From Prediction to Practice: A Rapid Review on the Opportunities and Challenges of using Artificial Intelligence in the Neonatal Intensive Care Unit (NICU)

Samantha Tudor, Urooj Raza Khan, Risha Bhatia, Michael Liem, Tafheem Ahmad Wani, James Boyd

Submitted to: Journal of Medical Internet Research
on: June 14, 2024

Table of Contents

From Prediction to Practice: A Rapid Review on the Opportunities and Challenges of using Artificial Intelligence in the Neonatal Intensive Care Unit (NICU)

Samantha Tudor[1] BBIS, MDH; Urooj Raza Khan[1] PhD; Risha Bhatia[2] PhD; Michael Liem[1] PhD; Tafheem Ahmad Wani[1] PhD; James Boyd[1] PhD

[1]La Trobe University Melbourne AU
[2]Monash University Melbourne AU

Corresponding Author:
Samantha Tudor BBIS, MDH
La Trobe University
Plenty Rd Bundoora
Melbourne
AU

Abstract

Background: The use of Artificial Intelligence (AI) in healthcare has been steadily increasing for over two decades. Integrating AI into Neonatal Intensive Care Units (NICUs) has the potential to reshape neonatal care and improve outcomes.

Objective: To analyse the current AI research landscape for predicting clinical outcomes and length of stay in the NICU, and to explore the benefits and challenges of utilising AI in the NICU for these predictions.

Methods: A rapid review was conducted across 6 databases, PubMed, Embase, CIHANL, Cochrane Library, Informit, and La Trobe Library, to identify Englished-language peer reviewed articles published between January 2017 and March 2023 that focused on the use of AI for predicting length of stay and clinical outcomes for NICU patient. A thematic analysis of AI application in NICUs from the articles identified was conducted.

Results: A total of 24 articles were included in the review, with AI applied in NICU settings to predict comorbidities (18/24), mortality (4/24), and length of stay (2/24). Sixteen of the studies were in the exploration stage, lacking a cohesive AI strategy, while the remaining eight were in the emerging stage, where the exploration of AI had been conducted systematically. None of the studies reported a fully integrated AI solution in a NICU setting. This review also identified several critical challenges, including data quality, clinical interpretability, model generalisation, and ethical considerations. Despite the lack of maturity and the challenges of AI in the NICU, the thematic analysis revealed four themes of potential AI application in enhancing NICU care: data-driven insights and predictive models, advancements in medical imaging, improved risk stratification, and personalised neonatal care and interventions.

Conclusions: AI provides opportunities, particularly in medical imaging and data-driven insights, offering the potential for improved diagnostic accuracy and personalised care. Data quality and ethical considerations are challenges to be addressed. Bridging these gaps can harness the transformative potential of AI to enhance neonatal care and healthcare delivery. Future research should prioritise practical implementation and ethical guidelines to fully realise the benefits of AI in the NICU.

(JMIR Preprints 14/06/2024:63175)
DOI: https://doi.org/10.2196/preprints.63175

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Rapid Review

Name	Designation	Affiliation
S Tudor	MDH	La Trobe University
U Khan	PhD	La Trobe University
R Bhatia	PhD	Monash University
M Liem	PhD	La Trobe University
T Wani	PhD	La Trobe University
J Boyd	PhD	La Trobe University

S Tudor, MDH; U Khan, PhD; R Bhatia, MBBCh, MA, MRCPCH, FRACP, AFRACMA, PhD; M Liem, PhD; T Wani, PhD; J H Boyd PhD

From Prediction to Practice: A Rapid Review on the Opportunities and Challenges of using Artificial Intelligence in the Neonatal Intensive Care Unit (NICU)

Abstract

Background: The use of Artificial Intelligence (AI) in healthcare has been steadily increasing for over two decades. Integrating AI into Neonatal Intensive Care Units (NICUs) has the potential to reshape neonatal care and improve outcomes.

Objective: To analyse the current AI research landscape for predicting clinical outcomes and length of stay in the NICU, and to explore the benefits and challenges of utilising AI in the NICU for these predictions.

Method: A rapid review was conducted across 6 databases, PubMed, Embase, CIHANL, Cochrane Library, Informit, and La Trobe Library, to identify Englished-language peer reviewed articles published between January 2017 and March 2023 that focused on the use of AI for predicting length of stay and clinical outcomes for NICU patient. A thematic analysis of AI application in NICUs from the articles identified was conducted.

Results: A total of 24 articles were included in the review, with AI applied in NICU settings to predict comorbidities (18/24), mortality (4/24), and length of stay (2/24). Sixteen of the studies were in the exploration stage, lacking a cohesive AI strategy, while the remaining eight were in the emerging stage, where the exploration of AI had been conducted systematically. None of the studies reported a fully integrated AI solution in a NICU setting. This review also identified several critical challenges, including data quality, clinical interpretability, model generalisation, and ethical considerations. Despite the lack of maturity and the challenges of AI in the NICU, the thematic analysis revealed four themes of potential AI application in enhancing NICU care: data-driven insights and predictive models, advancements in medical imaging, improved risk stratification, and personalised neonatal care and interventions.

Conclusion: AI provides opportunities, particularly in medical imaging and data-driven insights, offering the potential for improved diagnostic accuracy and personalised care. Data quality and ethical considerations are challenges to be addressed. Bridging these gaps can harness the transformative potential of AI to enhance neonatal care and healthcare delivery. Future research should prioritise practical implementation and ethical guidelines to fully realise the benefits of AI in the NICU.

Keywords: Artificial Intelligence; Deep Learning; Machine Learning; Neonatal Intensive Care Unit (NICU); Predictive Modelling; Length of Stay; Clinical Outcomes; Rapid Review.

Introduction

Background

Neonatal Intensive Care Units (NICUs) are specialised hospital units providing intensive medical care for critically ill newborns, particularly those born prematurely or with medical and surgical conditions that require close monitoring and specialised interventions and treatment. One in ten babies is born prematurely or sick, emphasising the essential role of neonatal care in promoting their well-being and survival (1). With an estimated 15 million premature births annually and around 1 million child deaths due to preterm complications each year (2) the significance of specialised care is evident. Surviving infants may encounter lifelong challenges, highlighting the need for effective neonatal care to ensure healthy development and improved long-term outcomes.

Within the complex ecosystem of the NICU, healthcare providers rely on the expertise of skilled clinicians and essential medical devices to deliver specialised care. The implementation of Electronic Medical Records (EMR) has made vast volumes of data more accessible in recent years, contributing to significant advances in medical research and technology, leading to improved survival rates for NICU patients (3). However, these advancements come with increased demands on healthcare resources, including specialised staff, equipment, and facilities. Innovative approaches are essential to address these challenges and enhance overall healthcare efficacy.

One avenue for improvement is the prediction of length of stay in the NICU. Accurate length of stay predictions are crucial for resource allocation, discharge planning, and optimising care pathways (4). Prolonged stays can have significant implications on neonatal development, parent-newborn interactions, and family well-being (5, 6). Additionally, predicting clinical outcomes is equally important, as it allows for monitoring essential developmental milestones, potential delays or improvements, and the lasting effects of medical interventions on NICU patients (7-12).

Healthcare providers face several challenges when predicting length of stay or clinical outcomes in the NICU setting. The vast amounts of data, including patient vitals, laboratory results, imaging data, and clinical notes, can be overwhelming to process manually (13). Limited availability of human resources in high-stress NICU settings further compounds the challenge, where teams must balance heavy workloads and rapid decision-making (14).

Furthermore, the dynamic nature of neonatal care, marked by rapid changes in health status, adds another layer of complexity to accurate predictions. In this context, Artificial Intelligence (AI) technology emerges as a potential solution. AI refers to the capability of computer systems to carry out tasks that typically require human intelligence, including prediction, learning, and decision-making (15). While traditional software and other computer systems rely on predetermined rules and instructions to accomplish specific tasks, AI sets itself apart by its capacity to learn from data and continually enhance performance without explicit programming (16).

In healthcare, predictive analytics powered by AI has proven effective for disease diagnosis and risk stratification, offering insights from patient data, medical records, and imaging results (17, 18). However, AI integration in the NICU environment is not without its challenges. Concerns related to data privacy, algorithm bias, and the need for transparent and interpretable models raise ethical considerations (19, 20). Understanding these challenges is essential for the responsible and successful implementation of AI for predictive analytics in neonatal care.

Research Focus & Aims

This review aimed to address the research question, "What are the opportunities and challenges in using Artificial Intelligence in the Neonatal Intensive Care Unit (NICU) to predict length of stay and clinical outcomes?".

The objective was to analyse the current AI research landscape for predicting clinical outcomes and length of stay in the NICU, exploring the benefits and challenges of utilising AI in the NICU for these predictions. This will help identify gaps in AI applications for predicting clinical outcomes and length of stay in the NICU and offer recommendations for future research.

Method

A rapid review was conducted to explore the opportunities and challenges associated with using AI technology in the NICU for predicting length of stay and clinical outcomes. This methodology, a condensed form of a systematic review, can provide valuable recommendations to inform policy and systems decisions (21). The rapid review method suits the growing digital health context, focusing on precise interventions and emerging technologies. It facilitates timely insights for decision-making, guiding future research and indicating AI integration gaps in neonatal care, following the guidelines outlined in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Statement (22).

Database Search

English language peer-reviewed studies between January 2017 and March 2023 were searched in six databases (Embase, Medline, CINAHL, Cochrane, Informit and La Trobe Library). The search utilised relevant search strings and Medical Subject Headings (MeSH). The keywords included Artificial Intelligence, NICU, length of stay, and Outcome (refer to Table 1).

Eligibility Criteria

This inclusion criteria focused on NICU patients and the adoption of AI technology for predicting outcomes and length of stay. Both retrospective and prospective studies were included for a comprehensive view. Retrospective studies offer historical insights through EHRs (Electronic Health Records), while prospective studies provide real-time data for dynamic observations. The combination ensured a comprehensive evaluation, considering AI's opportunities and challenges in the NICU. Studies within the last five years, in English, and with full-text availability were included, aiming to understand AI's role in predicting NICU outcomes and length of stay.

Study Selection and Quality Appraisal

A total of 811 studies were obtained and imported into Covidence Systematic Review Software (23). Among them, 85 duplicates were removed, resulting in 726 studies for screening. Out of the initial 726 studies, 169 were retained following a title and abstract review. The final selection comprised 24 articles deemed appropriate and included following full review. A PRISMA flow diagram (Figure 1) illustrates the process of article selection and exclusion.

