sophIA: Platform to assist healthcare professionals in monitoring neonatal health

Flávio Leandro de Morais[1], Stephany Paula da Silva Canejo[2], Maria Eduarda Ferro de Mello[1], Katia Maria Mendes[3], Waldemar Brandão Neto[2,3], Patricia Takako Endo[1]

[1]Programa de Pós-Graduação em Engenharia da Computação (PPGEC), Universidade de Pernambuco, Recife, Pernambuco, Brasil

[2]Programa Associado de Pós-Graduação em Enfermagem (PAPGenf), Universidade de Pernambuco, Recife, Pernambuco, Brasil

[3]Centro Universitário Integrado de Saúde Amaury de Medeiros (CISAM) Universidade de Pernambuco, Recife, Pernambuco, Brasil

flavio.leandromorais@upe.br, stephany.canejo@upe.br, mefm@ecomp.poli.br, katia.mendes@upe.br, waldemar.neto@upe.br, patricia.endo@upe.br

Abstract. *The neonatal period is marked by high vulnerability and significant morbidity and mortality, especially in the first days of life, requiring innovative strategies to improve the quality of care. This study presents sophIA, a platform under development based on Artificial Intelligence (AI) to support clinical decision-making in the Neonatal Intensive Care Unit. Its architecture integrates real clinical data, rigorous preparation, and predictive modeling using traditional and hybrid techniques, in addition to explainability mechanisms. The interface includes an interactive dashboard for indicator visualization and a support chatbot. The sophIA estimates risks for adverse neonatal outcomes and generates interpretable clinical alerts. Preliminary results indicate technical feasibility and methodological consistency, demonstrating improved performance and reliability. It is expected to contribute to more timely interventions, reduction of preventable complications, and strengthening of neonatal health management.*

1. Introduction

The neonatal period, defined as the first 28 days after birth, is characterized by high vulnerability and a high risk of health complications in the newborn. Several complications may arise due to factors occurring during fetal development, during childbirth, or related to clinical conditions such as prematurity and low birth weight [Grillo et al. 2022]. This vulnerability may lead to the development of diseases or clinical deterioration in the first seven days of life (early neonatal period), such as Transient Tachypnea of the Newborn (TTN), Respiratory Distress Syndrome (RDS), Persistent Pulmonary Hypertension of the Newborn (PPHN), sepsis, jaundice, and others. In the late neonatal period, which extends from the 7th to the 28th day of life, the neonate remains susceptible to conditions such as necrotizing enterocolitis, bronchopulmonary dysplasia, late-onset sepsis, and other complications [Zinjani 2023].

According to the United Nations Children's Fund (UNICEF), in 2023, there were approximately 2.3 million neonatal deaths worldwide, of which about 75% occurred in the early neonatal period and 25% in the late neonatal period [UNICEF 2024]. In 2024, the neonatal mortality rate in Brazil was approximately 8.42 deaths per thousand live births, according to data from the Sistema de Informação sobre Mortalidade (SIM) and the Sistema de Informação sobre Nascidos Vivos (SINASC). In the same year, the state of Pernambuco presented a rate higher than the national average, with approximately 9.01 deaths per thousand

live births [BRASIL-SIM 2024, BRASIL-SINASC 2024]. These rates indicate that, despite the advances observed in recent decades, neonatal mortality remains a relevant public health problem in the Brazilian context, especially in states historically marked by socioeconomic vulnerabilities and limitations in timely and qualified access to maternal and child health services, such as Pernambuco [Areco et al. 2025].

In these circumstances, investments and incorporation of preventive technologies have the potential to improve care during pregnancy and in the neonatal period by supporting healthcare professionals in clinical management, reducing risks, and promoting better maternal and neonatal outcomes. The scientific literature has highlighted Artificial Intelligence (AI) models as promising tools to support clinical practice and expand the diagnostic and prognostic effectiveness for various diseases [Kwok et al. 2022]. Such models have been used to predict complications during and after pregnancy, which include conditions such as preeclampsia, gestational diabetes, sepsis, fetal death, prematurity, neonatal death, among other health complications [De Morais et al. 2025]. In this context, this study presents sophIA, an AI-based decision-making platform under development to provide support to professionals through alerts generated from model predictions, in order to enhance the quality of care and promote more timely and accurate interventions in neonatal care.

2. sophIA

SophIA is a platform under development that emerged with the objective of generating alerts to assist healthcare professionals in improving the quality of care provided to neonates requiring admission to the Neonatal Intensive Care Unit (NICU). The initiative is aligned with the Sustainable Development Goals (SDGs) of the United Nations (UN), especially SDG 3, which aims to ensure healthy lives and promote well-being for all, including the reduction of neonatal mortality and the strengthening of health systems [United Nations 2015].

