Exploring the use of machine learning to assess the respiratory function of preterm infants

Vith Ketheeswaranathan

Pembroke College

University of Oxford

Abstract

Background: Preterm birth is associated with a number of pathologies that affect respiratory function. The identification of reduced lung function could aid with clinical diagnosis, earlier intervention and improved clinical outcomes for preterm infants. We explored the use of inter-breath intervals to predict the age at which healthy preterm infants' lungs are functioning. This can provide the groundwork to build a clinical tool that can assess preterm lung function in clinical practice.

Methods: Inter-breath intervals, measured through vital signs monitoring, were longitudinally recorded in healthy preterm infants with a post-menstrual age (PMA) < 37 weeks. Dataset 1, consisting of data from 32 infants, was analysed to compute 49 respiratory and statistical features. The relationship of these features with PMA was assessed through linear regression models. All features were used as inputs to train selected machine learning models to produce a predicted PMA. Mean Absolute Error (MAE) values were used to assess model accuracy.

Machine learning models with higher levels of accuracy were selected for the next stage of analysis. Inter-breath interval data from dataset 2 were analysed, consisting of 66 infants across 161 recordings. 50 features were extracted and used as inputs to train the selected models to output a predicted PMA. The most accurate model was used for further analysis to assess whether performance is affected by sex (paired t-test), ventilation method (ANOVA), and post-natal age (linear regression).

Results: 12 features from dataset 1 and 31 features from dataset 2 had a significant relationship with PMA ($p < 0.05$). The most accurate model was the bagged trees model trained on all 50 features, with a MAE of 1.30 weeks. Sex ($p = 0.17$), ventilation method ($p = 0.79$) and post-natal age ($p = 0.99$) did not affect model performance.

Conclusion: Inter-breath interval data offers novel directions for assessing the respiratory function of preterm infants.

Table of Contents

Introduction

Preterm birth is defined by The World Health Organisation (WHO) as births before the completion of the 37th gestational week of pregnancy.(1) Preterm birth can be expected even in low-risk pregnancies with healthy mothers, with WHO suggesting the global preterm birth rate ranged from 4-16% in 2020.(2) For example, studies such as the Multicountry Foetal Growth Longitudinal Study (INTERGROWTH) involved 4321 pregnancies with a preterm birth rate of 5%.(3) A number of risk factors have been associated with preterm births, although its exact aetiology is yet to be fully understood.(4) Factors such as black ethnicity, adolescent and advanced maternal age pregnancies, smoking and previous preterm pregnancies have all been shown to be risk factors, amongst others. (5-7) Preterm birth directly increases mortality rates in neonates, with it accounting for 35% of deaths among newborns.(8) It also may lead to complications such as respiratory distress syndrome, necrotising enterocolitis and cerebral palsy, as well as impacting longer term developmental outcomes.(9, 10)

Respiratory disorders are some of the most common pathologies affecting preterm infants. (11) Examples include respiratory distress syndrome, apnoea of prematurity and bronchopulmonary dysplasia.(12-14) It is estimated that 60-95% of infants born who have a gestational age of < 32 weeks will require some sort of respiratory support during their neonatal period.(15, 16) Understanding respiratory activity and how it evolves during a preterm infant's development is therefore essential in interpreting their overall health and development.

It is widely accepted that respiratory function in preterm infants changes with age. A number of studies have shown that as preterm infants develop, the number of apnoeic events decrease, less time is spent with hypoxic oxygen saturations and breathing patterns become more regular.(17-20) Respiratory activity has also been linked to broader aspects of preterm health and development. More frequent episodes of tachypnoea have been linked with slower rates of weight gain, suggesting that more irregular respiratory activity may contribute to delayed physiological growth.(21) Respiratory activity has been shown to be a predictor of common pathologies and declines in preterm infant health. Increased respiratory rate variability and apnoea rates can increase prior to diagnosis of late-onset sepsis, necrotising enterocolitis and escalations in respiratory support.(18, 21) It can therefore be deduced that more rigorous assessments of preterm infant respiratory development can lead to a better understanding of their overall clinical health and could encourage earlier diagnosis and interventions of pathologies, leading to better clinical outcomes. In particular, a method of monitoring and assessing respiratory activity in relation to post-menstrual age (PMA) could be a useful tool utilised in clinical neonatology.

Machine learning methods can be used to accurately predict the PMA of preterm infants based on real-time measures such as heart rate variability and EEG activity.(22-25) These age predictions, termed functional autonomic age (from heart rate variability data) and functional brain age (from EEG data) were compared with the true PMA, and large deviations between the measures are linked with long-term poorer developmental outcomes.(22, 23) The use of continuous respiratory activity monitoring to produce a similar age prediction (respiratory age) is yet to be considered. Such a tool has the potential to be used in clinical practice to better understand preterm respiratory development.

I hypothesise that continuous respiratory vital signs data can be used to train machine learning models to accurately assess the respiratory age of healthy preterm infants. This study will focus on infants expected to have normal respiratory development according to their clinical history, meaning that the respiratory age predicted should be similar to the true PMA. The study will provide the foundations for future work aiming to develop a machine learning tool that can identify infants with abnormal respiratory development through their respiratory age prediction differing from their true PMA, thus leading to earlier diagnosis, intervention, and improved neonatal care outcomes.

Before presenting my research study, further detail on the background to the study, including machine learning, foetal lung development and preterm respiratory pathologies, is provided.

Foetal Lung Development

A zygote is a single cell containing a combination of DNA from both gametes, formed during fertilisation.(26) The zygote undergoes multiple stages of mitosis to form a blastocyst, a structure consisting of two tissue layers called the trophoblast and the inner cell mass.(27) The trophoblast is the outermost layer that aids in implantation in the uterus lining, provides nutrients for the developing embryo, and is involved in placenta formation.(28) The inner cell mass gives rise to the formation of the foetus itself.(29) In the second week of gestation, the inner cell mass differentiates into the bilaminar embryonic disc, consisting of the epiblast (columnar epithelial cells) and hypoblast (cuboidal epithelial cells).(30)

During the third week of gestation, the bilaminar embryonic disc is transformed into a trilaminar embryonic disc in a process known as gastrulation. The three layers formed are the endoderm, mesoderm and ectoderm.(28) This process starts with the formation of a narrow groove called the primitive streak on the surface of the epiblast.(31) Cells of the epiblast migrate towards the primitive streak and move beneath it in a process known as invagination. Some of these invaginated cells displace the hypoblast to create the endoderm, some fill the space between the epiblast and newly formed endoderm to form the mesoderm, and the remaining cells in the epiblast form the ectoderm.(32) These three germ cell layers give rise to all subsequent embryonic tissues and the respiratory system is one of the structures that develops from the endoderm and mesoderm.(33)

The endoderm gives rise to the primitive foregut, and this is the site of formation of the primordial lung buds which defines the embryonic stage of the development of the respiratory system.(33) The pseudoglandular stage occurs during the 5th-17th gestational week. During this stage, the lung bud epithelia repeatedly grow and bifurcate to form the bronchial tree in a process called branching morphogenesis.(34) At around the 10th week of gestation, the first foetal breathing movements start and these have been shown to promote the differentiation and proliferation of these epithelial cells through the release of serotonin.(35) By 16 weeks, many of the key components of the respiratory system have started to develop except for those involved in gas exchange, such as the alveoli, making respiration impossible at this stage. Foetuses born during this period are therefore unable to survive.(31)

The canalicular stage occurs during the 16th-25th gestational weeks.(35) During this period, the lumina of the bronchi and the terminal bronchioles expand and give rise to the alveolar ducts and terminal alveolar sacs.(36)Vascularisation also occurs at this stage through angiogenesis, resulting in the first air-blood barriers forming.(37) It is at this point where the foetus is able to first carry out gas exchange and thus may survive if given intensive care.

The saccular stage, occurring between the 24th and 36th gestational week, is an intermediate stage where branching morphogenesis has terminated and alveolarisation is yet to start(38). The alveolar sacs formed in the canalicular stage are lined by type I alveolar epithelial cells which thin to allow intimate connections with the endothelial cells of the surrounding capillaries thus expanding the blood-air barrier.(33) Type II alveolar epithelial cells are also found in this region and start to produce surfactant, a phospholipid rich fluid that lowers surface tension and increases chances of survival of an early premature infant(39).

The alveolar stage occurs from 36 weeks gestation into childhood.(32) There are three forms of the alveolar stage – classical alveolarization, continued alveolarization and microvascular maturation.(34) During the saccular stage, a primary immature septa forms between the airspaces which gives rise to a new secondary septa during classical alveolarization. This new septa grows and subdivides the distal alveolar saccules into smaller units, the alveoli, hence expanding the surface area for gas exchange.(40) This process continues postnatally till childhood, but at a slower rate, and is known as continued alveolarisation.(41) The third phase, known as microvascular maturation, involves the fusion of capillary layers within the secondary septa into a single capillary layer, thus making a more efficient surface for gas exchange(36, 42). Alveoli continue to form during the first decade of postnatal life, increasing by at least 6-fold.(32) This is mainly through the growth of the secondary septa that subdivides existing primordial alveolar sacs. Early foetal breathing movements are important for the conditioning of respiratory muscles and the stimulation of lung development, but also result in the aspiration of amniotic fluid into the lungs.(43) This fluid is cleared at birth through the mouth and nose during vaginal delivery, and absorbed by the pulmonary vasculature and lymphatic system.(31)

Preterm births result in respiratory development in an altered environment.(44) Specifically, the in utero environment of a foetus is significantly more hypoxic in comparison to the atmosphere in which preterm babies develop in after birth.(45) This hypoxic environment is what drives key development processes such as bronchial branching morphogenesis and vascularisation.(44, 46) The hyperoxic environment has been shown to cause developmental abnormalities such as immature microvascular development, leading to pathologies such as bronchopulmonary dysplasia.(47)

Pathologies and factors that affect preterm lung function

Respiratory Distress Syndrome

Respiratory Distress Syndrome in neonates is a disorder due to surfactant deficiency and immaturity of the respiratory system, resulting in pulmonary dysfunction following birth.(48) Surfactant is normally produced by type II alveolar epithelial cells (also known as type II pneumocytes) which reduces surface tension and prevents alveolar collapse. In respiratory distress syndrome, the lack of surfactant production leads to increased alveolar surface tension and collapse, causing atelectasis and reduced gas exchange(49). Its incidence is reported to be approximately 1% of newborns, with rates decreasing as gestational age increases.(50) Risk factors include prematurity (<34 weeks gestation), maternal diabetes, birth asphyxia and low birth weight.(51) It is also more common in the white male population.(52) Symptoms typically present within the first 24 hours of birth and include tachypnoea, grunting, intercostal and subcostal retractions, nasal flarings and cyanosis.(53) Therefore, vital signs monitoring of measures such as respiratory rate and oxygen saturations are key for the diagnosis of respiratory distress syndrome. Diagnosis is usually made based on clinical symptoms and confirmed with evidence of oxygen requirement. The Vermont Oxford Network (VON) defined this as arterial partial pressures of lower than 6.6 kPa, central cyanosis in room air and the need for oxygen supplementation to maintain saturations of > 85%.(50, 51) Diagnosis can also be confirmed through imaging, with examples of typical x-ray findings being a diffuse ground-glass appearance and hypoexpansion of the lungs. Management of respiratory distress syndrome starts immediately following delivery with neonatal resuscitation.(54) Oxygen saturations should be maintained between 90% and 94%, initially through non-invasive respiratory support such as continuous positive airway pressure (CPAP), and then mechanical ventilation if this fails. Surfactant should also be administered, firstly through less-invasive means such as the use of a flexible catheter, then through more invasive methods such as intubation.(55) Antenatal corticosteroids may be given between 24 and 36 weeks gestation to stimulate surfactant production and decrease the risk of respiratory distress syndrome.(55)

Bronchopulmonary Dysplasia

Bronchopulmonary dysplasia is a chronic lung disease that most commonly occurs in infants who have needed mechanical ventilation and oxygen therapy as a result of respiratory distress.(56, 57) The National Institute of Child Health and Human Development (NICHD) defined bronchopulmonary dysplasia with the following criteria: gestational age of < 32 weeks, persistent parenchymal lung disease, radiographic evidence of parenchymal lung disease, and oxygen requirements at 36 weeks PMA for at least 3 days.[18] The aetiology of bronchopulmonary dysplasia is multifactorial. It has shown to be linked with oxygen toxicity, as this increases production of cytotoxic oxygen free radicals thus causing lung injury.(58, 59) Ventilator-associated lung injury due to volutrauma has been linked with bronchopulmonary dysplasia, with studies showing its inverse relationship with low PaCO2 levels.(60) It has also been linked to infants whose mothers had chorioamnionitis and its suspected that this may be due to increased endotoxin exposure whilst in an amniotic fluid environment, disrupting alveolar development. Other factors linked to bronchopulmonary dysplasia development include patent ductus arteriosus, concurrent infection and fluid overload causing pulmonary oedema.(53, 61, 62) Premature infants with bronchopulmonary dysplasia may present with tachypnoea and shallow breathing, intercostal and subcostal muscle retractions and wheezes heard on auscultation.(63) Infants often have a reduction in lung compliance as a result of small airway narrowing, interstitial fibrosis, oedema and atelectasis.(56) There is also an increase in dead space ventilation and decrease in ventilation-perfusion damage due to unequal distributions of lung damage in bronchopulmonary dysplasia.(56) Management includes strategies to reduce volutrauma such as permissive hypercapnia, where high levels of CO_2 are allowed but closely monitored by clinicians.(64) Other ventilatory strategies include patient triggered ventilation and high-frequency oscillatory ventilation and continuous positive airway pressure.(65-67) Corticosteroid therapy may also be used to reduce the inflammatory pathways and improve gas exchange in bronchopulmonary dysplasia.(68, 69) However, it is recommended to use corticosteroids with caution and in low doses due to adverse effects such as increased risk of cerebral palsy, growth suppression and immune deficiencies.(70, 71)