The included studies underwent quality assessment using modified WHO guidelines (21) and preferred QCC (Quality Criteria Checklist) (24) due to varied study designs. Appendix Table 2 contains detailed ratings, evaluating bias and quality.

Data Extraction

Relevant data from the selected studies were systematically extracted and recorded in Microsoft Excel (25). A customised template, designed by the research team specifically for this rapid review, was utilised for data extraction. The template included key features, such as study characteristics (background, methods, results, conclusions, impact), study details (type of study, keywords, country, number of participants), technology-related information (technology type, algorithm/model details), outcome (length of stay/outcome), evaluation measures, as well as insights on potential opportunities, challenges, discussed gaps, further research needs, and suggested improvement areas. The extracted data were organised and analysed for synthesising key findings and identifying and coding emerging themes relating to the research objectives.

Assessment of Maturity Levels in AI Adoption

The maturity levels of AI Adoption in the included studies were assessed using a structured framework defining "Exploring," "Emerging/Activating," or "Integrated Ecosystem" stages (26). The evaluation considered national strategies, systematic exploration, and policy support. Articles were categorised based on their alignment with these dimensions, ensuring consistent and reliable assessments.

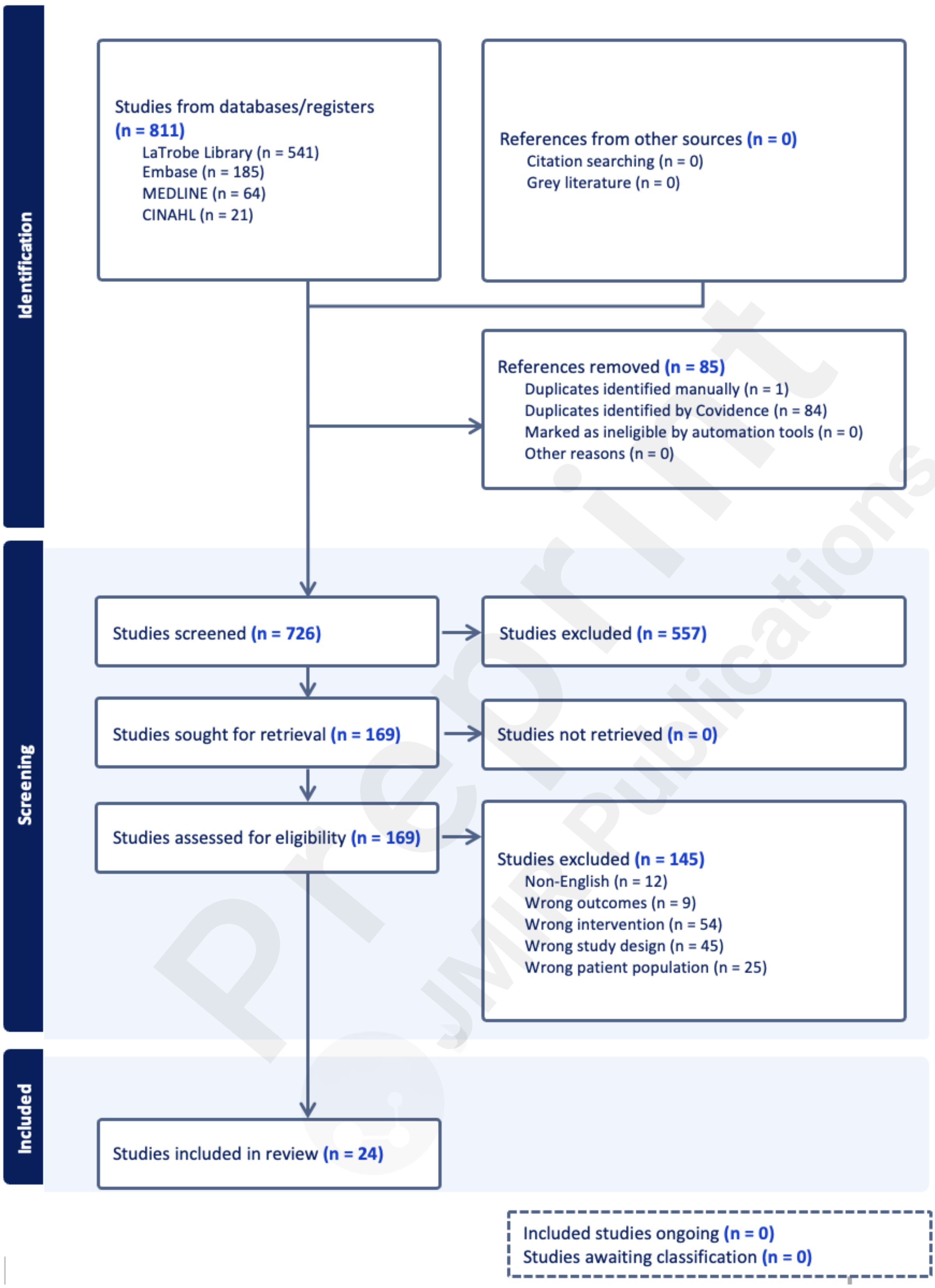

Figure 1 – PRISMA Flow Diagram

Results

Study Characteristics

The studies, as seen in Figure 2, include various technology interventions in NICU research: seven using Machine Learning (27-33), three combining Machine and Deep Learning (34-36), and fourteen exclusively focusing on Deep Learning (37-50). Geographically, thirteen studies are from the USA (27, 28, 32, 36, 37, 39, 41, 42, 44, 46-49), with others from Austria (two) (29, 40), Taiwan (two) (43, 45), and six in different countries including Argentina (31), China (35), Denmark (33), Iran (50), Italy (34), and Tanzania (30), along with one multinational dataset (38). Study designs include fifteen retrospective, six prospective single-centre, and three multi-centre studies. Participant cohorts vary, with seven studies featuring small cohorts, eight with medium, and nine with large cohorts.

Of the included studies, 16 fall under the "Exploring" category, lacking a cohesive AI strategy (27, 28, 30, 31, 35, 37-39, 41, 43-46, 48-50). Eight articles were classified as "Emerging/Activating," showing systematic exploration but without a fully established AI ecosystem (29, 32-34, 36, 40, 42, 47). None were categorised under "Integrated Ecosystem", indicating a mature and fully established AI ecosystem in neonatal care remains a future aspiration.

Studies primarily analyse three key outcomes: length of stay (two instances), morbidities (eighteen studies), and mortality (four studies). These characteristics offer insights into the diverse range of studies in AI's current NICU research.

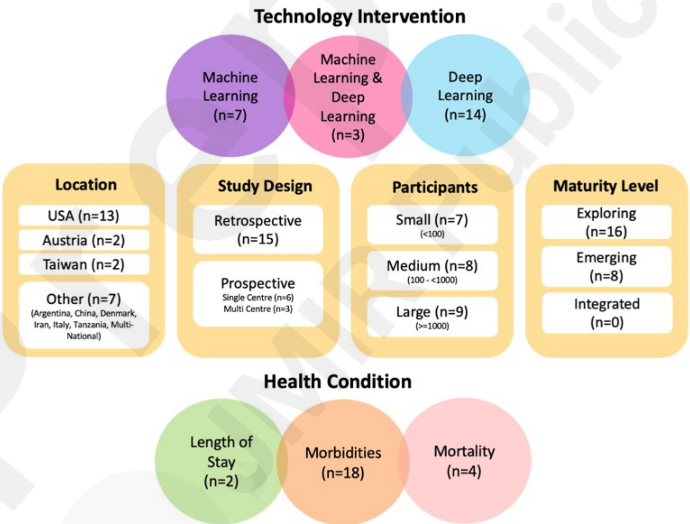

Figure 2 - Study Characteristics

Characteristics of AI predictions for NICU patients (n=24), including technology interventions, location, type of study design, number of participants, the AI maturity level, and the outcome predicted. USA: United States of America.

Temporal Technological Trends. The included literature depicts temporal technology trends (Figure 3), revealing shifts in preference within neonatal care research. In 2020 and 2021, five studies emphasised Deep Learning (38-40, 42-44, 46, 48-50), signifying its growing recognition in predicting outcomes. This trend continues in 2022, with three studies opting for Deep Learning (37, 45, 47). Post-2021, studies integrate Machine Learning and Deep Learning (34-36) methods, reflecting evolving research methodologies and the importance of advanced AI, particularly Deep Learning, in neonatal care predictions.

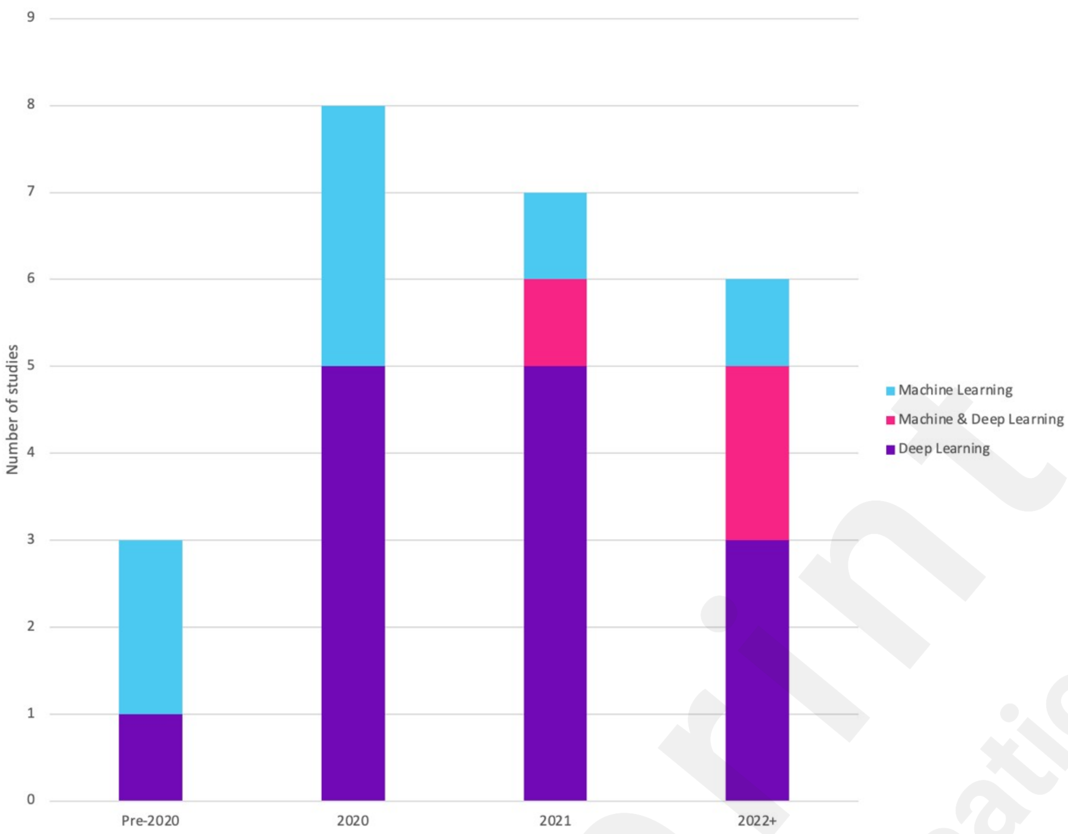

Figure 3 - Temporal Technology Trends

Illustrates the temporal distribution of included studies (totalling 24) across different technologies, including Machine Learning, Deep Learning, and a combination of both, based on their publication years.

