

Review

Exploring Computational Techniques in Preprocessing Neonatal Physiological Signals for Detecting Adverse Outcomes: Scoping Review

Jessica Rahman¹, BSc, PhD; Aida Brankovic², PhD; Mark Tracy³, PhD; Sankalp Khanna², PhD

¹Commonwealth Scientific and Industrial Research Organisation (CSIRO) Australian e-Health Research Centre, Australia, Sydney, Australia

²Commonwealth Scientific and Industrial Research Organisation (CSIRO) Australian e-Health Research Centre, Australia, Brisbane, Australia

³Neonatal Intensive Care Unit, Westmead, Sydney, Australia

Corresponding Author:

Jessica Rahman, BSc, PhD

Commonwealth Scientific and Industrial Research Organisation (CSIRO) Australian e-Health Research Centre, Australia

160 Hawkesbury Road

Sydney, 2145

Australia

Phone: 61 02 9325 3016

Email: jessica.rahman@csiro.au

Abstract

Background: Computational signal preprocessing is a prerequisite for developing data-driven predictive models for clinical decision support. Thus, identifying the best practices that adhere to clinical principles is critical to ensure transparency and reproducibility to drive clinical adoption. It further fosters reproducible, ethical, and reliable conduct of studies. This procedure is also crucial for setting up a software quality management system to ensure regulatory compliance in developing software as a medical device aimed at early preclinical detection of clinical deterioration.

Objective: This scoping review focuses on the neonatal intensive care unit setting and summarizes the state-of-the-art computational methods used for preprocessing neonatal clinical physiological signals; these signals are used for the development of machine learning models to predict the risk of adverse outcomes.

Methods: Five databases (PubMed, Web of Science, Scopus, IEEE, and ACM Digital Library) were searched using a combination of keywords and MeSH (Medical Subject Headings) terms. A total of 3585 papers from 2013 to January 2023 were identified based on the defined search terms and inclusion criteria. After removing duplicates, 2994 (83.51%) papers were screened by title and abstract, and 81 (0.03%) were selected for full-text review. Of these, 52 (64%) were eligible for inclusion in the detailed analysis.

Results: Of the 52 articles reviewed, 24 (46%) studies focused on diagnostic models, while the remainder ($n=28$, 54%) focused on prognostic models. The analysis conducted in these studies involved various physiological signals, with electrocardiograms being the most prevalent. Different programming languages were used, with MATLAB and Python being notable. The monitoring and capturing of physiological data used diverse systems, impacting data quality and introducing study heterogeneity. Outcomes of interest included sepsis, apnea, bradycardia, mortality, necrotizing enterocolitis, and hypoxic-ischemic encephalopathy, with some studies analyzing combinations of adverse outcomes. We found a partial or complete lack of transparency in reporting the setting and the methods used for signal preprocessing. This includes reporting methods to handle missing data, segment size for considered analysis, and details regarding the modification of the state-of-the-art methods for physiological signal processing to align with the clinical principles for neonates. Only 7 (13%) of the 52 reviewed studies reported all the recommended preprocessing steps, which could have impacts on the downstream analysis.

Conclusions: The review found heterogeneity in the techniques used and inconsistent reporting of parameters and procedures used for preprocessing neonatal physiological signals, which is necessary to confirm adherence to clinical and software quality management system practices, usefulness, and choice of best practices. Enhancing transparency in reporting and standardizing procedures will boost study interpretation and reproducibility and expedite clinical adoption, instilling confidence in the research findings and streamlining the translation of research outcomes into clinical practice, ultimately contributing to the advancement of neonatal care and patient outcomes.

(*Interact J Med Res 2024;13:e46946*) doi: [10.2196/46946](https://doi.org/10.2196/46946)

KEYWORDS

physiological signals; preterm; neonatal intensive care unit; morbidity; signal processing; signal analysis; adverse outcomes; predictive and diagnostic models

Introduction

Background

Premature infants are those born at <37 weeks gestational age, ranging from extreme preterm (23 weeks' gestation) to late preterm (37 weeks' gestation), and are defined as having very low birth weight of <1500 g. These extremely premature infants have a higher risk of death, and surviving infants are highly prone to physical, cognitive, and emotional impairment [1]. The patients usually have a long length of stay, ranging from <10 to >120 days [2], in the neonatal intensive care unit (NICU), where high-fidelity physiological changes are monitored to observe their health status and signs of deterioration. During this long length of stay, a large amount of data from infants are generated and not typically electronically aggregated for permanent storage [3]. With the advent of electronic health records, relevant patient information is easily available for advanced data analytics that can be used to improve health outcomes. The records contain demographics, etiology, pathology, medication, and physiology information. Physiological changes are regularly monitored in preterm infants, notably, electrocardiogram (ECG), oxygen saturation (SpO_2), heart rate (HR), respiratory rate, arterial blood pressure, electroencephalography (EEG), and temperature. Some advanced centers around the world have started linking the information derived from the electronic health records data with the continuously monitored physiological information for permanent storage, more frequently in lower resolution, which facilitates various data analytics [4-6]. Compared with intermittent assessment and review, continuous capturing and analysis of the physiological data from the standard bedside monitors allow for a better understanding of trends and have been shown to improve outcomes of infants in the NICU [5].

Clinical decision support systems (CDSSs) can integrate clinical and physiological information to provide automated support in patient care planning to facilitate the diagnostic process and therapy planning, generate critical alerts and reminders, and predict the risk of patient deterioration. CDSSs have the potential for a positive impact in improving clinical and economic measures in the health care system [7-9]. The technological advancement that allowed storing big data, as well as the advancement of artificial intelligence (AI), has given rise to machine learning (ML)- and AI-based CDSSs aiming to build data-driven models to predict adverse outcomes in premature infants ahead of clinical diagnosis time [10-12].

The steps of building the ML pipeline to predict adverse outcomes involve several intermediate computational steps using the physiological data, of which data preprocessing is the first indispensable step. Namely, in the NICU, physiological signals are collected using a diverse range of devices, which introduce a number of artifacts such as *environmental artifacts*

(eg, device connection failure, equipment noise, electrosurgical noise, and power line interferences); *experimental or human error* due to patient movement during data acquisition, incorrect or poor contact of the electrodes, and other contact noise; and *artifacts* due to muscle contraction, cardiac signals, and blinking [13,14]. These noises distort signals and may adversely affect model generalization capability and predictive power [10].

Although recently much progress has been made in building ML models using neonatal physiological data, there are limitations in the detailed reporting of the preprocessing techniques of these signals [15], which in turn hinder the reproducibility of the methods and results. In AI-powered software as a medical device (SaMD), this is especially important as the implementation of a software quality management system (QMS) is only possible by following the best practices and adhering to relevant regulatory standards and guidelines for medical devices, such as ISO 13485, IEC 62304, and IEC 82304-1. Beyond market access considerations, the ongoing international discourse on the regulation of medical software is specifically concentrated on AI and ML. This focus is a response to their growing applications, demanding increased attention from regulatory bodies such as the Australian Therapeutic Goods Administration and the US Food and Drug Administration [16]. Thus, it is crucial to adhere to a standardized protocol following clinical principles guided by domain experts and regulatory requirements while preprocessing the signals and reporting these techniques in detail; this ensures the reproducibility of the methods, allowing transparency in their clinical adoption.

Objectives

As the first step in bridging the gap in their reproducibility for clinical adoption, this review aims to identify studies that used computational methods to analyze premature infants' physiological signals for detecting adverse outcomes. The review describes different tools and techniques used to preprocess physiological signals and provides recommendations on what aspects need further details for the clinical adoption of the techniques. The remainder of the paper is organized as follows: the Methods section explains the detailed search and screening process, while the Results section begins with an overview of the reviewed studies, followed by a detailed analysis. The Discussion section highlights the key reporting patterns identified in this review along with their shortcomings and provides recommendations for transparent reporting of future studies as it allows for accurate reproduction of the results and makes them usable in the clinical setting [17]. A summary of the work concludes the paper.

Methods

Search Strategy

The database searches and study screening were conducted following the recommendations of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [18] and the Centre for Reviews and Dissemination guidance for undertaking reviews in health care [19].

Database and Search Strategies

A systematic database search was conducted on 5 databases: PubMed, IEEE, Web of Science, Scopus, and ACM Digital Library. The keywords were categorized into four concepts, which were then merged using the “AND” operator: concept 1—neonates or preterm infants; concept 2—vital signs or physiological signals; concept 3—computational techniques or signal processing; and concept 4—outcomes relating to neonates. Within each of these concepts, a combination of keywords and MeSH (Medical Subject Headings) terms were used to conduct the search process. The keywords under each

concept were combined by the “OR” operator. The searches were limited to only the titles and abstracts. Table 1 shows the list of keywords and Medical Subject Headings terms used to search the database.

The search was done on January 9, 2023, and the publication year of the papers was limited to 2013 to 2023. The reason for choosing the 10-year range was to report on recent techniques and tools, as the devices and computational tools used >10 years ago may be obsolete. Scopus, Wiley Online Library, and Web of Science have an additional filter for choosing the subject area. This was used to restrict the subject areas to multidisciplinary, engineering, computing, and statistics. This was done to identify more papers on multidisciplinary areas through these databases, as PubMed covers all the major medical and health informatics databases. The combination of the 5 databases ensured that all medical, information technology, and multidisciplinary research papers were included in the database search. The search was restricted to English-language articles. Finally, review articles were excluded from the search.

Table 1. List of keywords and MeSH (Medical Subject Headings) terms used to conduct the database search.

Concepts	Search strategy
Concept 1: neonates or preterm babies	
MeSH terms	“Infant, Premature”
Keywords	“premature” OR “preterm” OR “neonat*” OR “newborn” OR “infant” OR “nicu” OR “neonatal intensive care unit”
Concept 2: physiological signals or vital signs	
MeSH terms	“Vital Signs” OR “Physiology”
Keywords	“physiolog*” OR “ecg” OR “heart rate” OR “electrocardiography” OR “vital sign*” OR “physiomarker” OR “biomarker” OR “hrv”
Concept 3: computational techniques or signal processing	
MeSH terms	“Signal Processing, Computer-Assisted”
Keywords	“signal *” OR “predict*” OR “detect*” OR “comput*”
Concept 4: outcomes	
MeSH terms	None
Keywords	“sepsis” OR “mortality” OR “length of stay” OR “intraventricular hemorrhage” OR “hypoxi*” OR “apnea” OR “necrotising enterocolitis” OR “necrotizing enterocolitis”

Screening and Study Selection

The initial screening of the databases led to 3585 papers. Of these, 590 (16.46%) papers were manually identified as duplicates and excluded from the analysis. One paper was identified as a duplicate by the automation tool and removed. The remaining 2994 (83.51%) papers were subjected to title and abstract screening using the Rayyan Intelligent Systematic Review application (Qatar Computing Research Institute) [20].

Several inclusion criteria were set to select papers for full-text review. The criteria are mentioned in [Textbox 1](#).

After screening the titles and abstracts, 81 articles were selected for full-text review; 29 (36%) papers were excluded during this stage as they did not align with the inclusion criteria, leaving 52 (64%) papers eligible for detailed synthesis and analysis.

The title and abstract screening was done by 1 reviewer, while 2 reviewers independently checked for paper eligibility against the inclusion criteria at the full-text review stage. When both reviewers were not in agreement on any papers, a third reviewer assessed them to provide a final decision on the inclusion and exclusion of the papers. Data charting was done using Microsoft Excel, and the following variables were recorded in line with related review papers [10,21]: title, year, journal, authors, digital object identifier, data set, participant number, participant demographic, signals used, data set size, sample rate, other data (if applicable), outcome metric, device software, programming language, preprocessing methods, algorithms, other techniques, features, models, model type, results (quantified), and key findings. Data synthesis was done using a narrative approach by summarizing findings based on the similarities in the data sets and techniques used. The detailed search queries,

bibliography files of all databases, all included papers, metadata of all papers and metadata of all papers included for full-text

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Article type: articles must be peer-reviewed publications in a journal, conference, or workshop
- Data: articles must conduct an analysis on premature human infant data; articles must use physiological responses in some form
- Outcome: articles discuss applications relating to adverse neonatal outcomes such as mortality, length of stay, sepsis, necrotizing enterocolitis, intraventricular hemorrhage, hypoxic-ischemic encephalopathy, apnea, bradycardia, and other poor health outcomes, also known as morbidity. The disease outcomes were chosen based on the commonly researched outcome metric using preterm infant data and the search terms used in McAdams et al [10] that investigated artificial intelligence and machine learning techniques used to predict clinical outcomes in the neonatal intensive care unit
- Analysis: articles reported some form of computational techniques in their analysis
- Language: English

Exclusion criteria

- Article type: review papers are excluded
- Data: nonhuman data (eg, piglet infant data would not be considered); videos and images that do not look at the physiological responses and articles solely using demographic data for analysis were excluded
- Outcome: articles not focusing on these specified neonatal adverse outcomes were excluded
- Analysis: articles that only reported responses in their raw format were excluded
- Language: any languages other than English

Results

Overview of the Included Studies

Figure 1 shows the full process of database search and study selection using a PRISMA flow diagram.