Apnoea of prematurity

Apnoea of prematurity (AOP) is a common pathology, prevalent in greater than 50% of all preterm infants and all infants born before 29 weeks' gestation.(72, 73) It is commonly defined as a cessation in breathing at less than 37 weeks gestation lasting at least 20 seconds, or at least 10 seconds followed by bradycardia and hypoxaemia.(74) These apnoeic episodes can result in cerebral deoxygenation, suppression of brain activity, and have been linked to poorer long-term neurodevelopmental outcomes.(75-77) AOP arises from immaturity of the lungs, brainstem, and brain-respiratory drivers and occurs as the preterm neonate transitions from non-breathing gas exchange within the womb to breathing ambient air. Specifically, AOP has been linked to aberrant activity of both central and peripheral chemoreceptors and poor neuromuscular control causing reduced upper airway patency.(19, 78) There has also been some evidence linking AOP to underlying genetic mechanisms. Studies have shown that polymorphisms in the A1 and A2A adenosine receptor genes have been associated with a higher risk of AOP and varying response to methylxanthine treatment.(18, 79) Treatment used in clinical practice includes methylxanthines such as caffeine. The mechanism behind this is still unclear, however mechanisms thought to be involved include the stimulation of adenosine A1 receptors in the brainstem which regulates the hypoxic ventilator drive through increased hypercapnia sensitivity, decreased hypoxic suppression of respiration and increased diaphragmatic contractility.(80)

Sepsis

Sepsis is defined as a severe systemic inflammatory response due to a documented or suspected infection.(81) Neonatal sepsis can be classified based on the timing of presentation into either early-onset sepsis or late-onset sepsis. Early onset sepsis is due to vertical transmission from the mother and presents within the first 72 hours of life.(82) Late onset sepsis presents after 72 hours of life and is typically acquired from the neonates surrounding environment, and in more rare cases, late-presenting from vertical transmission.(83) The most common causative pathogens of early onset sepsis are Group B streptococcus and Escherichia coli (E Coli), which account for almost 70% of cases.(84) Late-onset sepsis is typically caused by a range of pathogens, including gram-positive agents such as coagulase-negative staphylococci, gram-negative agents such as E coli, and fungal pathogens including candida albicans.(85, 86) Sepsis typically effects multiple organ systems as it develops, beyond the original site of infection. These systems include the neurologic, cardiovascular, respiratory, gastrointestinal and genitourinary systems.(85, 87, 88) Example respiratory symptoms include apnoea, tachypnoea, desaturations, grunting and intercostal muscle retractions which all effect normal respiratory functioning of the neonate.(89, 90) Management is mainly through empirical antibiotic therapy and is narrowed following culture results to target specific pathogens based on sensitivities to varying antibiotics.(91, 92)

Morphine sulfate

Morphine sulfate is an opioid agonist that selectively binds to mu receptors. This occurs when morphine sulfate binds and activates g-protein coupled receptors, leading to a reduction in cyclic AMP (cAMP), reduced calcium influx at the pre-synaptic terminal, and hyperpolarisation of the post-synaptic neuron due to an efflux of potassium ions. As a result, ascending neuronal pain pathways are inhibited.(93, 94) Morphine sulfate also causes a number of adverse effects including sedation, respiratory depression, constipation, nausea and vomiting.(95) It is common practice to prescribe morphine to neonates, typically when they undergo painful procedures or are put on mechanical ventilation.(93) Morphine causes respiratory depression through binding to mu opioid receptors at key centres such as the pre-Botzinger complex (which normally plays a role in driving ventilatory rhythm), and depressing chemoreceptor detection of high CO2 and low O2 levels.(96, 97)

These pathologies and medications are examples of prevalent factors that alter respiratory function in preterm neonates. Following the development of our machine-learning model to accurately assess the respiratory age of healthy preterm infants, future work may focus on identifying infants with altered respiratory function, potentially as a result of these factors, through giving them a respiratory age which differs from their PMA. This could aid clinicians as it may prompt earlier investigations to identify which specific pathology is leading to this abnormal respiratory functioning, hence leading to earlier diagnosis, intervention and improved clinical outcomes.

Artificial Intelligence and Machine Learning

Machine Learning is based on models that can learn, analyse, and evaluate data to make predictions or decisions.(98) The general approach of machine learning models is that they learn patterns between the data and outcomes. They then are tested on a separate subset with the same type of data to evaluate how effective the model is at predicting outcomes based on the patterns it learned from the training data.(99) The use of machine learning in clinical practice is a rapidly growing field.(100) It has been used to learn, reason and assess clinical outcomes, including in neonatology.

A review on the use of artificial intelligence and machine-learning in neonatology reflected on the importance of continuous vital signs monitoring, particularly in the neonatal intensive care unit (NICU).(94) Machine learning algorithms have been incorporated into continuous vital signs monitoring devices to reduce the rate of false alarms for incidences such as apnoea and earlier recognition of disease states, sometimes hours before clinical signs start to show.(101-103) Vital signs such as heart rate are currently used in clinical practice, incorporated into scoring systems, to aid with early diagnosis and intervention of specific pathologies.(104) As an example, the HERO score uses heart rate variability from continuous ECG monitoring to aid with earlier sepsis diagnosis. Studies have shown that infants with continuous ECG monitoring had lower mortality rates, and that the HERO score elevated at earlier time points compared to other commonly used sepsis.(105, 106)

These models learn patterns in the training dataset via three main learning approaches: a) supervised b) unsupervised and c) reinforcement learning. Supervised learning is when the model is trained on a dataset with labels, e.g. a clinician has identified and labelled the clinical outcomes in question. It is then tested on an unlabelled dataset to predict the outcomes in new cases.(107) Supervised learning can be used for discrete outputs (classification) or continuous outputs (regression). Unsupervised learning is an approach where the model tries to identify patterns or groups within the data without using labelled outcomes.(108) For example, it may be used clinically to identify patients with similar clinical or physiological data and relate this to whether they have a specific pathology.(99) Reinforcement learning involves the model defining rewards and punishment functions for actions within a time-series; if the correct action is taken, the model is rewarded.(109) In a clinical context, this may be used to train a model to diagnose a pathology via a number of steps that maximises the possible rewards.(110)

The current study aims to assess models using the supervised learning approach, with the data labelled according to the PMA of the baby. Models such as linear SVM, bagged trees, Gaussian Process Regression, Linear Regression and Wide Neural Networks will be used with the aim to produce a predicted respiratory age based on features derived from the respiratory vital signs activity. Each data point consists of a set of input respiratory features and a known target label – the true infant PMA. The models will learn relationships between the input features and the true PMA. In order to prevent the models from learning patterns specific to the particular training dataset (known as overfitting bias), 5-fold cross-validation will be used. In this approach, the dataset is divided into 5 subsets, and the models are trained on 4 of these. The remaining subset is used as a test dataset where the model is not provided with the true PMA – instead, it must produce a predicted PMA based on the patterns learned in the other folds. This process is repeated 5 times, so each fold serves as the test set once. The model's accuracy is determined by the average difference (in weeks) between the predicted PMA and the true PMA across all folds. Future steps would be to then test these models on an independent dataset which the model has not come across yet, to assess how accurate the predicted PMAs are.

Aims and hypotheses

The aims of the current study are twofold. The first aim is to determine whether machine learning models can use respiratory activity and derived features to accurately predict the post-menstrual age of healthy preterm infants. The second, is to determine whether clinical factors such as ventilation, and demographic factors such as sex, influence the performance of the model. It is hypothesised that preterm infant respiratory activity can be used to train an accurate machine learning model to predict post-menstrual age.

Methods

Data from a total of 98 infants were used across two cohorts. Dataset 1 comprises of retrospective vital signs data obtained from the research database of our group, the Paediatric Neuroimaging Group, collected as part of two previous cross-sectional studies. (111) Dataset 2 is made up of two subsets. 2a consists of vital signs data I collected for 24 hours prospectively as part of this study. Dataset 2b consists of data from a separate ongoing study where vital signs are recorded continuously throughout the infant's stay in the neonatal unit. For the purpose of this analysis, a 24-hour period from each week of these continuous vital signs recordings for each infant was selected.

These data were collected at the Newborn Care Unit of the John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK between 2021 to 2023 for dataset 1, 2024 to 2025 for dataset 2a, and 2019 to 2025 for dataset 2b. Studies were approved by the National Research Ethics Service (ethics references: 12/SC/0447; 11/LO/0350; 23/NW/0155). Data were only collected from infants whose parents or legal guardians had given verbal and written consent before participation in the relevant research study. All studies complied with the Declaration of Helsinki and guidelines on Good Clinical Practice.

The inclusion and exclusion criteria for both datasets are outlined in Table 1.

	Dataset 1	**Datasets 2a and 2b**
Inclusion Criteria	• Infants with continuous vital signs recordings • Infants aged <37 weeks	• Infants with continuous vital signs recordings • Infants aged <37 weeks

	PMA*	PMA
Exclusion Criteria	• Grade III or IV intraventricular haemorrhage, • Hypoxic ischaemic encephalopathy, • Congenital malformations, • On mechanical ventilation, • Receiving opioid analgesics at the time of study • History of maternal substance misuse during pregnancy. • Receiving antibiotics at time of study * • Repeat recordings on the same infant *	• Grade III or IV intraventricular haemorrhage, • Congenital abnormalities affecting respiratory function • Receiving opioid analgesics at the time of study • Receiving antibiotics at time of study

Table 1: Inclusion and exclusion criteria for both dataset 1 (retrospective analysis) and dataset 2 (prospective data collection). The ''s denote criteria added as part of the current study analysis.*

Inclusion and exclusion criteria for infants from dataset 1 were applied during the original data collection phase for a study that examined EEG and vital signs. The original exclusion criteria for this dataset were chosen as these factors may have acted as confounders when interpreting the EEG recordings as they have been shown to impact brain activity. The exclusion criteria added as part of this retrospective analysis are labelled in Table 1 with '*'s. Infants ≥ 37 weeks at the time of recording were excluded to mirror the preterm population of dataset 2. Infants receiving antibiotics at the time of study were excluded to minimise the inclusion of infants with an active infection. This is because infection has been shown to alter lung function and therefore could confound the relationship between respiratory activity and PMA.(89, 90)

Exclusion criteria for infants in dataset 2 were chosen due to their effects on respiratory system functioning.(89, 90, 97) Whether a congenital abnormality was deemed to affect the respiratory system was determined on a case-by-case basis by the clinical team.

Factors	Dataset 1	Dataset 2a	Dataset 2b
Total Infants	32	30	36
Total Recordings	32	30	131
Median Recording Length (Hours)	2.0	24.7	24.0
Age			
Gestational age at birth (weeks)	29.7 (23.6-35)	29.6 (24.1-34.9)	30.2 (28.1 – 32.7)
Postmenstrual age at recording (weeks)	33.1 (28.9-36.9)	31.5 (26.4-36.0)	32.1 (29.6-34.1)
Postnatal age at recording (weeks)	3.5 (0.9-8.9)	1.9 (0.4-5.6)	1.9 (0.4-4.7)
Birthweight (g)	1336 (550-3,075)	1324.5 (550-2970)	1425.8 (635.0-2120.0)
Sex			
Females	17 (53.1)	11 (36.7)	13 (36.1)
Males	15 (46.9)	19 (63.3)	23 (63.9)
Mode of delivery			
Normal vaginal delivery	8 (25.0)	1 (3.3)	6 (16.7)
Elective Caesarean section	5 (15.6)	3 (10.0)	6 (16.7)
Emergency Caesarean section	19 (59.4)	24 (80.0)	21 (58.3)
Breech vaginal delivery	0 (0)	1 (3.3)	1 (2.8)
Ventouse/Forceps	0 (0)	1 (3.3)	2 (5.6)
Apgar scores			
Apgar at 1 minute	6.5 (0-10)	7.8 (4.0–9.0)	7.1 (1.0 -10.0)
Apgar at 5 minutes	8.9 (3-10)	8.6 (5.0-10.0)	8.9 (5.0-10.0)
Apgar at 10 minutes	9.7 (6-10)	9.0 (5.0-10.0)	9.6 (7.0-10.0)

Ethnicity			
White	18 (56.3)	23 (76.7)	23 (63.9)
Asian	3 (9.4)	2 (6.6)	3 (8.4)
Black	0 (0)	0 (0)	1 (2.8)
Mixed	5 (15.6)	3 (10.0)	1 (2.8)
Other	1 (3.1)	2 (6.6)	0 (0)
Information not recorded	5 (15.6)	0 (0)	8 (22.2)
Respiratory distress syndrome			
Yes	26 (81.2)	25 (83.3)	32 (88.9)
No	6 (18.8)	5 (16.7)	4 (11.1)
Patent ductus arteriosus			
Past treated	3 (9.4)	0 (1.9)	0 (0)
Ongoing Condition	7 (21.9)	8 (26.7)	2 (5.6)
No	22 (68.9)	22 (73.3)	34 (94.4)
Jaundice			
Past treated	29 (90.6)	16 (53.3)	20 (55.6)
Ongoing Condition	0 (0.0)	1 (3.3)	2 (5.6)
No	3 (9.4)	13 (43.3)	12 (33.3)
Level of care			
Intensive Therapy Unit	5 (15.6)	13 (43.3)	4 (11.1)
High Dependency Unit	12 (37.5)	15 (50.0)	24 (66.7)
Low Dependency Unit	15 (46.9)	1 (3.3)	8 (22.2)
Information Not Recorded	0	1 (3.3)	0 (0)
Respiratory support			
High flow therapy	12 (37.5)	10 (33.3)	29 (22.1)
Low flow therapy	2 (6.2)	0 (0)	20 (15.3)
Spontaneous ventilation	18 (56.2)	13 (43.3)	72 (55.0)
SIPPV		1 (3.3)	1(0.8)
Information not recorded	0 (0.0)	6 (20.0)	9 (6.9)

Table 2: Infant demographics across all three datasets analysed. Reported values are the mean (range) or number (%).