Morbidities Studied. The studies in this review outlines AI technology's applications in predicting specific NICU clinical outcomes (Figure 4), revealing prevalent trends. 'Growth & Development' leads with seven studies, mostly utilising Deep Learning (29, 37, 40-42, 48, 49), emphasising AI's role in assessing infant development. 'Ophthalmological' outcomes are studied in three exclusively Deep Learning-driven articles (38, 39, 43). 'Respiratory' outcomes involve four studies, equally leveraging Deep Learning and Machine Learning studies (31, 33-35), underscoring their relevance in neonatal respiratory health. The 'Other' category spans a variety of outcomes, with different AI techniques employed. 'Mortality' is studied by four articles, mostly using Deep Learning (30, 44, 47, 50). Lastly, 'Length of Stay' involves two studies, evenly split between Deep Learning and Machine Learning (27, 45). Overall, Deep Learning emerges as the predominant approach across categories, showcasing its diverse applicability in neonatal care. This comprehensive analysis accentuates AI's complex role in enhancing infant health and well-being.

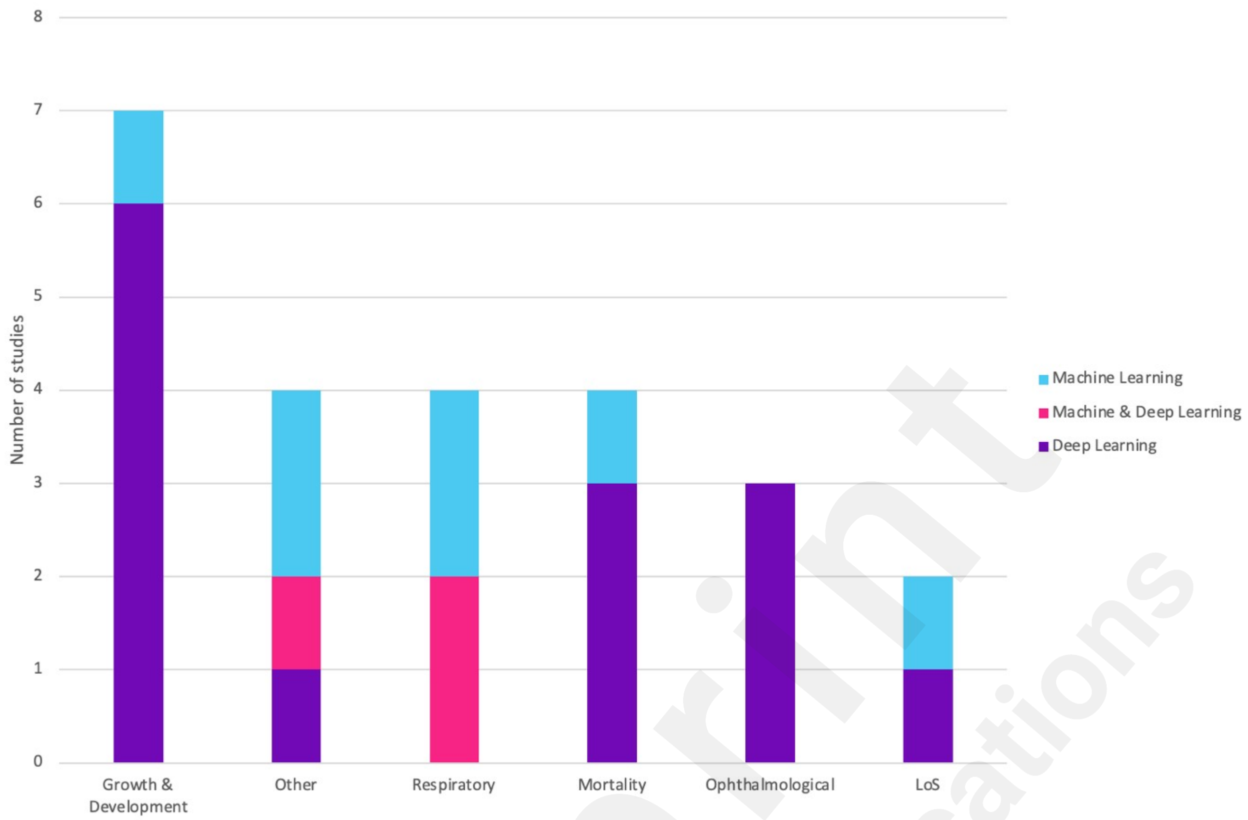

Figure 4 - Outcome Type by Technology Type

Illustrates the distribution of 24 included studies across different outcome types (Growth & development, Other, Respiratory, Mortality, Ophthalmological, LoS: Length of Stay) and technology types (Machine Learning, Both Machine Learning and Deep Learning, Deep Learning). The chart provides insights into the focus areas of these studies and the technology used to address specific neonatal care outcomes.

Predicting Outcomes Evaluation Measures. The studies employ diverse predictive model evaluation measures (Table 4) to assess machine or deep learning model accuracy in predicting outcomes or classifications in neonatal care. Figure 5 shows AI model performance across outcome categories, labelled 'Excellent,' 'Good,' 'Moderate,' or 'Not Reported' (NR). In the 'Excellent' category, Deep Learning excelled in 'Other,' 'Ophthalmological,' and 'Mortality' outcomes, while Machine Learning performed well in 'Respiratory' outcomes. 'Good' models, primarily Deep Learning, demonstrated strength in 'Growth & Development' and 'Mortality,' while Machine & Deep Learning showed promise in 'Other' and 'Respiratory' outcomes. 'Moderate' performance was observed in Deep Learning for 'Growth & Development' and 'Length of Stay,' and in Machine Learning for 'Mortality.' Notably, five studies did not report performance. This figure highlights AI strengths in various outcomes, indicating research directions and areas needing improvement.

Table 4 - Evaluation Measures

Evaluation Measure	Number of Studies	Study
AUROC	15	(27, 28, 34-38, 41, 42, 44-47, 49, 50)
Correlation (r)	1	(29)
Error Percentiles	1	(40)

Normalized average Root Mean Squared Error (RMSE)	1	(43)
Sensitivity & Specificity	3	(30, 32, 33)
Not Reported	3	(31, 39, 48)

Table 4 demonstrates various evaluation measures for predicting neonatal outcomes. Among these, the most common metric is AUROC, used in 15 studies. Other measures like Correlation (r), Error Percentiles, and Normalized Average RMSE were less frequent, each found in a single study. Sensitivity & Specificity were the focus of three studies. However, specific evaluation measures were not reported in three instances, showcasing diverse approaches to assessing AI models in neonatal care studies.

Performance by Outcome and Type of AI. Figure 7 provides an overview of the performance of AI models in predicting neonatal outcomes (refer Appendix 2 Table 8), categorised as 'Excellent,' 'Good,' 'Moderate,' or 'Not Reported' (NR), and further differentiated by the type of technology employed. In the 'Excellent' category, Deep Learning models demonstrated exceptional predictive abilities across various outcomes, including 'Other,' 'Ophthalmological,' and 'Mortality,' while Machine Learning models were effective in predicting 'Respiratory' outcomes. The 'Good' category saw Deep Learning models excelling in 'Growth & Development' and 'Mortality,' with Machine & Deep Learning models performing well in 'Other' and 'Respiratory' outcomes. Machine Learning models also demonstrated proficiency in predicting 'Length of Stay.' 'Moderate' AI models, predominantly Deep Learning, showed moderate performance in 'Growth & Development' and 'Length of Stay,' while Machine Learning models displayed moderate performance in one 'Mortality' study.

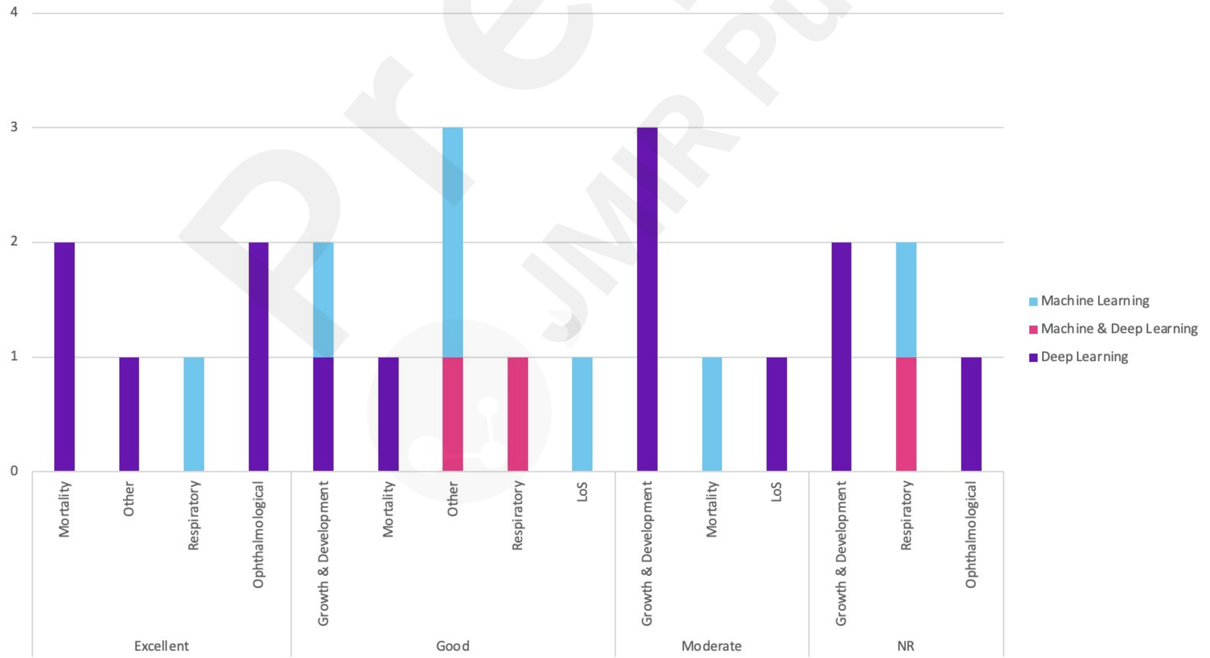

Figure 5 - Performance by Outcome by Type of AI

Displays the performance assessment of AI models (Machine Learning, Deep Learning and a combination of both) across different neonatal care outcome types (Growth & Development, Other, Respiratory, Mortality, Ophthalmological, LoS: Length of Stay). Performance is categorised into four levels: Excellent, Good, Moderate, and NR (Not Reported). The chart provides insights into the quality of AI models in predicting specific neonatal care outcomes across different AI types and outcome categories.

