	14	Yes	Yes	Not
Sun et al. (2022) [41]	China	403	275	128 (not described)	No	Not balancet	49 other types + chest X-ray	Yes	Not	Not
Verder et al. (2021) [42]	Denmark	72 (61)	35	26 (moderate and severe)	No	Not balancet	14	Yes	Not	Not
Dai et al. (2021) [43]	China	370 (245)	114	131 (mild and moderate = 64, severe = 67)	Yes	Not balancet	31other types + 51 genes	Yes	Yes	Not
Khurshid et al. (2021) [44]	Canada	12,990 (9,006)	5,818 (<33 weeks (day1 = 5,818, day7 = 5,527, day14 = 5,413), (<29 weeks (day1 = 1,736, day7 = 2,982, day14 = 2,899)	3,188 (<33 weeks (day1 = 3,188, day7 = 2,897, day14 = 2,783), (<29 weeks (day1 = 2,510, day7 = 2,264, day 14 = 2,163)	No (done by day and based on gestational age week)	Not balancet	20	Yes	Yes	Not
Ciora et al. (2024) [45]	Germany and European	1,039 (530)	174 (dataset 1 = 59, dataset 2 = 115)	356 (dataset 1 (mild = 58, moderate = 24, severe = 30), dataset 2 (mild = 94, moderate = 39, severe = 111))	Yes	Not balancet	16 other types + 500 genes	Not described	Yes	Yes
Moreira et al. (2023) [46]	Poland	97	35	62 (mild and moderate = 47, severe = 15)	Yes	Not balancet	33,252 genes per patient and 9 other types	Yes	Yes	Yes
Ochab et al. (2016) [47]	Poland	109	63	46 (not described)	No	Not balancet	14	Yes	Not	Not
Hu et al. (2023) [48]	Gene Expression Omnibus data base (GEO)	353	146 (dataset 1 = 112, dataset 2 = 34)	207 (dataset 1 = 187, dataset 2 = 20)	No (done by dataset)	Not balancet	Not described	Yes	Not	Not
Leigh et al. (2022) [49]	USA	1,191 (847 e 689)	dataset 1 = 428, dataset 2 = 352	dataset 1 = 419, dataset 2 = 337	No (done by dataset)	Not balancet	10	Yes	Yes	Not
He et al. (2023) [50]	China	1,604 (1,503)	1.032	471 (mild = 279, moderate = 147, severe = 45)	Yes	Not balancet	28	Yes	Yes	Not
Hwang et al. (2023) [51]	South Korea	16,384(11,177)	3.724	7,453 (mild = 3,383, moderate = 1,375, severe = 2,695)	Yes	Not balancet	416	Yes	Yes	Yes
Xing et al. (2022) [52]	China	121 patient (506 xhest X-ray)	0	121 (mild = 43, moderate = 47, severe = 31)	Yes	Not balancet	Chest X-ray	Not	Not	Not
Sharma et al. (2020) [53]	USA	323 (317)	70	121 (mild = 38, moderate = 78, severe = 77)	Yes	Not balancet	30	Yes	Yes	Not
Jaskari et al. (2020) [54]	Finland	977	531	275 (not described)	No	275	117 (111 sensor measurements)	Yes	Yes	Not
Wu et al. (2023) [55]	Taiwan	4,103 (3,200)	Not described	Not described	Yes	Not balancet	24	Yes	Yes	Not
Mairhörmann et al. (2023) [56]	Germany	107	33	70 (mild = 39, moderate = 11, severe = 20)	Yes	Not balancet	87	Yes	Yes	Not

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

2) RQ7: What are the characteristics of the datasets used by studies using artificial intelligence models in predicting diseases, conditions and neonatal mortality?

The analyzed studies exhibit heterogeneity in geographical locations, sample sizes and methodological approaches. Wang et al. [38] and [39] (China) highlight samples with 46 and 47 attributes, including analyses with both balanced and unbalanced data in both studies. Studies with large samples include Khurshid et al. [44] (Canada) with 12,990 records and Hwang et al. [51] (South Korea) with 16,384 records. The smallest datasets were used by Verder et al. [42] from Denmark, Moreira et al. [46] from Poland and Mairhörmann et al. [56] from Germany, with 72, 97 and 107 records, respectively.

The imbalance of data was a relevant issue in several studies, as unbalanced datasets can compromise the effectiveness of models by favoring the majority class. Only three studies performed data balancing; of these, two [38], [39] used the original (unbalanced) data and the balanced data, applying oversampling techniques such as SMOTE and SVMSMOTE for balancing. The third study (Jaskari et al. [54]) used the undersampling technique.

Unfortunately, in the other studies (16 out of 19), no data balancing techniques were reported. Data balancing is relevant because it helps to ensure that the artificial intelligence models are not biased toward one class more than another, especially in cases where some outcomes or categories are more frequently represented in the dataset than others. The absence of data balancing could potentially affect the fairness and accuracy of the conclusions drawn in those studies. Furthermore, the lack of reporting raises significant concerns about scientific reproducibility. When studies do not provide complete information on crucial aspects such as data balancing, it hampers the ability of other researchers to accurately reproduce and verify the results.