Of the 52 selected articles, 24 (46%) studies focused on diagnostic models, while the rest ($n=28$, 54%) focused on prognostic models. These included journal articles ($n=34$, 65%), conference articles ($n=17$, 33%) and a workshop article ($n=1$, 2%). The most prominent physiological signals analyzed were ECG ($n=36$, 69%), SpO₂ ($n=21$, 40%), HR ($n=16$, 31%), respiration ($n=16$, 31%), BP ($n=6$, 12%), EEG ($n=4$, 8%), and temperature ($n=3$, 6%). While 8 (15%) studies used a combination of programming languages; others used MATLAB ($n=6$, 12%), Python ($n=6$, 12%), and R software ($n=1$, 2%), while the remaining studies ($n=31$, 60%) did not report what language was used. Physiological data monitoring and capturing was done using a range of systems, which subsequently impacted the sampling rate and quality of the data, thus leading to heterogeneity of the studies. The most commonly used devices for data capturing were Phillips Intellivue MP20, MP70, MP450, and MX800 machines [74] ($n=14$, 27%). Some other notable devices and software were BedMaster Ex System [75], NicoletOne EEG system [76], ixTrend, Phillips Data Warehouse connect [77], and Vuelogger patient monitoring system. The most commonly analyzed outcomes of interest were sepsis ($n=20$, 38%), apnea ($n=17$, 33%), bradycardia ($n=13$, 25%), mortality ($n=7$, 13%), and hypoxic-ischemic encephalopathy ($n=5$, 10%). It should be noted that 14 (27%) of the reviewed studies analyzed a combination of adverse outcomes.

review are provided in [Multimedia Appendices 1-5](#) [22-73].

As the studies were found to be heterogeneous in their study design and analysis techniques, a narrative approach was taken to summarize the studies and their key findings. The studies were grouped according to the homogeneity in terms of the data sets used and sorted by the publication year. This approach was inspired by the review article by Mann et al [78].

One of the noticeable patterns identified through the results reported in [Table 2](#) is that the groups publishing studies using the same data set followed similar preprocessing techniques, although not at every step. For instance, studies using the ECG data from Cork University Maternity Hospital all used the same algorithm for QRS complex detection. However, they were diverse in their selection of filtering techniques and segmentation duration. Furthermore, they systematically failed to report detailed parameter settings for the QRS complex detection. While the approach of using similar preprocessing techniques helps maintain consistency to some extent, they do not confirm adhering to clinical practices identified from domain expert knowledge.

The QRS complex characteristics and RR intervals for neonates are different from those of adults and as such require an appropriate adjustment for QRS detection algorithms. This is a necessary first step for HR variability (HRV) analysis in neonates. However, a review published on neonatal HRV by Latremouille et al [15] revealed that given a lack of clear guidelines on neonatal vital signs and HRV analysis, several studies followed HRV analysis guidelines for adults published by the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology [79]. Our review found that 16 (44%) out of the 36 studies analyzing ECG signals used the Pan-Tompkins algorithms for QRS complex detection. The original implementation of the

algorithm was based on the ECG characteristics of the adult population and therefore was preprocessed accordingly. Only 4 (25%) of those 16 studies reported adjustment of the original algorithm to adapt to neonates, of which only 2 provided specific modification details. In the absence of detailed reporting on the parameter settings, it is difficult to determine whether the settings adhered to neonatal waveform morphology. Incomplete reporting and lack of transparency hinder the understanding of the strengths and weaknesses of a study and limit its reproducibility and usability. Moreover, transparent and detailed reporting is required to confirm the adherence to regulatory compliance and is crucial for the clinical adoption of these methods.

Similar to the QRS complex in ECG signals, the acceptable ranges of physiological signals for neonates are also different from those of the adult population. This review found that no studies reviewed the acceptable ranges of the analyzed signals against any published guidelines, which could pose several limitations in the clinical adoption of the methods. This is consistent with another review looking into physiological vital sign ranges from 34 weeks gestational age, and it identified that several studies reported the means of vital signs instead of ranges, which makes the interpretation into clinical practice difficult [80]. Here, we recommend clear reporting and the use of physiological signal ranges that are clinically validated through published studies and textbooks [81-83].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the database search and study selection.

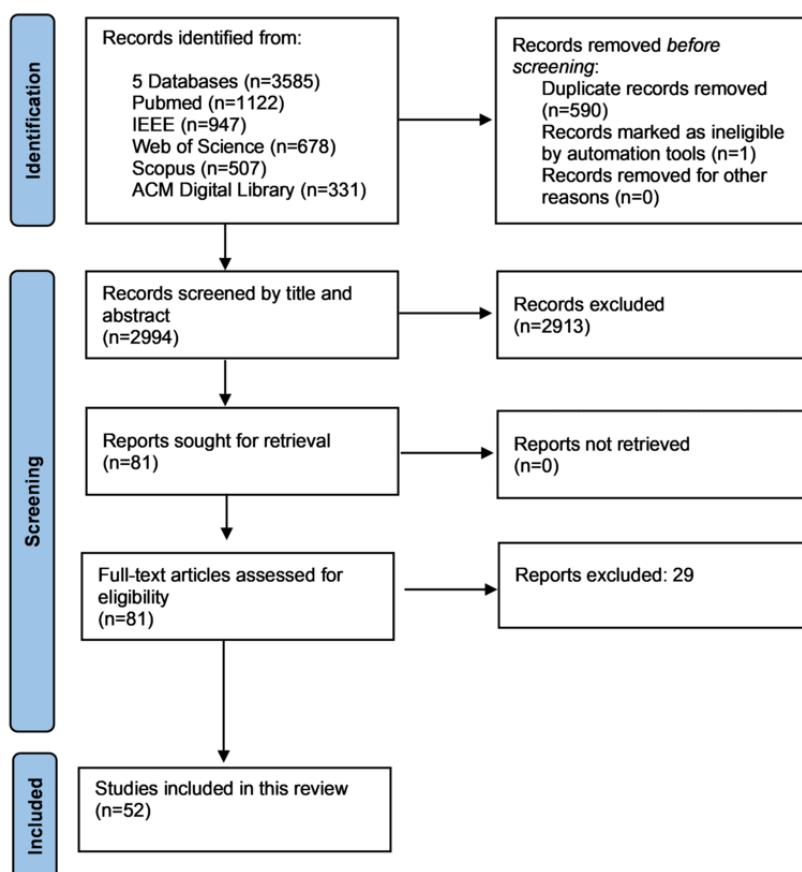


Table 2. Summary of the articles reviewed in this study, grouped according to the homogeneity in terms of the data sets used and sorted by the publication year.

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
National CHIME ^a database [84]	Cohen and de Chazal [22], 2013	Participants: n=288; data size: NR ^b ; model: diagnostic; outcome metric: sleep apnea	ECG ^c from single channel at 100 Hz, SpO ₂ ^d at 1 Hz	SpO ₂ values <65% and changes in saturation exceeding 4% per second were discarded. ECG QRS complex was detected using the Pan-Tompkins algorithm [85] to generate RR intervals. QRS complexes were filtered using a technique from Chazal et al [86]. Filtered intervals were time aligned with SpO ₂ using 30-second epochs	Eleven features were extracted from the signals. A combination of features from both signals resulted in 88.8% accuracy, 94.3% specificity, and 73.4% sensitivity in detecting sleep apnea
CHIME	Cohen and de Chazal [23], 2014	Participants: n=402; data size: NR; model: diagnostic; outcome metric: sleep apnea	ECG from single channel at 100 Hz, SpO ₂ at 1 Hz, actigraphy signals at 50 Hz	Actigraphy signals artifact rejection was done using the technique described by Lewicke et al [87]. SpO ₂ values <65% and changes in saturation exceeding 4% per second were discarded. ECG data were passed through a QRS detection algorithm (NR) to produce RR intervals, which were filtered using a previously outlined method [86]	Fourteen features were extracted from the signals. A linear discriminant classifier achieved an accuracy of 74.1%, a sensitivity of 82.0%, and a specificity of 60.9% in detecting sleep apnea
CHIME	Cohen and de Chazal [24], 2015	Participants: n=394; data size: NR; model: diagnostic; outcome metric: sleep apnea	ECG from single channel at 100 Hz, SpO ₂ at 1 Hz	SpO ₂ and ECG signals were time aligned to 30-second epochs. SpO ₂ values <65% and changes in saturation exceeding 4% per second were discarded. ECG QRS complex detected using the Pan-Tompkins algorithm [85] to generate RR intervals. QRS complexes were filtered using a technique from Chazal et al [86]	Eleven features were extracted from both signals. A linear discriminant model achieved 66.7% accuracy, 67% specificity, and 58.1% sensitivity using features from both signals
PICS ^e database [25,88]	Gee et al [26], 2016	Participants: n=10; data size: ~20-70 hours each; model: diagnostic; outcome metric: bradycardia	3-lead ECG at 500 Hz, respiration signal at 50 Hz	RR intervals from ECG were extracted using a modified Pan-Tompkins algorithm (modification details NR). Analysis was done on a 3-minute window before each bradycardia. No processing was reported for respiration signals	Bradycardia severity estimation accuracy was improved by an average of 11% using a point process model of heart rate and respiration

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
PICS	Gee et al [25], 2017	Participants, 10; data size: ~20-70 hours each; model: prognostic (+116 seconds); outcome metric: bradycardia	3-lead ECG at 500 Hz, respiration signal at 50 Hz	RR intervals from ECG were extracted using a modified Pan-Tompkins algorithm (modification details NR). The artifacts, due to movement, disconnection, or erroneous peaks, were removed by visual inspection. No processing was reported for respiration signals. Additional analysis on the frequency content of the RR time series was done using Morlet wavelet transform [89]	A point process model-based prediction algorithm achieved a mean AUROC ^f of 0.79 for >440 bradycardic events and was able to predict bradycardic events on an average of 116 seconds before onset (FPR ^g =0.15)
PICS	Das et al [27], 2019	Participants; n=10; data size: ~20-70 hours each; model: prognostic (time NR); outcome metric: bradycardia	3-lead ECG at 500 Hz	Baseline wander was removed using a high-pass filter with a cutoff frequency between 0.5 and 0.6 Hz. Motion and disconnection artifacts were removed by visual inspection. QRS complexes were detected using Pan-Tompkins algorithm [85]. Signals were segmented 5 minutes before and 2 minutes after a bradycardic event	Nonparametric modeling using kernel density estimation achieved a 5% false alarm rate in predicting the onset of bradycardia events
PICS	Mahmud et al [28], 2019	Participants; n=11; data size: ~20-70 hours each for 10 and 10 weeks for 1 participant; model: prognostic (time NR); outcome metric: bradycardia	3-lead ECG at 500 Hz	QRS complex was detected using an algorithm (NR). RR intervals were calculated from the detected peaks	Time and frequency domain features were extracted. An extreme gradient boosting model achieved an average AUROC of 0.867. HRV ^h results showed a significant variation between a healthy infant and an infant prone to bradycardia

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
PICS	Gee et al [29], 2019	Participant, n=10; data size: ~20-70 hours each; model: diagnostic; outcome metric: bradycardia, apnea of prematurity	3-lead ECG at 500 Hz, respiration signal at 50 Hz	Respiration signals were clipped into 60-second segments and normalized to 0-mean, unit variance. RR intervals from ECG signals were extracted using a Morlet wavelet transformation. An open-source peak finder (name NR) was applied to the wavelet scale ranging from 0.01 to 0.04, which is related to the QRS complex formation in the spectrogram. ECG signals were segmented to 15 seconds with the event in the middle. The segments were bandpassed filtered from 3 to 45 Hz, scaled to 0-mean unit variance, and scaled to the median QRS complex amplitude. Waveforms were visually inspected to remove segments with no distinguishable QRS complex or respiratory peaks	An autoencoder-prototype model was proposed, which achieves 93.1% (SD 0.4%) accuracy in predicting bradycardia and 82.3% (SD 3.8%) accuracy in classifying apnea
MIMIC-III ⁱ database from Beth Israel Deaconess Medical Center [90]	Song et al [30], 2020	Participants: 2819 (21 sepsis, 2798 control); data size: NR; model: prognostic (+48 hours); outcome metric: sepsis	HR ^j , SBP ^k , DBP ^l , MBP ^m , SpO ₂ , respiration, temperature, other (sampling rate NR)	Data quality was assessed by missing value filter and 3-sigma rule. The final observation carried forward was applied to vital signs not meeting data quality. Zero imputation was performed if calculation could not be performed (eg, divided by 0)	Several statistical features were extracted at 3-, 6-, 12-, and 24-hour window. Linear model, naive Bayes, decision tree, ensemble method, and neural network models were evaluated. The AUROC of the 48-hour prediction model achieved 0.861 and that of the onset detection model was 0.868
MIMIC-III	Baker et al [31], 2021	Participants; n=179 for 3-day and n=181 for 14-day model; data size, NR; model: prognostic (+3 days); outcome metric: mortality	HR, respiration signal, sampled hourly	Values <0 and flatline cases were eliminated	Several statistical features were extracted from the signals. CNN-LSTM ⁿ model using a 3-day scheme achieved AUROC of 0.9336 (SD 0.0337) across 5-fold cross-validation
MIMIC-III	Juraev et al [32], 2022	Participants; n=3133; data size: 24 hours from each; model: prognostic (time NR); outcome metric: mortality and LOS ^o	HR, respiration signal, SpO ₂ , BP, temperature (sampling rate NR)	Missing data were filled by forward and backward filling, using the mean value. For participants >24 measurements, they were reduced by taking the average of the nearest records. For <24 measurements, values were generated using filling algorithm	A dynamic ensemble KNN ^p method reached 0.988 (SD 0.001) F_1 -score in mortality classification. Voting of static ensemble regression models achieved an RMSE ^q of 12.509 (SD 0.079) in LOS prediction