Insights into potential opportunities for AI in NICU

Opportunities for AI applications in NICU have been themed into five (summarised in Appendix 3) from enhancing NICU care through data-driven insights and predictive models to utilising advancements in medical imaging, improving risk stratification, and personalising neonatal care and interventions.

Advancements in Medical imaging. Several studies within the selection highlight medical imaging advancements. Two studies focus on Retinopathy of Prematurity (ROP), showcasing the role of deep learning in standardising disease assessment and predicting visual outcomes post-treatment (38, 43). Choi et al. (2020) also explored deep learning for ROP severity assessment, highlighting broader medical implications. He et al. (2020) introduced a multi-task deep transfer learning model for early neurodevelopmental outcome prediction in preterm infants, emphasising imaging's predictive role. Saha et al. (2020) used deep learning CNNs to forecast motor outcomes via brain diffusion MRI data. Lastly, Lure et al. (2021) employed machine learning to distinguish neonatal conditions, enhancing clinical decision-making. These studies highlight medical imaging's potential in early diagnosis and monitoring of neonatal conditions.

Data-Driven Insights and Predictive Models. Data-driven insights and predictive models in neonatal care have the potential to enhance outcomes (27-29, 32, 38, 44, 45, 47, 48). They contribute to care planning by using machine learning to predict factors like length of stay, disease severity, and mortality risk, aiding healthcare providers (29, 44, 45). Some explore risk assessment, detecting and mitigating severe neonatal morbidities (27, 28, 49). Two introduce adaptive machine learning algorithms for dynamic patient status adaptation (34, 47). Lastly, Ruixiang et al. (2021) highlight data-driven insights' transformative role in advancing neonatal research and medical care by improving predictions, care planning, and risk assessment.

Improving Understanding and Risk Stratification. The theme presents a cohesive effort toward understanding neonatal conditions and individual risk stratification through data-driven approaches. Three articles focus on machine learning-driven risk assessment for conditions like Bronchopulmonary Dysplasia (BPD), in-hospital length of stay, and mortality risk (41, 44, 45). Leveraging extensive datasets, these studies offer critical insights into associated risk factors, enabling early intervention. Additionally, three articles delve into NICU utilisation trends (27), severe morbidity risk (28), and chronic lung disease determinants in Very Low Birth Weight (VLBW) infants (31). Two studies emphasise dynamic machine learning algorithms for mortality prediction and Pulmonary Hypertension (PH) (34, 47). Together, these studies deepen our understanding of neonatal conditions, reshaping disease paradigms and paving the way for more effective interventions.

Personalised Neonatal Care and Intervention. Numerous studies contribute to the theme of 'Personalised Neonatal Care and Intervention', extending the horizons of individualised care for neonates. These studies explore personalised care plans and interventions. They investigate trends in NICU utilisation (27), assess risk factors for severe neonatal morbidity (28), predict motor outcomes (49), and differentiate between critical neonatal conditions (46). By leveraging predictive models and continuous monitoring, these studies seek to tailor care strategies to the specific needs of each neonate. This data-driven approach holds the potential to revolutionise neonatal care by optimising interventions and potentially leading to better clinical outcomes.

Insights into potential challenges for AI in NICU

Literature indicated some concerns that span multiple facets of AI application in NICU, ranging from data quality and quantity issues to clinical interpretability, model generalisation, clinical and diagnostic variability, and ethical and regulatory considerations. Appendix 3, Table 5 provides a condensed overview of the key thematic challenges encountered across the various included studies.

Data Quality and Quantity. Several studies face recurring challenges with healthcare data in the NICU, including issues related to quality, quantity, and availability (33, 35, 37, 39, 41, 43, 44, 48-50). Addressed by multiple articles, these challenges span data quality's influence on predictive models, the need for larger datasets, variability due to data quality, and concerns about both data quality and quantity. The authors highlight these challenges' significance in developing effective machine learning solutions for neonatal care. Overcoming these hurdles is crucial for successful machine learning implementation, ultimately enhancing healthcare outcomes for neonates.

Clinical Interpretability and Usability. Challenges in clinical interpretability and usability arise as machine-generated insights need integration into clinical practice. Studies emphasise the need for standardised measures for neurodevelopmental assessment, ethical data sharing, complex functional connectivity estimation, user-friendly tools in low-resource settings, and distinguishing specific illnesses (27, 29, 30, 32, 36, 38, 40, 43). Overcoming these challenges is crucial for meaningful AI integration in the NICU and its adoption by healthcare professionals, and improvement healthcare outcomes these patients.

Model Generalisation and Validation. The challenge of ensuring AI model reliability across diverse populations in neonatal care is evident. He et al. (2022) advocate for external validation of AI models for Bronchopulmonary Dysplasia (BPD) severity prediction across various populations. Chen et al. (2021) stress the need for adaptable and generalisable models, particularly for diverse patient demographics and different camera systems in Retinopathy of Prematurity (ROP). Similarly, Choi et al. (2020) focus on creating versatile deep learning scales for ROP applicable across various centres. Saha et al. (2020) emphasise the importance of generalising predictions for diverse neonatal populations in motor outcome prediction. Verder et al. (2021) highlight the necessity of generalising and validating models for accurate predictions in diverse clinical scenarios for BPD. Sheikhtaheri et al. (2021) discuss the validation of predictive models for neonatal deaths in NICUs, recognising the need for external validation. Ofman et al. (2019) note the challenge of studying disease determinants across multiple centres in chronic lung disease. Amodeo et al. (2021) explore the complexities of predicting Pulmonary Hypertension (PH) outcomes in a diverse patient population.

Clinical and Diagnostic Variability. Handling clinical and diagnostic variability is a critical challenge in the NICU, with particular significance in fields such as ophthalmology and critical care, where even subtle variations can have a profound impact on outcomes (29, 36, 39, 44, 46, 49). Iyer et al. (2022) highlight the need for standardised, objective, and scalable measures for neurodevelopmental assessment, contrasting with current subjective and non-scalable methods. Similarly, Lee et al. (2021) note the necessity for a larger training dataset and consideration of site-specific differences to improve model performance. In ophthalmology, Choi et al. (2020) highlight the diagnostic variability and subjective quantification of severity levels, hindering the interpretation of clinical trial data. Saha et al. (2020) raise concerns about the limited sample size and heterogeneity of brain injuries, which may lead to overfitting and poor prediction performance. Furthermore, Shalish et al. (2017) point out practice variability and uncertainty in defining extubation failure, complicating clinical decision-making.

Addressing imbalanced data and the need for standardised physiological testing are highlighted by Ofman et al. (2019) and Sheikhtaheri et al. (2021), respectively. Moreover, Amodeo et al. (2021) underscore the challenges in measuring lung vascularisation accurately, given the techniques' reproducibility issues and the impact of hemodynamic changes during neonatal transition. These challenges collectively illustrate the complexity and variability inherent in neonatal care, underscoring the importance of developing robust AI models and data-driven solutions to enhance precision and reliability in clinical decision-making.

Ethical and Regulatory Challenges. The integration of machine learning in neonatal care introduces ethical and regulatory complexities. Chen et al. (2021) emphasise the ethical challenge of data sharing, weighing the practicality against privacy concerns in multi-institutional datasets. The study highlights regulatory hurdles in validating models across diverse populations and camera systems. Patel et al. (2022) note the regulatory risks of overfitting single-site datasets and ethical concerns around addressing missing data. They also note the need to evaluate model performance with evolving clinical cohorts, which has regulatory implications. Ofman et al. (2019) underline ethical and regulatory issues linked to non-standardised physiologic testing in NICUs, impacting research and clinical practices. Kovacs et al. (2021) note the ethical concern of deploying tools in resource-limited settings, balancing computational limitations with clinicians' interpretability. These articles emphasise the intricate ethical and regulatory challenges in AI integration in neonatal care, requiring careful navigation for ethical standards and regulatory compliance.

Discussion

The objective of this rapid review was to analyse the current AI research landscape for predicting clinical outcomes and length of stay in the NICU, exploring benefits and challenges of utilising AI in the NICU for these predictions. The surge in studies between 2020-2021 reflects a growing recognition of AI's potential to reshape neonatal care, possibly accelerated by the COVID-19 pandemic's impact on healthcare systems and driving innovative solutions to mitigate disruptions. The sustained research momentum, including studies in 2022, signifies an enduring commitment to exploring AI's role in the NICU.

Deep Learning emerges as a dominant technology across various clinical outcomes, highlighting its effectiveness in automatically identifying relevant features from raw data, especially in medical imaging and clinical analysis. However, the field's maturity in AI integration within neonatal care remains in the early exploration stage, lacking a mature and cohesive ecosystem.