The name of the platform, sophIA, originated from the Greek word *sophía*, meaning "wisdom," combined with the acronym AI, referring to Artificial Intelligence. The platform is being developed to integrate existing systems of the Sistema Único de Saúde (SUS). Figure 1 presents the development modules of the platform.

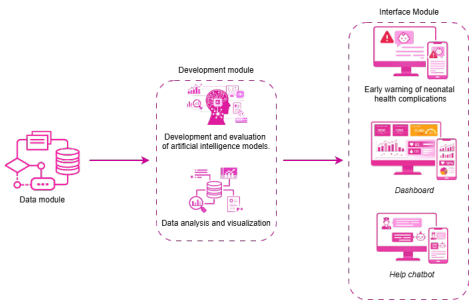

Figura 1. Modules and functionalities of the sophIA platform.

2.1. Data module

This module constitutes one of the fundamental layers of the sophIA architecture, being responsible for the ingestion, integration, and governance of neonatal clinical data used in the development and application of AI models aimed at predicting outcomes in neonatal health. It is composed of a real clinical database from the Centro Universitário Integrado de Saúde Amaury de Medeiros (CISAM), a reference hospital in the care of high-risk pregnancies, childbirth assistance, and neonatal care, including services provided in the NICU [de Pernambuco 2025].

The database includes information from the gestational period, the childbirth process, birth, and the daily follow-up of the neonate during hospitalization. Among the in-

cluded attributes are demographic data, maternal conditions, neonatal clinical parameters, laboratory results, and records of assessments performed by the multidisciplinary healthcare team. The data used for the development of the sophIA platform models were submitted to and approved by the Research Ethics Committee (REC) of CISAM, under CAAE No. 78547524.8.0000.5191, opinion No. 6,891,723, dated June 17, 2024, ensuring ethical and regulatory compliance.

This module also includes a strategic stage aimed at defining the clinical outcomes to be predicted. This definition is conducted through meetings with professionals from the NICU of CISAM, in collaboration with faculty members and researchers from the Faculdade de Enfermagem Nossa Senhora das Graças (FENSG). In this process, the clinical relevance of the selected outcomes and the availability of the attributes required to adequately support the training and evaluation of the AI models are assessed.

2.2. Development module

2.2.1. Data preparation

Rigorous procedures are adopted, initially involving the exploratory data analysis stage, in which the type of attributes, degree of completeness, quality of records, and statistical distribution of attributes, among other aspects, are evaluated. Subsequently, attribute selection is performed using different approaches, including clinical expertise, evidence from the scientific literature, and mathematical and computational methods. Finally, preprocessing is applied, including data cleaning, standardization, handling of missing data, class balancing, and, when necessary, appropriate splitting of training and testing sets. The careful execution of these phases, aligned with the specific characteristics and needs of the database, together with validation by multidisciplinary teams, is essential to ensure the methodological validity, reliability, generalizability, and responsibility of the developed models.

2.2.2. Development and evaluation of AI models

Different AI models are explored and compared, including traditional models and more robust approaches, with systematic hyperparameter tuning performed through cross-validation strategies. The models are evaluated across multiple experimental scenarios, considering different combinations of attributes, class balancing strategies, and training configurations, with the objective of identifying solutions that demonstrate better predictive performance, stability, and generalization capacity.

In addition to individual models, hybrid models are developed based on the best approaches identified in the experiments. Different hybridization strategies may be employed, including voting techniques, stacking, or decision fusion, seeking to exploit the complementarity between models and improve predictions. Performance is evaluated using metrics appropriate to the type of problem and data distribution. For classification tasks, metrics considered include, but are not limited to, sensitivity, specificity, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC), prioritizing metrics sensitive to the minority class in imbalanced scenarios. For regression problems, metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2) are used. To ensure a fair and statistically grounded comparison among models, appropriate statistical tests are applied according to the characteristics of the data and the experiment, allowing the verification of significant differences in performance between models.

The Explainable Artificial Intelligence (XAI) methodology will be used to develop inherently interpretable models that systematically present the model's decision-making processes, clarifying how predictions are generated and increasing transparency and confidence in the identification of these health conditions. To this end, interpretability techniques are employed to understand the overall behavior of the model, identifying the most relevant

attributes for predicting clinical outcomes and analyzing the direction and magnitude of their influence. The explainability stage contributes to the clinical validation of the results, allowing healthcare professionals to understand the factors associated with the predictions and to use the system as a decision-support tool rather than a substitute for clinical judgment.