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
University Hospitals in France	Ghahjaverestan et al [33], 2015	Participants; n=32; data size: 105 segments of ECG with 250 seconds duration; model: diagnostic; outcome metric: apnea-bradycardia	One lead ECG at 400 Hz and respiration signals (sampling rate NR)	Baseline and noise of 50 Hz were removed from ECG signals, QRS complexes were detected using Pan-Tompkins algorithm [85]. RR intervals were further downsampled to 10 Hz for 1 prediction model	A Kalman-filter-based method achieved sensitivity and specificity of 94.74% and 94.17%, respectively, in predicting apnea-bradycardia episodes
University Hospitals in France	Navarro et al [34], 2015	Participants; n=51; data size: testing cohort mean duration—2.4 hours; model: diagnostic; outcome metric: sepsis	Respiration signals at 400 Hz, downsampled to 64 Hz	Frequency content >32 Hz from breathing signals was removed using a seventh-order Butterworth low-pass filter. After rejecting artifacts due to gross movements, a fourth-order Butterworth filter with a cutoff frequency between 0.5 and 20 Hz was applied. Smoothing filtering using an SG ^r filter [91] was applied. A simple extrema detector is then applied to detect respiratory cycles	14 features, computed in 10-second sliding excerpts, were extracted from the breathing signals. A logistic regression classifier automatically rejects artifacts to 86% sensitivity and specificity, which is used in the proposed framework for neonatal sepsis detection
University Hospitals in France	Ghahjaverestan et al [35], 2016	Participants; n=32; data size: real (236 segments Synthetic) 200 sequences of 400 seconds; model: diagnostic (0.59-second delay); outcome metric: apnea-bradycardia	ECG at 400 Hz. Synthetic signals at 10 Hz	Baseline and noise of 50 Hz were removed from ECG signals using a combination of low-pass and notch filters; QRS complexes were detected using Pan-Tompkins algorithm [85]. Three features were extracted using a wavelet-based beat delineator [92]. Features were transformed to 10 Hz using interpolation (technique NR)	A CHMM ^s achieved 95.74% sensitivity and 91.88% specificity in detecting apnea-bradycardia episodes, with a detection delay of -0.59 seconds
University Hospitals in France	León et al [36], 2021	Participant, n=49; data size: NR; model: prognostic (+6 hours); outcome metric: sepsis	ECG at 500 Hz	RR intervals were detected using modified Pan-Tompkins algorithms, and filter coefficients were adapted for newborns [93]. A sliding window of 30 minutes, with no overlaps, was applied to extract HRV parameters from the RR time series. 30-minute segments with a maximum RR >1 second or a minimum RR of <0.19 seconds were excluded,	Time, frequency, and non-linear features were extracted from the HRV parameters. A logistic regression model using visibility graph features achieved 0.877 AUROC in predicting sepsis 6 hours before the start of antibiotics

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
University Hospitals in France	León et al [37], 2021	Participants; n=259; data size: NR; model: prognostic (+6 hours); outcome metric: sepsis	ECG at 500 Hz	RR intervals were detected using modified Pan-Tompkins algorithms and filter coefficients were adapted for neonates [93]. RR time series were extracted and segmented into 5-minute segments. The 5-minute periods corresponding to 30 continuous minutes were grouped by calculating the median of each corresponding HRV feature	Time, frequency, nonlinear, and visibility graph features were extracted from the HRV parameters. An RNN ^t model achieved 0.904 AUROC in predicting sepsis 6 hours before the time of infection and >80% accuracy 24 hours before the onset of infection
University Hospitals in France	Doyen et al [38], 2021	Participants; n=52; data size: 8 hours of recording from each; model: diagnostic (+2.9-second delay); outcome metric: bradycardia	3-lead ECG at 300 Hz	QRS complexes were detected using a multi-feature probabilistic real-time detector [93]	A high rate of false alarms (64%) was observed in real life. The proposed optimal decentralized fusion of 3 detection methods had a significant detection delay of 2.9 seconds, sensitivity of 97.6% and false alarm rate of 63.7%
University Hospitals in France	Sadoughi et al [39], 2021	Participants; n=32; data size: 233 episodes with a duration of 21.48 (SD 16.07) seconds; model: diagnostic (+5.05-second delay); outcome metric: apnea-bradycardia	One lead ECG at 400 Hz	The same preprocessing techniques as reported in Ghahjaverestan et al [35]. QRS complexes were identified using Pan-Tompkins method [85]. The RR time series were uniformly upsampled to 10 Hz using a linear interpolation technique	A proposed layered HMM ^u model achieved 97.14% (SD 0.31%) accuracy in detecting apnea-bradycardia episodes, with a detection delay of -5.04 (SD 0.41) seconds
Cork University Maternity Hospital	Ahmed et al [40], 2015	Participants: NR; data size: 54 1-hour recordings; model: diagnostic; outcome metric: HIE	2-lead ECG, EEG ^v (sampling rate NR)	Artifacts were manually removed. R-peaks from raw ECG signals were extracted using Pan-Tompkins method [85]. The timing of the peaks was adjusted and uniformly sampled to 256 Hz using Hermite spline quadratic interpolation. Then, HRV features were extracted from a 1-minute window with 30-second overlap using the normalized RR interval	Seven time and frequency domain HRV features were extracted. A Gaussian supervector approach with SVM ^w achieved 0.81 AUROC in classifying HIE ^x

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Cork University Maternity Hospital	Temko et al [41], 2015	Participant, n=38; data size: 1 hour of EEG and ECG recordings from each; model: diagnostic; outcome metric: HIE	ECG and video EEG at 256 Hz	The 1-hour EEG segments were downsampled to 32 Hz with an antialiasing filter set to 16 Hz. The filtered EEG was segmented into a 60-second epoch with no overlap. QRS complexes from ECG signals were extracted using the algorithm reported in [94]. The resulting peaks were manually inspected to correct ectopic beats or mark artifacts. Then, signals were segmented into 60-second epochs	An SVM classifier using a subset of 9 EEG, 2 hours, and 1 clinical feature achieved 87% AUROC and 84% accuracy in predicting HIE
Cork University Maternity Hospital	Lloyd et al [42], 2016	Participant, n=43; data size: mean recording duration 41 hours 40 minutes; model: diagnostic; outcome metric: future adverse outcome in infants	EEG at 256 Hz, SpO ₂ and HR at 1 Hz	EEG recordings were visually checked for quality, and poor-quality data were discarded. 1-hour epochs of EEG at 12 and 2 hours of age were then extracted from each recording. 1-hour epochs of HR and SpO ₂ were extracted at 12- and 24-hour time point.	A logistic regression model predicted a 2-year poor outcome with an AUROC of 0.83
Cork University Maternity Hospital	Semenova et al [43], 2018	Participants; n=35 with 23 used; data size: 824 hours; prognostic (time NR); outcome metric: short-term adverse outcome	ECG at 256 or 1024 Hz, BP at 1 Hz	Diastolic and systolic pressures every second were used to calculate MAP ^y . ECG signals were segmented to nonoverlapping 5-minute epochs. QRS complexes were extracted by the Pan-Tompkins method [85]. Abnormal RR intervals were corrected by moving average. Periods of clear movement of artifacts were automatically discarded (method NR)	Fifteen time, frequency, and nonlinear features were extracted from HRV. An XGBoost decision tree using all features achieved an AUROC of 0.97 in predicting short-term outcomes in infants

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Cork University Maternity Hospital	Semenova et al [44], 2019	Participants; n=43 with 23 used; data size: total 831 hours; prognostic (time NR); outcome metric: 5 adverse outcomes	ECG at 256 or 1024 Hz, BP at 1 Hz	DBP and SBP every second were used to calculate MAP. Segments with MAP<10 mm Hg were discarded due to disconnection of the pressure transducer or movements. The MAP was segmented into 1-hour windows. Values outside 3 SD were discarded. ECG signals were segmented into nonoverlapping 5-minute epochs. QRS complexes were extracted using the Pan-Tompkins method [85]. ECG signal was band-pass filtered with 4-30 Hz cutoff frequency. Abnormal values of RR intervals were corrected by the moving average filter	Time, frequency, and nonlinear features were extracted from HRV. An XG-Boost decision tree using a single HRV feature achieved 0.87 AUROC, while multiple features reached 0.97 AUROC in predicting adverse outcomes
Máxima Medical Center NICU ^z	Joshi et al [45], 2020	Participants; n=49; data size: ~144 hours each; model: prognostic (+0-24 h); outcome metric: sepsis	ECG at 250 Hz, CI ^{aa} at 62.5 Hz	Respiration waveforms were bandpass filtered between 0.45 and 1.45 Hz. QRS complexes from ECG were extracted using a DT-CWT ^{ab} -based method described in [95]. IBIs ^{ac} were detected from the CI signal peaks using an algorithm (NR). Features were extracted from every 3-hour data	Twenty-two features were extracted from the signals. A naive bayes classifier reached up to 0.78 AUROC and 3 hours leading up to sepsis
Máxima medical Center NICU	Varisco et al [46], 2021	Participants; n=20; data size: ~570 hours; model: prognostic (+6 hours); outcome metric: central apnea preceding late-onset sepsis	ECG at 240 or 250 Hz, CI at 60 or 62.5 Hz, SpO ₂ at 0.5 or 1 Hz	A filtered respiration signal without cardiac artifacts was generated using algorithms reported in studies by Lee et al [96], Mohr et al [97], and Vergales et al [98]. Steps include Fourier transformation and integer frequencies filtered out, then resampled to 60 Hz and high-pass filtered with a cutoff frequency of 0.4 Hz, and a low-pass filter with a very low cutoff frequency optimized to fit apnea annotations by clinical experts (value NR)	An optimization of the algorithm was proposed to detect central apnea, which achieved 90.5% recall, 19.7% precision, and 30.8% F ₁ -score

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Máxima medical Center NICU	Cabrera-Quiros et al [47], 2021	Participants; n=64; data size: NR; model: prognostic (+3 hours); outcome metric: sepsis	ECG at 250 Hz, CI at 62.5 Hz	QRS complexes from ECG were extracted using a DT-CWT-based method described the same as Joshi et al [45]. CI signal was filtered to remove cardiac artifacts, and peaks were detected using methods similar to those in previous works (NR). Features were extracted from every 1-hour signal	Time domain features were extracted from HRV. Classification using a combination of all features and logistic regression model reached a mean accuracy of 0.79 (SD 0.12) and mean precision of 0.82 (SD 0.18), 3 hours before the onset of sepsis
Máxima Medical Center NICU	Varisco et al [48], 2022	Participants; n=20; data size: 960 hours of data from 20 infants, 7818 event extracted; model: diagnostic; outcome metric: central apnea	ECG at 250 Hz, CI at 62.5 Hz, SpO ₂ at 1 Hz	QRS complexes were detected using the same method as reported in Joshi et al [45] and Cabrera-Quiros et al [47]. From ECG, SII ^{ad} was calculated by applying a bandpass filter (0.001-0.40 Hz) using 10-second segments and then computing a kernel density estimate to return patient motion measurement every second. RR intervals were resampled at 250 Hz. CI signal was processed using the method by Redmond et al [99] to calculate RRE ^{ac} . No preprocessing was done on SpO ₂ . Each feature was extracted using 30-second windows. z score normalization was applied to the feature matrix	47 features were extracted from the vitals. A logistic regression model achieved 0.9 AUROC in detecting central apnea
Máxima Medical Center NICU	Peng et al [49], 2022	Participants; n=128; data size: ~24 hours each; model: prognostic (+24 hours); outcome metric: sepsis	ECG at 250 Hz	QRS complexes from ECG were extracted using a DT-CWT-based method described by Rooijakkens et al [95]. RR intervals from the complexes were divided into nonoverlapping 1-hour segments. The segments were centered, and missing values in the segments were filled by zero padding on the 2 ends	A ResNet-based neural network, DeepLOS, was proposed, which achieved a 0.72 <i>F</i> ₁ -score in predicting late-onset sepsis

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Máxima Medical Center NICU	Peng et al [50], 2022	Participants; n=127; data size: ~48 hours each; model: prognostic (+6 hours); outcome metric: sepsis	ECG at 250 Hz, CI at 62.5 Hz	QRS complexes from ECG were extracted using a DT-CWT-based method described in [95]. CI signal was filtered to remove cardiac artifacts (method NR). Peaks were detected using the method reported by Lee et al [96]. SII was calculated from ECG and CI waveforms using a CWT-based method, as reported by Zuzarte et al [100]. Signals were divided into 1-hour-long nonoverlapping segments. Features were calculated in both 1-hour segments and 5-minute subsegments	60 Features were extracted from the signals. An XGB model using the features achieved an AUROC of 0.88 in predicting late-onset sepsis 6 hours preceding the onset.
Royal Infirmary of Edinburgh NICU	Stanculescu et al [51], 2014	Participants; n=24; data size: 30 hours each; model: prognostic (+3-6 hours); outcome metric: sepsis	ECG-derived HR, PR ^{af} (sampling rate NR)	An extension of the forward-backward algorithm [101] is developed for missing data inference	An autoregressive HMM model achieved up to 0.80 AUROC in predicting sepsis
Royal Infirmary of Edinburgh NICU	Stanculescu et al [52], 2014	Participants; n=24; data size: 540 hours; model: diagnostic; outcome metric: sepsis	ECG-derived HR, PR core and peripheral temperature and SpO ₂ at 1 Hz	An automated oximeter error detection algorithm was applied on the basis of the method described by Stanculescu et al [51]. Rows containing missing data on the observation matrix are set to 0	An HSLDS ^{ag} was able to predict sepsis with up to 0.65 F ₁ -score
Kasturba hospital NICU, Manipal, India	Shirwaikar et al [53], 2016	Participant: NR; data size: 229 examples; model: diagnostic; outcome metric: apnea	HR (sampling rate NR)	Visualization technique was applied to identify issues in data. Missing values were not treated due to low percentage. For categorical features, 0 was added for missing values. Minimum-maximum normalization and z score normalization were done	An RF ^{ah} model using HR features achieved 0.88 accuracy and 0.72 κ in detecting apnea
Kasturba hospital NICU, Manipal, India	Shirwaikar et al [54], 2019	Participants; n=367 (315 used); data size: NR; model: diagnostic; outcome metric: apnea	ECG (sampling rate NR)	No preprocessing techniques were reported on the raw signals. Observations with missing features were discarded. Other features (continuous values) that had missing values were converted to discrete with the addition of the group name “not known”	Statistical features were extracted from the signals. A Multilayer Perceptron model and a deep autoencoder model reached 0.82 and 0.83 AUROC, respectively, in detecting apnea