The growing presence of various national guidelines and policies highlight the progress AI implementation within healthcare (51-53). However, there is a need to prioritize the development of a national AI strategy tailored for neonatal care due to the unique challenges of this environment. Specific considerations for neonatal care include the need for specialised models that account for the unique physiology and comorbidities of premature infants, as well as the heightened need of ensuring patient safety and data privacy in a patient population that have continuous and complete medical records from birth.

The study reveals substantial opportunities for the use of AI in the NICU to predict length of stay and patient outcomes. These cover diverse areas: advancements in medical imaging, data-driven insights and predictive models, improved understanding and risk stratification, and personalised neonatal care and intervention.

The first opportunity focuses on AI's potential to transform neonatal care in the NICU by leveraging machine learning to perform predictive analytics to help inform the clinician during the clinical decision-making process, aiding in earlier interventions. However, there are research gaps in translating AI's theoretical effectiveness into real-world NICU settings (3-6, 54-58). This brings Model Generalization and Validation, the third challenge theme, that highlights the importance of ensuring AI models perform reliably across diverse patient populations. External validation across different populations is necessary for reliability and adaptability. The unique patient populations and scenarios in the NICU demand tailored approaches for model generalization and validation, vital for trustworthy predictions.

The second opportunity highlights the use of AI into NICU medical imaging, demonstrated by six studies. These studies primarily employ deep learning to analyse medical images, notably for conditions like Retinopathy of Prematurity (RoP). AI-driven imaging enhances accuracy, facilitating early intervention and personalized care plans. These findings not only affirm anticipated AI benefits in medical imaging but also underline its relevance in the NICU, particularly focusing on neonatal conditions. AI-powered imaging is key for precise diagnosis, standardized disease assessment, and predicting critical outcomes, aligning closely with literature expectations (15, 59, 60). It enables tailored, timely care plans for individual neonatal needs, a significant advantage in this setting. The challenge here is handling clinical and diagnostic variability in these settings, where even subtle variations can have a significant impact on outcomes. Addressing this challenge involves the development of models and algorithms that can adapt to diverse clinical contexts, accounting for nuances that might otherwise be overlooked. It calls for an ongoing effort to refine models and diagnostic tools to accommodate the inherent variability in neonatal care, ultimately improving the precision and reliability of clinical decision-making.

The third and fourth opportunities highlight the transformative power of data-driven insights and predictive models in the NICU. Thirteen studies within these themes leverage AI to translate healthcare data into actionable insights, aiding decision-making. They encompass predictive modelling, spanning length of stay, disease severity, and mortality risk, contributing to a deeper understanding of neonatal conditions. These studies deepen the comprehension of neonatal conditions and their determinants, aligning with the literature's focus on advancing scientific understanding and medical care in the NICU (43, 61-66). By employing data-driven approaches, they pave the way for redefining disease paradigms, ultimately enhancing interventions and care strategies. The challenge here is data quality and quantity, as it involves acquiring high-quality, sufficient data for AI model training. Data incompleteness and inconsistency, common issues in healthcare data, align with existing literature. Overcoming these challenges requires advanced algorithms and robust data management strategies within the NICU (65).

The final opportunity, Personalised Neonatal Care and Intervention, signals a significant shift in NICU healthcare. Studies focusing on tailoring care plans and interventions for individual neonatal needs explore personalised care plans, risk assessment, predicting motor outcomes, and distinguishing critical neonatal conditions. These efforts, utilizing predictive models and continuous monitoring, aim to optimize interventions. Personalised care plans carry substantial implications, potentially reducing unnecessary treatments, cutting healthcare costs, and improving clinical outcomes. These align with literature stressing the importance of individualised neonatal care (55, 67). This is closely related to challenges of Clinical Interpretability and Usability as well as managing clinical and diagnostic variability. Clinical Interpretability and Usability, focuses on the transparency and practicality of AI-driven models within the NICU. Healthcare providers in the NICU demand accurate predictions and clear understanding of AI outputs.

Regardless of any application, ethical and regulatory considerations and challenges remains in all applications. It emphasizes the complex landscape requiring careful navigation. Key ethical concerns include data sharing dilemmas, safeguarding patient data privacy, and security. Regulatory hurdles, especially in validating AI models across diverse populations and settings, require adaptable frameworks. Only four studies directly mention these ethical and regulatory challenges, suggesting a potential research gap in the NICU field of AI, requiring more exploration and consideration. Ensuring responsible AI adoption is pivotal for equity in healthcare access and maintaining high standards of care.

Future directions for NICU opportunities present key research areas. Firstly, integrating AI with medical imaging, especially for conditions like ROP, demands refined technology, larger datasets, and real-world validation. Secondly, enhancing data-driven insights and predictive models requires broader clinical scenarios and improved data quality via collaborative NICU efforts. Understanding neonatal conditions better, exploring diverse risk factors, and fostering multidisciplinary collaborations are key priorities. Practical AI application in NICUs, optimizing resource allocation and care, requires real-world implementation and thorough assessment. Lastly, refining scalable personalised care plans and interventions, while ethically considering AI personalisation, remains essential for expanding AI's role and improving neonatal outcomes.

Conclusion

The rapid review has highlighted substantial opportunities for AI in the NICU. Advancements in medical imaging, combined with AI, have the potential to improve diagnostic accuracy and enable early intervention in this field. Furthermore, data-driven insights and predictive models offer the opportunity to enhance clinical decision-making and deepen our understanding of neonatal conditions. Additionally, personalised neonatal care could optimise healthcare delivery for individual neonates.

Simultaneously, the study has revealed crucial challenges in integrating AI in the NICU. Issues related to data quality emphasise the need for robust data management strategies. Ensuring clinical interpretability and usability is essential to ensure AI tools align smoothly with clinical workflows, especially in the high-stress NICU environment. Moreover, achieving model generalisation and validation across diverse patient populations and addressing clinical and diagnostic variability are essential considerations. Ethical and regulatory challenges, including data privacy, security, and model validation, underscore the importance of responsible AI adoption in the NICU.

These findings align with existing literature, revealing the unique complexities of the NICU context. The identified research gaps include the need for practical AI implementations within the NICU, considering resource constraints and clinical requirements. Additionally, there is a need for the development of ethical frameworks and regulatory guidelines tailored specifically to the NICU environment. Future research should focus on practical implementation, ethical frameworks, and regulatory guidelines to realise AI's potential in the NICU.

This research holds potential for neonatal care, benefiting a range of stakeholders. Healthcare professionals, including neonatologists, nurses, and clinicians, can gain directly from the insights provided in this study. Healthcare institutions that manage NICUs also stand to benefit, as the research highlights the importance of practical AI implementation within the NICU, enabling them to optimise resource allocation and improve patient care. This research serves as an important resource for advancing AI technology in neonatal care, fostering a future of improved healthcare delivery and enhanced well-being for NICU patients and their families.

Limitations

The exploration of the potential of AI in the NICU to predict clinical outcomes and length of stay comes with a recognition of strengths and limitations that shape the scope and methodology of this study. The search strategy in this rapid review was designed to be comprehensive, incorporating a range of keywords and subject headings to capture relevant literature. The search terms were carefully chosen in collaboration with experienced research librarians to ensure inclusivity; however, the complexity of the AI field and evolving terminologies might have introduced limitations. It is possible that studies utilising emerging AI-related terminology were not captured. Furthermore, the inclusion criteria focused on academic research articles, which inadvertently excluded insights from private organisations and the grey literature. This limitation could result in missing valuable perspectives and data that could enhance understanding of AI applications in the NICU. Moreover, the decision to restrict the study to English full-text articles introduced a language bias, primarily focusing on studies from English-speaking countries. As a result, research conducted in other languages may have been excluded, limiting the diversity and global representation in this review. Additionally, this study's exclusion criteria specifically focused on interventions emphasising predictive aspects of AI, which could have led to an incomplete portrayal of AI's multifaceted role in the NICU. Non-predictive aspects of AI, while relevant, were not the central focus of this rapid review.

Despite these potential limitations, this study provides valuable insights into the challenges and opportunities in the emerging field of AI predictions in the NICU. The decision to focus on academic research articles ensured a rigorous examination of peer-reviewed literature, lending credibility to the findings. While recognising these limitations, the study serves as a foundation for further exploration and analysis in this dynamic and evolving field.

Future Directions

The rapid review process has exposed several research gaps. Firstly, there is a noticeable gap concerning the maturity level of AI implementation in NICUs. While existing studies have laid the foundation, further research is needed to extend AI maturity to the next stages of emerging and integration. Additionally, the scarcity of studies addressing ethical and regulatory aspects of AI in NICUs is evident. These areas, such as liability in cases of adverse outcomes resulting from AI predictions, warrant comprehensive investigation. Furthermore, the absence of studies reporting unfavourable or null results suggests potential reporting bias within the field. Finally, collaborative partnerships and patient and family engagement are instrumental in advancing AI research in NICUs.

Overall, these future directions and improvements collectively contribute to the ongoing evolution of AI research in NICUs, fostering a more comprehensive understanding of the opportunities and challenges while refining the methodology for more effective and robust reviews in the future.

Acknowledgements

ST was responsible for the conception of this study, literature review, and drafting of the manuscript. UK and RB contributed to the concept and review. ML, TW, and JB revised the manuscript. All authors contributed to the manuscript and approved the submitted version.

Conflicts of Interest

None declared.