2.2.3. Data Analytics and Visualization

Descriptive and exploratory analyses are conducted to characterize the maternal-infant population profile. Epidemiological investigations are performed to understand the frequency and distribution of neonatal conditions, the association between maternal risk factors and adverse outcomes and temporal trends stratified by clinical profiles. These analyses enable the identification of patterns relevant to clinical practice, providing a foundation for strategic decision-making in neonatal health.

The creation of diverse graphical representations is fundamental to this stage. It encompasses univariate, bivariate and multivariate visualizations, as well as correlation analyses and the mapping of clinical and care indicators. These visual tools facilitate data interpretation for healthcare professionals and managers, while also guiding feature selection and the definition of modeling strategies. Additionally, inferential statistical analyses are applied where appropriate.

2.3. User Interface Module

2.3.1. Operational functions

The risk prediction module constitutes the primary architectural component of the sophIA platform. Through this interface, clinicians can generate predictive profiles for various adverse health conditions, such as neonatal sepsis, bronchopulmonary dysplasia, PPHN, RDS and neonatal mortality. Access to sophIA requires user registration and authentication to maintain data integrity, security and a clear audit trail of all actions. The platform is designed to integrate with existing hospital systems and Electronic Medical Records (EMR), establishing interoperability for the automated import of necessary clinical variables, presenting predictive results directly to the professional, eliminating the burden of manual information entry.

In settings lacking integrated information systems, users can input data via a structured form that captures maternal and neonatal variables. Based on these variables, carefully selected through scientific evidence and validated by maternal-infant specialists, the models estimate the probability of adverse neonatal health outcomes. SophIA features an interactive dashboard designed for the real-time visualization and monitoring of neonatal indicators. Within this environment, managers, researchers and healthcare practitioners can access comprehensive data regarding maternal health, neonatal status and regional demographics. The panel provides interactive charts, geospatial mapping and performance metrics, facilitating a rapid and strategic analysis of the clinical landscape at the platform's implementation site.

SophIA also features an interactive chatbot powered by generative AI, designed to provide continuous support to healthcare professionals throughout their use of the platform. Functioning as a virtual assistant, this tool guides users through system navigation, clarifies inquiries and assists in the interpretation of predictive model outputs. Furthermore, it offers simplified explanations of the factors influencing specific risk estimates, thereby enhancing transparency and the overall understanding of results.

2.4. Relevance and Impact Scientific and Technological

The technological potential of sophIA resides in its integrated digital ecosystem, which combines accessible epidemiological visualization, predictive tools for adverse neonatal outcomes, and an interactive chatbot tailored to healthcare professionals' needs. As a decision-

support tool, it facilitates timely interventions within NICU settings, directly contributing to the reduction of neonatal morbidity and mortality. By aligning with daily clinical progression and laboratory results, the platform serves as a vital adjunct to clinical decision-making, strengthening multidisciplinary critical thinking while mitigating subjectivity in care.

The scientific contribution of this work lies in the systematic evaluation and comparison of AI models for neonatal outcome prediction using real-world clinical data, including the investigation of hybrid approaches and the application of explainability techniques to support clinical interpretation. These elements provide a methodological contribution by advancing the understanding of how different modeling strategies perform in neonatal health scenarios and how their predictions can be effectively integrated into clinical decision-making. By providing strategic data to clinicians, researchers and policymakers, the system promotes evidence-based public health strategies and acts as an instrument for health equity, addressing regional disparities in socioeconomically vulnerable settings.

While initially designed for hospitals within the Recife metropolitan area, sophIA holds significant potential for scalability. Its ability to bridge the gap between academic research and public reference hospitals offers a replicable framework for other regions across Northeastern Brazil, ensuring that technological innovation translates into improved maternal-infant care.

The scientific contribution of this work lies in the systematic evaluation and comparison of AI models for neonatal outcome prediction using real-world clinical data, including the investigation of hybrid approaches and the application of explainability techniques to support clinical interpretation. These elements provide a methodological contribution by advancing the understanding of how different modeling strategies perform in neonatal health scenarios and how their predictions can be effectively integrated into clinical decision-making.

2.5. Current Results and Ongoing Work

This project seeks to design, evaluate, validate and implement an integrated digital platform composed of predictive models grounded in diverse AI approaches. The overarching objective is to support healthcare professionals in clinical decision-making related to neonatal health from the earliest hours of life. From a technological perspective, the goal is to deliver a robust, interoperable and scalable solution capable of integrating clinical data, generating individualized risk estimates and presenting results in an interpretable and accessible manner.