Six studies did not report issues with missing data, while 13 faced this problem and adopted different strategies to address it. Wang et al. [38] and [39] used mean imputation for numerical values and the most frequent category for categorical values, while Dai et al. [43] applied manual imputation. Kostekci et al. [40] mentioned that machine learning methods can handle missing data without the need for exclusion, but did not specify the technique used. Khurshid et al. [44] employed iterative imputation (IterativeImputer), adjusting missing values based on regressions on other variables, while Moreira et al. [46] and Hwang et al. [51] used multiple imputation. Jaskari et al. [54] transformed non-synchronous data into a synthetic set, filling in missing values with the closest observation before the start of the defined time interval. Studies [49], [50], [55], [56] opted to exclude records with missing data.

Among the analyzed studies, most (16 out of 19) did not mention the presence of outliers, which may indicate the absence of explicit treatment or the assumption that the analyzed data did not contain extreme values. Only three studies directly addressed this issue. Ciora et al. [45] applied a strategy based on clustering morbidity profiles to identify subgroups of patients and morbidity burdens, normalizing disease severity scales for clustering analysis. Meanwhile, Moreira et al. [46] and Hwang et al. [51] used the Winsorization technique, which limits extreme values to a specific percentile to reduce the impact of outliers without completely removing them.

Furthermore, 17 studies performed feature selection, with automatic methods being the most commonly used (13 out of 19 analyzed studies). Wang et al. [38] and [39] used the SelectKBest technique with the f_regression function, which selects the most relevant features based on their statistical correlation with the target feature. Studies [42], [43], [46], [48], [51] applied the Least Absolute Shrinkage and Selection Operator (LASSO), a regularization method that automatically eliminates irrelevant features, improving model interpretability and performance. Jaskari et al. [54] used Random Forest to evaluate feature importance, while Sun et al. [41] proposed the CisiV technique, a method based on instrumental variables and a causal structure graph to select relevant variables. On the other hand, the least used selection methods were manual, involving selection through literature [40], [44], [49] and selection by neonatology experts [45], [50].

3) RQ8: What artificial intelligence models are used to predict neonatal diseases, conditions and mortality?

Figure 11 summarizes the artificial intelligence models used in the studies analyzed for the prediction of bronchopulmonary dysplasia, providing a comprehensive view of the application of these techniques. Among the most frequent models, Logistic Regression stands out, being used in 13 studies [38]–[40], [44], [46]–[51], [53]–[55]. Tree-based models, such as Random Forest, appear in 11 studies [38], [40], [44], [46], [49]–[51], [54]–[56], while Support Vector Machines were employed in 10 studies [40]–[42], [44], [46]–[48], [51], [54], [55]. Deep learning techniques, such as artificial neural networks and advanced architectures (ResNet, AlexNet and Inception-V3), were proposed in five studies [38], [44], [52], [55], [56]. Other approaches, including XGBoost and Gradient Boosting, were less prominent, being applied in four studies [41], [50], [51], [55].

Among the 35 models presented, 13 correspond to machine learning techniques, such as Random Forest, Support Vector Machines, K-Nearest Neighbors, XGBoost and Decision Tree. Additionally, 10 models are from Deep Learning, including architectures like ResNet, AlexNet, DenseNet-201 and U-Net Convolutional Neural Network, demonstrating their growing relevance in the prediction of neonatal diseases. Finally, the remaining six models belong to other categories, such as probabilistic methods (Gaussian Processes and Poisson Regression), ensemble algorithms (Soft Voting Ensemble) and clustering strategies (Agglomerative Hierarchical Clustering).

Among the studies analyzed, a significant portion (11 out of 19) did not describe the use of any hyperparameter optimization techniques [38]–[40], [43], [44], [46], [48], [50], [52], [53], [55]. Of the studies that explicitly mentioned hyperparameter optimization, Grid Search was the most com-

