Brain activity as a biomarker for personalised caffeine treatment in premature neonates

Fatima Usman[1#], Coen S. Zandvoort[1#], Shellie Robinson[1], Mariska Peck[1], Maria M. Cobo[1,2], Tricia Adjei[1], Luke Baxter[1], Ria Evans Fry[1], Annalisa G. V. Hauck[1], Richard Rogers[3], Gabriela Schmidt Mellado[1], Alexandra Scrivens[1], Marianne van der Vaart[1], Maarten De Vos[4,5], Eleri Adams[6], John van den Anker[1,7] and Caroline Hartley[1]*

1. Department of Paediatrics, University of Oxford, Oxford, UK
2. Universidad San Francisco de Quito USFQ, Colegio de Ciencias Biologicas y Ambientales, Quito, Ecuador
3. Nuffield Division of Anaesthetics, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
4. Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium
5. Department of Development and Regeneration, University Hospitals Leuven, Child Neurology, KU Leuven, Leuven, Belgium
6. Newborn Care Unit, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
7. Division of Clinical Pharmacology, Children's National Hospital, Washington DC, USA

# These authors contributed equally

* Corresponding author: Caroline Hartley, Department of Paediatrics, University of Oxford, UK.
caroline.hartley@paediatrics.ox.ac.uk

Abstract

Medication in hospitalised infants is often prescribed using a ‘one-size-fits-all’ approach due to lack of clinical biomarkers. Caffeine is one of the most frequently administered medicines in neonatology - prescribed for the management of apnoea of prematurity, to aid extubation and increasingly for conditions such as bronchopulmonary dysplasia. Caffeine guidelines for the management of apnoea of prematurity indicate use based on age of the infant, but this does not account for individual variation in apnoea rate. Consequently, infants may risk caffeine undertreatment or adverse events due to over-exposure. Apnoea in preterm infants is related to brainstem immaturity, hence as an essential first step to assessing whether brain maturity may be a useful biomarker for caffeine treatment, we tested the hypothesis that apnoea rate is related to brain maturity. Using electroencephalography (EEG), we demonstrate that apnoea rate in late preterm infants is dependent on brain maturity, not postmenstrual age. We provide initial evidence that when caffeine is discontinued, infants with immature brain activity have more frequent apnoeas and desaturations compared with those with more mature brain function. These findings indicate that brain maturity (assessed automatically using machine learning) is a candidate biomarker for personalised caffeine treatment in preterm infants.

Main text

Apnoea of prematurity (AOP) is a developmental respiratory disorder affecting more than half of the 15 million infants born prematurely each year worldwide[1,2]. Apnoea is a breathing cessation of variable duration which can be life threatening and may impact long-term neurodevelopmental outcomes[3,4], so it is essential to optimise treatment. AOP is a consequence of nervous system immaturity of preterm infants[5] and will usually resolve by late prematurity[6].

Caffeine is the primary pharmacological treatment for AOP, reducing apnoea rates, duration of ventilatory support and extubation failure, and improving neurodevelopmental outcomes[7]. Clinical guidelines for the treatment and prevention of AOP indicate caffeine requirement based on gestational age or weight of an infant, for example, the World Health Organisation recommends all infants born before 34 weeks' gestation are given caffeine[8]. However, prescription based on age does not account for individual variability in caffeine requirement nor directly relate to the underlying cause of apnoea (i.e., immaturity of the brain) (Figure 1). This exposes some infants to caffeine-related adverse effects, such as tachycardia and seizures[9,10]. Conversely, some infants discontinue treatment too early and experience substantial apnoeas[11]. An individualised approach is needed[12] and can be achieved through identification of biomarkers for caffeine requirement. Currently, no such biomarkers have been identified.

Given the direct link between brain development and apnoea, here we propose brain maturity as a candidate biomarker (Figure 1). We define brain maturity as the difference between an infant's 'brain age' and their postmenstrual age (PMA). Thus, brain maturity is a marker of whether an individual's brain activity is more or less mature relative to their PMA, with negative values of brain maturity indicating that their brain activity is immature compared with their PMA and positive values indicating that their brain activity is more mature than their PMA. Brain age can be assessed using electroencephalography (EEG), combined with recently developed and validated machine learning models, and uses features in the EEG compared with the model training data to ascertain an infant's brain age[13,14]. These models capitalise on the fact that the EEG changes rapidly in preterm infants – the EEG of a very preterm infant is characterised by discontinuous activity interspersed with high amplitude bursts. Brain activity becomes increasingly continuous and lower in amplitude with increasing age[15]. Similarly, sensory-evoked responses change with PMA [16–18]. Very immature brain activity relative to PMA may be associated with poorer later life neurodevelopmental outcomes[13,14], yet there will also be normal variation in brain maturity – just as older children learn to walk and talk at different ages, brain developmental trajectories exhibit individual variation modulated by factors such as genetics and the environment[19,20]. We postulate that this individual variation in brain maturity (i.e., individual variation in EEG activity) will explain variation in apnoea rate and consequently could be used to guide treatment with caffeine. EEG is currently used clinically to assess seizures to determine the need for anticonvulsants, and can be used within depth of anaesthesia monitors, thus, guiding the level of anaesthetics[21,22]. However, it is not currently routinely used to assess other medication in infants, including caffeine.