Although the sophIA platform is still under development, several results have already been achieved, supporting its technical and scientific feasibility. A comprehensive literature study was conducted to identify existing gaps and current advances in the application of AI in neonatal health, providing a structured foundation for the platform [De Morais et al. 2025]. An experimental study evaluating AI models for neonatal sepsis prediction has been carried out, including multiple experiments and analyses, as well as investigations into model learning behavior and explainability aspects [Morais et al. 2026]. A web-based platform website [1] is already online, designed to host information about the system, its functionalities, and ongoing developments, and will, in future versions, provide predictive models for neonatal outcome prediction. The project has also generated multiple scientific outputs, including abstracts published in local and regional conferences addressing predictions of neonatal sepsis, neonatal mortality, and pulmonary diseases. Ongoing efforts include the submission of manuscripts to international journals and the continued development and evaluation of predictive models for pulmonary hypertension, bronchopulmonary dysplasia, and neonatal mortality, as well as the development of the platform's user interface

[1] Available at: https://sophia.dotlabbrazil.com.br/

for integration with CISAM.

The sophIA platform is expected to integrate multiple predictive models, including hybrid approaches, combined with explainability mechanisms and generative AI support, functioning as a clinical decision-support system that enhances early identification of adverse neonatal outcomes, risk stratification and case prioritization without replacing professional autonomy. Additionally, sophIA' visualization dashboard will support the monitoring of epidemiological and care indicators, enabling comparative analyses and identification of regional vulnerabilities thereby strengthening governance within the maternal and child health network and supporting public health policies. From a scientific perspective, the sophIA project contributes to advancing AI in neonatology, fostering the translation of research into practical applications within the public health system, with the potential to reduce preventable complications, optimize care processes and improve neonatal outcomes while bridging the gap between research, clinical practice and health system strengthening.

Referências

Areco, K. N., Bandiera-Paiva, P., Balda, R. C., Costa-Nobre, D. T., Marinonio, A. S. S., Sanudo, A., Kawakami, M. D., Miyoshi, M. H., e Oliveira, C. N. V., Freitas, R. M., et al. (2025). Development and validation of an epidemiological risk score for neonatal death in a middle-income country. *Frontiers in Public Health*, 13:1675040.

BRASIL-SIM (2024). Sistema de informação sobre mortalidade. Available at: `http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sim/cnv/obt10uf.def`. Accessed: January 05, 2026. Data updated in December 2025.

BRASIL-SINASC (2024). Sistema de informações sobre nascidos vivos. Available at: `http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinasc/cnv/nvuf.def`. Accessed: January 05, 2026. Data updated in December 2025.

De Morais, F. L., Da Silva, R. C. L., Silva, A. B., Simão, E. P., De Mello, M. E. F., da Silva Canejo, S. P., Mendes, K. M., Neto, W. B., Da Silva, J. R. F., da Silva, M. H. L. F., et al. (2025). On usage of artificial intelligence for predicting neonatal diseases, conditions and mortality: A bibliometric review. *IEEE Access*.

de Pernambuco, U. (2025). Cisam - centro universitário integrado de saúde amaury de medeiros. Disponível em: `https://www.upe.br/uh-cisam`. Acesso em: 03 de junho de 2026.

Grillo, M. A., Mariani, G., and Ferraris, J. R. (2022). Prematurity and low birth weight in neonates as a risk factor for obesity, hypertension, and chronic kidney disease in pediatric and adult age. *Frontiers in medicine*, 8:769734.

Kwok, T. C., Henry, C., Saffaran, S., Meeus, M., Bates, D., Van Laere, D., Boylan, G., Boardman, J. P., and Sharkey, D. (2022). Application and potential of artificial intelligence in neonatal medicine. In *Seminars in Fetal and Neonatal Medicine*, volume 27, page 101346. Elsevier.

Morais, F. L. d., Canejo, S. P. d. S., Mello, M. E. F. d., Silva, R. C. L. d., Barros, M. H. L. F. d. S., Rocha, E. d. S., Mendes, K. M., Neto, W. B., and Endo, P. T. (2026). On the usage of artificial intelligence for identifying main attributes and predicting neonatal sepsis. *Scientific Report*. Preprint, not peer reviewed.

UNICEF (2024). Levels and trends in child mortality 2024. `https://data.unicef.org/resources/levels-and-trends-in-child-mortality-2024/`. Accessed: November 13, 2025.

United Nations (2015). Transforming our world: the 2030 agenda for sustainable development. Resolution adopted by the General Assembly on 25 September 2015 (A/RES/70/1). Accessed: 2026-02-25.

Zinjani, S. (2023). Common medical conditions in the neonates. In *Clinical Anesthesia for the Newborn and the Neonate*, pages 49–70. Springer.