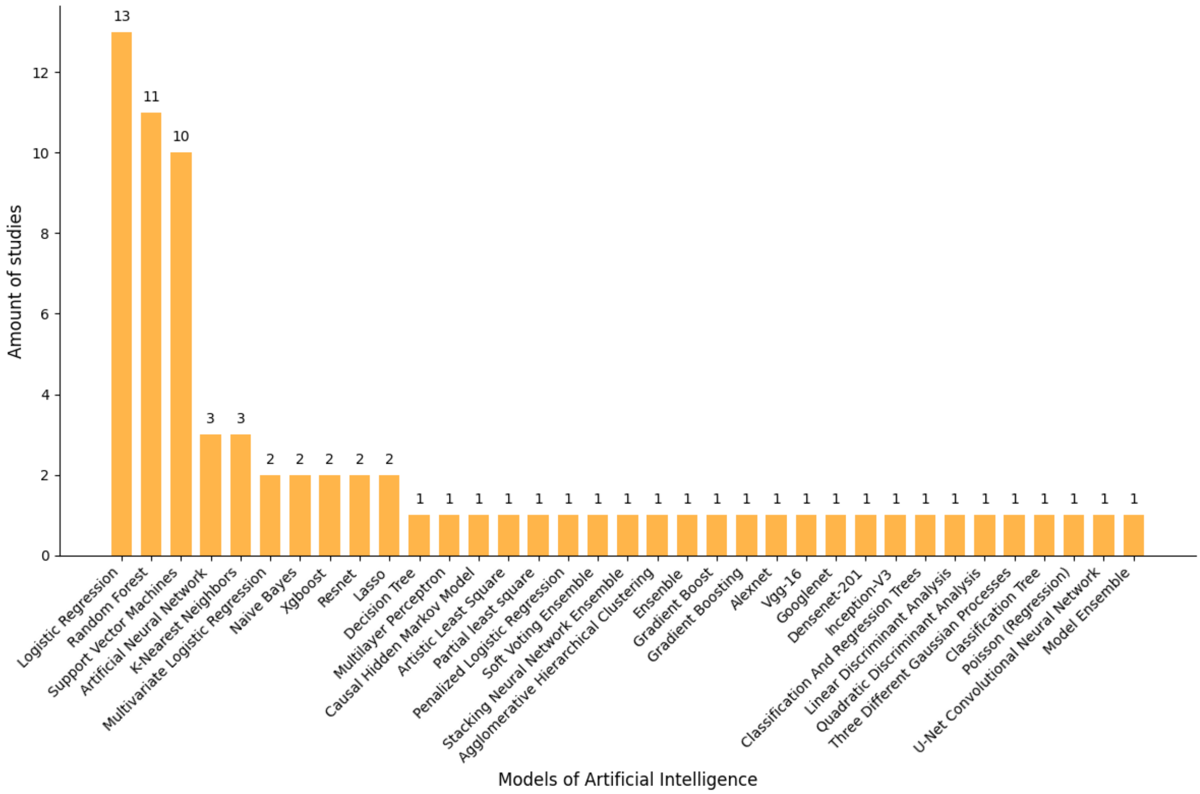


FIGURE 11. **Models used for the prediction of bronchopulmonary dysplasia.**

monly applied method, being used in four studies [42], [47], [49], [51]. Sun et al. [41] used the Adam optimizer, while Mairhörmann et al. [56] employed a randomized search algorithm. Finally, Ciora et al. [45] did not apply any technique.

The k-fold cross-validation approach to validate the models was applied in 14 out of the 19 studies analyzed. The most common configuration was the use of 10 folds, adopted by 10 papers [40], [43], [44], [46]–[49], [51], [55], [56], suggesting a preference for a traditional setup that balances the number of partitions and the size of the test set. Other studies used different values of k; Verder et al. [42] used 7 folds, Jaskari et al. [54] used 8 folds and the studies by He et al. [50] and Xing et al. [52] used 5 folds each. However, five studies chose not to employ k-fold as a validation method [38], [39], [41], [45], [53].

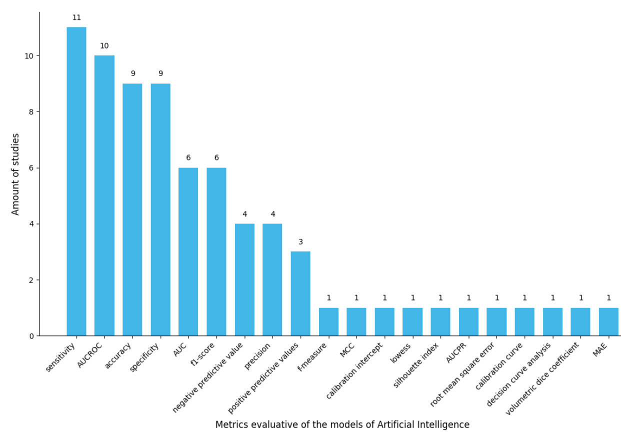


FIGURE 12. **Metrics used to evaluate the performance of artificial intelligence models in predicting bronchopulmonary dysplasia.**

4) RQ9: How is the performance of artificial intelligence models in predicting neonatal diseases, conditions and mortality evaluated?

Figure 12 presents the evaluation metrics used to measure the performance of artificial intelligence models, highlighting the frequency of use of each metric among the analyzed studies.

The analysis identified 20 evaluation metrics, with sensitivity standing out as the most commonly used metric, employed in 11 studies [38]–[40], [42], [44], [47]–[49], [51], [52], [54]. The area under the Receiver Operating Characteristic curve (AUCROC) appears in 10 studies [39], [40], [43], [46], [49]–[51], [53]–[55], followed by accuracy, which was used in 9 studies [38]–[41], [46], [47], [49], [52], [54]. The metrics of specificity and negative predictive value are mentioned in 9 studies [38], [42], [44], [47]–[49], [51], [52], [54] and 6 studies [38], [44], [48], [49], [54], [55], respectively. Additionally, metrics such as the F1-score [41], [46], [51], [52], [54], precision [40], [48], [51], [52] and positive predictive value [44], [49], [54] also appear, though with less frequency.