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
University of Massachusetts Memorial Healthcare NICU	Williamson et al [55], 2013	Participants; n=6; data size: ~5-8 hours for each patient; model: prognostic (+5.5 minutes); outcome metric: apnea	ECG, SpO ₂ , respiratory signal, pulse plethysmogram (sampling rate NR)	IBIs were extracted from abdominal respiratory movements (method NR), and RR intervals were extracted from ECG signals (method NR). Physically implausible IBI and RR interval values were automatically removed (range NR). Values were resampled to 10 Hz using shape-preserving piecewise cubic interpolation. Signals were then log transformed and converted to 0 mean, unit variance	Features were extracted from all signals. A GMM ^{ai} model reached 0.8 AU-ROC in predicting apnea
Jackson Memorial Hospital NICU	Schiavonato et al [56], 2013	Participants; n=20; data size: 1186 minutes; model: diagnostic; outcome metric: periods of high distress or pain	ECG at 1000 Hz	Pan-Tompkins algorithm [85] was modified to detect QRS complexes. ECG was filtered using a bandpass filter with a 16-26 Hz cutoff frequency. A low-pass filter by an order 120 FIR ^{aj} filter with a corner frequency of 25 Hz and a high-pass filter by an order 160 FIR filter with a corner frequency of 25 Hz were applied. Then, a polynomial filter of order 21 was applied as the differentiator filter. Finally, a 111-order moving average filter was used, and QRS complex was detected using an adaptive threshold. Lomb-Scargle LMS ^{ak} spectral estimation [102] was used for missing and irregular RR intervals	The proposed framework provided real-time analysis and HRV extraction to identify the characteristics correlated to periods of high distress or pain
Montreal Children's Hospital	Rubles-Rubio et al [57], 2014	Participants; n=24; data size: 9.0 (SD 2.2) hours for each; model: diagnostic; outcome metric: apnea	SpO ₂ , RIP ^{al} (sampling rate NR)	Signals were low-pass filtered with a cutoff frequency of 10 Hz, with an 8-pole Bessel antialiasing filter digitized and sampled at 50 Hz	A linear Gaussian discriminant classifier detected the episodes with a 0.73 probability of detection and 0.22 probability of false alarm
University of Alabama at Birmingham	Amperayani et al [58], 2017	Participants; n=18; data size: 24 hours each; model: prognostic (+23 hours); outcome metric: bradycardia, hypoxemia	ECG at 500 Hz and HR at 1 Hz	HR data were converted to interbeat RR intervals using RR=60/HR. No processing on ECG signals was reported	A point process model using RR intervals showed a strong correlation with bradycardia events and a modest correlation with hypoxemia events

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Monash Children's Hospital NICU, Australia	Hu et al [59], 2018	Participants; n<80; data size: 407 patient-day; model: prognostic (+24 hours); outcome metric: sepsis	HR, SpO ₂ , respiration signal at 1 Hz	Data were scaled down to 1 record per minute. Data blocks with invalid values were deleted. Then, the sliding window was set to 60 minutes to feed to the ML ^{am} models	Features were extracted from all signals. A gradient boosting decision tree achieved up to 0.97 AUROC and 0.92 weighted F_1 -score in patient-based cross-validation in predicting sepsis
University of Virginia and Columbia University NICU	Sullivan et al [60], 2018	Participants; n=78; data size: NR; model: prognostic (+12 hours); outcome metric: death, sIVH ^{an} (severe), BPD ^{ao} , treated ROP, ^{ap} late-onset sepsis, and NEC ^{aq}	HR, SpO ₂ at 0.5 Hz	Infants with <6 hours of data within 12 hours of birth were discarded. Cross-correlation of HR and SpO ₂ was calculated over 10-minute windows using the XCORR function of MATLAB with a lag time of -30 to +30 seconds	A POPS ^{ar} was developed and fit a multivariate logistic regression model, which performed well in predicting death, sIVH, and BPD, but not tROP, sepsis, and NEC
9 NICUs in the United States	Zimmet et al [61], 2020	Participants; n=2989; data size: 121 data points per infant; model: prognostic (+2 days); outcome metric: mortality, sepsis	HRC ^{as} index from ECG	Infants with missing data on either end of the total duration were extrapolated to the window edge by repeating the most proximal HRC index values. Interior missing values were updated using linear interpolation. A fifth-order B-spline with equally spaced knots was used to capture information from independent samples (HRC indexes 12 samples apart)	An unsupervised ensemble of clustering techniques was proposed to cluster infants to different levels of risk
Children's National Hospital, Washington	Kota et al [103], 2020	Participants; n=95; data size: median recording duration of 75.78 hours; model: diagnostic; outcome metric: HIE	EEG at 200 or 256 Hz	EEG contamination from EEG was detected using the method described by Govindan et al [104]. EEG signals with amplitude>500 μ V or SD<0.01 μ V were discarded as artifacts. The volume conduction was attenuated by calculating the global average of EEG voltages from all electrodes and subtracting the global average from the EEG value of every electrode in the frequency domain [62]. The values were then transformed to the time domain for spectral analysis. EEG was segmented into 10-minute nonoverlapping artifact and seizure-free epochs. Spectral analysis was done using a Welch periodogram approach [105,106] using 3-second epochs	EEG delta power was identified to be a crucial biomarker for predicting neonates with HIE who died with those who survived

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Akbar Abadi Hospital NICU, Iran	Mirnia et al [63], 2021	Participants; n=5; data size: ~24 hours each; model: diagnostic, outcome metric: sepsis	ECG at 200 Hz	RR intervals were calculated from ECG using HeRO ^{at} model	Features were extracted from HRV. HeRO model was tested using this data set. HeRO score was able to distinguish between healthy and septic newborns
St Louis Children's Hospital NICU	Lee et al [64], 2021	Participants; n=275; data size: 4, 01,33,460 data points; model: prognostic (+6 hours); outcome metric: mortality	HR, respiration signal and SpO ₂ at 1 Hz	Missing or out-of-range values were replaced with NaN and then imputed using mean values for that variable across all training and testing data. Data were downsampled to every 10 seconds to extract features. Dynamic variables were calculated as rolling means, SD, and absolute z score on 5- and 30-minute windows to reduce the influence of outliers	Thirty-four features were extracted from the signals. An RF model achieved 88% sensitivity and 0.93 AUROC in predicting mortality
University of Virginia Children's Hospital, Morgan Stanley Children's Hospital, and St Louis Children's Hospital	Sullivan et al [65], 2021	Participants; n=408, (266 used); data size: NR; model: diagnostic; outcome metric: sepsis	HR and SpO ₂ at 0.5 Hz	HR and SpO ₂ values of 0 were removed. Eight features were extracted in 10-minute windows and averaged hourly. Cross-correlation between HR and SpO ₂ was calculated in 10-minute windows of signals normalized to have 0 mean and SD of 1. Cross-correlation was done using the XCORR function of MATLAB with a lag time of -30 to +30 seconds	A logistic regression model using clinical and physiological features achieved an AUC ^{au} of 0.821 in predicting late-onset sepsis

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
St Louis Children's Hospital NICU	Feng et al [66], 2021	Participants; n=285; data size: ~80 hours each; model: prognostic (+6 hours); outcome metric: mortality	HR, respiration, SpO ₂ , and ART-M ^{av} or NIBP-M ^{aw} at 1 Hz	Infants with data >80 hours were truncated, and <80 hours were padded with 0s. Mean, median, mode, and Bayesian ridge data imputation techniques were explored. Bayesian ridge was used to sample 5 data sets by sampling different posteriors each time. Then, the average was reported using 4-fold cross-validation. The rolling mean of each vital sign with a range of 5 minutes was used to reduce noise. Finally, the end of each sample was padded with 1 segment where all features equaled 0. Features were extracted from 5-minute segments	A deep learning model using LSTM named DeepPBSMonitor was developed to predict mortality with 0.888 accuracy, 0.78 recall, and 0.897 AUC
University of Massachusetts Memorial Healthcare	Zuzarte et al [67], 2021	Participants; n=10; data size: 241.34 hours; model: prognostic (+310 seconds); outcome metric: apnea-bradycardia-hypoxia	ECG at 500 Hz, PPG ^{ax} at 125 Hz, SpO ₂ , HR, respiration signals from pneumogram at 50 Hz	PPG signals were filtered using a wavelet-based algorithm to remove gross body movements. A binary marker sampled at 25 Hz was obtained to indicate the presence or absence of movement. QRS complexes were detected using a modified Pan-Tompkins algorithm (modification NR). IBIs were detected using automated peak detection from LabChart Software RR intervals, and IBI values were then interpolated at 10 Hz	The prediction framework using GMM and logistic regression model achieved 75% accuracy in predicting bradycardia severity during the apnea-bradycardia-hypoxia event
University of Virginia NICU	Niestroy et al [68], 2022	Participants; n=5957; data size: random daily 10 minutes segments from each; model: prognostic (+1-7 days); outcome metric: mortality	HR and SpO ₂ at 0.5 Hz	No preprocessing was reported on the vitals. They were grouped to calculate the average in 10-minute nonoverlapping windows	Features were extracted from all signals. A multi-variable logistic regression model using 5 features achieved the AUROC of 0.83 in predicting mortality
University of Virginia Children's Hospital, Morgan Stanley Children's Hospital and St Louis Children's Hospital	Kausch et al [69], 2023	Participants; n=2494; data size: NR; model: prognostic (+24 hours); outcome metric: sepsis	HR and SpO ₂ at 0.5 Hz	HR and SpO ₂ were pre-processed by removing the values containing 0. Features were calculated in 10-minute nonoverlapping windows. Windows with >50% missing data were excluded from subsequent analysis	Several features were extracted from the vitals. An XGB model achieved training AUROC of 0.834 using the data from NICU 1, and 0.792 and 0.807 testing AUROC using data from NICU 2 and NICU 3, respectively

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
Karolinska University Hospital Solna and Huddinge NICU, Stockholm, Sweden	Honoré et al [70], 2023	Participants; n=325; data size: 2866 hospitalization days; model: prognostic (+24 hours); outcome metric: sepsis	IBI from ECG, respiration from CI, SpO ₂ (sampling rate 1-500 Hz)	All signals were resampled to 1 Hz. Segments with at most 15 seconds missing were linearly interpolated. All signals were filtered with a moving mean filter of width 3. IBI signals were further filtered to remove ectopic beats and strong nonlinearities with a moving median filter of width 3 and Butterworth bandpass filter of order 6 with low-cut and high-cut frequencies of 0.0021 and 0.43 Hz. Signals were divided into 45-minute segments. Features were calculated using a sliding time frame with 50% overlap	A naive bayes classifier achieved an AUROC of 0.82 up to 24 hours before clinical suspicion of sepsis. Adding respiratory signals improved the performance compared with only using heart rate features
Simulated and real data (NICU name NR)	Masoudi et al [71], 2013	Participants; n=32; data size: 233 episodes, ~7 seconds each; model: diagnostic (+2.32-second delay); outcome metric: apnea-bradycardia	2-channel ECG	No preprocessing techniques were reported. Signals were sampled in 7-second intervals	A coupled HMM model achieved 84.92% sensitivity, 94.17% specificity with a time detection delay of 2.32 (SD 4.82) seconds in detection apnea-bradycardia episodes
Simulated and real signals from NICU (NICU name NR)	Altuve et al [72], 2015	Participants; n=32; data size: 148 RR intervals with a mean duration of 26.25 (SD 11.37) minutes; model: diagnostic (+1.73-second delay); Outcome metric: apnea-bradycardia	ECG (sampling rate NR)	Hidden semi-Markov models to represent the temporal evolution of RR intervals. A preprocessing method that includes quantization and a delayed version of the observation vector is proposed. RR time series was resampled at 10 Hz and segmented at a 7-second interval	The proposed model achieved up to 93.84 (SD 0.79) in specificity and 89.66 (SD 0.71) in sensitivity with a detection delay of 1.59 (SD 0.24) seconds

Data set used	Author, year	Study settings	Physiological signal analyzed	Signal processing and computational techniques	Key findings
NICU (name NR)	Honoré et al [73], 2020	Participants; n=22; data size: 3501 time series, 1200 samples in each; model: prognostic (+72 hours); outcome metric: sepsis	SpO ₂ , respiratory frequency, and RR interval from ECG at 1 Hz	Data were segmented into 20-minute time frames. Time frames with missing data were discarded	Features were extracted from all signals. A combined GMM-HMM model achieved 0.74% (SD 0.05%) accuracy in detecting sepsis. The model was compared with HeRO model, which underperformed using this data set

^aCHIME: Collaborative Home Infant Monitoring Evaluation.^bNR: not reported.^cECG: electrocardiogram.^dSpO₂: oxygen saturation.^ePICS: Preterm Infant Cardio-Respiratory Signals.^fAUROC: area under receiver operating characteristic curve.^gFPR: false positive rate.^hHRV: heart rate variability.ⁱMIMIC-III: Medical Information Mart for Intensive Care.^jHR: heart rate.^kSBP: systolic blood pressure.^lDBP: diastolic blood pressure.^mMBP: mean blood pressure.ⁿLSTM: convolutional neural network-Long Short-Term Memory Network.^oLOS: length of stay.^pKNN: k-nearest neighbor.^qRMSE: root mean square error.^rSG: Savitzky-Golay.^sCHMM: coupled Hidden Markov Model.^tRNN: recurrent neural network.^uHMM: Hidden Markov Model.^vEEG: electroencephalography.^wSVM: support vector machine.^xHIE: hypoxic-ischemic encephalopathy.^yMAP: mean arterial pressure.^zNICU: neonatal intensive care.^{aa}CI: chest impedance.^{ab}DT-CWT: Discrete Time Continuous Wavelet Transform.^{ac}IBI: interbreath variable.^{ad}SII: Signal Instability Index.^{ae}RRE: ribcage respiratory effort.^{af}PR: pulse oximeter.^{ag}HSLDS: Hierarchical Switching Linear Dynamical System.^{ah}RF: random forest.^{ai}GMM: Gaussian Mixture Model.^{aj}FIR: Finite impulse response.^{ak}LMS: least-mean-square.^{al}RIP: respiratory inductive plethysmograph.^{am}ML: machine learning.^{an}IVH: intraventricular hemorrhage.^{ao}BPD: bronchopulmonary dysplasia^{ap}ROP: retinopathy of prematurity.^{aq}NEC: necrotizing enterocolitis

^{ar}POPS: pulse oximetry predictive score.