Data Collection

Data collection for Dataset 2A

Dataset 2A was collected specifically for this study. I recruited the infants in this dataset and recorded the data myself. Due to the lack of intervention in this study, with us only recording vital signs data that is already being monitored, over 90% of parents were happy for their infant to take part in the study as shown in table 3. A small number of challenges were faced during data collection. There was no lower age limit for our target age range (<37 weeks PMA), however, the number of babies with low PMAs, specifically <30 weeks, was noticeably lower. This can be put down simply to the lower frequency of preterm infants born at this age range, as well as the higher rate of pathologies such as infection and the higher likelihood of infant's receiving opioids at lower PMAs. The other challenge was navigating through the higher number of research studies happening on the Newborn Care Unit at the John Radcliffe Hospital. There would be occasions where an infant would be eligible for a number of research studies so a balance needed to be found between overwhelming the parents with research and trying to increase recruitment rates for each study. Daily neonatal recruitment meetings aim to ensure relevant studies are discussed across research teams and considered so that parents are not overwhelmed. As several other studies are currently running on the neonatal unit relevant to infants born at early gestational ages this may have also impacted the number of infants studied at lower PMAs. Overall, however, once parents were approached, the vast majority were happy to take part in the current study.

Families Approached	33
Consent Given	30
Declined	3
Infants Recorded	30
Total Recordings (Including Repeats)	33
Withdrawals	0

Table 3: Recruitment numbers for dataset 2a, conducted on the Neonatal Unit at the John Radcliffe Hospital, Oxford.

Recording methods

Infants in both datasets had continuous vital signs monitoring as the standard of care by a Phillips Intellivue MX800 or MX750 monitor. Recordings were continuously downloaded from the monitor onto a research recording laptop using iXtrend (ixitos, Germany), an electronic data capture software. The median vital signs recording lengths for all datasets are shown in Table 2. Three electrodes placed on the infants' chest recorded electrocardiograph (ECG) data at a sampling rate of 250 Hz for heart rate, and impedance pneumography (IP) traces at 62.5 Hz for respiratory rate. A probe on the infant's hand or foot recorded photoplethysmography (PPG) data at 125 Hz for oxygen saturations and pulse. The pre-processed monitor data of heart rate, respiratory rate, oxygen saturations and pulse rate were also downloaded at a sampling rate of 0.97 Hz.

Analysis

MATLAB (ver. R2024b; MathWorks Inc., Natick, USA) was used to analyse vital signs data.

Inter-breath intervals (IBIs) were automatically extracted from the impedance pneumograph (IP) traces using an algorithm developed previously.(112) (1). To briefly summarise, the algorithm consists of three main steps. First, artefacts in the IP trace due to infant movement or cardiac electrical activity, and not respiration, are filtered and removed. The second step of the algorithm uses an adaptive threshold to identify individual breaths and take into account variations due to factors such as shallow breathing and changes in electrode positioning. A breath is identified when the IP signal crosses the threshold. The threshold is 0.4 times the standard deviation of the previous 15 breaths and thus varies with variations of IP signal amplitude over time. (The threshold of 0.4 times the standard deviation was used as this was found to be the most optimal after comparing breaths detected by the algorithm for different thresholds with ones identified by clinical members of staff visually observing the infant's breathing in the algorithm validation data set.(112) The final step of the algorithm uses a linear SVM classifier to identify which of the central apnoeas (defined as IBIs ≥ 15 s) identified in step 2 are true and which are artifactually low signals and thus false apnoeas. Another machine learning model was then applied to identify true and false short pauses in IBIs (5 to 15 seconds). Following the application of the two SVM classifiers, the dataset consisted of IBIs without falsely identified pauses due to shallow signal amplitude. Figure 1 shows an example apnoea of

25.6 seconds. Regular breathing was observed before and after the apnoea (as indicated by the red dots representing individual breaths).

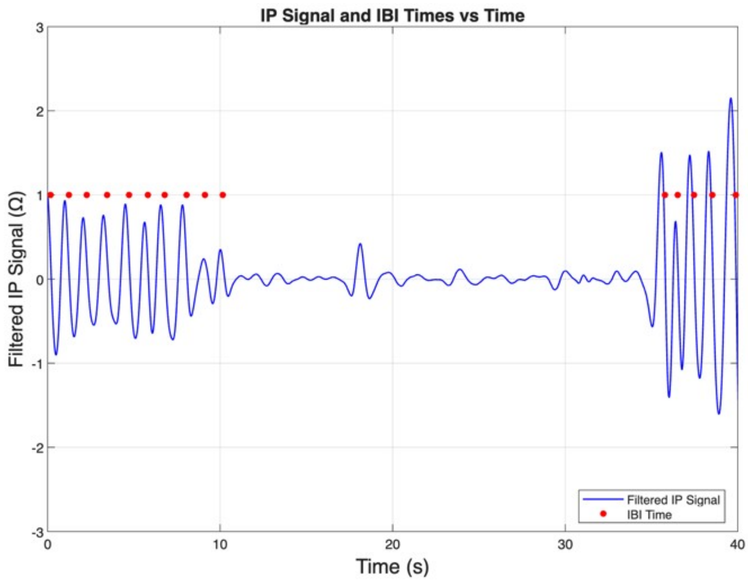


Figure 1: Example apnoea. The filtered IP signal (blue) and individual breaths detected by the algorithm (red dots) across time. Inter-breath intervals are the time between breaths, and the large gap in breaths denotes an apnoea (temporary cessation in breathing).

The IBI data from both datasets were analysed to obtain features to be used as the inputs into the selected machine learning models. The output of these models was respiratory age.

Feature selection

To develop a machine learning model that produces an accurate respiratory age, multiple features of the IBI distribution were considered. Obvious metrics to use were those that are currently used clinically to assess respiration – the respiratory rate and apnoea rate. To identify other possible features existing literature was searched. A study by Iyer et al. (2022) used heart rate variability data from preterm infants, obtained from ECG recordings, to produce a functional autonomic age which was strongly correlated with PMA. They used 50 computed heart rate variability features as inputs into a machine learning model, specifically a Gaussian Process Regression Model, to produce the functional autonomic age. Specifically, these 50 features were computed from NN intervals that were determined by the intervals between the R peaks of the ECG recordings. This compares similarly to the IBI data of our current study as both forms of data are measures of time intervals of physiological rhythms (see figure 2). We therefore applied similar feature extraction methods to our IBI data by adapting the code provided by Iyer et al. (2022) which can be found here: https://github.com/brain-modelling-group/Estimate-FAA.(22)

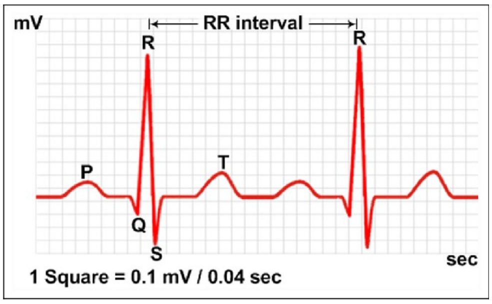


Figure 2: RR interval of an electrocardiogram(113)

The features included common statistical measures of distribution, spectral measures and entropy measures. The specific respiratory measures of respiratory rate and pauses in breathing (defined as an IBI of > 5 seconds) were added to the list obtained from the HRV study. Respiratory rate was calculated by dividing 60 (seconds) by the median of the IBIs. An extra feature, apnoea (defined as an IBI of >15 seconds), was added for analysis of dataset 2. It was left out of dataset 1 analysis due to the low frequency of apnoeic events lasting >15 seconds due to the shorter recording periods. The final list of features used to analyse datasets 1 and 2 are listed in table 4.

Respiratory Measures	Basic Statistical Features	Spectral Measures	Entropy Measures	Power Law Exponents	Non-Linear Measures
Respiratory Rate	Mean	Band Power 1*	Sample Entropy*	BA Average*	Median SNLEO*
Apnoea Rate (5 seconds)	Median	Band Power 2*	Short-Term IBI Variability*	BN*	Hurst
Apnoea Rate (15 seconds) ^	IBI 5th Percentile	LF Normalised*	Long-Term IBI Variability*	BN*BA Average*	
	IBI 95th Percentile	HF Normalised*	Fractal Dimension	BWHc*	
	Minimum IBI	HF:LF Ratio*	IBI MSE Scale 1*	BWHm*	
	Maximum IBI	Frequency Width at Half Maximum	IBI MSE Scale 2*	Power Law (Alpha)	
	Range	Frequency Width	IBI MSE Scale 3*	PLBSKc*	
	Area*	Instantaneous Autospectrum	IBI MSE Scale 4*	PLBSKm*	
	Standard Deviation		Median Spectral Entropy*	PLXmin*	
	Skewness		Max PSE*		
	Kurtosis		Shannon Entropy		
			Renyri Entropy		
			Max Perm Ent*		
			Approximate Entropy		
			Fuzz Entropy		
			Conditional Entropy		
			LZC*		

Table 4: List of all features extracted from the IBI data of both dataset and dataset 2.
Features with ‘^’s denotes ones only used in dataset 2.
Features with *s are defined in table
A burst: Sustained period of longer IBIs as seen when plotting IBIs against time

Feature Name	**Abbreviation**	**Definition**
Area	-	Total sum of IBIs over time, calculated using trapezoidal integration. This represents the area underneath the curve when plotting time vs IBI.
Low Frequency Spectral Power	LF Normalised	Comparison of the slow IBI changes (low frequency spectral power) with the overall variability in IBIs (standard deviation)
High Frequency Spectral Power	HF Normalised	Comparison of the fast IBI changes (high frequency spectral power) with the overall variability in IBIs (standard deviation)
Frequency Power Ratio	HF:LF Ratio	The balance between the fast and slow changes in IBI signal
Lempel-Ziv Complexity	LZC	A measure of how many different patterns appear in the IBI sequence. Higher values indicate more variation.
Burst Area Average	BA Average	The mean area of each burst, calculated by adding the IBIs for every burst in a recording
Burst Number	BN	The Total Number of Bursts
Burst number * Burst Area Average	BN*BA Average	The product of the Burst Number and the Burst Area Average
Burst Width Height Intercept	BWHc	The intercept of a simple linear regression line for the height and width of each burst
Burst Width Height Gradient	BWHm	The gradient of a simple linear regression line for the height and width of each burst
Burst Skewness Kurtosis Intercept	BSKc	The intercept of a simple linear regression line for the skewness vs kurtosis of each burst
Burst Skewness Kurtosis Gradient	BSKm	The gradient of a simple linear regression line for the skewness vs kurtosis of each burst
Power Law (alpha)	Alpha	A measure of the spread of burst areas across a recording. A large alpha value indicates fewer larger burst areas.
Minimum Burst Value	Xmin	The smallest burst area value

Band Power 1	BP1	The average statistical power of periods of slow changes in IBI across time
Band Power 2	BP2	The average statistical power of periods of fast changes in IBI across time
Instantaneous Autospectrum	Inst Spec	A measure of the overall fluctuation of IBIs by looking at the speed of breathing rhythms and when they occur
Sample Entropy	-	A measure of how predictable the IBIs are by looking at how regular breathing patterns occur. A low value indicates more regular breathing
Short-term IBI variability	SD1	A measure of how consecutive IBIs differ from each other
Long-term IBI variability	SD2	A measure of how IBIs change over a many breaths
Fractal Dimension	-	A measure of how complex the IBI patterns are, with higher values indicating more irregular and intricate breathing patterns
IBI Multiscale Entropy 1-4	IBI MSE 1-4	A measure of the unpredictability of breathing patterns across different scales
Median Spectral Entropy	-	Spectral Entropy is the measure of the spread of IBI signal across different breathing patterns. The median gives a representation of this across the whole recording.
Maximum Permutation Spectral Entropy	Max PSE	The point at which breathing rhythms were most evenly mixed, with no single rhythm dominating
Maximum Permuation Entropy	Max Perm Ent	The most irregular point in the breathing patterns during the recording
Median Signal Nonlinear Energy Operator	Median SNLEO	SNLEO is a measure of sudden changes in IBI amplitude; the median represents a typical size of these changes

Table 5: Selected features defined along with abbreviations used throughout text

Dataset 1

Each respiratory feature's relationship with PMA was analysed. Simple linear regression models, looking at the correlation of each feature with PMA, were used to identify statistical correlations. Simple linear regression models were chosen as dataset 1 consisted of only infants with one test occasion and no repeats. This exploratory analysis provided insights into the relationship between PMA and the selected features. Additionally, visualisation of the data allowed for an initial assessment of the trends in respiratory features with PMA.

Machine Learning model training was carried out on the MATLAB Regression Learner App. The specific models used were linear SVM, bagged trees, Gaussian process regression, linear regression models and wide neural networks. All 50 features were inputted into the selected machine learning models which gave an output of predicted PMA i.e. respiratory age. Features were then ranked in ascending order of p-values, computed from the simple linear regression analysis. Selected models were then trained again on only features with significant correlations ($p < 0.05$) as inputs, thus providing a data driven approach for feature selection. Model accuracy was measured through the mean absolute error values produced through 5-fold cross-validation. Scatter plots were also produced for each machine learning model, plotting the true PMA against the predicted respiratory age.

The correlation between the predicted PMA and the true PMA in the 3 most accurate models was assessed using Pearson's correlation coefficients. This exploratory approach gave us an initial insight into the predictive accuracy of the models.

Dataset 2

The types of machine learning models that performed best in dataset 1 were retrained with data from dataset 2. As before, simple linear regression models were used to rank features in order of p-values. The models were trained on 3 sets of features: a) on all 50 respiratory features b) on the statistically significant features ($p < 0.05$) from the regression analysis on dataset 2 c) the statistically significant features from dataset 1. Scatter plots of true PMA versus predicted PMA were plotted to visualise model performance, and accuracy was assessed through the mean absolute error from 5-fold cross validation. The correlation between the predicted and true PMA was assessed using Pearson's correlation.

The model with the lowest MAE was then selected for further analysis, specifically looking at variability of the prediction errors between different clinical subgroups within dataset 2. The subgroups compared were males versus females, and the different ventilation methods during recordings (high-flow, low-flow and self-ventilation). Infants on synchronised intermittent positive pressure ventilation (SIPPV) were excluded from this part of the analysis due to low numbers within the dataset (2 infants). The prediction error for each infant was calculated through subtracting the model's output (predicted PMA) from the true PMA and converting these all to positive error values giving the absolute error. The differences in prediction errors were assessed with statistical significance through a paired t-test for sex and one-way ANOVA test for ventilation mode.