Less explored metrics point to usage in more specific analyses, such as clustering and regression, or for models that were trained and tested with imbalanced data. For instance, Kostekci et al. [40] used the Matthews correlation coefficient (MCC) to evaluate models trained and tested with imbalanced data. The silhouette index, employed exclusively by Ciora et al. [45], stands out as a metric for evaluating the cohesion and separation of clusters applied to clustering models. Studies by Khurshid et al. [44] and Hu et al. [48] used calibration-related metrics, such as calibration intercept and calibration curve, indicating a focus on the reliability of predictive probabilities in the models. Other metrics, like root mean square error (RMSE) and mean absolute error (MAE), appeared only in Hu et al. [48] and Mairhörmann et al. [56], being particularly relevant for regression and volumetric segmentation models. The use of the volumetric Dice coefficient by Mairhörmann et al. [56] reinforces the need for specialized metrics in studies dealing with image segmentation.

V. CONCLUSION

Neonatal diseases, conditions and mortality present a global public health issue, requiring constant efforts and investments to reduce their rates and improve the health of neonates in vulnerable situations. In this context, the literature highlights artificial intelligence models as promising and low-cost tools to assist healthcare professionals in early identification and intervention.

Our bibliometric review aimed to provide an overview of current research in artificial intelligence for predicting neonatal diseases, conditions and mortality, based on 629 publications found in six scientific databases. As a result, we can state that research in this field has grown over the past decade, with a particularly rapid increase in the last five years, accounting for 69.50% of the total rise in scientific production during this period. This trend suggests a growing recognition of artificial intelligence potential in neonatal healthcare and an expectation of further expansion in the coming years.

We identified the leading contributing countries and academic institutions, with the United States, China and the United Kingdom emerging as major research hubs, demonstrating strong international collaborations. However, we highlight that there is an underrepresentation of low- and middle-income countries, pointing to inequalities in access to advanced technologies and scientific production on the topic.

The analyzed studies also show that machine learning and Deep Learning models have been widely applied to predict diseases such as infections, bronchopulmonary dysplasia, necrotizing enterocolitis and pulmonary hypertension, as well as health conditions such as hypoxic-ischemic encephalopathy, prematurity, perinatal asphyxia, bradycardia, hypothermia and jaundice. These models have also been applied to identify patterns related to neonatal outcomes, including mortality, morbidity, cerebral palsy, blindness and sepsis. The most commonly used models include logistic regression, Random Forest, convolutional neural networks and boosting algorithms, which stand out for their high performance in classifying and prognosticating neonatal outcomes.

Additionally, artificial intelligence-based prediction of neonatal diseases predominantly relies on clinical and demographic data, such as mechanical ventilation, sepsis, gestational age and birth weight, as well as laboratory tests, genetic data and medical images (X-rays and MRIs). Class imbalance, missing data and outliers are common challenges, addressed by value imputation, exclusion of incomplete records and feature selection techniques such as LASSO and Random Forest. Given the advances and challenges observed, it is clear that artificial intelligence has the potential to assist healthcare professionals in making earlier and more accurate diagnoses, as well as optimizing clinical decision-making. The results of this research highlight a growing interest among researchers in the application of artificial intelligence to neonatal healthcare, with more than 50% of publications conducted in the last 3 years. The increase in studies using artificial intelligence to predict neonatal diseases, conditions, and outcomes such as bronchopulmonary dysplasia, sepsis, necrotizing enterocolitis, pulmonary hypertension, and retinopathy of prematurity demonstrates the urgency for tools that facilitate medical and neonatal clinical practices.

With scientific and technological advancements, it is essential to consider the ethical implications of applying artificial intelligence models in neonatal healthcare. It is important to emphasize that the performance of these models directly depends on the quality of the data and treatments used. Moreover, for the generalization of the models, it is crucial to have data that encompass various population characteristics, including biological and social aspects.

In this regard, regulating interdisciplinary collaboration between neonatal healthcare professionals, researchers, and data scientists is essential to establish guidelines that prioritize neonatal patient safety and equitable access to these technologies. The development of tools using Machine Learning and Deep Learning techniques with clinical, demographic, genetic, and imaging data could be crucial in hospital routines and may directly impact the life of the baby. Strengthening scientific and technological research in this field will be essential to consolidate artificial intelligence as a reliable and accessible tool capable of improving clinical outcomes and ultimately reducing neonatal mortality on a global scale.

ACKNOWLEDGMENTS

We would like to thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) and Universidade de Pernambuco (UPE).

REFERENCES

[1] M. da Saúde, "Guia de orientações para o método canguru na atenção básica: Cuidado compartilhado," 2016, acesso em: 17 nov. 2024. [Online]. Available: http://bvsms.saude.gov.br/bvs/publicacoes/guia_orientacoes_metodo_canguru.pdf

[2] UNICEF, "Levels and trends in child mortality - report 2023," 2023, acesso em: 17 nov. 2024. [Online]. Available: https://data.unicef.org/resources/levels-and-trends-in-child-mortality-2024/

[3] D. You, *Levels and trends in child mortality: report 2022*. UNICEF, 2023.