^{as}HRC: heart rate characteristics.

^{at}HeRO: heart rate observation.

^{au}AUC: area under the curve.

^{av}ART-M: arterial mean blood pressure.

^{aw}NIBP-M: noninvasive blood pressure.

^{ax}PPG: photoplethysmography.

Preprocessing Steps

Overview

Preprocessing of physiological data typically involves several steps, including the handling of missing data, filtering,

segmentation, and waveform analysis for feature extraction. Here, we define 5 required preprocessing steps (based on the steps outlined in Berkaya et al [13]) and identify the steps reported by each of the studies in this review (Table 3). The definition of each of the steps is given in subsequent sections.

Table 3. Required physiological signal preprocessing steps reported by each of the studies in this review.

Author, year	Required preprocessing step reported				
	Handling of missing data	Artifact removal	Resampling, normalization	Waveform feature extraction	Data segmentation
Cohen and de Chazal [22], 2013	✓	✓		✓	✓
Cohen and de Chazal [23], 2014	✓	✓	✓	✓	✓
Cohen and de Chazal [24], 2015	✓	✓		✓	✓
Gee et al [26], 2016				✓	✓
Gee et al [25], 2017		✓		✓	✓
Das et al [27], 2019		✓		✓	✓
Mahmud [28], 2019					
Gee et al [29], 2019		✓	✓	✓	✓
Song et al [30], 2020	✓	✓		N/A ^a	
Baker et al [31], 2021		✓	✓	N/A	✓
Juraev et al [32], 2022	✓		✓	N/A	
Montazeri Ghahjaverestan et al [33], 2015		✓	✓	✓	
Navarro et al [34], 2015		✓	✓	✓	✓
Montazeri Ghahjaverestan et al [35], 2016		✓	✓	✓	✓
León et al [36], 2021				✓	✓
León et al [37], 2021				✓	✓
Doyen et al [38], 2021				✓	✓
Sadoughi et al [39], 2021		✓	✓	✓	
Ahmed et al [40], 2015	✓	✓	✓	✓	✓
Temko et al [41], 2015		✓	✓	✓	✓
Lloyd et al [42], 2016		✓		N/A	✓
Semenova et al [43], 2018		✓		✓	✓
Semenova et al [44], 2019		✓	✓	✓	✓
Joshi et al [45], 2020		✓		✓	✓
Varisco et al [46], 2021		✓	✓	✓	
Cabrera-Quiros et al [47], 2021				✓	✓
Varisco et al [48], 2022		✓	✓	✓	✓
Peng et al [49], 2022				✓	✓
Peng et al [50], 2022		✓		✓	✓
Stanculescu et al [51], 2014	✓	✓		N/A	
Stanculescu et al [52], 2014	✓	✓		N/A	
Shirwaikar et al [53], 2016	✓	✓	✓	N/A	✓
Shirwaikar et al [54], 2019	✓			N/A	✓
Williamson et al [55], 2013			✓		✓
Schiavenato et al [56], 2013	✓	✓		✓	
Robles-Rubio et al [57], 2014		✓	✓	N/A	
Amperayani et al [58], 2017					
Hu et al [59], 2018	✓	✓	✓	N/A	✓
Sullivan et al [60], 2018	✓			N/A	✓

Author, year	Required preprocessing step reported				
	Handling of missing data	Artifact removal	Resampling, normalization	Waveform feature extraction	Data segmentation
Zimmet et al [61], 2020	✓			N/A	✓
Kota et al [62], 2020		✓		✓	✓
Mirnia et al [63], 2021				N/A	
Lee et al [64], 2021	✓	✓	✓	N/A	✓
Sullivan et al [65], 2021	✓	✓		N/A	✓
Feng et al [66], 2021	✓	✓	✓	N/A	✓
Zuzarte et al [67], 2021		✓	✓	✓	✓
Niestroy et al [68], 2022				N/A	✓
Kausch et al [69], 2023	✓	✓		N/A	✓
Honoré et al [70], 2023	✓	✓	✓	N/A	✓
Masoudi et al [71], 2013					✓
Altuve et al [72], 2015		✓	✓		✓
Honoré et al [73], 2020	✓			N/A	✓

^aN/A: Not applicable.

Handling of Missing Data

During neonatal physiological monitoring, instances of missing data may arise due to sensor disconnection, improper placements, or signal dropouts. To tackle this issue, methodologies like data imputation or interpolation are applied. For example, if gaps exist in a neonate's HR monitoring data, interpolation methods can estimate the missing values by considering neighboring data points. Widely used interpolation techniques include linear interpolation, spline interpolation, and time-based interpolation. In addition, common data imputation methods involve forward fill, backward fill, and imputation using mean or median values. Methods such as forward fill [30], moving average [44], mean imputation [64,66], and interpolation [67] were used by some studies reviewed in this paper.

Artifact Removal

Neonatal signals can be affected by artifacts, such as those from muscle movements or electrical interference. Commonly used techniques, such as bandpass or notch filters, along with moving averages, are used to effectively eliminate these disturbances. For instance, in neonatal EEG signals, adaptive filters prove beneficial in eliminating artifacts caused by muscle movements, resulting in a clearer representation of the baby's brain activity. Some methods used by the reviewed papers were high-pass filter [27,46] bandpass filter [29,33,44,45,56].

Resampling and Normalization

Overview

Resampling is a technique that standardizes data intervals, involving either upsampling (increasing data point frequency) or downsampling (decreasing frequency) to create a regular time series. This aligns signals from different devices or physiological sources. Normalization ensures uniformity and reliability across these standardized sampling rates. For instance, if neonatal HR signals from different devices have varied

sampling rates, resampling achieves a common rate, while normalization, using techniques such as minimum-maximum, z score, or log scale, ensures consistent amplitude scaling for accurate comparative analysis. In the reviewed studies, normalization techniques such as minimum-maximum [53] and 0 mean normalization [29,59] were used. In terms of resampling, both downsampling [33,34,41] and upsampling [39] techniques were used.

Waveform Feature Extraction

Extracting relevant features from a signal's waveform is a fundamental step in signal preprocessing. This involves identifying key characteristics such as peaks, troughs, or other significant points in the signal. In the context of neonatal ECG, feature extraction may involve identifying key points such as R-peaks to analyze HRV, providing valuable insights into the infant's autonomic nervous system development. The Pan-Tompkins algorithm is a popular method chosen by multiple papers reviewed in this study that conducted R-peak detection from the QRS complex [22,24,27,33,35,39].

Data Segmentation

Segmenting data is the process of breaking down a continuous signal into smaller, more manageable sections to enable targeted analysis. This practice is especially beneficial when dealing with lengthy signals. Data segmentation is a common preprocessing step in ML workflows. For instance, in the analysis of neonatal sleep patterns using EEG, data segmentation can involve dividing the continuous EEG signal into epochs, allowing for the identification and study of sleep stages in shorter, more manageable segments. Commonly used segmentation techniques include fixed length, sliding window, and threshold- and feature-based segmentation. Some of the data segmentation sizes used in the reviewed studies were 30-second [22-24,45] and 1-minute [41] epochs and a sliding window of varied sizes [35,40,55,59,64].

In neonatal physiological signal processing, these preprocessing techniques contribute to the accurate interpretation of signals, aiding health care professionals in monitoring and providing appropriate care in the NICU or other clinical settings.

It can be seen from [Table 3](#) that only 7 (13%) out of the 52 reviewed studies reported all the recommended preprocessing steps. This could have several impacts on the downstream analysis. For instance, several papers missed reporting on how they segmented the data for feature extraction and classification, although it is essential for clinical validation in cases where the segment duration is dependent on the adverse outcome prediction performance. In HRV analysis, it is important to indicate whether it is a short-term (~5 minutes) or a long-term (≥ 24 hours) analysis as they reflect different underlying physiological processes and thus demonstrate different predictive power [107]. Along with the segment duration, additional information such as the sampling rate of the signals will provide a clear reflection of the data set size. Downsampling the data to a low sampling rate (eg, 50 Hz) has also shown a significant impact on HRV analysis [108]. Although all the reviewed studies mentioned the participant number, and majority of them (n=39) reported the sampling rate of the signals, very few provided details on the sample size or data set duration or whether the data set was resampled for subsequent analysis. These elements provide a clearer picture of the computational time and resources required for clinical validation and adoption. Although physiological recordings collected in the NICU environment suffer greatly from missing data due to similar factors that introduce artifacts [109], reporting how missing data are handled is scarce. Different methods for dealing with missing values could cause different results, and not all might be suitable for a particular problem. Therefore, it is important to report all the details related to the adopted approach.

The incomplete or partial reporting found in these studies has significant implications for the implementation of QMS in using these techniques for clinical adoption. A good implementation of QMS requires a comprehensive reporting of each intermediary step involved in constructing an AI and ML pipeline. The International Medical Device Regulators Forum offers guidance on the clinical evaluation required for any product intended for use as a medical device [110]. According to the International Medical Device Regulators Forum guidelines, during clinical evaluation, relevant research articles are reviewed to identify clinical evidence supporting the product [111]. The guideline encourages manufacturers to follow these recognized standards and best practices in the development, validation, and manufacturing processes. Clinical evaluations are required by the European Union medical device regulation, and it is also mentioned in the ISO 13485 (the quality management standard for medical devices). Thus, detailed reporting is crucial as it can be used by regulatory bodies to evaluate future SaMD products clinically. Steps such as the missing data handling procedures are also required by the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) checklist for model development and validation, which assesses the risk of bias and clinical usefulness of the prediction model [112]. Another example is a questionnaire prepared by the German

Notified Body Interest Group, and it was adopted to assess some AI-powered medical products in the European Union. This questionnaire includes inquiries about data management, including data collection, labeling, preprocessing procedures, and relevant documentation. Transparent and detailed reporting of these steps is essential to ensure the safety, efficacy, and reliability of SaMD.

Discussion

Principal Findings

This review aimed to summarize the computational methods used for preprocessing preterm infants' physiological data as a first step in developing data-driven predictive models for adverse outcomes related to clinical decision support. This is an important step, especially from a clinician's perspective, because it increases the trustworthiness of the developed models by allowing for the verification and reproduction of the results. In addition, it aids in achieving regulatory compliance and ensures the safety, efficacy, and ethical use of AI-based health care devices. Furthermore, it allows us to recognize the shortcomings in the current state-of-the-art studies and recommend guidelines for transparent reporting. The review found that the studies were heterogeneous in terms of their methods and applications. Therefore, a narrative approach to reporting the results was taken instead of a quantitative approach. Through the analysis we identified several key components that were incomplete or partially reported by the included studies, which are summarized in [Table 3](#). To ensure transparent reporting for any future studies in this area, we recommend detailed reporting of all preprocessing steps listed in [Table 3](#), which will allow revealing their strengths and weaknesses and ultimately make them usable and reproducible. Reproducible research allows clinicians to make more informed decisions about patient care and treatment based on the evidence that has been thoroughly assessed.

Comparison With Prior Work

The reviews published in recent years have highlighted the potential of big data and AI in supporting clinical decision-making in the neonatal health care domain [10,15,21,113,114], particularly in using physiological data for detecting or predicting neonatal health outcomes. However, appropriate preprocessing of these data is a prerequisite for developing clinically deployable models. A systematic review by McAdams et al [10] reported different ML models used to predict different clinical outcomes in neonates. However, their primary focus was on 5 neonatal morbidities, and they did not focus on reporting the preprocessing methods applied before building the ML models. Furthermore, they did not include studies using real-time continuous physiological data; 28 out of their 68 studies were based on physiological data (not continuous), and the rest were based on electronic medical records and imaging data. Latremouille et al [15] performed a review on HRV analysis for neonates. The primary limitation of the work was the lack of reporting in detail about the preprocessing steps of ECG signals before HRV analysis, such as ECG handling and segmentation, R-wave (QRS complex) identification technique, software and parameters, and ranges of all HRV features. They identified these components as

incomplete or missing in the studies they reviewed and thus recommended clear reporting of these aspects for future studies in this area. These limitations served as a motivation for our review to focus on the preprocessing techniques of neonatal physiological signals in a broader sense, which serves as the preliminary step for any big data-based approaches.

Limitations

There are several limitations to this review. Screening of all the included studies was conducted independently by 1 reviewer, which may have introduced bias. In addition, this review did not include a quantitative or comparative analysis of the reviewed studies, as the techniques used to analyze the physiological signals were diverse. Future work could include a quantitative evaluation of the studies that were homogeneous in design.

Conclusions

This review explores the computational methods used by the current state-of-the-art ML-driven clinical decision support approaches to preprocess physiological signals collected from infants treated in the neonatal setting. A summary of the studies identified heterogeneity in the techniques used for analysis and revealed a lack of consistent and detailed reporting, which is important for building robust, transparent, and clinically deployable prediction models. The availability of powerful hardware and software resources in the NICU environment and growing interest in big data and AI are driving strong demand for clinical decision support applications. We recommend clear reporting of the different steps in the preprocessing of the neonatal physiological signals to ensure transparency in clinical validation and accelerate the adoption of developed models in the clinical setting. This will further enhance the delivery and adoption of reliable, regulatory-compliant, safe, and effective products in health care.

Acknowledgments

The work is supported by the Cerebral Palsy Alliance and the University of Sydney (ethics approval number 2021/ETH00824).