Post-natal age was defined as the difference between the PMA at time of recording to the PMA at birth. The impact of post-natal age on the absolute prediction accuracy of the most accurate machine learning model was analysed through a scatter plot. This plotted the post-natal age against the absolute prediction errors for each infant (converting the prediction errors to all positive values), along with a line of best fit computed through a simple linear regression model. This analysis aided in visualising if the post-natal age had any impact on model performance. Correlation between the absolute prediction error and post-natal age was assessed through a Pearsons correlation test to get an r and p value.

Results

Respiratory features correlate with PMA

Dataset 1 consisted of 32 infants with a PMA range of 28.9 to 36.9 weeks.

The relationship between the respiratory feature and PMA varied greatly depending on the respiratory feature. 24.49% (12 out of 49) of respiratory features had a significant relationship with PMA (Figure 3 and Appendix Figure 1). The minimum IBI data was approximately 0.3 seconds for all infants, meaning that the range and maximum IBI data were almost identical. Respiratory features that had a significant relationship with age included measures from basic statistical features (respiratory rate, mean IBI, skewness, kurtosis) and various entropy measures (such as Renyri entropy). While the remaining 37 features (Appendix Figure 1) were not significantly correlated with PMA at a significance level of 5%, this should be interpreted with caution due to the sample size, with some features demonstrating a clear trend with PMA.

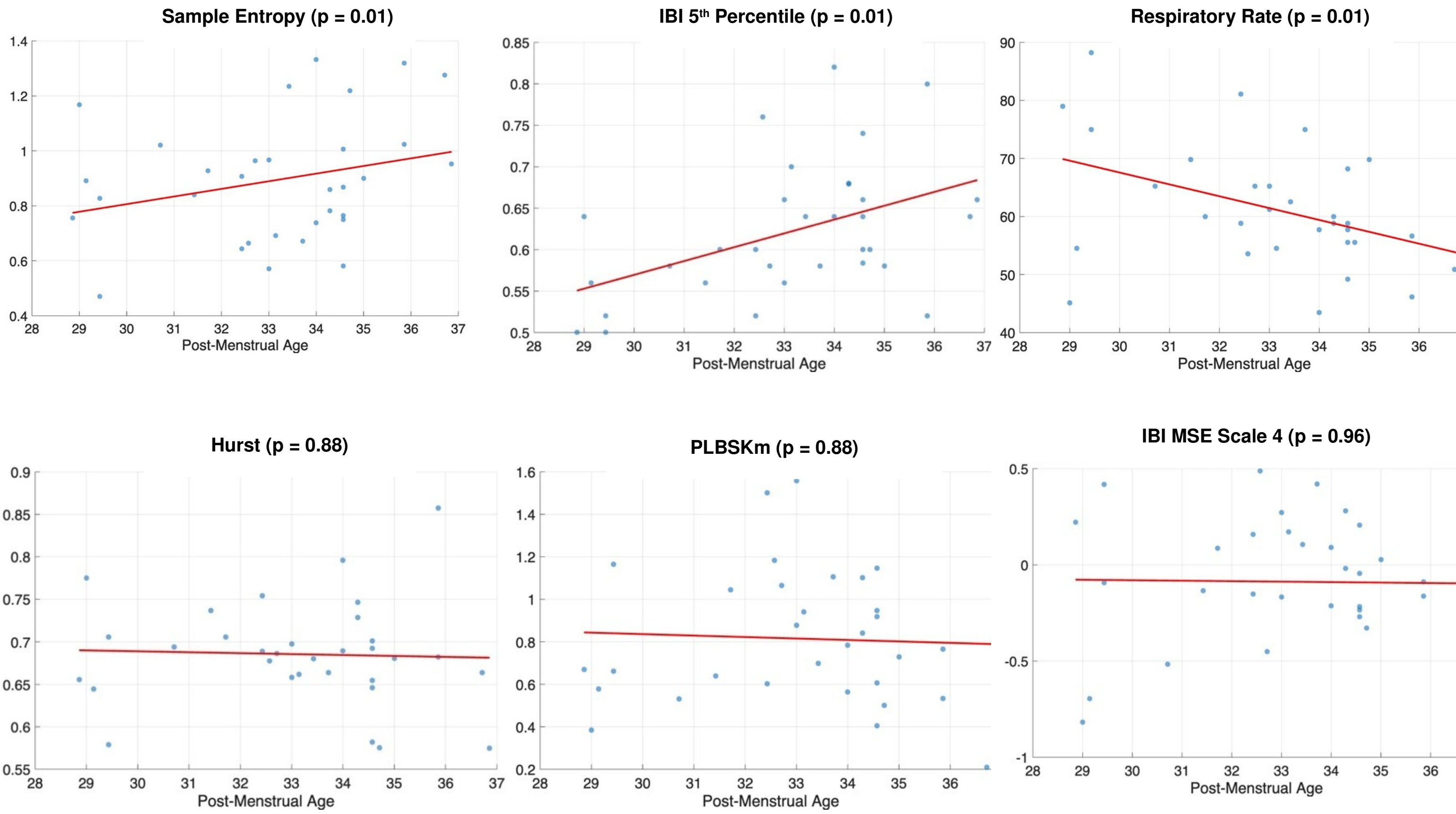


Figure 3: Scatter plots showing the relationship between 6 respiratory features and postmenstrual age (PMA) for dataset 1. The features with the 3 highest and 3 lowest p values have been selected. The red line represents the line of best fit from simple linear regression models. Plots for all features are shown in Appendix Figure 1 .

Can respiratory features predict age? Initial model training (Dataset 1)

Selected machine learning models were trained with all 49 of dataset 1's features used as inputs and predicted PMA as the output. The scatter plots of the true versus predicated PMAs from the machine learning models are shown in figure 4. The mean absolute errors assessing the accuracy of these models are shown in table 6. Following training with all features, the linear SVM, bagged trees and Gaussian process regression models all performed to a higher degree of accuracy (MAEs of 1.74, 1.71 and 2.08 weeks respectively) when compared to the wide neural network and linear regression models (MAEs of 8.01 and 4.58 weeks respectively). The three more accurate model types were then trained using only the 12 statistically significant dataset 1 features and produced similar MAEs (Figure 5), with the bagged trees and Gaussian Process regression models having a slight decrease (1.66 and 1.86 weeks) and the linear SVM a small increase in MAE (1.76 weeks). It is worth noting that the correlation coefficients and p values for Linear SVM, Bagged Trees and Gaussian Process Regression models in dataset 1 indicated a poor correlation between the true and predicted PMAs (Table 6).

The mean absolute errors and correlation results from dataset 1 training suggests that this data cannot be used to accurately assess respiratory age. However, there were a number of limitations during this process. The sample size (32) and data length (median 2.0 hours) may have been too short, as indicated by the lack of apnoeas lasting >15 seconds. Indeed, apnoea rate is known to decrease with age, but the data length may have been too short to accurately assess this feature. Similarly, measures such as skewness and kurtosis may not have been accurately assessed in such a short recording length. These two factors (sample size and data length) are overcome in dataset 2. Nevertheless, the exploration here demonstrated the potential of addressing this research question given that some features correlate well with age.

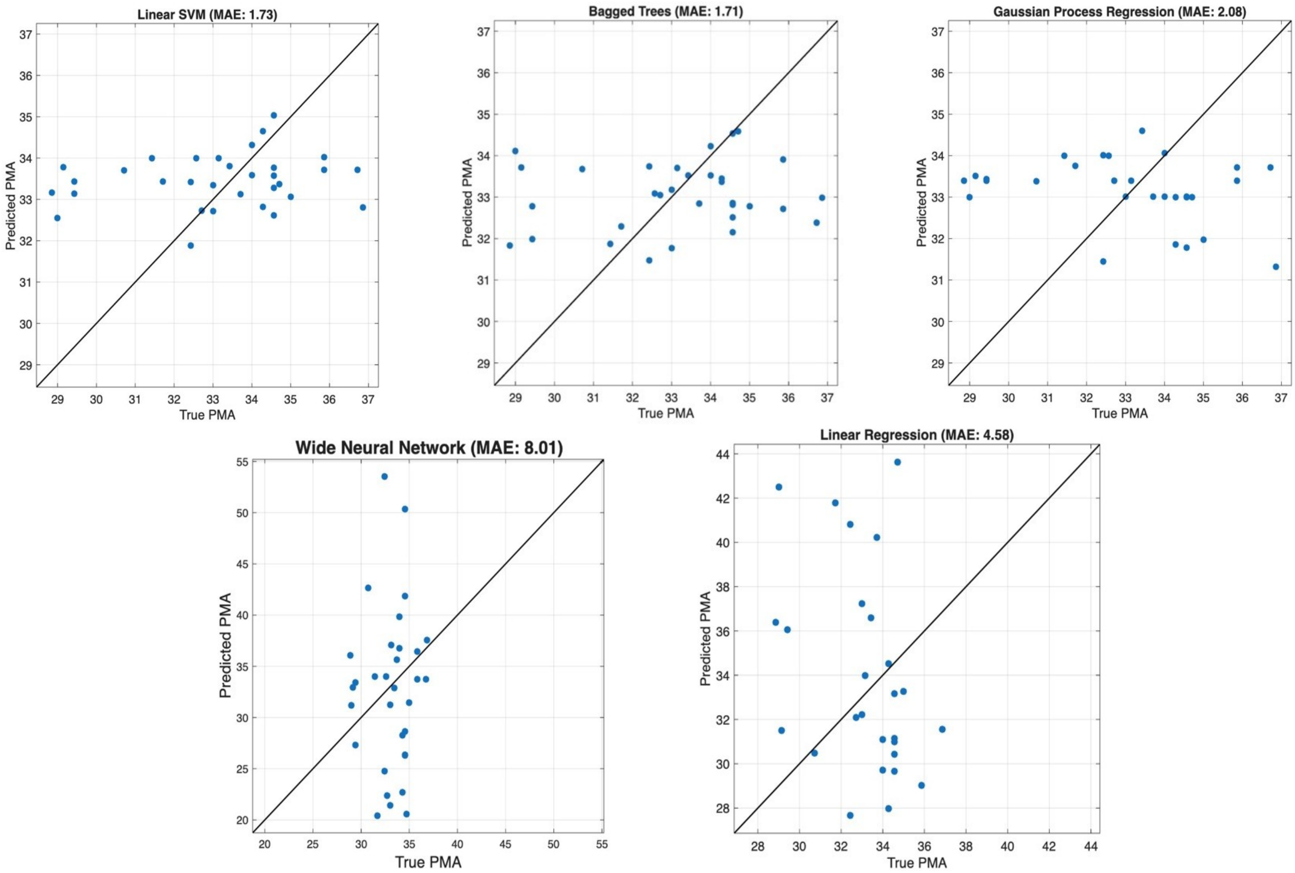


Figure 4: Predicted vs true postmenstrual age (PMA) for five machine learning models with 49 respiratory feature inputs Each scatter plot compares the PMA with the predicted PMA by 5 different machine learning models for all infants in dataset 1. The black line represents the perfect prediction line. The models included are a) Linear SVM b) Bagged Trees c) Gaussian process regression d) Wide neural networks e) Linear regression

	All Features			**12 Significant Features**		
Machine Learning Model	**MAE**	**r**	**p**	**MAE**	**r**	**p**
Gaussian Process Regression	2.08	-0.26	0.16	1.86	-0.08	0.66
Linear SVM	1.73	0.15	0.41	1.76	0.16	0.38
Bagged Trees	1.71	0.10	0.59	1.66	0.27	0.2
Wide Neural Network	8.01					
Linear Regression	4.58					

Table 6: Mean Absolute Errors (MAEs) and Pearson correlation r and p-values for machine learning models trained on

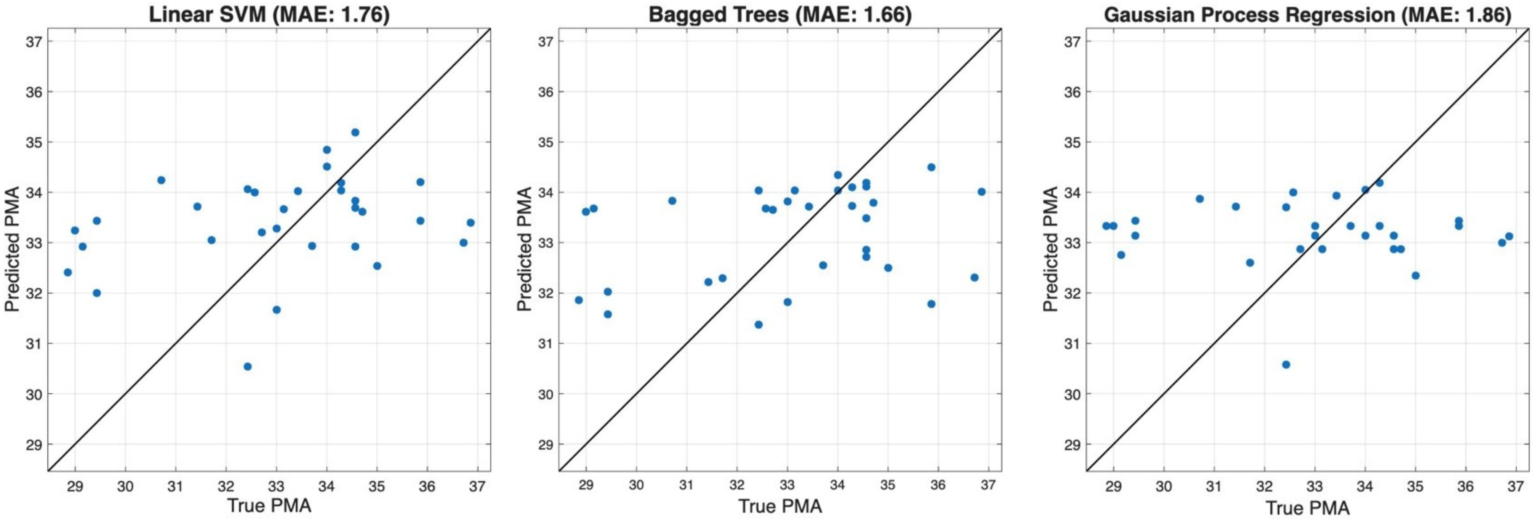


Figure 5: Predicted vs True postmenstrual age (PMA) for three machine learning models with 12 respiratory feature inputs

Each scatter plot compares the PMA with the predicted PMA for 3 machine learning models for all infants in dataset 1. The 12 respiratory feature inputs were selected as they had statistically significant p values ($p < 0.05$) from a simple linear regression model. The black line represents the line of best fit from simple linear regression. The models included are a) Linear SVM 2) Bagged Trees 3) Gaussian Process Regression

Does model accuracy improve with increased data length? (Dataset 2)

50 features were extracted from all participants in dataset 2 due to the addition of the apnoea rate (15 second breathing pauses) feature. 62% of the features (31 of 50) significantly correlated with PMA (Appendix Figure 1). All 3 models had the highest accuracy when trained on all 50 features (Table 7, Figure 6) in comparison to when only features that correlated significantly with age were used and when training the model on the dataset 2 according to the features that significantly correlated with age in dataset 1 (Table 7), though all models performed similarly well. In contrast to the models trained with dataset 1 (Table 7), all three models showed moderate positive correlations between predicted age and true PMA (Table 7, Figure 6, $r = 0.34 - 0.49$, $p < 0.0001$). Nevertheless, there is still a clear poorer performance in the models in the youngest infants (Figure 6).