Appendices

Appendix 1 – Search Method

Table 1 - Search Method and Number of Results per Database

DATABASE	SEARCH METHOD	RESULTS
OVID (MEDLINE, EMBASE)	(exp Artificial Intelligence/ OR AI OR "artificial intelligence" OR "machine learn*" OR "deep learn*" OR "neural network*") AND (exp Intensive Care, Neonatal/ OR exp Infant, Premature/ OR exp Infant, Premature, Diseases/ OR exp Intensive Care Units, Neonatal/) AND (("length of stay" OR "LoS").mp. OR ("outcome" OR "prognosis").mp.)	249 (64, 185)
CINAHL	((artificial intelligence OR ai OR a.i. OR "machine learning" OR "deep learning") AND (nicu OR "neonatal intensive care unit" OR "special care" OR "baby unit" OR "newborn intensive care")) AND (("length of stay" OR los OR "inpatient stay" OR "time in hospital" OR "time to discharge") OR (outcomes OR prognosis)) AND (English Language) AND (Published Date: 20170101-20230131)	21
COCHRANE	(MH "Artificial Intelligence") AND (MH "Intensive Care Units, Neonatal")	1
INFORMIT	"Artificial Intelligence" AND "Intensive Care Units, Neonatal"	0
LA TROBE LIBRARY	((Al OR "artificial intelligence" OR "machine learn*" OR "deep learn*" OR "neural network*") AND (NICU OR "neonatal intensive care unit*" OR "newborn intensive care unit*") AND ("length of stay" OR "LoS" OR "outcome" OR "prognosis")) AND (english language and last 5 years)	541

Table 2 - Quality Criteria Checklist

Study	Was the research question clearly stated?	**Was the selection of study subjects/ patients free from bias?**	**Were study groups compa rable?**	Was method of handling withdrawals described?	**Was blinding used to prevent introduc tion of bias?**	**Was the interventio n, compariso n, and any intervening factors described in detail**	Were outcome s clearly defined and the measure ments valid and reliable?	Was the statistical analysis appropriate for the study design and type of outcome indicators?	Do the conclusion s consider potential biases and limitations in the results?	Is bias due to study's funding or sponsorshi p unlikely?	**Qualit y Rating**
(Iyer et al., 2022)	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**High**
(He et al., 2022)	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	**High**
(Ali et al., 2022)	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	**Mediu m**
(Lin et al., 2022)	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear	Yes	Yes	**Mediu m**
(Lee et al., 2021)	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**High**
(Chen et al., 2021)	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**Mediu m**
(Gsch wandt ner et al.,	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	**Mediu m**

2020)											
(C. Y. Huang et al., 2020)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	**High**
(Braun et al., 2020)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	**High**
(Choi et al., 2020)	Yes	Yes	Unclear	No	Unclear	Yes	Yes	Unclear	Yes	Yes	**Medium**
(Saha et al., 2020)	Yes	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**Medium**
(C. Y. Huang et al., 2020)	Yes	Unclear	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**Medium**
(Hamilton et al., 2020)	Yes	Unclear	Yes	Unclear	No	Yes	Yes	Yes	Yes	Yes	**Medium**
(He et al., 2018)	Yes	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	**Medium**
(Kausch et al., 2022)	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**High**
(Verd	Yes	Unclear	Unclear	No	No	Yes	Yes	Unclear	Yes	Yes	**Mediu**

er et al., 2021)											**m**
(Patel et al., 2022)	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	Yes	**High**
(Ruixiang et al., 2021)	Yes	Unclear	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Unclear	Yes	**Medium**
(Shalish et al., 2017)	Yes	Unclear	Yes	No	Unclear	Yes	Yes	Unclear	Yes	Yes	**Medium**
(Sheikhtaheri et al., 2021)	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	**Medium**
(Ofman et al., 2019)	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes	Yes	**High**
(Amodeo et al., 2021)	Yes	Yes	Yes	Unclear	Unclear	Yes	Yes	Unclear	Yes	Yes	**High**
(Kovacs et al., 2021)	Yes	No	Yes	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes	**Medium**
(Lure et al.,	Yes	No	Yes	No	Yes	Unclear	Yes	Yes	Yes	Yes	**Medium**

2021)												

Table 1 - Included Studies

Title	**Published Year**	**Journal**	**Study**	**Type of study**	**Country**	**Number of participants**	**Type of technology**	**Outcome Category**
Bedside tracking of functional autonomic age in preterm infants.	2022	Pediatric research	Iyer 2022	Prospective study	Austria	67	Machine Learning	Growth & Development
Risk factors and machine learning prediction models for bronchopulmonary dysplasia severity in the Chinese population.	2022	World journal of pediatrics : WJP	He 2022	Retrospective study	China	471	Machine & Deep Learning	Respiratory
A self-training deep neural network for early prediction of cognitive deficits in very preterm infants using brain functional connectome data.	2022	Pediatric radiology	Ali 2022	Retrospective study	USA	343	Deep Learning	Growth & Development
Predicting in-hospital length of stay for very-low-birth-weight preterm infants using machine learning techniques.	2022	Journal of the Formosan Medical Association = Taiwan yi zhi	Lin 2022	Retrospective study	Taiwan	2940	Deep Learning	Length of Stay
Predicting mortality risk for preterm infants using random forest.	2021	Scientific reports	Lee 2021	Retrospective study	USA	275	Deep Learning	Mortality
Deep Learning for the Diagnosis of Stage in Retinopathy of Prematurity: Accuracy and Generalizability across Populations and Cameras.	2021	Ophthalmology. Retina	Chen 2021	Retrospective study	USA, India, Napal	1252	Deep Learning	Ophthalmological

Deep learning for estimation of functional brain maturation from EEG of premature neonates.	2020	Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference	Gschwandtner 2020	Retrospective study	Austria	43	Deep Learning	Growth & Development
A multi-task, multi-stage deep transfer learning model for early prediction of neurodevelopment in very preterm infants.	2020	Scientific reports	He 2020	Prospective study	USA	1226	Deep Learning	Growth & Development
Trends in Neonatal Intensive Care Unit Utilization in a Large Integrated Health Care System.	2020	JAMA network open	Braun 2020	Retrospective study	USA	39220	Machine Learning	Length of Stay
Variability in Plus Disease Identified Using a Deep Learning-Based Retinopathy of Prematurity Severity Scale.	2020	Ophthalmology. Retina	Choi 2020	Retrospective study	USA	871	Deep Learning	Ophthalmological
Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural	2020	NeuroImage	Saha 2020	Prospective study	USA	77	Deep Learning	Growth & Development

network (CNN) model.								
Prediction of visual outcomes by an artificial neural network following intravitreal injection and laser therapy for retinopathy of prematurity.	2020	The British journal of ophthalmology	Huang 2020	Retrospective study	Taiwan	60	Deep Learning	Ophthalmological
Estimating risk of severe neonatal morbidity in preterm births under 32 weeks of gestation.	2020	The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians	Hamilton 2020	Multicenter prospective study	USA	1039	Machine Learning	Other
Early prediction of cognitive deficits in very preterm infants using functional connectome data	2018	NeuroImage. Clinical	He 2018	Prospective study	USA	912	Deep Learning	Growth & Development

in an artificial neural network framework.								
Cardiorespiratory signature of neonatal sepsis: development and validation of prediction models in 3 NICUs	2023	Pediatric Research	Kausch 2023	Retrospective study	USA	2494	Machine & Deep Learning	Other
Bronchopulmonary dysplasia predicted at birth by artificial intelligence	2020	Acta paediatrica (Oslo, Norway : 1992)	Verder 2020	Multicenter prospective study	Denmark	61	Machine Learning	Respiratory
The criticality Index-mortality: A dynamic machine learning prediction algorithm for mortality prediction in children cared for in an ICU	2022	Frontiers in Pediatrics	Patel 2022	Retrospective study	USA	8399	Deep Learning	Mortality
Early Physical Linear Growth of Small-for-Gestational-Age Infants Based on Computer Analysis Method	2021	Journal of Healthcare Engineering	Ruixiang 2021	Retrospective study	USA		Deep Learning	Growth & Development
Prediction of Extubation readiness in extremely preterm infants by the automated analysis of cardiorespiratory behavior: study protocol.	2017	BMC Pediatrics	Shalish 2017	Multicenter prospective study	USA	170	Machine Learning	Other
Prediction of neonatal deaths in NICUs: development and validation of machine learning models.	2021	BMC Medical Informatics & Decision Making	Sheikhtaheri 2021	Retrospective study	Iran	1762	Deep Learning	Mortality
The discovery BPD (D-BPD) program: study protocol of a prospective translational	2019	BMC Pediatrics	Ofman 2019	Prospective study	Argentina	325	Machine Learning	Respiratory

multicenter collaborative study to investigate determinants of chronic lung disease in very low birth weight infants.								
A maChine and deep Learning Approach to predict pulmoNary hyperteNsIon in newbornS with congenital diaphragmatic Hernia (CLANNISH): Protocol for a retrospective study	2021	PLoS One	Amodeo 2021	Retrospective study	Italy	56	Machine & Deep Learning	Respiratory
Developing practical clinical tools for predicting neonatal mortality at a neonatal intensive care unit in Tanzania	2021	BMC pediatrics	Kovacs 2021	Prospective study	Tanzania	165	Machine Learning	Mortality
Using machine learning analysis to assist in differentiating between necrotizing enterocolitis and spontaneous intestinal perforation: A novel predictive analytic tool	2021	J Pediatr Surg	Lure 2021	Retrospective study	USA	40	Deep Learning	Other

Appendix 2 - Clinical Outcomes

Table 2 - Clinical Outcomes Description

Category	Outcome	Description	Study
Growth & Development	Cognitive Development	The mental processes and abilities that involve the acquisition, processing, and utilisation of information from the environment. It includes various aspects of thinking, reasoning, memory, attention, problem-solving, decision-making, language comprehension, and other higher-order mental activities (68).	(37) (40) (41) (42)
	Growth	The physiological process of physical development and maturation in preterm infants. It involves the increase in size, weight, and overall development of various body structures and	(29) (48)

		functions as the infants progress through their neonatal period (69).	
	Motor Outcomes	Motor outcomes refer to the assessment and evaluation of an individual's motor skills and abilities, which encompass movements, coordination, muscle strength, and overall physical function (70).	(49)
Length of Stay	Length of Stay	These studies focus on investigating and predicting the length of hospital stays for NICU patients.	(27) (45)
Mortality	Mortality	Mortality studies focus on understanding and predicting the likelihood of death among extremely and very preterm infants.	(30) (44) (47) (50)
Ophthalmological	Retinopathy of Prematurity (ROP)	A potentially sight-threatening eye condition that affects two thirds of premature infants with very low birth weight (<1250 g at birth). It is characterised by abnormal blood vessel growth in the retina, which can lead to scarring and retinal detachment, potentially causing vision impairment or blindness if left untreated (71).	(38) (39) (43)
Other	Extubation Readiness	Extubation readiness refers to the state of a patient being prepared and deemed suitable for the removal of an endotracheal tube that has been providing mechanical ventilation. Medical professionals assess various physiological and clinical parameters to determine if an infant is ready for extubation, aiming to ensure a successful transition to spontaneous breathing and avoid potential complications (72).	(32)
	Multiple (Mortality, Periventricular Leukomalacia (PVL), Intraventricular Haemorrhage (IVH), BPD,	Intraventricular Haemorrhage (IVH) is a medical condition characterised by bleeding that occurs within the brain's ventricular system, particularly in the ventricles, which are fluid-filled cavities in the brain. IVH is most commonly seen in premature infants, especially those born very prematurely, and is considered a serious medical complication in neonatal care. (73) Periventricular Leukomalacia (PVL) is a neurological disorder	(28)