There is limited evidence to date to suggest that brain activity in human preterm infants is related to apnoea rate: Henderson-Smart and colleagues[23] demonstrated prolonged auditory brainstem conduction times in infants with AOP. We hypothesised that, in late preterm infants, apnoea rate is related to brain maturity and that this relationship is stronger than the relationship between apnoea rate and PMA. Indeed, demonstrating that there is a significant correlation between apnoea rate and brain maturity, and that this is stronger than the relationship between apnoea rate and PMA, is an essential first step to suggest that brain maturity may be a useful biomarker for caffeine treatment; if this were not the case then treatment based on an alternative biomarker would be a better approach. For this study, we focused on late preterm infants from 31–36 weeks PMA to investigate whether brain maturity can be used to assess caffeine requirement. It is possible that some of even the youngest infants in this age range will not need to start caffeine treatment and that brain maturity could be used to guide this decision. Moreover, brain maturity may be a candidate biomarker to indicate when to discontinue treatment in those that have received it, which is an important step in their care, expedites discharge and could shift the balance in an individual from adverse effects and towards benefit.

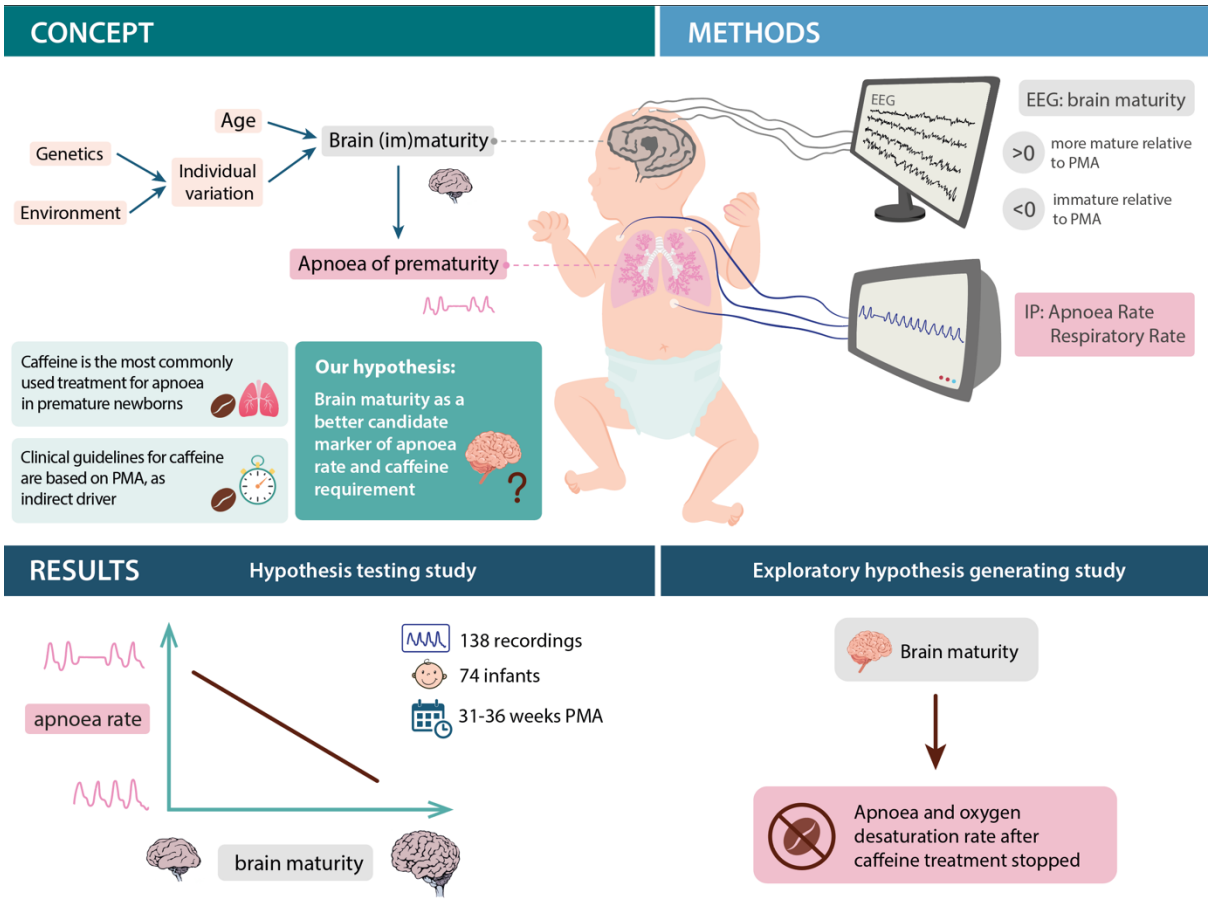

Figure 1. Schematic of study concept, design and main results. EEG: electroencephalography, IP: impedance pneumography, PMA: postmenstrual age. Note that other factors such as infection can lead to apnoea in premature infants, however, here we focus on the relationship between apnoea and brain development.

We included 74 infants on 138 separate occasions (demographic information is provided in Table 1, note: some infants were receiving caffeine at the time of study according to clinical requirement). We included preterm infants between 31 – 36 weeks PMA who were stable and without any significant respiratory or central nervous system complication at the time of the EEG recording (additional details in the Methods). EEG and vital signs were recorded for one to two hours, with continued vital signs recording in some infants up to the time of discharge (see Methods). Sensory-evoked (visual and tactile) and resting state EEG was used to calculate brain maturity by applying previously developed and validated models[13,14]. We used both sensory-evoked and resting state models to improve accuracy by assessing these different (i.e. sensory-evoked and resting state) aspects of functional brain maturity; results using each model independently are presented in the supplementary material and demonstrate similar trends to those presented in the main results (Figure S1). The model accuracies in this data set were similar to those in the previously published data indicating that brain maturity could be accurately predicted in the data presented here (Figure S2). Apnoeas were identified as pauses in breathing of at least 15 seconds using impedance pneumography. To identify individual breaths from the impedance pneumography we used an algorithm that we previously developed and validated in an independent cohort of preterm infants[24].