	All Features			32 Significant Features (Dataset 2)			12 Significant Features (Dataset 1)		
Machine Learning Model	**MAE**	**r**	**p**	**MAE**	**r**	**p**	**MAE**	**r**	**p**
Linear SVM	1.35	0.45	<0.0001	1.45	0.43	<0.0001	1.39	0.34	<0.0001
Bagged Trees	1.30	0.47	<0.0001	1.37	0.47	<0.0001	1.41	0.47	<0.0001
Gaussian Process Regression	1.37	0.49	<0.0001	1.42	0.47	<0.0001	1.40	0.46	<0.0001

Table 7: Mean Absolute Errors (MAEs) and Pearson's correlation r and p-values for machine learning models trained on dataset 2
The mean absolute errors (MAEs) and Pearson's correlation r and p-values are presented for three machine learning models applied to dataset 2. Models were trained using: (1) all 50 features from dataset 2, (2) only

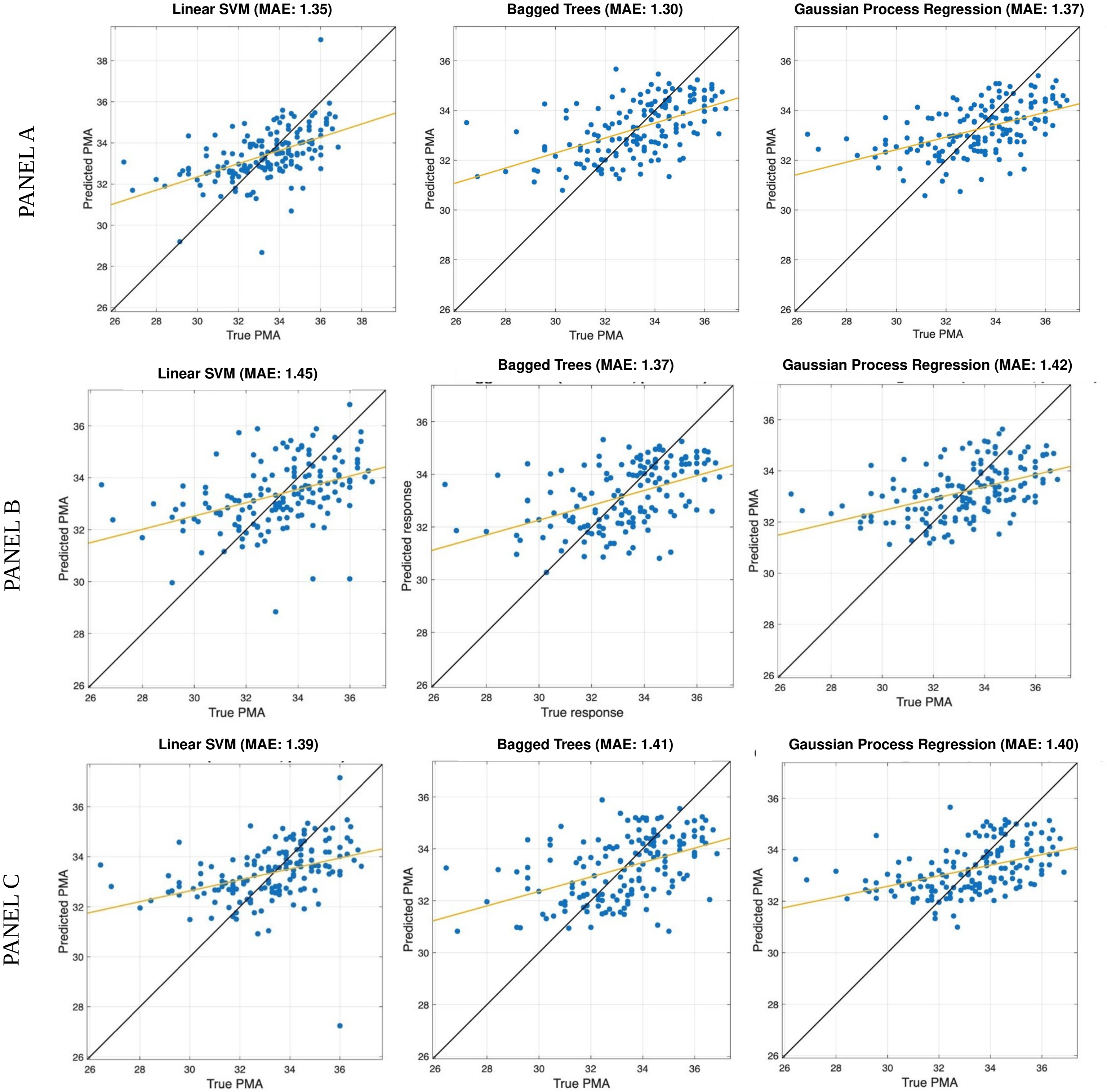


Figure 6: Predicted vs True PMA for 3 machine learning models for 3 rounds of training

Panel A – Models trained on all 50 respiratory features from dataset 2

Panel B – Models trained on 31 respiratory features with statistically significant p-values from simple linear regression models with true PMA from dataset 2

Panel C – Models trained on 12 respiratory features with statistically significant p-values from simple linear regression models with true PMA from dataset 1

The bagged trees model trained on all features was selected for further analysis as it had the lowest MAE (1.30). Key characteristics were investigated to determine whether they impacted model results. There was no significant difference in absolute prediction errors between male and female infants (p = 0.17), or across mode of ventilations (p = 0.79, figure 7). Similarly, there was no significant relationship between post-natal age and absolute prediction error (r = -0.01, p = 0.96, figure 8).

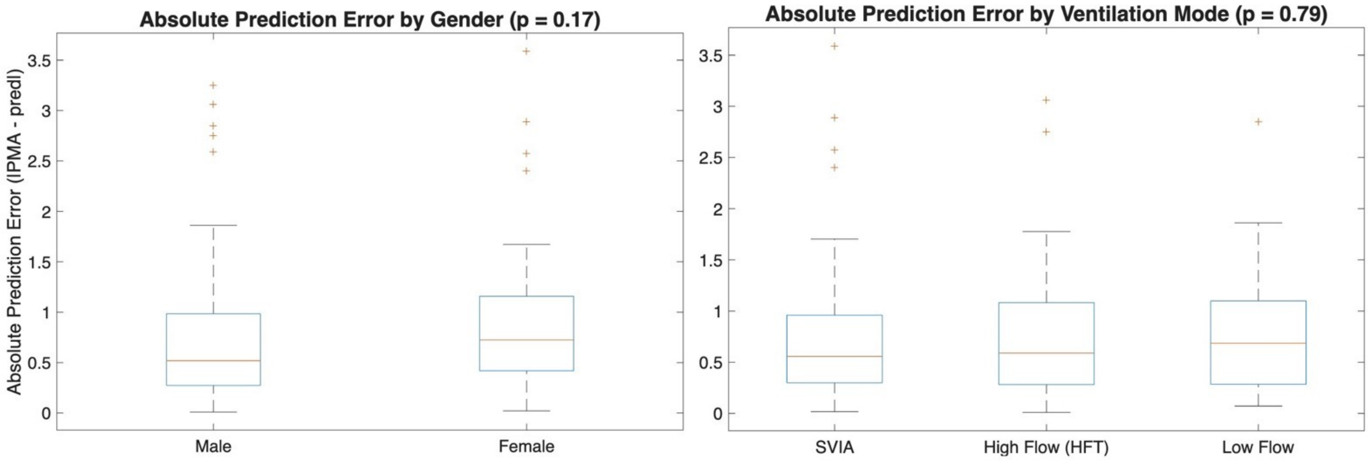


Figure 7: Box and whisker plots comparing differences in predication accuracy of the bagged trees model trained on all 50 features from dataset 2; plot A compares males versus females, and plot B compares the different modes of ventilation at time of recording. Outliers are marked with a '+'.

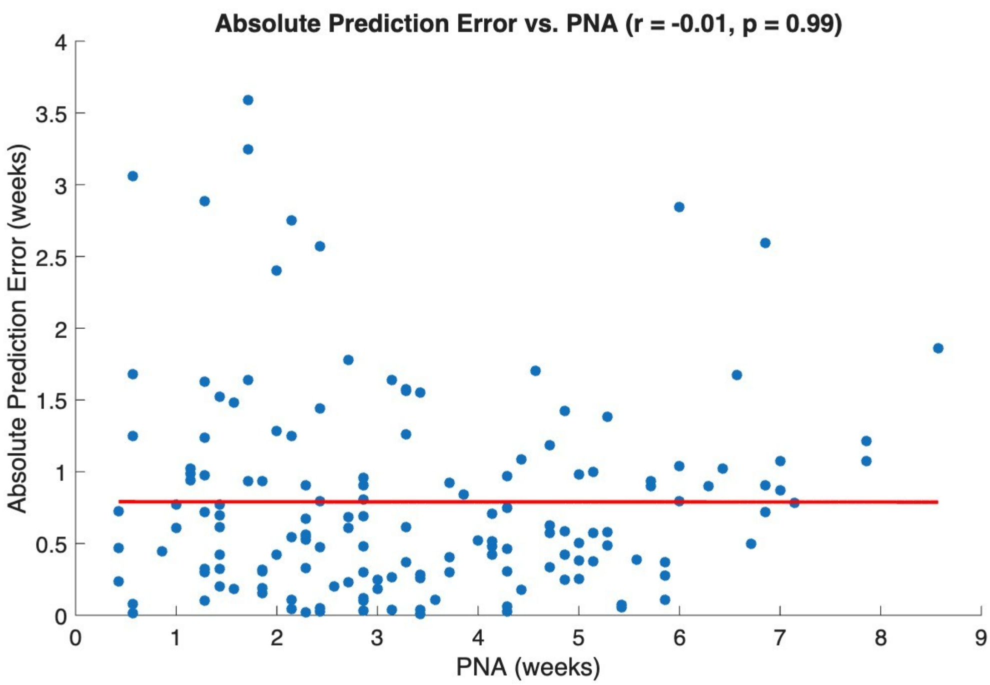


Figure 8: Scatter graph comparing the postnatal age with prediction errors of the bagged trees model trained on all 50 features of dataset 2. Line of best fit in red is plotted via a simple linear regression model.

Discussion

Our study aimed to assess whether the respiratory activity in premature infants could be used to estimate the age at which their respiratory system is functioning, using features extracted from inter-breath intervals. The correlations of these features with PMA were individually assessed using three datasets – one retrospective (dataset 1) and one prospective (with data from dataset 2a and 2b). Multiple machine learning models were then trained and tested to predict the age of premature infants based on these features, and these predictions were compared with the true PMA. The effect of clinical subgroups such as sex and ventilation modes on model accuracy was then assessed, as well as the impact of post-natal age.

A number of features have a significant correlation with PMA across both datasets, emphasising the developing respiratory function in preterm infants over time. The simple linear regression analysis showed that features such as respiratory rate, mean IBI and certain entropy measures had significant correlations with PMA, matching the underlying physiological development of the respiratory system. Specifically, entropy measures such as Shannon and Renyi entropy increasing with age may reflect more regulated and complex breathing patterns due to maturation of central and peripheral mechanisms. Respiratory centres within the brainstem, including the pre-Bötzinger complex, develop more synaptic connections enabling more regulated and responsive breathing to physiological states such as hypoxia and hypercapnia. The simultaneous development of pulmonary structures during the post-canalicular and saccular stages at approximately 24 gestational weeks onwards, increases the surface area for gas exchange and compliance of the alveoli. As a result, tidal volumes start to become more regular and less unstable. In combination, the neural and structural developments of the central respiratory control system (in the brainstem) and the pulmonary system (the lungs and associated gas exchange structures) during this period facilitates a more feedback-driven respiratory rhythm. This leads to an increase in controlled and regulated variability, reflected in the inter-breath intervals, and matches the rise in certain entropy measures with increasing PMA.

Previous studies have shown that apnoea rate decreases with age in preterm infants.(17, 18) Our results also indicated that features such as apnoea rate (defined as breathing pauses of either 5 or 15 seconds), showed significant decreasing patterns with increasing PMA. This likely reflects more consistent baseline respiratory activity and fewer erratic pathological disruptions as commonly seen in preterm infants. Notably, more respiratory features were found to be statistically significant in dataset 2 (31 out of 50) in comparison to dataset 1 (12 out of 49). This most likely can be attributed to the longer recording lengths in datasets 2a and 2b versus dataset 1 (median length 24.0 and 24.7 hours versus 2.0 hours respectively). This extended recording period in dataset 2 means more behavioural and physiological states are captured for each infant, ranging from periods of wakefulness to sleep. These states are reflected in changing respiratory activity and inter-breath intervals. The longer recording periods in dataset 2 likely improved feature extraction accuracy and detection of more subtle changes in inter-breath intervals, leading to a greater number of statistically significant correlations with PMA. It is also interesting to note that 11 of the 12 features with significant correlations in dataset 1 with PMA were also significant in dataset 2, further supporting the robustness of the selected features.

The predictive performance of the machine learning models further illustrated the significant relationship between a number of respiratory features and PMA. Models predicted the age of the infant with an accuracy of just over 1 week, which is similar to prediction models which use EEG and heart rate activity to predict the age of the infant. Future work may want to assess model significance through permutation testing, as done in the EEG study, to assess the statistical significance of our models' accuracy.