[4] W. H. Organization, "Every newborn: an action plan to end preventable deaths," 2014.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

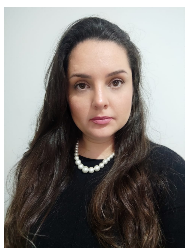
[5] ——, "Newborn mortality," 2024, acesso em: 17 nov. 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/newborn-mortality
[6] ——, "Newborn infections," 2024, acesso em: 17 nov. 2024. [Online]. Available: https://www.who.int/teams/maternal-newborn-child-adolescent-health-and-ageing/newborn-health/newborn-infections
[7] ——, "[maternal and newborn] - mortality/causes of death," 2016, acesso em: 17 nov. 2024. [Online]. Available: https://www.who.int/data/gho/data/themes/topics/topic-details/mca/maternal-and-newborn---mortality-causes-of-death
[8] ——, "Congenital disorders," 2023, acesso em: 17 nov. 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/birth-defects/
[9] U. Nations, "The sustainable development goals report 2019," New York, 2019, acesso em: 17 nov. 2024. [Online]. Available: https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019.pdf
[10] C. Henry *et al.*, "Application and potential of artificial intelligence in neonatal medicine," in *Seminars in Fetal and Neonatal Medicine*. WB Saunders, 2022, p. 101346.
[11] C. B. Collin *et al.*, "Computational models for clinical applications in personalized medicine guidelines and recommendations for data integration and model validation," *Journal of Personalized Medicine*, vol. 12, no. 2, p. 166, jan 2022.
[12] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2019.
[13] D. Lestari *et al.*, "Bibliometric analysis of intelligent techniques for obstetric complication prediction in the last 20 years," in *2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA)*. IEEE, 2024.
[14] Q. Zhou *et al.*, "Bibliometric analysis of bronchopulmonary dysplasia in extremely premature infants in the web of science database using citespace software," *Frontiers in Pediatrics*, vol. 9, p. 705033, 2021.
[15] Y. Zhang *et al.*, "Research hotspots and trends in retinopathy of prematurity from 2003 to 2022: a bibliometric analysis," *Frontiers in Pediatrics*, vol. 11, p. 1273413, 2023.
[16] S. H. Z., S. W. Al-Jabi, M. J. Shahwan, and A. A. Jairoun, "Global research production in neonatal abstinence syndrome: A bibliometric analysis," *World Journal of Clinical Pediatrics*, vol. 11, no. 3, p. 307, 2022.
[17] F. I. Maulana, D. Lestari, P. D. P. Adi, M. Hamim, A. Rahayu, and V. P. Widartha, "Direction of machine learning research in obstetric care in the last 10 years," in *2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA)*. IEEE, August 2024, pp. 1–7.
[18] PRISMA, "Prisma statement," 2024, accessed October 30, 2024. [Online]. Available: https://www.prisma-statement.org/
[19] ACM Digital Library, "Acm digital library," 2025. [Online]. Available: https://dl.acm.org/
[20] IEEE Xplore, "Ieee xplore advanced search," 2025. [Online]. Available: https://ieeexplore.ieee.org/search/advanced
[21] Clarivate Web of Science, "Web of science - advanced search," 2025. [Online]. Available: https://access.clarivate.com/login?app=wos
[22] PubMed, "Pubmed advanced search," 2025. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/advanced/
[23] ScienceDirect, "Sciencedirect - online scientific research," 2025. [Online]. Available: https://www.sciencedirect.com/
[24] Scopus, "Scopus - abstract and citation database," 2025. [Online]. Available: https://www.scopus.com/home.uri
[25] I. Zupic and T. Čater, "Bibliometric methods in management and organization," *Organizational Research Methods*, vol. 18, no. 3, pp. 429–472, 2015.
[26] N. Coulter, I. Monarch, and S. Konda, "Software engineering as seen through its research literature: A study in co-word analysis," *Journal of the American Society for Information Science*, vol. 49, no. 13, pp. 1206–1223, 1998.
[27] M. Callon, J.-P. Courtial, W. A. Turner, and S. Bauin, "From translations to problematic networks: An introduction to co-word analysis," *Social Science Information*, vol. 22, no. 2, pp. 191–235, 1983.
[28] H. Small, "Co-citation in the scientific literature: A new measure of the relationship between two documents," *Journal of the American Society for Information Science*, vol. 24, no. 4, pp. 265–269, 1973.
[29] F. L. De Morais, R. C. L. Da Silva, A. B. Da Silva, E. P. Simão, M. E. F. De Mello, S. P. Da Silva Canejo, and P. T. Endo, "Usage of artificial intelligence predicting neonatal diseases, conditions and mortality: A bibliometric review," *OSF*, February 16 2025. [Online]. Available: https://doi.org/10.17605/OSF.IO/2F6GQ