Data Availability

The data sets generated during and analyzed during this study are available from the corresponding author on reasonable request or are available in the Multimedia Appendices.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed search queries.

[[DOCX File , 18 KB-Multimedia Appendix 1](#)]

Multimedia Appendix 2

Bibliography files for all databases.

[[ZIP File \(Zip Archive\), 2585 KB-Multimedia Appendix 2](#)]

Multimedia Appendix 3

All included papers.

[[ZIP File \(Zip Archive\), 61626 KB-Multimedia Appendix 3](#)]

Multimedia Appendix 4

Metadata of all papers.

[[XLSX File \(Microsoft Excel File\), 2654 KB-Multimedia Appendix 4](#)]

Multimedia Appendix 5

Metadata of all papers in the full-text review.

[[XLSX File \(Microsoft Excel File\), 78 KB-Multimedia Appendix 5](#)]

Multimedia Appendix 6

PRISMA-ScR checklist.

[[DOCX File , 84 KB-Multimedia Appendix 6](#)]

References

1. Platt MJ. Outcomes in preterm infants. *Public Health*. May 2014;128(5):399-403. [[FREE Full text](#)] [doi: [10.1016/j.puhe.2014.03.010](https://doi.org/10.1016/j.puhe.2014.03.010)] [Medline: [24794180](#)]
2. Seaton SE, Barker L, Draper ES, Abrams KR, Modi N, Manktelow BN. Estimating neonatal length of stay for babies born very preterm. *Arch Dis Child Fetal Neonatal Ed*. Mar 2019;104(2):F182-F186. [[FREE Full text](#)] [doi: [10.1136/archdischild-2017-314405](https://doi.org/10.1136/archdischild-2017-314405)] [Medline: [29588296](#)]
3. De Georgia MA, Kaffashi F, Jacono FJ, Loparo KA. Information technology in critical care: review of monitoring and data acquisition systems for patient care and research. *ScientificWorldJournal*. 2015;2015:727694. [[FREE Full text](#)] [doi: [10.1155/2015/727694](https://doi.org/10.1155/2015/727694)] [Medline: [25734185](#)]
4. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ. Big data for health. *IEEE J Biomed Health Inform*. Jul 2015;19(4):1193-1208. [[FREE Full text](#)] [doi: [10.1109/JBHI.2015.2450362](https://doi.org/10.1109/JBHI.2015.2450362)] [Medline: [26173222](#)]
5. Kumar N, Akangire G, Sullivan B, Fairchild K, Sampath V. Continuous vital sign analysis for predicting and preventing neonatal diseases in the twenty-first century: big data to the forefront. *Pediatr Res*. Jan 2020;87(2):210-220. [[FREE Full text](#)] [doi: [10.1038/s41390-019-0527-0](https://doi.org/10.1038/s41390-019-0527-0)] [Medline: [31377752](#)]
6. Baljak V, Ljubovic A, Michel J, Montgomery M, Salaway R. A scalable realtime analytics pipeline and storage architecture for physiological monitoring big data. *Smart Health*. Dec 2018;9-10:275-286. [[FREE Full text](#)] [doi: [10.1016/j.smhl.2018.07.013](https://doi.org/10.1016/j.smhl.2018.07.013)]
7. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Effect of clinical decision-support systems: a systematic review. *Ann Intern Med*. Jul 03, 2012;157(1):29-43. [[FREE Full text](#)] [doi: [10.7326/0003-4819-157-1-201207030-00450](https://doi.org/10.7326/0003-4819-157-1-201207030-00450)] [Medline: [22751758](#)]
8. Comito C, Falcone D, Forestiero A. Current trends and practices in smart health monitoring and clinical decision support. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. 2020. Presented at: BIBM 2020; December 16-19, 2020; Seoul, Korea (South). [doi: [10.1109/BIBM49941.2020.9313449](https://doi.org/10.1109/BIBM49941.2020.9313449)]
9. Yu KH, Kohane IS. Framing the challenges of artificial intelligence in medicine. *BMJ Qual Saf*. Mar 2019;28(3):238-241. [doi: [10.1136/bmjqqs-2018-008551](https://doi.org/10.1136/bmjqqs-2018-008551)] [Medline: [30291179](#)]
10. 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*. Dec 2022;42(12):1561-1575. [doi: [10.1038/s41372-022-01392-8](https://doi.org/10.1038/s41372-022-01392-8)] [Medline: [35562414](#)]
11. Mahadevaiah G, Rv P, Bermejo I, Jaffray D, Dekker A, Wee L. Artificial intelligence-based clinical decision support in modern medical physics: selection, acceptance, commissioning, and quality assurance. *Med Phys*. Jun 2020;47(5):e228-e235. [[FREE Full text](#)] [doi: [10.1002/mp.13562](https://doi.org/10.1002/mp.13562)] [Medline: [32418341](#)]
12. Sloane EJ, Silva RJ. Chapter 83 - Artificial intelligence in medical devices and clinical decision support systems. In: Iadanza E, editor. *Clinical Engineering Handbook*. Cambridge, MA. Academic Press; 2019:556-568.
13. Kaplan Berkaya S, Uysal AK, Sora Gunal E, Ergin S, Gunal S, Gulmezoglu MB. A survey on ECG analysis. *Biomed Signal Process Control*. May 2018;43:216-235. [[FREE Full text](#)] [doi: [10.1016/j.bspc.2018.03.003](https://doi.org/10.1016/j.bspc.2018.03.003)]
14. Sweeney KT, Ward TE, McLoone SF. Artifact removal in physiological signals—practices and possibilities. *IEEE Trans Inform Technol Biomed*. May 2012;16(3):488-500. [[FREE Full text](#)] [doi: [10.1109/titb.2012.2188536](https://doi.org/10.1109/titb.2012.2188536)]
15. Latremouille S, Lam J, Shalish W, Sant'Anna G. Neonatal heart rate variability: a contemporary scoping review of analysis methods and clinical applications. *BMJ Open*. Dec 21, 2021;11(12):e055209. [[FREE Full text](#)] [doi: [10.1136/bmjopen-2021-055209](https://doi.org/10.1136/bmjopen-2021-055209)] [Medline: [34933863](#)]
16. Baird P, Cobbaert K. Software as a medical device: a comparison of the EU's approach with the US's approach. *British Standards Institution*. URL: <https://www.bsigroup.com/globalassets/meddev/localfiles/fr-fr/whitepapers/md-wp-software.pdf> [accessed 2024-07-24]
17. Simera I. Get the content right: following reporting guidelines will make your research paper more complete, transparent and usable. *J Pak Med Assoc*. Feb 2013;63(2):283-285. [Medline: [23894916](#)]
18. 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*. Mar 29, 2021;372:n71. [[FREE Full text](#)] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](#)]
19. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. York, UK. Centre for Reviews and Dissemination; 2009.
20. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev*. Dec 05, 2016;5(1):210. [[FREE Full text](#)] [doi: [10.1186/s13643-016-0384-4](https://doi.org/10.1186/s13643-016-0384-4)] [Medline: [27919275](#)]
21. Mangold C, Zoretic S, Thallapureddy K, Moreira A, Chorath K, Moreira A. Machine learning models for predicting neonatal mortality: a systematic review. *Neonatology*. 2021;118(4):394-405. [[FREE Full text](#)] [doi: [10.1159/000516891](https://doi.org/10.1159/000516891)] [Medline: [34261070](#)]
22. Cohen G, de Chazal P. Detection of sleep apnoea in infants using ECG and oximetry signals. In: Proceedings of the Computing in Cardiology 2013. 2013. Presented at: CinC 2013; September 22-25, 2013; Zaragoza, Spain.

23. Cohen G, de Chazal P. A multi-modal approach to sleep-wake classification in infants using minimally invasive sensors. In: Proceedings of the Computing in Cardiology. 2014. Presented at: CinC 2014; September 7-10, 2014; Cambridge, MA.
24. Cohen G, de Chazal P. Automated detection of sleep apnea in infants: a multi-modal approach. Comput Biol Med. Aug 2015;63:118-123. [FREE Full text] [doi: [10.1016/j.combiomed.2015.05.007](https://doi.org/10.1016/j.combiomed.2015.05.007)] [Medline: [26073098](https://pubmed.ncbi.nlm.nih.gov/26073098/)]
25. Gee AH, Barbieri R, Paydarfar D, Indic P. Predicting bradycardia in preterm infants using point process analysis of heart rate. IEEE Trans Biomed Eng. Sep 2017;64(9):2300-2308. [FREE Full text] [doi: [10.1109/tbme.2016.2632746](https://doi.org/10.1109/tbme.2016.2632746)]
26. Gee AH, Barbieri R, Paydarfar D, Indic P. Improving heart rate estimation in preterm infants with bivariate point process analysis of heart rate and respiration. In: Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016. Presented at: EMBC 2016; August 16-20, 2016; Orlando, FL. [doi: [10.1109/embc.2016.7590851](https://doi.org/10.1109/embc.2016.7590851)]
27. Das S, Moraffah B, Banerjee A, Gupta SK, Papandreou-Suppappola A. Bradycardia prediction in preterm infants using nonparametric kernel density estimation. In: Proceedings of the 53rd Asilomar Conference on Signals, Systems, and Computers. 2019. Presented at: ACSSC 2019; November 3-6, 2019; Pacific Grove, CA. [doi: [10.1109/ieeeconf44664.2019.9049007](https://doi.org/10.1109/ieeeconf44664.2019.9049007)]
28. Mahmud MS, Wang H, Kim Y. Accelerated prediction of bradycardia in preterm infants using time-frequency analysis. In: Proceedings of the International Conference on Computing, Networking and Communications. 2019. Presented at: ICNC 2019; February 18-21, 2019; Honolulu, HI. [doi: [10.2172/10116953](https://doi.org/10.2172/10116953)]
29. Gee AH, Garcia-Olano D, Ghosh J, Paydarfar D. Explaining deep classification of time-series data with learned prototypes. CEUR Workshop Proc. Aug 2019;2429:15-22. [FREE Full text] [Medline: [33867901](https://pubmed.ncbi.nlm.nih.gov/33867901/)]
30. Song W, Jung SY, Baek H, Choi CW, Jung YH, Yoo S. A predictive model based on machine learning for the early detection of late-onset neonatal sepsis: development and observational study. JMIR Med Inform. Jul 31, 2020;8(7):e15965. [FREE Full text] [doi: [10.2196/15965](https://doi.org/10.2196/15965)] [Medline: [32735230](https://pubmed.ncbi.nlm.nih.gov/32735230/)]
31. Baker S, Xiang W, Atkinson I. Hybridized neural networks for non-invasive and continuous mortality risk assessment in neonates. Comput Biol Med. Jul 2021;134:104521. [FREE Full text] [doi: [10.1016/j.combiomed.2021.104521](https://doi.org/10.1016/j.combiomed.2021.104521)] [Medline: [34111664](https://pubmed.ncbi.nlm.nih.gov/34111664/)]
32. Juraev F, El-Sappagh S, Abdulkhamidov E, Ali F, Abuhmed T. Multilayer dynamic ensemble model for intensive care unit mortality prediction of neonate patients. J Biomed Inform. Nov 2022;135:104216. [FREE Full text] [doi: [10.1016/j.jbi.2022.104216](https://doi.org/10.1016/j.jbi.2022.104216)] [Medline: [36208833](https://pubmed.ncbi.nlm.nih.gov/36208833/)]
33. Montazeri Ghahjaverestan N, Shamsollahi MB, Ge D, Hernández AI. Switching Kalman filter based methods for apnea bradycardia detection from ECG signals. Physiol Meas. Sep 2015;36(9):1763-1783. [FREE Full text] [doi: [10.1088/0967-3334/36/9/1763](https://doi.org/10.1088/0967-3334/36/9/1763)] [Medline: [26235524](https://pubmed.ncbi.nlm.nih.gov/26235524/)]
34. Navarro X, Porée F, Beuchée A, Carrault G. Artifact rejection and cycle detection in immature breathing: application to the early detection of neonatal sepsis. Biomed Signal Processing Control. Feb 2015;16:9-16. [FREE Full text] [doi: [10.1016/j.bspc.2014.10.007](https://doi.org/10.1016/j.bspc.2014.10.007)]
35. Montazeri Ghahjaverestan N, Masoudi S, Shamsollahi MB, Beuchée A, Pladys P, Ge D, et al. Coupled hidden Markov model-based method for apnea bradycardia detection. IEEE J Biomed Health Inform. Mar 2016;20(2):527-538. [FREE Full text] [doi: [10.1109/JBHI.2015.2405075](https://doi.org/10.1109/JBHI.2015.2405075)] [Medline: [25706937](https://pubmed.ncbi.nlm.nih.gov/25706937/)]
36. Leon C, Carrault G, Pladys P, Beuchée A. Early detection of late onset sepsis in premature infants using visibility graph analysis of heart rate variability. IEEE J Biomed Health Inform. Apr 2021;25(4):1006-1017. [FREE Full text] [doi: [10.1109/JBHI.2020.3021662](https://doi.org/10.1109/JBHI.2020.3021662)] [Medline: [32881699](https://pubmed.ncbi.nlm.nih.gov/32881699/)]
37. León C, Pladys P, Beuchée A, Carrault G. Recurrent neural networks for early detection of late onset sepsis in premature infants using heart rate variability. In: Proceedings of the Computing in Cardiology. 2021. Presented at: CinC 2021; September 13-15, 2021; Brno, Czech Republic. [doi: [10.23919/cinc53138.2021.9662715](https://doi.org/10.23919/cinc53138.2021.9662715)]
38. Doyen M, Hernández AI, Flamant C, Defontaine A, Favrais G, Altuve M, et al. Early bradycardia detection and therapeutic interventions in preterm infant monitoring. Sci Rep. May 18, 2021;11(1):10486. [FREE Full text] [doi: [10.1038/s41598-021-89468-x](https://doi.org/10.1038/s41598-021-89468-x)] [Medline: [34006917](https://pubmed.ncbi.nlm.nih.gov/34006917/)]
39. Sadoughi A, Shamsollahi MB, Fatemizadeh E, Beuchée A, Hernández AI, Montazeri Ghahjaverestan N. Detection of apnea bradycardia from ECG signals of preterm infants using layered hidden Markov model. Ann Biomed Eng. Sep 2021;49(9):2159-2169. [FREE Full text] [doi: [10.1007/s10439-021-02732-z](https://doi.org/10.1007/s10439-021-02732-z)] [Medline: [33638031](https://pubmed.ncbi.nlm.nih.gov/33638031/)]
40. Ahmed R, Temko A, Marnane WP, Boylan G, Lightbody G. Classification of hypoxic-ischemic encephalopathy using long term heart rate variability based features. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2015. Presented at: EMBC 2015; August 25-29, 2015; Milan, Italy. [doi: [10.1109/embc.2015.7318866](https://doi.org/10.1109/embc.2015.7318866)]
41. Temko A, Doyle O, Murray D, Lightbody G, Boylan G, Marnane W. Multimodal predictor of neurodevelopmental outcome in newborns with hypoxic-ischaemic encephalopathy. Comput Biol Med. Aug 2015;63:169-177. [FREE Full text] [doi: [10.1016/j.combiomed.2015.05.017](https://doi.org/10.1016/j.combiomed.2015.05.017)] [Medline: [26093065](https://pubmed.ncbi.nlm.nih.gov/26093065/)]
42. Lloyd RO, O'Toole JM, Livingstone V, Hutch WD, Pavlidis E, Cronin AM, et al. Predicting 2-y outcome in preterm infants using early multimodal physiological monitoring. Pediatr Res. Sep 2016;80(3):382-388. [FREE Full text] [doi: [10.1038/pr.2016.92](https://doi.org/10.1038/pr.2016.92)] [Medline: [27089498](https://pubmed.ncbi.nlm.nih.gov/27089498/)]