The statistically significant MAEs of the linear SVM and bagged trees models (1.3 – 1.4 weeks) in dataset 2, are similar to the accuracies seen by Iyer et al. In their study, a similar group of features were extracted from heart rate variability data and a Gaussian Process Regression model was used to produce a functional autonomic age with a MAE of 1.66 weeks from the true PMA. Moreover, Zandvoort et al. used sensory evoked EEG activity and support vector regression to predict the brain age of preterm infants with a mean absolute error of 1.4-1.5 weeks. Our findings therefore supplement previous work using physiological measurements to assess the development of organ systems in preterm infants. Nevertheless, the line of best fit (correlation between the predicted and true PMA) in each plot does however deviate from the perfect prediction line, particularly at the extreme age weeks. This suggests that predictions across the three models were less accurate for the youngest and oldest infants. This is likely due to the uneven distribution of ages of the infants tested, with most infants in dataset 2 clustered around a more central narrow PMA window. The models have more training data for babies in this range, leading to better predictive performance in comparison to the extremes on either side of the central age range. The correlation coefficients (r values) suggest a moderate predictive strength across all models, despite the consistently significant p-values (table 6). The models are able to capture meaningful and reliable patterns between the respiratory features and true PMA, however the r values ranging from 0.34-0.49 suggests further research is needed to improve model accuracy. Model performance was similar when only using features which correlated with PMA (both from dataset 1 and dataset 2), suggesting that the inclusion of a wider range of features does not significantly degrade model performance. This indicates that the selected models have a high robustness to dimensionality reduction and may be focussing more on the most useful feature inputs. Further work should consider feature selection methods in more detail and focus on data collection of infants younger than 30 weeks PMA.

A number of scoring systems and models within paediatric medicine are tailored specifically for clinical subgroups such as sex. Examples of these include the WHO growth charts, bone age assessments and blood pressure percentiles.(114-116) Looking specifically at respiratory measures, the GLI-2012 score is an example of a paediatric tool that takes into account variation between males and females when measuring lung function.(117) Taking this into account, it is interesting to note that the model performance of the most accurate model, the bagged trees model trained on all dataset 2 features, was not affected by sex, ventilation mode (figure 7) or post-natal age (figure 8). This supports the idea of having one model for all preterm infants over a tailored model for each clinical subgroup or by postnatal age range.

In order to improve our models' performances, particularly across a wider age range, future work should have a focus on obtaining inter-breath interval data from both extremes of the age range. Once a more uniform and significant predictive accuracy is achieved across a wider age range of preterm infants, work could then focus on assessing the predicted PMA in infants with known pathologies such as infections, bronchopulmonary dysplasia and receiving medications known to suppress the respiratory system like morphine. It would be interesting to investigate whether our model's respiratory ages largely deviate from the true PMA of these infants, reflecting the impact of the pathology or medication on respiratory function. The long-term development of such infants with these deviations could be measured to see if large gaps between true and predicted respiratory ages during the neonatal period results in long-term developmental issues. Similar work was done in the Zandvoort et al. paper where large brain age gaps were found to impact long-term neurodevelopment. (23)

As well as the narrow age range of recorded infants, other limitations of our study should be acknowledged. The hyperparameters set by the Regression Learner App on MATLAB for the machine-learning models used were not altered. Ideally, the mean absolute errors from cross-validation would be used as a measure of accuracy as hyperparameters are altered during model training. Once the MAE is optimised, the models with the specifically developed hyperparameters, will then be tested on an independent dataset to better assess generalisability. A fully independent test set was not available for this study under the time constraints set for data collection, so the mean absolute errors may be more optimistic due to the lack of testing. Another limitation is that all data used was from a single hospital meaning all data is from the same respiratory monitoring devices and practices. Data from different hospitals would be useful to assess the model's abilities to be applied for broader clinical application.

The long-term clinical aim of this work is to add to the current vital signs tools central to clinical decision-making in neonatology today. In particular, measures such as heart rate variability have been used to establish tools such as the HeRO score to detect sepsis at earlier stages in preterm infants and therefore reduce mortality. My work forms the foundation blocks for the use of real-time respiratory vital signs data, specifically inter-breath intervals, to aid with risk stratification and encourage earlier intervention for specific pathologies such as bronchopulmonary dysplasia, sepsis, apnoea of prematurity and respiratory depression as a result of opioid use. With further validation and testing on independent datasets and preterm populations with respiratory pathologies, a respiratory age prediction tool could be incorporated into bedside monitoring systems to drive physiology-driven neonatal care.

References

1. WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976. Acta Obstet Gynecol Scand. 1977;56(3):247-53.
2. Preterm birth 2023 [Available from: https://www.who.int/news-room/fact-sheets/detail/preterm-birth#:~:text=Across%20countries%2C%20the%20rate%20of,of%20babies%20born%20in%202020.
3. Papageorghiou AT, Ohuma EO, Altman DG, Todros T, Cheikh Ismail L, Lambert A, et al. International standards for fetal growth based on serial ultrasound measurements: the Fetal Growth Longitudinal Study of the INTERGROWTH-21st Project. Lancet. 2014;384(9946):869-79.
4. Vogel JP, Chawanpaiboon S, Moller AB, Watananirun K, Bonet M, Lumbiganon P. The global epidemiology of preterm birth. Best Pract Res Clin Obstet Gynaecol. 2018;52:3-12.
5. Schaaf JM, Liem SM, Mol BW, Abu-Hanna A, Ravelli AC. Ethnic and racial disparities in the risk of preterm birth: a systematic review and meta-analysis. Am J Perinatol. 2013;30(6):433-50.
6. Waldenstrom U, Aasheim V, Nilsen ABV, Rasmussen S, Pettersson HJ, Shytt E. Adverse pregnancy outcomes related to advanced maternal age compared with smoking and being overweight. Obstet Gynecol. 2014;123(1):104-12.
7. Ruiz M, Goldblatt P, Morrison J, Kukla L, Svancara J, Riitta-Jarvelin M, et al. Mother's education and the risk of preterm and small for gestational age birth: a DRIVERS meta-analysis of 12 European cohorts. J Epidemiol Community Health. 2015;69(9):826-33.
8. UNICEF W, World Bank Group and United Nations. Levels and Trends in Child Mortality Report 2017. 2017.
9. Mwaniki MK, Atieno M, Lawn JE, Newton CR. Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review. Lancet. 2012;379(9814):445-52.

10. Araujo BF, Zatti H, Madi JM, Coelho MB, Olmi FB, Canabarro CT. Analysis of neonatal morbidity and mortality in late-preterm newborn infants. J Pediatr (Rio J). 2012;88(3):259-66.
11. Mahoney AD, Jain L. Respiratory disorders in moderately preterm, late preterm, and early term infants. Clin Perinatol. 2013;40(4):665-78.
12. Ramachandrappa A, Rosenberg ES, Wagoner S, Jain L. Morbidity and mortality in late preterm infants with severe hypoxic respiratory failure on extra-corporeal membrane oxygenation. J Pediatr. 2011;159(2):192-8 e3.
13. Roth-Kleiner M, Wagner BP, Bachmann D, Pfenninger J. Respiratory distress syndrome in near-term babies after caesarean section. Swiss Med Wkly. 2003;133(19-20):283-8.
14. Hack M, Fanaroff AA, Klaus MH, Mendelawitz BD, Merkatz IR. Neonatal respiratory distress following elective delivery. A preventable disease? Am J Obstet Gynecol. 1976;126(1):43-7.
15. Serenius F, Ewald U, Farooqi A, Holmgren PA, Hakansson S, Sedin G. Short-term outcome after active perinatal management at 23-25 weeks of gestation. A study from two Swedish tertiary care centres. Part 2: infant survival. Acta Paediatr. 2004;93(8):1081-9.
16. Walsh MC, Morris BH, Wrage LA, Vohr BR, Poole WK, Tyson JE, et al. Extremely low birthweight neonates with protracted ventilation: mortality and 18-month neurodevelopmental outcomes. J Pediatr. 2005;146(6):798-804.
17. Henderson-Smart DJ. The effect of gestational age on the incidence and duration of recurrent apnoea in newborn babies. Aust Paediatr J. 1981;17(4):273-6.
18. Fairchild K, Mohr M, Paget-Brown A, Tabacaru C, Lake D, Delos J, et al. Clinical associations of immature breathing in preterm infants: part 1-central apnea. Pediatr Res. 2016;80(1):21-7.
19. Erickson G, Dobson NR, Hunt CE. Immature control of breathing and apnea of prematurity: the known and unknown. J Perinatol. 2021;41(9):2111-23.
20. Hunt CE, Corwin MJ, Weese-Mayer DE, Ward SL, Ramanathan R, Lister G, et al. Longitudinal assessment of hemoglobin oxygen saturation in preterm and term infants in the first six months of life. J Pediatr. 2011;159(3):377-83 e1.
21. Warburton A, Monga R, Sampath V, Kumar N. Continuous pulse oximetry and respiratory rate trends predict short-term respiratory and growth outcomes in premature infants. Pediatr Res. 2019;85(4):494-501.
22. Iyer KK, Leitner U, Giordano V, Roberts JA, Vanhatalo S, Klebermass-Schrehof K, et al. Bedside tracking of functional autonomic age in preterm infants. Pediatr Res. 2023;94(1):206-12.
23. Zandvoort CS, van der Vaart M, Robinson S, Usman F, Schmidt Mellado G, Evans Fry R, et al. Sensory event-related potential morphology predicts age in premature infants. Clin Neurophysiol. 2024;157:61-72.
24. Stevenson NJ, Oberdorfer L, Tataranno ML, Breakspear M, Colditz PB, de Vries LS, et al. Automated cot-side tracking of functional brain age in preterm infants. Ann Clin Transl Neurol. 2020;7(6):891-902.
25. Ansari A, Pillay K, Arasteh E, Dereymaeker A, Mellado GS, Jansen K, et al. Resting state electroencephalographic brain activity in neonates can predict age and is indicative of neurodevelopmental outcome. Clin Neurophysiol. 2024;163:226-35.
26. Condic ML. Totipotency: what it is and what it is not. Stem Cells Dev. 2014;23(8):796-812.
27. Zhang HT, Hiiragi T. Symmetry Breaking in the Mammalian Embryo. Annu Rev Cell Dev Biol. 2018;34:405-26.
28. Muhr J AT, Ackerman KM. Embryology, Gastrulation.2023.