[30] V. Larivière, Y. Gingras, C. R. Sugimoto, and A. Tsou, "Team size matters: Collaboration and scientific impact since 1900," *Journal of the Association for Information Science and Technology*, vol. 66, no. 7, pp. 1323–1332, 2015.
[31] J. M. Brown, J. P. Campbell, A. Beers, K. Chang, S. Ostmo, R. P. Chan, J. Dy, D. Erdogmus, S. Ioannidis, J. Kalpathy-Cramer *et al.*, "Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks," *JAMA ophthalmology*, vol. 136, no. 7, pp. 803–810, 2018.
[32] A. J. Gasparrini, B. Wang, X. Sun, E. A. Kennedy, A. Hernandez-Leyva, I. M. Ndao, P. I. Tarr, B. B. Warner, and G. Dantas, "Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome," *Nature microbiology*, vol. 4, no. 12, pp. 2285–2297, 2019.
[33] A. H. Ansari, P. J. Cherian, A. Caicedo, G. Naulaers, M. De Vos, and S. Van Huffel, "Neonatal seizure detection using deep convolutional neural networks," *International journal of neural systems*, vol. 29, no. 04, p. 1850011, 2019.
[34] S. Mani, A. Ozdas, C. Aliferis, H. A. Varol, Q. Chen, R. Carnevale, Y. Chen, J. Romano-Keeler, H. Nian, and J.-H. Weitkamp, "Medical decision support using machine learning for early detection of late-onset neonatal sepsis," *Journal of the American Medical Informatics Association*, vol. 21, no. 2, pp. 326–336, 2014.
[35] G. Ball, P. Aljabar, T. Arichi, N. Tusor, D. Cox, N. Merchant, P. Nongena, J. V. Hajnal, A. D. Edwards, and S. J. Counsell, "Machine-learning to characterise neonatal functional connectivity in the preterm brain," *Neuroimage*, vol. 124, pp. 267–275, 2016.
[36] C. D. Smyser, N. U. Dosenbach, T. A. Smyser, A. Z. Snyder, C. E. Rogers, T. E. Inder, B. L. Schlaggar, and J. J. Neil, "Prediction of brain maturity in infants using machine-learning algorithms," *NeuroImage*, vol. 136, pp. 1–9, 2016.
[37] A. J. Masino, M. C. Harris, D. Forsyth, S. Ostapenko, L. Srinivasan, C. P. Bonafide, F. Balamuth, M. Schmatz, and R. W. Grundmeier, "Machine learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data," *PloS one*, vol. 14, no. 2, p. e0212665, 2019.
[38] D. Wang, S. Huang, J. Cao, Z. Feng, Q. Jiang, W. Zhang, J. Chen, S. Kutty, C. Liu, W. Liao *et al.*, "A comprehensive study on machine learning models combining with oversampling for bronchopulmonary dysplasia-associated pulmonary hypertension in very preterm infants," *Respiratory research*, vol. 25, no. 1, p. 199, 2024.
[39] D. Wang, S. Huang, J. Cao, Z. Feng, Q. Jiang, W. Zhang, J. Chen, C. Liu, W. Liao, L. Zhang *et al.*, "A multivariate logistic regression model with over sampling for bronchopulmonary dysplasia associated with pulmonary hypertension in very preterm infants: A multi-center retrospective study," in *2023 IEEE International Conference on Medical Artificial Intelligence (MedAI)*. IEEE, 2023, pp. 459–460.
[40] Y. E. Kostekci, B. Bakırarar, E. Okulu, O. Erdeve, B. Atasay, and S. Arsan, "An early prediction model for estimating bronchopulmonary dysplasia in preterm infants," *Neonatology*, vol. 120, no. 6, pp. 709–717, 2023.
[41] S. Sun, W. Ji, L. Tang, X. Zhang, Y. Zhang, and R. Feng, "An interpretable causal approach for bronchopulmonary dysplasia prediction," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2022, pp. 1193–1200.
[42] H. Verder, C. Heiring, R. Ramanathan, N. Scoutaris, P. Verder, T. E. Jessen, A. Höskuldsson, L. Bender, M. Dahl, C. Eschen *et al.*, "Bronchopulmonary dysplasia predicted at birth by artificial intelligence," *Acta Paediatrica*, vol. 110, no. 2, pp. 503–509, 2021.
[43] D. Dai, H. Chen, X. Dong, J. Chen, M. Mei, Y. Lu, L. Yang, B. Wu, Y. Cao, J. Wang *et al.*, "Bronchopulmonary dysplasia predicted by developing a machine learning model of genetic and clinical information," *Frontiers in Genetics*, vol. 12, p. 689071, 2021.
[44] F. Khurshid, H. Coo, A. Khalil, J. Messiha, J. Y. Ting, J. Wong, and P. S. Shah, "Comparison of multivariable logistic regression and machine learning models for predicting bronchopulmonary dysplasia or death in very preterm infants," *Frontiers in Pediatrics*, vol. 9, p. 759776, 2021.
[45] O.-A. Ciora, T. Seegmüller, J. S. Fischer, T. Wirth, F. Häfner, S. Stoecklein, A. W. Flemmer, K. Förster, A. Kindt, D. Bassler *et al.*, "Delineating morbidity patterns in preterm infants at near-term age using a data-driven approach," *BMC pediatrics*, vol. 24, no. 1, p. 249, 2024.
[46] A. Moreira, M. Tovar, A. M. Smith, G. C. Lee, J. A. Meunier, Z. Cheema, A. Moreira, C. Winter, S. B. Mustafa, S. Seidner *et al.*, "Development of