43. Semenova O, Carra G, Lightbody G, Boylan G, Dempsey E, Temko A. Heart rate variability during periods of low blood pressure as a predictor of short-term outcome in preterms. In: Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2018. Presented at: EMBC 2018; July 18-21, 2018; Honolulu, HI. [doi: [10.1109/EMBC.2018.8513600](https://doi.org/10.1109/EMBC.2018.8513600)]

44. Semenova O, Carra G, Lightbody G, Boylan G, Dempsey E, Temko A. Prediction of short-term health outcomes in preterm neonates from heart-rate variability and blood pressure using boosted decision trees. Comput Methods Programs Biomed. Oct 2019;180:104996. [[FREE Full text](#)] [doi: [10.1016/j.cmpb.2019.104996](https://doi.org/10.1016/j.cmpb.2019.104996)] [Medline: [31421605](#)]

45. Joshi R, Kommers D, Oosterwijk L, Feijis L, van Pul C, Andriessen P. Predicting neonatal sepsis using features of heart rate variability, respiratory characteristics, and ECG-derived estimates of infant motion. IEEE J Biomed Health Inform. Mar 2020;24(3):681-692. [[FREE Full text](#)] [doi: [10.1109/JBHI.2019.2927463](https://doi.org/10.1109/JBHI.2019.2927463)] [Medline: [31295130](#)]

46. Varisco G, Kommers D, Long X, Zhan Z, Nano MM, Cottaar W, et al. Optimized detection of central apneas preceding late-onset sepsis in premature infants. In: Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 2021. Presented at: EMBC 2021; November 1-5, 2021; Virtual Event. [doi: [10.1109/EMBC46164.2021.9629528](https://doi.org/10.1109/EMBC46164.2021.9629528)]

47. Cabrera-Quiros L, Kommers D, Wolvers MK, Oosterwijk L, Arents N, van der Sluijs-Bens J, et al. Prediction of late-onset sepsis in preterm infants using monitoring signals and machine learning. Crit Care Explor. Jan 27, 2021;3(1):e0302. [[FREE Full text](#)] [doi: [10.1097/CCE.0000000000000302](https://doi.org/10.1097/CCE.0000000000000302)] [Medline: [33532727](#)]

48. Varisco G, Peng Z, Kommers D, Zhan Z, Cottaar W, Andriessen P, et al. Central apnea detection in premature infants using machine learning. Comput Methods Programs Biomed. Nov 2022;226:107155. [[FREE Full text](#)] [doi: [10.1016/j.cmpb.2022.107155](https://doi.org/10.1016/j.cmpb.2022.107155)] [Medline: [36215858](#)]

49. Peng Z, Varisco G, Liang RH, Kommers D, Cottaar W, Andriessen P, et al. DeepLOS: deep learning for late-onset sepsis prediction in preterm infants using heart rate variability. Smart Health. Dec 2022;26:100335. [[FREE Full text](#)] [doi: [10.1016/j.smhl.2022.100335](https://doi.org/10.1016/j.smhl.2022.100335)]

50. Peng Z, Varisco G, Long X, Liang RH, Kommers D, Cottaar W, et al. A continuous late-onset sepsis prediction algorithm for preterm infants using multi-channel physiological signals from a patient monitor. IEEE J Biomed Health Inform. Jan 2023;27(1):550-561. [doi: [10.1109/JBHI.2022.3216055](https://doi.org/10.1109/JBHI.2022.3216055)] [Medline: [36264730](#)]

51. Stanculescu I, Williams CK, Freer Y. Autoregressive hidden Markov models for the early detection of neonatal sepsis. IEEE J Biomed Health Inform. Sep 2014;18(5):1560-1570. [[FREE Full text](#)] [doi: [10.1109/JBHI.2013.2294692](https://doi.org/10.1109/JBHI.2013.2294692)] [Medline: [25192568](#)]

52. Stanculescu I, Williams CK, Freer Y. A hierarchical switching linear dynamical system applied to the detection of sepsis in neonatal condition monitoring. In: Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence. 2014. Presented at: UAI'14; July 23-27, 2014; Quebec City, Quebec.

53. Shirwaikar RD, Acharya UD, Makkithaya K, Surulivelr M, Lewis LE. Machine learning techniques for neonatal apnea prediction. J Artif Intell. 2016;9(1-3):33-38. [[FREE Full text](#)] [doi: [10.3923/jai.2016.33.38](https://doi.org/10.3923/jai.2016.33.38)]

54. Shirwaikar RD, Acharya UD, Makkithaya K, Srivastava S, Lewis UL, M S. Optimizing neural networks for medical data sets: a case study on neonatal apnea prediction. Artif Intell Med. Jul 2019;98:59-76. [[FREE Full text](#)] [doi: [10.1016/j.artmed.2019.07.008](https://doi.org/10.1016/j.artmed.2019.07.008)] [Medline: [31521253](#)]

55. Williamson JR, Bliss DW, Browne DW, Indic P, Bloch-Salisbury E, Paydarfar D. Individualized apnea prediction in preterm infants using cardio-respiratory and movement signals. In: Proceedings of the IEEE International Conference on Body Sensor Networks. 2013. Presented at: BSN 2013; May 6-9, 2013; Cambridge, MA. [doi: [10.1109/BSN.2013.6575523](https://doi.org/10.1109/BSN.2013.6575523)]

56. Schiavenato M, Oliu C, Bello E, Bohorquez J, Claure N. Development of a system for the assessment of heart rate variability in the NICU. In: Proceedings of the 29th Southern Biomedical Engineering Conference. 2013. Presented at: SBEC 2013; May 3-5, 2013; Miami, FL. [doi: [10.1109/SBEC.2013.21](https://doi.org/10.1109/SBEC.2013.21)]

57. Robles-Rubio CA, Brown KA, Bertolizio G, Kearney RE. Automated analysis of respiratory behavior for the prediction of apnea in infants following general anesthesia. In: Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2014. Presented at: EMBC 2014; August 26-30, 2014; Chicago, IL. [doi: [10.1109/EMBC.2014.6943579](https://doi.org/10.1109/EMBC.2014.6943579)]

58. Amperayani VA, Indic P, Travers CP, Barbieri R, Paydarfar D, Ambalavanan N. An algorithm for risk stratification of preterm infants. In: Proceedings of the Computing in Cardiology. 2017. Presented at: CinC 2017; September 24-27, 2017; Rennes, France. [doi: [10.22489/CinC.2017.333-127](https://doi.org/10.22489/CinC.2017.333-127)]

59. Hu Y, Lee VC, Tan K. Prediction of clinicians' treatment in preterm infants with suspected late-onset sepsis — an ML approach. In: Proceedings of the 13th IEEE Conference on Industrial Electronics and Applications. 2018. Presented at: ICIEA 2018; May 31-June 2, 2018; Wuhan, China. [doi: [10.1109/ICIEA.2018.8397888](https://doi.org/10.1109/ICIEA.2018.8397888)]

60. Sullivan BA, Wallman-Stokes A, Isler J, Sahni R, Moorman JR, Fairchild KD, et al. Early pulse oximetry data improves prediction of death and adverse outcomes in a two-center cohort of very low birth weight infants. Am J Perinatol. Nov 2018;35(13):1331-1338. [[FREE Full text](#)] [doi: [10.1055/s-0038-1654712](https://doi.org/10.1055/s-0038-1654712)] [Medline: [29807371](#)]

61. Zimmet AM, Sullivan BA, Moorman JR, Lake DE, Ratcliffe SJ. Trajectories of the heart rate characteristics index, a biomarker of sepsis in premature infants, predict Neonatal ICU mortality. JRSM Cardiovasc Dis. Aug 05, 2020;9:2048004020945142. [[FREE Full text](#)] [doi: [10.1177/2048004020945142](https://doi.org/10.1177/2048004020945142)] [Medline: [33240492](#)]

62. Kota S, du Plessis A, Massaro AN, Chang T, Al-Shargabi T, Govindan RB. A frequency based spatial filter to mitigate volume conduction in electroencephalogram signals. *Annu Int Conf IEEE Eng Med Biol Soc.* Aug 2016;2016:4001-4004. [doi: [10.1109/EMBC.2016.7591604](https://doi.org/10.1109/EMBC.2016.7591604)] [Medline: [28269162](#)]
63. Mirmia K, Heidarzadeh M, Afjeh SA, Alizadeh P, Kashan AA, Bordbar A, et al. Signal processing of heart rate for predicting sepsis in premature neonates. *J Med Signals Sens.* Jul 21, 2021;11(3):222-226. [[FREE Full text](#)] [doi: [10.4103/jmss.JMSS_30_20](https://doi.org/10.4103/jmss.JMSS_30_20)] [Medline: [34466402](#)]
64. Lee J, Cai J, Li F, Vesoulis ZA. Predicting mortality risk for preterm infants using random forest. *Sci Rep.* Mar 31, 2021;11(1):7308. [[FREE Full text](#)] [doi: [10.1038/s41598-021-86748-4](https://doi.org/10.1038/s41598-021-86748-4)] [Medline: [33790395](#)]
65. Sullivan BA, Nagraj VP, Berry KL, Fleiss N, Rambhia A, Kumar R, et al. Clinical and vital sign changes associated with late-onset sepsis in very low birth weight infants at 3 NICUs. *J Neonatal Perinatal Med.* 2021;14(4):553-561. [[FREE Full text](#)] [doi: [10.3233/NPM-200578](https://doi.org/10.3233/NPM-200578)] [Medline: [33523025](#)]
66. Feng J, Lee J, Vesoulis ZA, Li F. Predicting mortality risk for preterm infants using deep learning models with time-series vital sign data. *NPJ Digit Med.* Jul 14, 2021;4(1):108. [[FREE Full text](#)] [doi: [10.1038/s41746-021-00479-4](https://doi.org/10.1038/s41746-021-00479-4)] [Medline: [34262112](#)]
67. Zuzarte I, Sternad D, Paydarfar D. Predicting apneic events in preterm infants using cardio-respiratory and movement features. *Comput Methods Programs Biomed.* Sep 2021;209:106321. [[FREE Full text](#)] [doi: [10.1016/j.cmpb.2021.106321](https://doi.org/10.1016/j.cmpb.2021.106321)] [Medline: [34380078](#)]
68. Niestroy JC, Moorman JR, Levinson MA, Manir SA, Clark TW, Fairchild KD, et al. Discovery of signatures of fatal neonatal illness in vital signs using highly comparative time-series analysis. *NPJ Digit Med.* Jan 17, 2022;5(1):6. [[FREE Full text](#)] [doi: [10.1038/s41746-021-00551-z](https://doi.org/10.1038/s41746-021-00551-z)] [Medline: [35039624](#)]
69. Kausch SL, Brandberg JG, Qiu J, Panda A, Binai A, Isler J, et al. Cardiorespiratory signature of neonatal sepsis: development and validation of prediction models in 3 NICUs. *Pediatr Res.* Jun 2023;93(7):1913-1921. [[FREE Full text](#)] [doi: [10.1038/s41390-022-02444-7](https://doi.org/10.1038/s41390-022-02444-7)] [Medline: [36593281](#)]
70. Honoré A, Forsberg D, Adolphson K, Chatterjee S, Jost K, Herlenius E. Vital sign-based detection of sepsis in neonates using machine learning. *Acta Paediatr.* Apr 2023;112(4):686-696. [[FREE Full text](#)] [doi: [10.1111/apa.16660](https://doi.org/10.1111/apa.16660)] [Medline: [36607251](#)]
71. Masoudi S, Montazeri N, Shamsollahi MB, Ge D, Beuchée A, Pladys P, et al. Early detection of apnea-bradycardia episodes in preterm infants based on coupled hidden Markov model. In: Proceedings of the IEEE International Symposium on Signal Processing and Information Technology. 2013. Presented at: ISSPIT 2013; December 12-15, 2013; Athens, Greece. [doi: [10.1109/ISSPIT.2013.6781887](https://doi.org/10.1109/ISSPIT.2013.6781887)]
72. Altuve M, Carrault G, Beuchée A, Pladys P, Hernández AI. Online apnea-bradycardia detection based on hidden semi-Markov models. *Med Biol Eng Comput.* Jan 2015;53(1):1-13. [[FREE Full text](#)] [doi: [10.1007/s11517-014-1207-1](https://doi.org/10.1007/s11517-014-1207-1)] [Medline: [25300402](#)]
73. Honoré A, Liu D, Forsberg D, Coste K, Herlenius E, Chatterjee S, et al. Hidden Markov models for sepsis detection in preterm infants. In: Proceedings of the ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. 2020. Presented at: ICASSP 2020; May 4-8, 2020; Barcelona, Spain. [doi: [10.1109/icassp40776.2020.9054635](https://doi.org/10.1109/icassp40776.2020.9054635)]
74. Continuous patient monitors. Phillips Healthcare. URL: <https://www.philips.com.au/healthcare/solutions/patient-monitoring/continuous-patient-monitoring-systems> [accessed 2024-07-21]
75. BedMasterEx. Anandic. URL: <https://www.anandic.com/produkte/produkt-detail/bedmasterex/> [accessed 2024-07-21]
76. NicoletOne™ EEG system. Natus. URL: <https://natus.com/neuro/nicoletone-eeg-system/> [accessed 2024-07-21]
77. Patient Information Center iX (PIC iX). Philips. URL: <https://www.philips.com.au/healthcare/product/HCNOCTN171/patient-information-center-ix-pic-ix> [accessed 2024-07-21]
78. Mann KD, Good NM, Fatehi F, Khanna S, Campbell V, Conway R, et al. Predicting patient deterioration: a review of tools in the digital hospital setting. *J Med Internet Res.* Sep 30, 2021;23(9):e28209. [[FREE Full text](#)] [doi: [10.2196/28209](https://doi.org/10.2196/28209)] [Medline: [34591017](#)]
79. Malik M, Bigger JT, Camm AJ, Kleiger RE, Malliani A, Moss AJ, et al. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Eur Heart J.* Mar 01, 1996;17(3):354-381. [[FREE Full text](#)] [doi: [10.1093/oxfordjournals.eurheartj.a014868](https://doi.org/10.1093/oxfordjournals.eurheartj.a014868)]
80. Paliwoda M, New K, Davies M, Bogossian F. Physiological vital sign ranges in newborns from 34 weeks gestation: a systematic review. *Int J Nurs Stud.* Jan 2018;77:81-90. [[FREE Full text](#)] [doi: [10.1016/j.ijnurstu.2017.10.004](https://doi.org/10.1016/j.ijnurstu.2017.10.004)] [Medline: [29055243](#)]
81. Gardner SL, Carter BS, Enzman-Hines M, Niermeyer S. Merenstein & Gardner's Handbook of Neonatal Intensive Care: An Interprofessional Approach. Amsterdam, The Netherlands. Elsevier; 2020.
82. Sinha S, Miali L, Jardine L. Essential Neonatal Medicine. Hoboken, NJ. John Wiley & Sons; 2012.
83. Paliwoda M, New K, Bogossian F, Ballard E. Physiological vital sign reference ranges for well late preterm newborns calculated during a typical 2 h newborn period between 2 h and 7 days of life. *Physiol Meas.* Sep 27, 2021;42(9):095001. [[FREE Full text](#)] [doi: [10.1088/1361-6579/ac155b](https://doi.org/10.1088/1361-6579/ac155b)] [Medline: [34271562](#)]