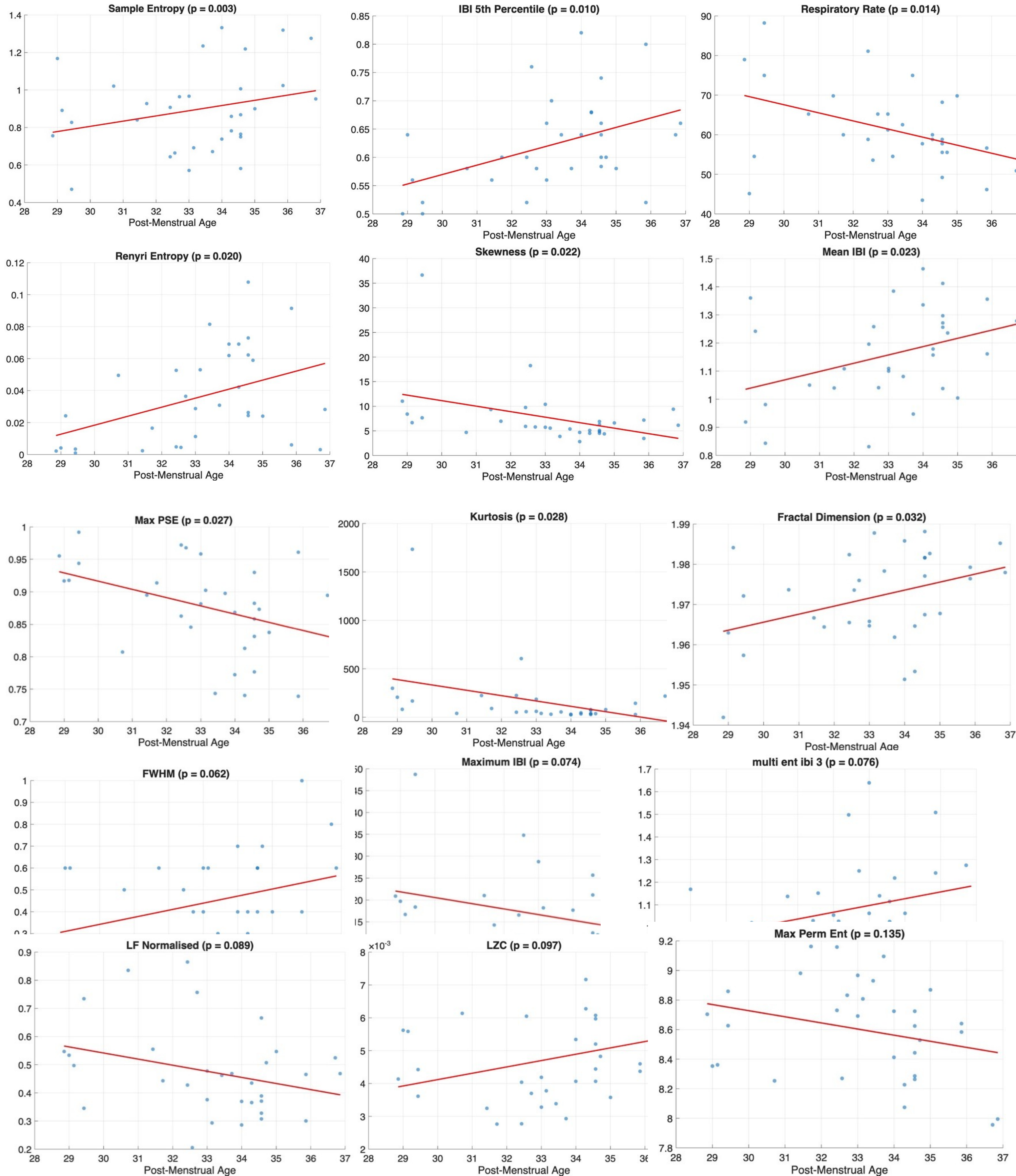
29. Marikawa Y, Alarcon VB. Establishment of trophectoderm and inner cell mass lineages in the mouse embryo. Mol Reprod Dev. 2009;76(11):1019-32.
30. Niu Y, Sun N, Li C, Lei Y, Huang Z, Wu J, et al. Dissecting primate early post-implantation development using long-term in vitro embryo culture. Science. 2019;366(6467).
31. Persaud TVNT, M. G. . The developing human : clinically oriented embryology.: Philadelphia, PA: Elsevier.; 2025.
32. Sadler TW. Langman's Medical embryology.: Philadelphia Wolters Kluwer.; 2024.
33. Caldeira I, Fernandes-Silva H, Machado-Costa D, Correia-Pinto J, Moura RS. Developmental Pathways Underlying Lung Development and Congenital Lung Disorders. Cells. 2021;10(11).
34. Schittny JC. Development of the lung. Cell Tissue Res. 2017;367(3):427-44.
35. Pan J, Copland I, Post M, Yeger H, Cutz E. Mechanical stretch-induced serotonin release from pulmonary neuroendocrine cells: implications for lung development. Am J Physiol Lung Cell Mol Physiol. 2006;290(1):L185-93.
36. Kitaoka H, Burri PH, Weibel ER. Development of the human fetal airway tree: analysis of the numerical density of airway endtips. Anat Rec. 1996;244(2):207-13.
37. Vu TH, Alemayehu Y, Werb Z. New insights into saccular development and vascular formation in lung allografts under the renal capsule. Mech Dev. 2003;120(3):305-13.
38. Herriges M, Morrisey EE. Lung development: orchestrating the generation and regeneration of a complex organ. Development. 2014;141(3):502-13.
39. Madurga A, Mizikova I, Ruiz-Camp J, Morty RE. Recent advances in late lung development and the pathogenesis of bronchopulmonary dysplasia. Am J Physiol Lung Cell Mol Physiol. 2013;305(12):L893-905.
40. Nikolic MZ, Sun D, Rawlins EL. Human lung development: recent progress and new challenges. Development. 2018;145(16).
41. Hyde DM, Blozis SA, Avdalovic MV, Putney LF, Dettorre R, Quesenberry NJ, et al. Alveoli increase in number but not size from birth to adulthood in rhesus monkeys. Am J Physiol Lung Cell Mol Physiol. 2007;293(3):L570-9.
42. Caduff JH, Fischer LC, Burri PH. Scanning electron microscope study of the developing microvasculature in the postnatal rat lung. Anat Rec. 1986;216(2):154-64.
43. Jobe AH, Ikegami M. Lung development and function in preterm infants in the surfactant treatment era. Annu Rev Physiol. 2000;62:825-46.
44. Colin AA, McEvoy C, Castile RG. Respiratory morbidity and lung function in preterm infants of 32 to 36 weeks' gestational age. Pediatrics. 2010;126(1):115-28.
45. Rodesch F, Simon P, Donner C, Jauniaux E. Oxygen measurements in endometrial and trophoblastic tissues during early pregnancy. Obstet Gynecol. 1992;80(2):283-5.
46. Groenman F, Rutter M, Caniggia I, Tibboel D, Post M. Hypoxia-inducible factors in the first trimester human lung. J Histochem Cytochem. 2007;55(4):355-63.
47. De Paepe ME, Mao Q, Powell J, Rubin SE, DeKoninck P, Appel N, et al. Growth of pulmonary microvasculature in ventilated preterm infants. Am J Respir Crit Care Med. 2006;173(2):204-11.
48. Reuter S, Moser C, Baack M. Respiratory distress in the newborn. Pediatr Rev. 2014;35(10):417-28; quiz 29.
49. Duggan M, Kavanagh BP. Pulmonary atelectasis: a pathogenic perioperative entity. Anesthesiology. 2005;102(4):838-54.
50. Sweet DG, Carnielli V, Greisen G, Hallman M, Ozek E, Plavka R, et al. European consensus guidelines on the management of neonatal respiratory distress syndrome in preterm infants--2013 update. Neonatology. 2013;103(4):353-68.

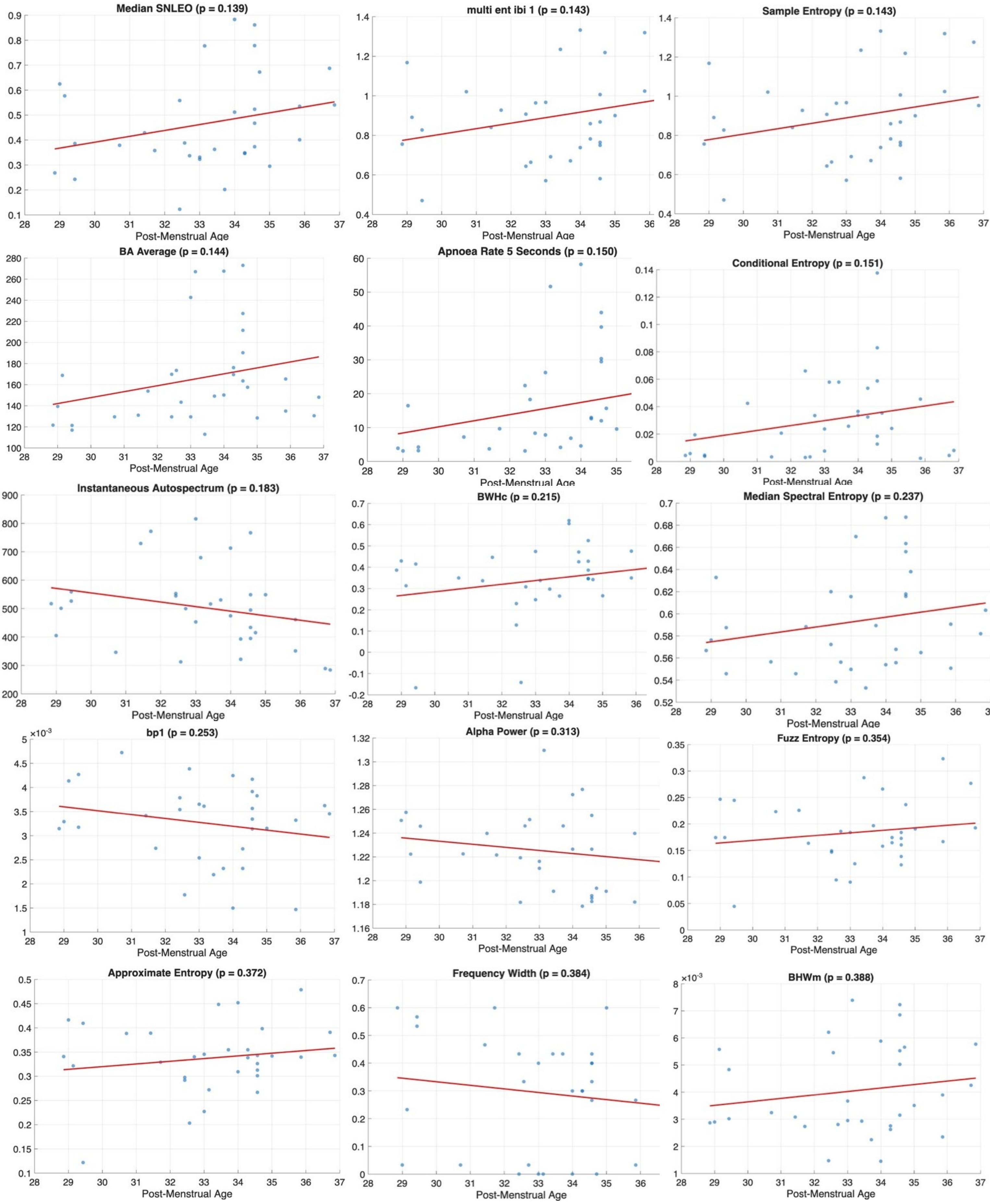
51. Liu J, Yang N, Liu Y. High-risk Factors of Respiratory Distress Syndrome in Term Neonates: A Retrospective Case-control Study. Balkan Med J. 2014;31(1):64-8.
52. Anadkat JS, Kuzniewicz MW, Chaudhari BP, Cole FS, Hamvas A. Increased risk for respiratory distress among white, male, late preterm and term infants. J Perinatol. 2012;32(10):780-5.
53. Dynamed. Respiratory Distress Syndrome (RDS) of the Newborn.: EBSCO Information Services; 2025 [Available from: https://www.dynamed.com/condition/respiratory-distress-syndrome-rds-of-the-newborn-23.
54. Dynamed. Neonatal Resuscitation 2025 [Available from: https://www.dynamed.com/management/neonatal-resuscitation.
55. Sweet DG, Carnielli VP, Greisen G, Hallman M, Klebermass-Schrehof K, Ozek E, et al. European Consensus Guidelines on the Management of Respiratory Distress Syndrome: 2022 Update. Neonatology. 2023;120(1):3-23.
56. Kinsella JP, Greenough A, Abman SH. Bronchopulmonary dysplasia. Lancet. 2006;367(9520):1421-31.
57. Northway WH, Jr., Rosan RC, Porter DY. Pulmonary disease following respirator therapy of hyaline-membrane disease. Bronchopulmonary dysplasia. N Engl J Med. 1967;276(7):357-68.
58. Bonikos DS, Bensch KG, Ludwin SK, Northway WH, Jr. Oxygen toxicity in the newborn. The effect of prolonged 100 per cent O2 exposure on the lungs of newborn mice. Lab Invest. 1975;32(5):619-35.
59. Crapo JD, Peters-Golden M, Marsh-Salin J, Shelburne JS. Pathologic changes in the lungs of oxygen-adapted rats: a morphometric analysis. Lab Invest. 1978;39(6):640-53.
60. Kraybill EN, Runyan DK, Bose CL, Khan JH. Risk factors for chronic lung disease in infants with birth weights of 751 to 1000 grams. J Pediatr. 1989;115(1):115-20.
61. Gonzalez A, Sosenko IR, Chandar J, Hummler H, Claure N, Bancalari E. Influence of infection on patent ductus arteriosus and chronic lung disease in premature infants weighing 1000 grams or less. J Pediatr. 1996;128(4):470-8.
62. Watterberg KL, Demers LM, Scott SM, Murphy S. Chorioamnionitis and early lung inflammation in infants in whom bronchopulmonary dysplasia develops. Pediatrics. 1996;97(2):210-5.
63. Bryan MH, Hardie MJ, Reilly BJ, Swyer PR. Pulmonary function studies during the first year of life in infants recovering from the respiratory distress syndrome. Pediatrics. 1973;52(2):169-78.
64. Woodgate PG, Davies MW. Permissive hypercapnia for the prevention of morbidity and mortality in mechanically ventilated newborn infants. Cochrane Database Syst Rev. 2001;2001(2):CD002061.
65. Greenough A, Rossor TE, Sundaresan A, Murthy V, Milner AD. Synchronized mechanical ventilation for respiratory support in newborn infants. Cochrane Database Syst Rev. 2016;9(9):CD000456.
66. Bollen CW, Uiterwaal CS, van Vught AJ. Cumulative metaanalysis of high-frequency versus conventional ventilation in premature neonates. Am J Respir Crit Care Med. 2003;168(10):1150-5.
67. Van Marter LJ, Allred EN, Pagano M, Sanocka U, Parad R, Moore M, et al. Do clinical markers of barotrauma and oxygen toxicity explain interhospital variation in rates of chronic lung disease? The Neonatology Committee for the Developmental Network. Pediatrics. 2000;105(6):1194-201.