	Necrotizing Enterocolitis (NEC)	that primarily affects premature infants, especially those born very prematurely. It is characterised by damage to the white matter of the brain, particularly in the periventricular region, which is the area surrounding the brain's fluid-filled cavities called ventricles. (74)	
	Necrotizing Enterocolitis (NEC) / Spontaneous Intestinal Perforation (SIP)	NEC and SIP are two distinct but serious gastrointestinal conditions that primarily affect premature infants. NEC involves the inflammation and potential death of intestinal tissue (75), while SIP refers to the spontaneous rupture or perforation of the intestinal wall (76). Both conditions pose significant health risks to neonates and require prompt medical attention.	(46)
	Sepsis	Neonatal sepsis is a serious medical condition characterised by a systemic infection that occurs within the first 28 days of life in newborn infants. It is caused by the presence of bacteria, viruses, fungi, or other pathogens in the bloodstream or other body tissues. Neonatal sepsis can lead to a range of complications and health issues, including organ dysfunction, shock, and even mortality if not promptly diagnosed and treated (77).	(36)
Respiratory	BPD	A chronic lung disorder that primarily affects premature infants, especially those born very preterm. It is characterised by inflammation, injury, and abnormal growth of the developing lung tissue, particularly the alveoli (air sacs) and bronchioles (small airways) (78). BPD often occurs as a result of mechanical ventilation and oxygen therapy used to support the immature lungs of premature infants.	(35) (31) (33)
	Pulmonary Hypertension	The elevated blood pressure within the arteries of the lungs in newborns. This condition can occur in premature infants and those with certain medical conditions, where the blood vessels in the lungs become narrow and constricted, leading to increased pressure in the pulmonary arteries. PH can strain the right side of the heart as it works harder to pump blood through the lungs, potentially compromising oxygenation, and overall cardiac	(34)

		function (79).	

Appendix 3 - Applications & Concerns

Table 3 - Opportunity Themes and Findings

Study	Application of AI in the NICU	Advancements in Medical Imaging	Data-Driven Insights and Predictive Models	Improving Understanding and Risk Stratification	Personalised Neonatal Care and Intervention
(29)	X		X		X
(35)	X				
(37)	X				
(45)	X		X		
(44)	X		X		
(38)	X	X	X		
(40)	X				
(43)	X	X			
(27)	X		X	X	X
(39)	X	X			
(49)	X	X			X
(43)	X	X			
(28)	X		X	X	
(41)	X			X	
(36)	X				
(33)	X				
(47)	X		X		

(48)	X		X		
(32)	X		X		
(50)	X				
(31)	X		X	X	
(34)	X		X	X	X
(30)	X				
(46)	X	X	X		X

Table 5 - Challenge Themes and Findings

Study	**Data Quality and Quantity Challenges**	**Clinical Interpretability and Usability**	**Model Generalisation and Validation**	**Clinical and Diagnostic Variability**	**Ethical and Regulatory Challenges**
(29)		X		X	
(35)	X		X		
(37)	X				
(45)					
(44)	X			X	
(38)		X	X		X
(40)		X	X	X	
(43)		X			
(27)		X			
(39)	X		X	X	
(49)	X		X	X	
(43)	X				
(28)					
(41)	X				

(36)		X			
(33)	X		X	X	
(47)					X
(48)	X				
(32)		X		X	
(50)	X		X	X	
(31)			X	X	X
(34)			X	X	
(30)		X			X
(46)					

References

1. WHO. WHO recommendations for care of the preterm or low-birth-weight infant: World Health Organization; 2022.
2. Ohuma EO, Moller A-B, Bradley E, Chakwera S, Hussain-Alkhateeb L, Lewin A, et al. National, regional, and global estimates of preterm birth in 2020, with trends from 2010: a systematic analysis. The Lancet. 2023;402(10409):1261-71.
3. Cheong JL, Olsen JE, Huang L, Dalziel KM, Boland RA, Burnett AC, et al. Changing consumption of resources for respiratory support and short-term outcomes in four consecutive geographical cohorts of infants born extremely preterm over 25 years since the early 1990s. BMJ open. 2020;10(9):e037507.
4. Gruenberg DA, Shelton W, Rose SL, Rutter AE, Socaris S, McGee G. Factors influencing length of stay in the intensive care unit. American journal of critical care. 2006;15(5):502-9.
5. Araki S, Saito T, Ichikawa S, Saito K, Takada T, Noguchi S, et al. Family-Centered Care in Neonatal Intensive Care Units: Combining Intensive Care and Family Support. Journal of UOEH. 2017;39(3):235-40.
6. Santos J, Pearce SE, Stroustrup A. Impact of hospital-based environmental exposures on neurodevelopmental outcomes of preterm infants. Current opinion in pediatrics. 2015;27(2):254.
7. Avila-Alvarez A, Solar Boga A, Bermudez-Hormigo C, Fuentes Carballal J. Extrauterine growth restriction among neonates with a birthweight less than 1,500 grams. Anales de Pediatria. 2018;89(6):325-32.
8. Bahado-Singh RO, Sonek J, McKenna D, Cool D, Aydas B, Turkoglu O, et al. Artificial intelligence and amniotic fluid multiomics: prediction of perinatal outcome in asymptomatic women with short cervix. Ultrasound Obstet Gynecol. 2019;54(1):110-8.
9. Hsu J-F, Chang Y-F, Cheng H-J, Yang C, Lin C-Y, Chu S-M, et al. Machine Learning Approaches to Predict In-Hospital Mortality among Neonates with Clinically Suspected Sepsis in the Neonatal Intensive Care Unit. J. 2021;11(8).
10. Irles C, Gonzalez-Perez G, Carrera Muinos S, Michel Macias C, Sanchez Gomez C, Martinez-Zepeda A, et al. Estimation of Neonatal Intestinal Perforation Associated with Necrotizing Enterocolitis by Machine Learning Reveals New Key Factors. International journal of environmental research and public health. 2018;15(11).
11. Juraev F, El-Sappagh S, Abdukhamidov E, Ali F, Abuhmed T. Multilayer dynamic ensemble model for intensive care unit mortality prediction of neonate patients. J Biomed Inform. 2022;135:104216.
12. McAdams RM, Kaur R, Sun Y, Bindra H, Cho SJ, Singh H. Predicting clinical outcomes using artificial intelligence and machine learning in neonatal intensive care units: a systematic review. J Perinatol. 2022;42(12):1561-75.
13. Klock P, Erdmann AL. Caring for newborns in a NICU: dealing with the fragility of living/surviving in the light of complexity. Revista da Escola de Enfermagem da USP. 2012;46:45-51.
14. Samra HA, McGrath JM, Rollins W. Patient Safety in the NICU: A Comprehensive Review.

The Journal of perinatal & neonatal nursing. 2011;25(2):123-32.
15. Goodfellow I, Bengio Y, Courville A. Deep learning: MIT press; 2016.
16. Van Zuylen H. Difference between artificial intelligence and traditional methods. Artificial Intelligence Applications to Critical Transportation Issues. 2012:3.
17. Szolovits P, Patil RS, Schwartz WB. Artificial intelligence in medical diagnosis. Annals of internal medicine. 1988;108(1):80-7.
18. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nature biomedical engineering. 2018;2(10):719-31.
19. Han JH, Yoon SJ, Lee HS, Park G, Lim J, Shin JE, et al. Application of Machine Learning Approaches to Predict Postnatal Growth Failure in Very Low Birth Weight Infants. Yonsei Medical Journal. 2022;63(7):640-7.
20. Panesar A. Ethics of Intelligence. In: Panesar A, editor. Machine Learning and AI for Healthcare : Big Data for Improved Health Outcomes. Berkeley, CA: Apress; 2019. p. 207-54.
21. Tricco AC, Langlois E, Straus SE, Organization WH. Rapid reviews to strengthen health policy and systems: a practical guide: World Health Organization; 2017.
22. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
23. Veritas Health Innovation. Covidence 2019 [Available from: www.covidence.org.
24. Duval D, Pearce-Smith N, Palmer JC, Sarfo-Annin JK, Rudd P, Clark R. Critical appraisal in rapid systematic reviews of COVID-19 studies: implementation of the Quality Criteria Checklist (QCC). Systematic Reviews. 2023;12(1):1-4.
25. Microsoft. Microsoft Excel. Redmond, WA2023.
26. Development BCfS. Reimagining Global Health through Artificial Intelligence: The Roadmap to AI Maturity. Broadband Commission for Sustainable Development; 2020.
27. Braun D, Braun E, Chiu V, Burgos AE, Gupta M, Volodarskiy M, Getahun D. Trends in Neonatal Intensive Care Unit Utilization in a Large Integrated Health Care System. JAMA netw. 2020;3(6):e205239.
28. Hamilton EF, Dyachenko A, Ciampi A, Maurel K, Warrick PA, Garite TJ. Estimating risk of severe neonatal morbidity in preterm births under 32 weeks of gestation. J Matern Fetal Neonatal Med. 2020;33(1):73-80.
29. Iyer KK, Leitner U, Giordano V, Roberts JA, Vanhatalo S, Klebermass-Schrehof K, Stevenson NJ. Bedside tracking of functional autonomic age in preterm infants. Pediatr Res. 2022;14:14.
30. Kovacs D, Msanga DR, Mshana SE, Bilal M, Oravcova K, Matthews L. Developing practical clinical tools for predicting neonatal mortality at a neonatal intensive care unit in Tanzania. BMC pediatrics. 2021;21(1):1-537.
31. Ofman G, Caballero MT, Alvarez Paggi D, Marzec J, Nowogrodzki F, Cho H-Y, et al. The discovery BPD (D-BPD) program: study protocol of a prospective translational multicenter collaborative study to investigate determinants of chronic lung disease in very low birth weight infants. BMC Pediatrics. 2019;19(1):N.PAG-N.PAG.
32. Shalish W, Kanbar LJ, Rao S, Robles-Rubio CA, Kovacs L, Chawla S, et al. Prediction of Extubation readiness in extremely preterm infants by the automated analysis of cardiorespiratory behavior: study protocol. BMC Pediatrics. 2017;17:1-15.