84. Crowell DH, Brooks LJ, Colton T, Corwin MJ, Hoppenbrouwers TT, Hunt CE, et al. Infant polysomnography: reliability. Collaborative Home Infant Monitoring Evaluation (CHIME) Steering Committee. *Sleep*. Jul 1997;20(7):553-560. [Medline: [9322271](#)]
85. Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng*. Mar 1985;BME-32(3):230-236. [[FREE Full text](#)] [doi: [10.1109/tbme.1985.325532](#)]
86. de Chazal P, Heneghan C, Sheridan E, Reilly R, Nolan P, O'Malley M. Automated processing of the single-lead electrocardiogram for the detection of obstructive sleep apnoea. *IEEE Trans Biomed Eng*. Jun 2003;50(6):686-696. [[FREE Full text](#)] [doi: [10.1109/tbme.2003.812203](#)]
87. Lewicke AT, Sazonov ES, Schuckers SA. Sleep-wake identification in infants: heart rate variability compared to actigraphy. In: Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2004. Presented at: IEMBS 2004; September 1-5, 2004; San Francisco, CA. [doi: [10.1109/iemb.2004.1403189](#)]
88. Goldberger AL, Amaral L, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. PhysioBank, PhysioToolkit, and PhysioNet : components of a new research resource for complex physiologic signals. *Circulation*. Jun 13, 2000;101(23). [[FREE Full text](#)] [doi: [10.1161/01.CIR.101.23.e215](#)]
89. Torrence C, Compo GP. A practical guide to wavelet analysis. *Bull Am Meteorol Soc*. 1998;79(1):61-78. [[FREE Full text](#)] [doi: [10.1175/1520-0477\(1998\)079<0061:APGTWA>2.0.CO;2](#)]
90. Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. May 24, 2016;3:160035. [[FREE Full text](#)] [doi: [10.1038/sdata.2016.35](#)] [Medline: [27219127](#)]
91. Savitzky A, Golay MJ. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 1964;36(8):1627-1639. [[FREE Full text](#)] [doi: [10.1021/ac60214a047](#)]
92. Dumont J, Hernandez AI, Fleureau J, Carrault G. Modelling temporal evolution of cardiac electrophysiological features using Hidden Semi-Markov Models. In: Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2008. Presented at: IEMBS 2008; August 20-25, 2008; Vancouver, BC. [doi: [10.1109/iemb.2008.4649116](#)]
93. Doyen M, Ge D, Beuchée A, Carrault G, Hernández AI. Robust, real-time generic detector based on a multi-feature probabilistic method. *PLoS One*. Oct 29, 2019;14(10):e0223785. [[FREE Full text](#)] [doi: [10.1371/journal.pone.0223785](#)] [Medline: [31661497](#)]
94. Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Trans Biomed Eng*. Dec 1986;33(12):1157-1165. [[FREE Full text](#)] [doi: [10.1109/tbme.1986.325695](#)] [Medline: [3817849](#)]
95. Rooijakkers MJ, Rabotti C, Oei SG, Mischi M. Low-complexity R-peak detection for ambulatory fetal monitoring. *Physiol Meas*. Jul 2012;33(7):1135-1150. [[FREE Full text](#)] [doi: [10.1088/0967-3334/33/7/1135](#)] [Medline: [22735075](#)]
96. Lee H, Rusin CG, Lake DE, Clark MT, Guin L, Smoot TJ, et al. A new algorithm for detecting central apnea in neonates. *Physiol Meas*. Jan 2012;33(1):1-17. [[FREE Full text](#)] [doi: [10.1088/0967-3334/33/1/1](#)] [Medline: [22156193](#)]
97. Mohr MA, Vergales BD, Lee H, Clark MT, Lake DE, Mennen AC, et al. Very long apnea events in preterm infants. *J Appl Physiol* (1985). Mar 01, 2015;118(5):558-568. [[FREE Full text](#)] [doi: [10.1152/japplphysiol.00144.2014](#)] [Medline: [25549762](#)]
98. Vergales BD, Paget-Brown AO, Lee H, Guin LE, Smoot TJ, Rusin CG, et al. Accurate automated apnea analysis in preterm infants. *Am J Perinatol*. Feb 2014;31(2):157-162. [[FREE Full text](#)] [doi: [10.1055/s-0033-1343769](#)] [Medline: [23592319](#)]
99. Redmond SJ, Heneghan C. Cardiorespiratory-based sleep staging in subjects with obstructive sleep apnea. *IEEE Trans Biomed Eng*. Mar 2006;53(3):485-496. [[FREE Full text](#)] [doi: [10.1109/TBME.2005.869773](#)] [Medline: [16532775](#)]
100. Zuzarte I, Indic P, Sternad D, Paydarfar D. Quantifying movement in preterm infants using photoplethysmography. *Ann Biomed Eng*. Feb 2019;47(2):646-658. [[FREE Full text](#)] [doi: [10.1007/s10439-018-02135-7](#)] [Medline: [30255214](#)]
101. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc IEEE*. Feb 1989;77(2):257-286. [[FREE Full text](#)] [doi: [10.1109/5.18626](#)]
102. Lomb NR. Least-squares frequency analysis of unequally spaced data. *Astrophys Space Sci*. Feb 1976;39(2):447-462. [[FREE Full text](#)] [doi: [10.1007/BF00648343](#)]
103. Kota S, Massaro AN, Chang T, Al-Shargabi T, Cristante C, Vezina G, et al. Prognostic value of continuous electroencephalogram delta power in neonates with hypoxic-ischemic encephalopathy. *J Child Neurol*. Jul 2020;35(8):517-525. [[FREE Full text](#)] [doi: [10.1177/0883073820915323](#)] [Medline: [32306827](#)]
104. Govindan RB, Kota S, Al-Shargabi T, Massaro AN, Chang T, du Plessis A. Effect of electrocardiogram interference on cortico-cortical connectivity analysis and a possible solution. *J Neurosci Methods*. Sep 01, 2016;270:76-84. [[FREE Full text](#)] [doi: [10.1016/j.jneumeth.2016.06.009](#)] [Medline: [27291356](#)]
105. Govindan RB, Massaro A, Vezina G, Tsujchida T, Cristante C, du Plessis A. Does relative or absolute EEG power have prognostic value in HIE setting? *Clin Neurophysiol*. Jan 2017;128(1):14-15. [[FREE Full text](#)] [doi: [10.1016/j.clinph.2016.10.094](#)] [Medline: [27866114](#)]
106. Halliday DM, Rosenberg JR, Amjad AM, Breeze P, Conway BA, Farmer SF. A framework for the analysis of mixed time series/point process data--theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. *Prog Biophys Mol Biol*. 1995;64(2-3):237-278. [[FREE Full text](#)] [doi: [10.1016/s0079-6107\(96\)00009-0](#)] [Medline: [8987386](#)]

107. Shaffer F, Meehan ZM, Zerr CL. A critical review of ultra-short-term heart rate variability norms research. *Front Neurosci.* Nov 19, 2020;14:594880. [[FREE Full text](#)] [doi: [10.3389/fnins.2020.594880](https://doi.org/10.3389/fnins.2020.594880)] [Medline: [33328866](#)]
108. Kwon O, Jeong J, Kim HB, Kwon IH, Park SY, Kim JE, et al. Electrocardiogram sampling frequency range acceptable for heart rate variability analysis. *Healthc Inform Res.* Jul 2018;24(3):198-206. [[FREE Full text](#)] [doi: [10.4258/DIR.2018.24.3.198](https://doi.org/10.4258/DIR.2018.24.3.198)] [Medline: [30109153](#)]
109. Gravel J, Opatrný L, Gouin S. High rate of missing vital signs data at triage in a paediatric emergency department. *Paediatr Child Health.* Apr 2006;11(4):211-215. [[FREE Full text](#)] [doi: [10.1093/pch/11.4.211](https://doi.org/10.1093/pch/11.4.211)] [Medline: [19030273](#)]
110. Software as a medical device (SaMD): clinical evaluation : guidance for industry and Food and Drug Administration staff. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Devices and Radiological Health. Dec 8, 2017. URL: <https://collections.nlm.nih.gov/catalog/nlm:nluid-101720008-pdf> [accessed 2024-02-27]
111. Introduction to SaMD, IEC 62304 and IEC 82304-1. Medical Device HQ. URL: <https://medicaldevicehq.com/introduction-to-samd-iec62304-and-iec-82304-1-online-course/> [accessed 2024-07-21]
112. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *BMC Med.* Jan 06, 2015;13:1. [[FREE Full text](#)] [doi: [10.1186/s12916-014-0241-z](https://doi.org/10.1186/s12916-014-0241-z)] [Medline: [25563062](#)]
113. Ramakrishnan R, Rao S, He JR. Perinatal health predictors using artificial intelligence: a review. *Womens Health (Lond).* 2021;17:17455065211046132. [[FREE Full text](#)] [doi: [10.1177/17455065211046132](https://doi.org/10.1177/17455065211046132)] [Medline: [34519596](#)]
114. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng.* Jun 21, 2017;19:221-248. [[FREE Full text](#)] [doi: [10.1146/annurev-bioeng-071516-044442](https://doi.org/10.1146/annurev-bioeng-071516-044442)] [Medline: [28301734](#)]

Abbreviations

- AI:** artificial intelligence
CDSS: clinical decision support system
ECG: electrocardiogram
EEG: electroencephalography
HR: heart rate
HRV: heart rate variability
MeSH: Medical Subject Headings
ML: machine learning
NICU: neonatal intensive care unit
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QMS: quality management system
SaMD: software as a medical device
SpO2: oxygen saturation
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by T de Azevedo Cardoso; submitted 02.03.23; peer-reviewed by I Provaznik, D Gopukumar, K Zhang, E Grooby; comments to author 16.01.24; revised version received 27.03.24; accepted 26.06.24; published 20.08.24

Please cite as:

Rahman J, Brankovic A, Tracy M, Khanna S
Exploring Computational Techniques in Preprocessing Neonatal Physiological Signals for Detecting Adverse Outcomes: Scoping Review
Interact J Med Res 2024;13:e46946
URL: <https://www.i-jmr.org/2024/1/e46946>
doi: [10.2196/46946](https://doi.org/10.2196/46946)
PMID:

©Jessica Rahman, Aida Brankovic, Mark Tracy, Sankalp Khanna. Originally published in the Interactive Journal of Medical Research (<https://www.i-jmr.org/>), 20.08.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Interactive Journal of Medical Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.i-jmr.org/>, as well as this copyright and license information must be included.