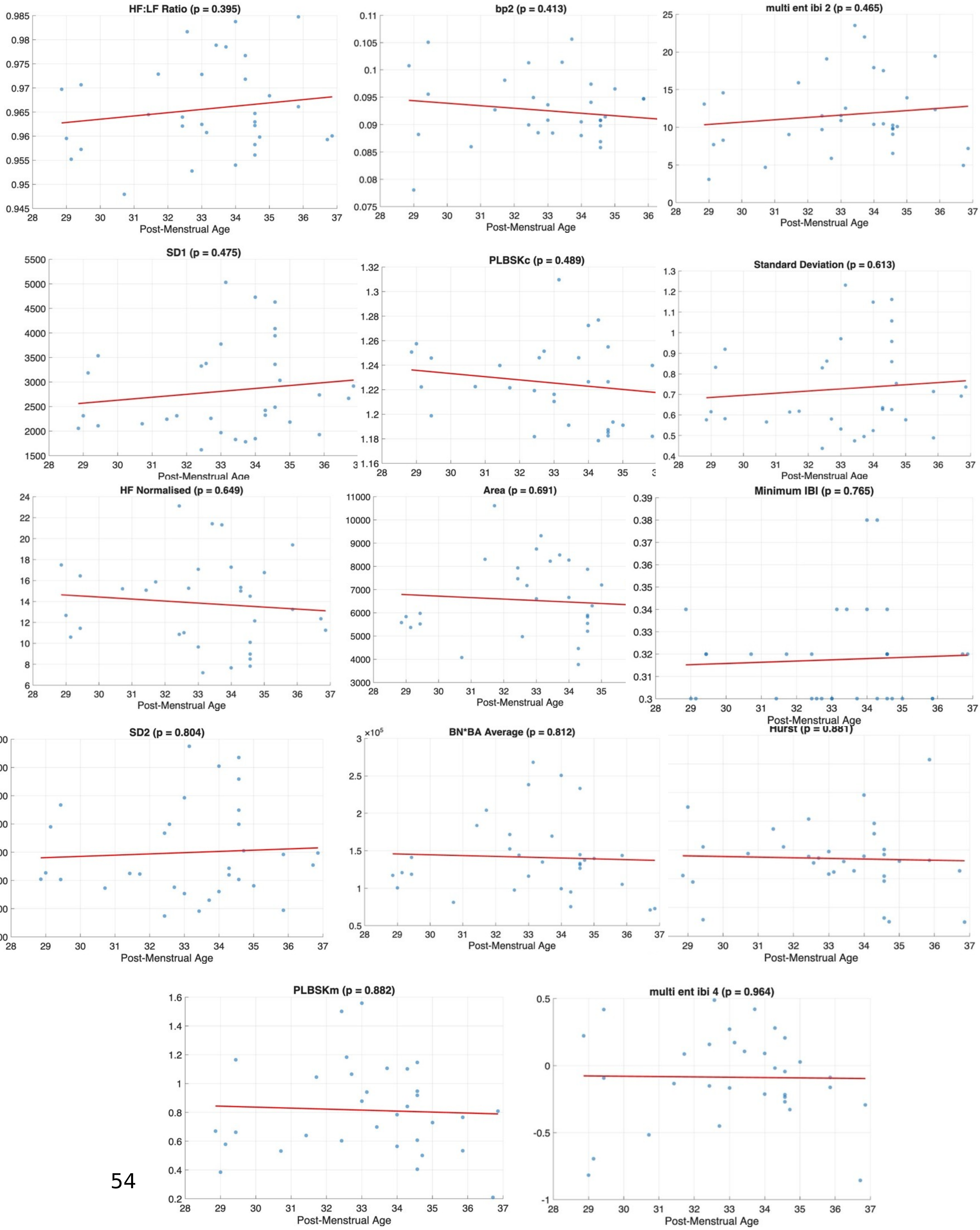
68. Yoder MC, Jr., Chua R, Tepper R. Effect of dexamethasone on pulmonary inflammation and pulmonary function of ventilator-dependent infants with bronchopulmonary dysplasia. Am Rev Respir Dis. 1991;143(5 Pt 1):1044-8.
69. Halliday HL. Clinical trials of postnatal corticosteroids: inhaled and systemic. Biol Neonate. 1999;76 Suppl 1:29-40.
70. Committee on F, Newborn. Postnatal corticosteroids to treat or prevent chronic lung disease in preterm infants. Pediatrics. 2002;109(2):330-8.
71. Doyle LW, Halliday HL, Ehrenkranz RA, Davis PG, Sinclair JC. Impact of postnatal systemic corticosteroids on mortality and cerebral palsy in preterm infants: effect modification by risk for chronic lung disease. Pediatrics. 2005;115(3):655-61.
72. Finer NN, Higgins R, Kattwinkel J, Martin RJ. Summary proceedings from the apnea-of-prematurity group. Pediatrics. 2006;117(3 Pt 2):S47-51.
73. Eichenwald EC, Aina A, Stark AR. Apnea frequently persists beyond term gestation in infants delivered at 24 to 28 weeks. Pediatrics. 1997;100(3 Pt 1):354-9.
74. Aranda JV, Turmen T. Methylxanthines in apnea of prematurity. Clin Perinatol. 1979;6(1):87-108.
75. Seo Hee Choi M, Juyoung Lee, MD, PhD1,2, Soo Kyung Nam, MD3, Yong Hoon Jun, MD, PhD1,. Cerebral Oxygenation during Apnea in Preterm Infants: Effects of Accompanying Peripheral Oxygen Desaturation. Neonatal Medicine. 2021;2021;28(1):14-21.
76. Janvier A, Khairy M, Kokkotis A, Cormier C, Messmer D, Barrington KJ. Apnea is associated with neurodevelopmental impairment in very low birth weight infants. J Perinatol. 2004;24(12):763-8.
77. Coen S. Zandvoorta AD, Luke Baxtera, Katrien Jansenb,c, Gunnar Naulaersb, Maarten de Vosc,d, Caroline Hartleya. Apnoea suppresses brain activity in infants. Imaging Neuroscience. 2024;Volume 2, 2024.
78. Martin RJ, Wilson CG. Apnea of prematurity. Compr Physiol. 2012;2(4):2923-31.
79. Marshall AP, Lim K, Ali SK, Gale TJ, Dargaville PA. Physiological instability after respiratory pauses in preterm infants. Pediatr Pulmonol. 2019;54(11):1712-21.
80. Di Fiore JM, Martin RJ, Gauda EB. Apnea of prematurity--perfect storm. Respir Physiol Neurobiol. 2013;189(2):213-22.
81. Goldstein B, Giroir B, Randolph A, International Consensus Conference on Pediatric S. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. Pediatr Crit Care Med. 2005;6(1):2-8.
82. Glaser MA, Hughes LM, Jnah A, Newberry D. Neonatal Sepsis: A Review of Pathophysiology and Current Management Strategies. Adv Neonatal Care. 2021;21(1):49-60.
83. Afonso EDP, Blot S. Effect of gestational age on the epidemiology of late-onset sepsis in neonatal intensive care units - a review. Expert Rev Anti Infect Ther. 2017;15(10):917-24.
84. Simonsen KA, Anderson-Berry AL, Delair SF, Davies HD. Early-onset neonatal sepsis. Clin Microbiol Rev. 2014;27(1):21-47.
85. Shane AL, Sanchez PJ, Stoll BJ. Neonatal sepsis. Lancet. 2017;390(10104):1770-80.
86. Stockmann C, Spigarelli MG, Campbell SC, Constance JE, Courter JD, Thorell EA, et al. Considerations in the pharmacologic treatment and prevention of neonatal sepsis. Paediatr Drugs. 2014;16(1):67-81.
87. Dong Y, Glaser K, Speer CP. Late-onset sepsis caused by Gram-negative bacteria in very low birth weight infants: a systematic review. Expert Rev Anti Infect Ther. 2019;17(3):177-88.

88. Boettiger M, Tyer-Viola L, Hagan J. Nurses' Early Recognition of Neonatal Sepsis. J Obstet Gynecol Neonatal Nurs. 2017;46(6):834-45.
89. Iroh Tam PY, Bendel CM. Diagnostics for neonatal sepsis: current approaches and future directions. Pediatr Res. 2017;82(4):574-83.
90. Das A, Shukla S, Rahman N, Gunzler D, Abughali N. Clinical Indicators of Late-Onset Sepsis Workup in Very Low-Birth-Weight Infants in the Neonatal Intensive Care Unit. Am J Perinatol. 2016;33(9):856-60.
91. Eichenwald EC HA, Martin CR, et al. Cloherty and Stark's Manual of Neonatal Care, 8e. Lippincott Williams & Wilkins, a Wolters Kluwer business2017.
92. Taketomo CK, Hodding, J. H., Kraus, D. M. Lexicomp pediatric & neonatal dosage handbook : an extensive resource for clinicians treating pediatric and neonatal patients Lexicomp/Wolters Kluwer.; 2018.
93. Schuurmans J, Benders M, Lemmers P, van Bel F. Neonatal morphine in extremely and very preterm neonates: its effect on the developing brain - a review. J Matern Fetal Neonatal Med. 2015;28(2):222-8.
94. Kwok TC, Henry C, Saffaran S, Meeus M, Bates D, Van Laere D, et al. Application and potential of artificial intelligence in neonatal medicine. Semin Fetal Neonatal Med. 2022;27(5):101346.
95. Formulary BN. Joint Formulary Committee. Morphine (indications and dose) 2024 [Available from: https://bnfc.nice.org.uk/drugs/morphine/#indications-and-dose.
96. Bateman JT, Saunders SE, Levitt ES. Understanding and countering opioid-induced respiratory depression. Br J Pharmacol. 2023;180(7):813-28.
97. Smith JC, Ellenberger HH, Ballanyi K, Richter DW, Feldman JL. Pre-Botzinger complex: a brainstem region that may generate respiratory rhythm in mammals. Science. 1991;254(5032):726-9.
98. Keles E, Bagci U. The past, current, and future of neonatal intensive care units with artificial intelligence: a systematic review. NPJ Digit Med. 2023;6(1):220.
99. McAdams RM, Kaur R, Sun Y, Bindra H, Cho SJ, Singh H. Predicting clinical outcomes using artificial intelligence and machine learning in neonatal intensive care units: a systematic review. J Perinatol. 2022;42(12):1561-75.
100. Hanson CW, 3rd, Marshall BE. Artificial intelligence applications in the intensive care unit. Crit Care Med. 2001;29(2):427-35.
101. Ostojic D, Guglielmini S, Moser V, Fauchere JC, Bucher HU, Bassler D, et al. Reducing False Alarm Rates in Neonatal Intensive Care: A New Machine Learning Approach. Adv Exp Med Biol. 2020;1232:285-90.
102. 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. 2020;87(2):210-20.
103. 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. 2021;3(1):e0302.
104. Sullivan BA, McClure C, Hicks J, Lake DE, Moorman JR, Fairchild KD. Early Heart Rate Characteristics Predict Death and Morbidities in Preterm Infants. J Pediatr. 2016;174:57-62.
105. Moorman JR, Carlo WA, Kattwinkel J, Schelonka RL, Porcelli PJ, Navarrete CT, et al. Mortality reduction by heart rate characteristic monitoring in very low birth weight neonates: a randomized trial. J Pediatr. 2011;159(6):900-6 e1.
106. Zeigler AC, Ainsworth JE, Fairchild KD, Wynn JL, Sullivan BA. Sepsis and Mortality Prediction in Very Low Birth Weight Infants: Analysis of HeRO and nSOFA. Am J Perinatol. 2023;40(4):407-14.

107. R. D. Shirwaikar NM, Dinesh Acharya U, K. Makkithaya and Govardhan Hegde K. Supervised Learning techniques for analysis of neonatal data. 2nd International Conference on Applied and Theoretical Computing and Communication Technology. 2016:25-31.
108. Shahriar RAAHS. Supervised and Unsupervised-Based Analytics of Intensive Care Unit Data. 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)2019, July.
109. Nemati S, Ghassemi MM, Clifford GD. Optimal medication dosing from suboptimal clinical examples: a deep reinforcement learning approach. Annu Int Conf IEEE Eng Med Biol Soc. 2016;2016:2978-81.
110. Aniruddh Raghu MK, Leo Anthony Celi, Peter Szolovits, Marzyeh Ghassemi. Continuous State-Space Models for Optimal Sepsis Treatment: a Deep Reinforcement Learning Approach. Proceedings of the 2nd Machine Learning for Healthcare Conference2017.
111. Fatima Usman CSZ, Shellie Robinson, Mariska Peck, Maria M. Cobo, Tricia Adjei, Luke Baxter, Ria Evans Fry, Annalisa G. V. Hauck, Richard Rogers, Gabriela Schmidt Mellado, Alexandra Scrivens, Marianne van der Vaart, Maarten De Vos, Eleri Adams, John van den Anker, Caroline Hartley. Brain activity as a biomarker for personalised caffeine treatment in premature neonates
. medRxiv. 2025.
112. Adjei T, Purdy R, Jorge J, Adams E, Buckle M, Evans Fry R, et al. New method to measure interbreath intervals in infants for the assessment of apnoea and respiration. BMJ Open Respir Res. 2021;8(1).
113. Lin C-C YC-Y, Zhou Z, Wu S. ntelligent health monitoring system based on smart clothing. International Journal of Distributed Sensor Networks. 2018;2018;14(8).
114. Team HP. UK-WHO growth charts - 0-4 years 2024 [Available from: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-0-4-years.
115. Minty R, Mahomed N, van Wyk N, Mndebele G, Lockhat Z, Ranchod A. Comparison of bone age assessment using manual Greulich and Pyle method versus automated BoneXpert method in South African children. SA J Radiol. 2025;29(1):3033.
116. What is High Blood Pressure? : National Heart, Lung, and blood Institute 2024 [Available from: https://www.nhlbi.nih.gov/health/high-blood-pressure.
117. GLI Tools - A collection of tools developed by the GLI project.: European Respiratory Society; [Available from: https://www.ersnet.org/science-and-research/ongoing-clinical-research-collaborations/the-global-lung-function-initiative/gli-tools/.


Sample Entropy (p = 0.003)
Post-Menstrual Age
IBI 5th Percentile (p = 0.010)
Post-Menstrual Age
Respiratory Rate (p = 0.014)
Post-Menstrual Age
Renyri Entropy (p = 0.020)
Post-Menstrual Age
Skewness (p = 0.022)
Post-Menstrual Age
Mean IBI (p = 0.023)
Post-Menstrual Age
Max PSE (p = 0.027)
Post-Menstrual Age
Kurtosis (p = 0.028)
Post-Menstrual Age
Fractal Dimension (p = 0.032)
Post-Menstrual Age
FWHM (p = 0.062)
Maximum IBI (p = 0.074)
multi ent ibi 3 (p = 0.076)
LF Normalised (p = 0.089)
Post-Menstrual Age
LZC (p = 0.097)
×10-3
Post-Menstrual Age
Max Perm Ent (p = 0.135)
Post-Menstrual Age


Median SNLEO (p = 0.139)
Post-Menstrual Age
multi ent ibi 1 (p = 0.143)
Post-Menstrual Age
Sample Entropy (p = 0.143)
Post-Menstrual Age
BA Average (p = 0.144)
Post-Menstrual Age
Apnoea Rate 5 Seconds (p = 0.150)
Post-Menstrual Age
Conditional Entropy (p = 0.151)
Post-Menstrual Age
Instantaneous Autospectrum (p = 0.183)
Post-Menstrual Age
BWHc (p = 0.215)
Post-Menstrual Age
Median Spectral Entropy (p = 0.237)
Post-Menstrual Age
×10^-3
bp1 (p = 0.253)
Post-Menstrual Age
Alpha Power (p = 0.313)
Post-Menstrual Age
Fuzz Entropy (p = 0.354)
Post-Menstrual Age
Approximate Entropy (p = 0.372)
Post-Menstrual Age
Frequency Width (p = 0.384)
Post-Menstrual Age
×10^-3
BHWm (p = 0.388)
Post-Menstrual Age


HF:LF Ratio (p = 0.395)
bp2 (p = 0.413)
multi ent ibi 2 (p = 0.465)
SD1 (p = 0.475)
PLBSKc (p = 0.489)
Standard Deviation (p = 0.613)
HF Normalised (p = 0.649)
Area (p = 0.691)
Minimum IBI (p = 0.765)
SD2 (p = 0.804)
BN*BA Average (p = 0.812)
×10^5
Hurst (p = 0.881)
PLBSKm (p = 0.882)
multi ent ibi 4 (p = 0.964)
Post-Menstrual Age

Appendix Figure 1: Scatter plots showing the relationship between 48 respiratory features and postmenstrual age (PMA) for dataset 1. The plot for range has been removed as it has almost identical data with the maximum IBI data. The plots are in order of ascending p-values derived from simple linear regression models. The red line represents the line of best fit from simple linear regression models.