a peripheral blood transcriptomic gene signature to predict bronchopulmonary dysplasia," *American Journal of Physiology-Lung Cellular and Molecular Physiology*, vol. 324, no. 1, pp. L76–L87, 2023.
[47] W. Marcin Ochab, "Expert system supporting an early prediction of the brochopulmonary dysplasia," *Computers in Biology and Medicine*, pp. 236–244, 2016.
[48] Z. Hu, C. Liu, Y. Mao, J. Shi, J. Xu, G. Zhou, and F. Jiang, "Integration of transcriptomics reveals ferroptosis-related signatures and immune cell infiltration in bronchopulmonary dysplasia," *Heliyon*, vol. 9, no. 10, 2023.
[49] R. M. Leigh, A. Pham, S. S. Rao, F. M. Vora, G. Hou, C. Kent, A. Rodriguez, A. Narang, J. B. Tan, and F.-S. Chou, "Machine learning for prediction of bronchopulmonary dysplasia-free survival among very preterm infants," *BMC pediatrics*, vol. 22, no. 1, p. 542, 2022.
[50] W. He, L. Zhang, R. Feng, W.-H. Fang, Y. Cao, S.-Q. Sun, P. Shi, J.-G. Zhou, L.-F. Tang, X.-B. Zhang *et al.*, "Risk factors and machine learning prediction models for bronchopulmonary dysplasia severity in the chinese population," *World Journal of Pediatrics*, vol. 19, no. 6, pp. 568–576, 2023.
[51] J. K. Hwang, D. H. Kim, J. Y. Na, J. Son, Y. J. Oh, D. Jung, C.-R. Kim, T. H. Kim, and H.-K. Park, "Two-stage learning-based prediction of bronchopulmonary dysplasia in very low birth weight infants: a nationwide cohort study," *Frontiers in Pediatrics*, vol. 11, p. 1155921, 2023.
[52] W. Xing, W. He, X. Li, J. Chen, Y. Cao, W. Zhou, Q. Shen, X. Zhang, and D. Ta, "Early severity prediction of bpd for premature infants from chest x-ray images using deep learning: A study at the 28th day of oxygen inhalation," *Computer Methods and Programs in Biomedicine*, vol. 221, p. 106869, 2022.
[53] A. Sharma, Y. Xin, X. Chen, and B. G. Sood, "Early prediction of moderate to severe bronchopulmonary dysplasia in extremely premature infants," *Pediatrics & Neonatology*, vol. 61, no. 3, pp. 290–299, 2020.
[54] J. Jaskari, J. Myllärinen, M. Leskinen, A. B. Rad, J. Hollmen, S. Andersson, and S. Särkkä, "Machine learning methods for neonatal mortality and morbidity classification," *Ieee Access*, vol. 8, pp. 123 347–123 358, 2020.
[55] T.-Y. Wu, W.-T. Lin, Y.-J. Chen, Y.-S. Chang, C.-H. Lin, and Y.-J. Lin, "Machine learning to predict late respiratory support in preterm infants: a retrospective cohort study," *Scientific reports*, vol. 13, no. 1, p. 2839, 2023.
[56] B. Mairhörmann, A. Castelblanco, F. Häfner, V. Koliogiannis, L. Haist, D. Winter, A. Flemmer, H. Ehrhardt, S. Stöcklein, O. Dietrich *et al.*, "Automated mri lung segmentation and 3d morphologic features for quantification of neonatal lung disease," *Radiology: Artificial Intelligence*, vol. 5, no. 6, p. e220239, 2023.

FLÁVIO LEANDRO DE MORAIS is a PhD student in the Graduate Program in Computer Engineering (PPGEC) at the University of Pernambuco (UPE), Brazil. He holds a Master's degree in Computer Engineering from the same program (2024), where he developed computational models to predict low birth weight. He earned his Bachelor's degree in Information Systems from the University of Pernambuco, Caruaru campus (2021). Currently, his doctoral research—under the supervision of Professor Patricia Takako Endo—focuses on the development of a platform to improve the quality of neonatal care using artificial intelligence models. Flávio is a member of the dotLAB Brazil research group, which collaborates with Dublin City University (DCU). His main research interests lie in machine learning, artificial intelligence, and data analysis, particularly applied to the fields of healthcare.

RAYSA CARLA LEAL DA SILVA is an undergraduate student in Information Systems at Universidade de Pernambuco (UPE, since 2022) and holds a Bachelor's degree (2021) in Pharmacy. Conducts research in data analysis, data visualization, and the application of artificial intelligence models in health.

ANNA BEATRIZ SILVA is a Master's student in Computer Engineering at the University of Pernambuco (UPE), researching the use of artificial intelligence models for predicting bronchopulmonary dysplasia. Holds a Bachelor's degree in Information Systems from UPE (2024).