33. Verder H, Heiring C, Ramanathan R, Scoutaris N, Verder P, Jessen TE, et al. Bronchopulmonary dysplasia predicted at birth by artificial intelligence. Acta paediatrica (Oslo, Norway : 1992). 2021;110(2):503-9.
34. Amodeo I, De Nunzio G, Raffaeli G, Borzani I, Griggio A, Conte L, et al. A maChine and deep Learning Approach to predict pulmoNary hyperteNsIon in newbornS with congenital diaphragmatic Hernia (CLANNISH): Protocol for a retrospective study. PLoS ONE. 2021;16(11):e0259724-e.
35. He W, Zhang L, Feng R, Fang WH, Cao Y, Sun SQ, et al. Risk factors and machine learning prediction models for bronchopulmonary dysplasia severity in the Chinese population. World J Pediatr. 2022.
36. Kausch SL, Brandberg JG, Qiu JJ, Panda A, Binai A, Isler J, et al. Cardiorespiratory signature of neonatal sepsis: Development and validation of prediction models in 3 NICUs. medRxiv. 2022((Kausch, Fairchild, Sullivan) Department of Pediatrics, University of Virginia, School of Medicine, Charlottesville, VA, United States(Brandberg, Qiu, Moorman, Lake) Department of Medicine, University of Virginia, School of Medicine, Charlottesville, VA).
37. Ali R, Li H, Dillman JR, Altaye M, Wang H, Parikh NA, He L. A self-training deep neural network for early prediction of cognitive deficits in very preterm infants using brain functional connectome data. Pediatr Radiol. 2022;52(11):2227-40.
38. Chen JS, Coyner AS, Ostmo S, Sonmez K, Bajimaya S, Pradhan E, et al. Deep Learning for the Diagnosis of Stage in Retinopathy of Prematurity: Accuracy and Generalizability across Populations and Cameras. Ophthalmology Retina. 2021;5(10):1027-35.
39. Choi RY, Brown JM, Kalpathy-Cramer J, Chan RVP, Ostmo S, Chiang MF, Campbell JP. Variability in Plus Disease Identified Using a Deep Learning-Based Retinopathy of Prematurity Severity Scale. Ophthalmol Retina. 2020;4(10):1016-21.
40. Gschwandtner L, Hartmann M, Oberdorfer L, Furbass F, Klebermas-Schrehof K, Werther T, et al. Deep learning for estimation of functional brain maturation from EEG of premature neonates. Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference. 2020;2020(101763872):104-7.
41. He L, Li H, Holland SK, Yuan W, Altaye M, Parikh NA. Early prediction of cognitive deficits in very preterm infants using functional connectome data in an artificial neural network framework. Neuroimage (Amst). 2018;18:290-7.
42. He L, Li H, Wang J, Chen M, Gozdas E, Dillman JR, Parikh NA. A multi-task, multi-stage deep transfer learning model for early prediction of neurodevelopment in very preterm infants. Sci. 2020;10(1):15072.
43. Huang CY, Kuo RJ, Li CH, Ting DS, Kang EY, Lai CC, et al. Prediction of visual outcomes by an artificial neural network following intravitreal injection and laser therapy for retinopathy of prematurity. Br J Ophthalmol. 2020;104(9):1277-82.
44. Lee J, Cai J, Li F, Vesoulis ZA. Predicting mortality risk for preterm infants using random forest. Sci. 2021;11(1):7308.
45. Lin WT, Wu TY, Chen YJ, Chang YS, Lin CH, Lin YJ. Predicting in-hospital length of stay for very-low-birth-weight preterm infants using machine learning techniques. J Formos Med Assoc. 2022;121(6):1141-8.
46. Lure AC, Du X, Black EW, Irons R, Lemas DJ, Taylor JA, et al. Using machine learning

analysis to assist in differentiating between necrotizing enterocolitis and spontaneous intestinal perforation: A novel predictive analytic tool. J Pediatr Surg. 2021;56(10):1703-10.
47. Patel AK, Trujillo-Rivera E, Morizono H, Pollack MM. The criticality Index-mortality: A dynamic machine learning prediction algorithm for mortality prediction in children cared for in an ICU. Frontiers in Pediatrics. 2022;10((Patel, Pollack) Department of Pediatrics, Division of Critical Care Medicine, Children's National Health System, George Washington University School of Medicine and Health Sciences, Washington, DC, United States(Trujillo-Rivera) Department of Bio-Informa):1023539.
48. Ruixiang L, Mingrong Y, Li C, Rongxiu Z. Early Physical Linear Growth of Small-for-Gestational-Age Infants Based on Computer Analysis Method. Journal of Healthcare Engineering. 2021;2021((Ruixiang) College of Clinical Medicine, Tianjin Medical University, Tianjin 300070, China(Mingrong, Li) Pediatric Department, Affiliated Hospital of Weifang Medical University, Shandong, Weifang 261031, China(Rongxiu) Pediatric Department, Tianjin Medica):7227928.
49. Saha S, Pagnozzi A, Bourgeat P, George JM, Bradford D, Colditz PB, et al. Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model. Neuroimage. 2020;215(cpp, 9215515):116807.
50. Sheikhtaheri A, Zarkesh MR, Moradi R, Kermani F. Prediction of neonatal deaths in NICUs: development and validation of machine learning models. BMC Medical Informatics & Decision Making. 2021;21(1):1-14.
51. Association AM. Artificial Intelligence in Healthcare 2023 [Available from: https://www.ama.com.au/articles/artificial-intelligence-healthcare.
52. Organization WH. Harnessing Artificial Intelligence for Health 2023 [Available from: https://www.who.int/teams/digital-health-and-innovation/harnessing-artificial-intelligence-for-health.
53. University M. Australian Alliance for Artificial Intelligence in Healthcare: National Agenda Roadmap. Macquarie University; 2023.
54. Chen M, Decary M, editors. Artificial intelligence in healthcare: An essential guide for health leaders. Healthcare management forum; 2020: SAGE Publications Sage CA: Los Angeles, CA.
55. Henry C, Saffaran S, Meeus M, Bates D, Van Laere D, Boylan G, et al., editors. Application and potential of artificial intelligence in neonatal medicine. Seminars in Fetal and Neonatal Medicine; 2022: Elsevier.
56. Hintz SR, Bann CM, Ambalavanan N, Cotten CM, Das A, Higgins RD, et al. Predicting time to hospital discharge for extremely preterm infants. Pediatrics. 2010;125(1):e146-e54.
57. Snowdon A. Digital health: a framework for healthcare transformation. HIMSS, Chicago. 2020.
58. White RD, Smith JA, Shepley MM. Recommended standards for newborn ICU design, eighth edition. J Perinatol. 2013;33(1):S2-S16.
59. Kim JH, Sampath V, Canvasser J. Challenges in diagnosing necrotizing enterocolitis. Pediatr Res. 2020;88(Suppl 1):16-20.
60. Leon C, Carrault G, Pladys P, Beuchee A. Early detection of late onset sepsis in premature infants using visibility graph analysis of heart rate variability. IEEE journal of biomedical and health informatics. 2020;PP.

61. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6(2):94-8.
62. Hartley C. Toward personalized medicine for pharmacological interventions in neonates using vital signs. Paediatric and Neonatal Pain. 2021;3(4):147-55.
63. Kwok TnC, Henry C, Saffaran S, Meeus M, Bates D, Van Laere D, et al. Application and potential of artificial intelligence in neonatal medicine. Seminars in Fetal and Neonatal Medicine. 2022;27(5):101346.
64. Leon C, Carrault G, Pladys P, Beuchee A. Early Detection of Late Onset Sepsis in Premature Infants Using Visibility Graph Analysis of Heart Rate Variability. IEEE journal of biomedical and health informatics. 2021;25(4):1006-17.
65. Lu Z-x, Qian P, Bi D, Ye Z-w, He X, Zhao Y-h, et al. Application of AI and IoT in clinical medicine: summary and challenges. Current Medical Science. 2021;41:1134-50.
66. Shaheen MY. Applications of Artificial Intelligence (AI) in healthcare: A review. ScienceOpen Preprints. 2021.
67. Rochefort CM, Rathwell BA, Clarke SP. Rationing of nursing care interventions and its association with nurse-reported outcomes in the neonatal intensive care unit: a cross-sectional survey. BMC nursing. 2016;15:1-8.
68. Peyton C, Einspieler C. General movements: A behavioral biomarker of later motor and cognitive dysfunction in NICU graduates. Pediatric annals. 2018;47(4):e159-e64.
69. WHO. WHO child growth standards: growth velocity based on weight, length and head circumference: methods and development: World Health Organization; 2009.
70. Lenke MC. Motor outcomes in premature infants. Newborn and infant nursing Reviews. 2003;3(3):104-9.
71. Kim SJ, Port AD, Swan R, Campbell JP, Chan RVP, Chiang MF. Retinopathy of prematurity: a review of risk factors and their clinical significance. Survey of ophthalmology. 2018;63(5):618-37.
72. Smith J, Pieper CH, Maree D, Gie RP. Compliance of the respiratory system as a predictor for successful extubation in very-low-birth-weight infants recovering from respiratory distress syndrome. SAMJ South African medical journal. 1999;89(10):1097-102.
73. Perlman JM, Volpe JJ. Prevention of neonatal intraventricular hemorrhage. Clinical neuropharmacology. 1987;10(2):126-42.
74. Volpe JJ. Neurobiology of periventricular leukomalacia in the premature infant. Pediatr Res. 2001;50(5):553-62.
75. Caplan MS. Neonatal necrotizing enterocolitis. Introduction. Seminars in perinatology. 2008;32(2):69-.
76. Neu J, Modi N, Caplan M. Necrotizing enterocolitis comes in different forms: Historical perspectives and defining the disease. Seminars in fetal & neonatal medicine. 2018;23(6):370-3.
77. Pace E, Yanowitz T. Infections in the NICU: Neonatal sepsis. Seminars in Pediatric Surgery. 2022;31(4):151200.
78. Thébaud B, Goss KN, Laughon M, Whitsett JA, Abman SH, Steinhorn RH, et al. Bronchopulmonary dysplasia. Nature reviews Disease primers. 2019;5(1):78.
79. Abman SH. Pulmonary hypertension: the hidden danger for newborns. Neonatology. 2021;118(2):211-7.