ESTEFANI PONTES SIMÃO holds a Ph.D. in Biological Sciences with an emphasis on Biotechnology and a postgraduate degree in Microbiology. She is a substitute professor at the Universidade Federal do Rio Grande do Norte in the bachelor's degree in pharmacy and a professor at the Universidade Potiguar in health courses. She is a postdoctoral researcher at DotLab Brazil working on health data analysis using artificial intelligence.

MARIA EDUARDA FERRO DE MELLO is a researcher in the DotLab Brazil research group. She holds a Master's degree in Computer Engineering from the University of Pernambuco (UPE), in 2025, where she conducted research on predictive models for maternal and child health. She holds a Bachelor's degree in Biomedicine from the Federal University of Pernambuco (UFPE), in 2022. She has experience in Medical Mycology, working with yeast identification, detection of virulence factors, and susceptibility profiling to antifungal agents and other bioactive substances. Currently, her research focuses on artificial intelligence applied to healthcare, with an emphasis on tree-based predictive models.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3582503

STEPHANY PAULA DA SILVA CANEJO is a Master's student in Nursing at Escola de Enfermagem Nossa Senhora de Fátima of the Universidade de Pernambuco (FENSG/UPE), with a research focus on predictive models for clinically significant neonatal outcomes, particularly neonatal pulmonary diseases and conditions. Holds a Bachelor's degree in Nursing from Centro Universitário dos Guararapes (2021) and postgraduate specializations in Occupational Nursing (2023) and Public Health Nursing with emphasis on Health Surveillance (2023) from DNA Pós-graduação, as well as in Nursing Education (2024) from FAVENI. Has experience in epidemiological surveillance at Real Hospital Português and in geriatric care at Pousada Geriátrica Qualy Vida. Currently teaches nursing at Escola Irmã Dulce (EID) and EENSF.

KATIA MARIA MENDES is the current nursing manager of the neonatal unit at the Amaury de Medeiros Integrated Health Center (CISAM-UPE), where she also coordinates the Neonatal Multiprofessional Residency Program. She has worked in neonatal care since 1992 and in neonatal nursing management since 1997. Professor of Neonatal and Child Health since 2015. Holds a Master's degree in Nursing (2015, UPE/UEPB) and a Bachelor's in Nursing with a focus on Public Health (1991, UFPE). Specialized in Pediatric Nursing (1997), Hospital Management (2008), Nursing Education (2005), Stomatherapy – TiSOBEST (2011), and Neuroeducation in Early Childhood (2017).

WALDEMAR BRANDÃO NETO is an assistant professor at the Universidade de Pernambuco (UPE), working in the field of pediatric and adolescent health nursing. He is a permanent faculty member of the joint graduate program in nursing (PAPGEnf UPE/UEPB) and Editor-in-Chief of Revista Enfermagem Digital Cuidado e Promoção da Saúde (REDCPS). Holds a Ph.D. (2017) in child and adolescent health from Universidade Federal de Pernambuco (UFPE), a Master's (2012) in nursing with a focus on health promotion from Universidade de Pernambuco and Universidade Estadual da Paraíba (UPE/UEPB), and has professional experience in neonatal intensive care. Coordinates and collaborates on research in parental empowerment in NICUs, child satisfaction with nursing care, school health nursing and violence prevention, predictive models for maternal-neonatal care, and adolescent participation in health education.

JACKSON RANIEL FLORENCIO DA SILVA is a professor of Software Engineering at the University of Pernambuco , in the Information Systems program. Holds a Master's in Computer Science (UFPE) and a Bachelor's in Information Systems (UPE), and is a Ph.D. in Computer Science at UFPE (2024). His research focuses on software engineering, information systems, and health informatics.

MAICON HERVERTON LINO FERREIRA DA SILVA BARROS is a Postdoctoral Researcher in Computer Engineering at the Universidade de Pernambuco (UPE), Brazil. He earned his Ph.D. in Computer Engineering from the Programa de Pós-graduação em Engenharia de Computação (PPGEC), UPE, in 2024, following a Master's in Applied Informatics from the Universidade Federal Rural de Pernambuco (UFRPE) in 2013, and a Bachelor's in Information Systems from UFRPE in 2011. Is a member of the dotLAB Brazil research group, in collaboration with Dublin City University (DCU), and has contributed as a scientific reviewer for journals including Frontiers, Heliyon (Elsevier), and iScience (Elsevier). He is an IEEE member (94484272), and his research interests focus on artificial intelligence, machine learning, and data analysis, particularly in healthcare contexts.

PATRICIA TAKAKO ENDO is a full professor at the Universidade de Pernambuco (UPE), leads the dotLAB Brazil research group at UPE, focusing on data science and artificial intelligence applied to health, especially in neglected tropical diseases and maternal-neonatal health. Holds a Ph.D. (2014) and a Master 92008) in Computer Science from Universidade Federal de Pernambuco, and a Bachelor's in Computer Engineering from the Universidade Federal do Pará (2005).

• • •