Apnoea rate was not significantly associated with PMA in the age range studied (*p*:0.58; β [95% CI]:-0.04 [-0.16; 0.09]; ρ:-0.0472; E_{point}:1.26; E_{CI}:1.00; Figure 2A). Whilst we would expect that apnoea rate decreases with PMA when considering a wider age range[6], this result demonstrates that apnoea rate is not strongly related to PMA in older preterm infants highlighting potential pitfalls of guidelines based on age. Consistent with our hypothesis, apnoea rate was significantly related to brain maturity (*p*:0.024; β [95% CI]:-0.22 [-0.41; -0.03]; ρ:-0.19; E_{point}:1.83; E_{CI}:1.22; Figure 2B, comparison of the relationship between apnoea rate and brain maturity versus apnoea rate and PMA: *p*:0.12; one-tailed bootstrapped 95% CI: [-0.2287, 1]). There was a clear relationship between brain maturity and the distribution of prolonged inter-breath intervals, with immature brain activity associated with longer inter-breath intervals (Figure 2C), suggesting the association between apnoea rate and brain maturity holds irrespective of the exact definition of apnoea used, which is variable within clinical practice[25]. In preterm infants, the development of central chemoreceptors is immature and there is a blunted central nervous system response to hypercapnia[26]. This immaturity of respiratory centres, which we postulate is related to general immaturity of the brain as assessed here, likely underlies these results.

Interestingly, when considering the average respiratory rate (i.e., normal breathing between apnoeas) we observed the opposite results - respiratory rate significantly decreased with PMA (*p*:0.0001; β [95% CI]:-1.87 [-2.81; -0.93]; ρ:-0.3186; E_{point}:2.37; E_{CI}: 1.75; Figure 2D,F) and respiratory rate did not significantly relate to brain maturity (*p*:0.18; β [95% CI]:1.17 [-0.54; 2.88]; ρ:0.1149; E_{point}:1.77; E_{CI}: 1.00; Figure 2E). After birth, critical changes in lung and chest wall dynamics play a major role in developing breathing mechanics[27]. With advancing PMA, lung alveoli mature both functionally and anatomically, improving lung-chest wall elasticity[28]. We speculate that these changes play

a major role in driving decreases in respiratory rate with advancing PMA. The contrasting results between apnoea and respiratory rate underlie the distinction between normal breathing and apnoea and further highlight the importance of utilising brain maturity as a biomarker for caffeine requirement in preterm infants.

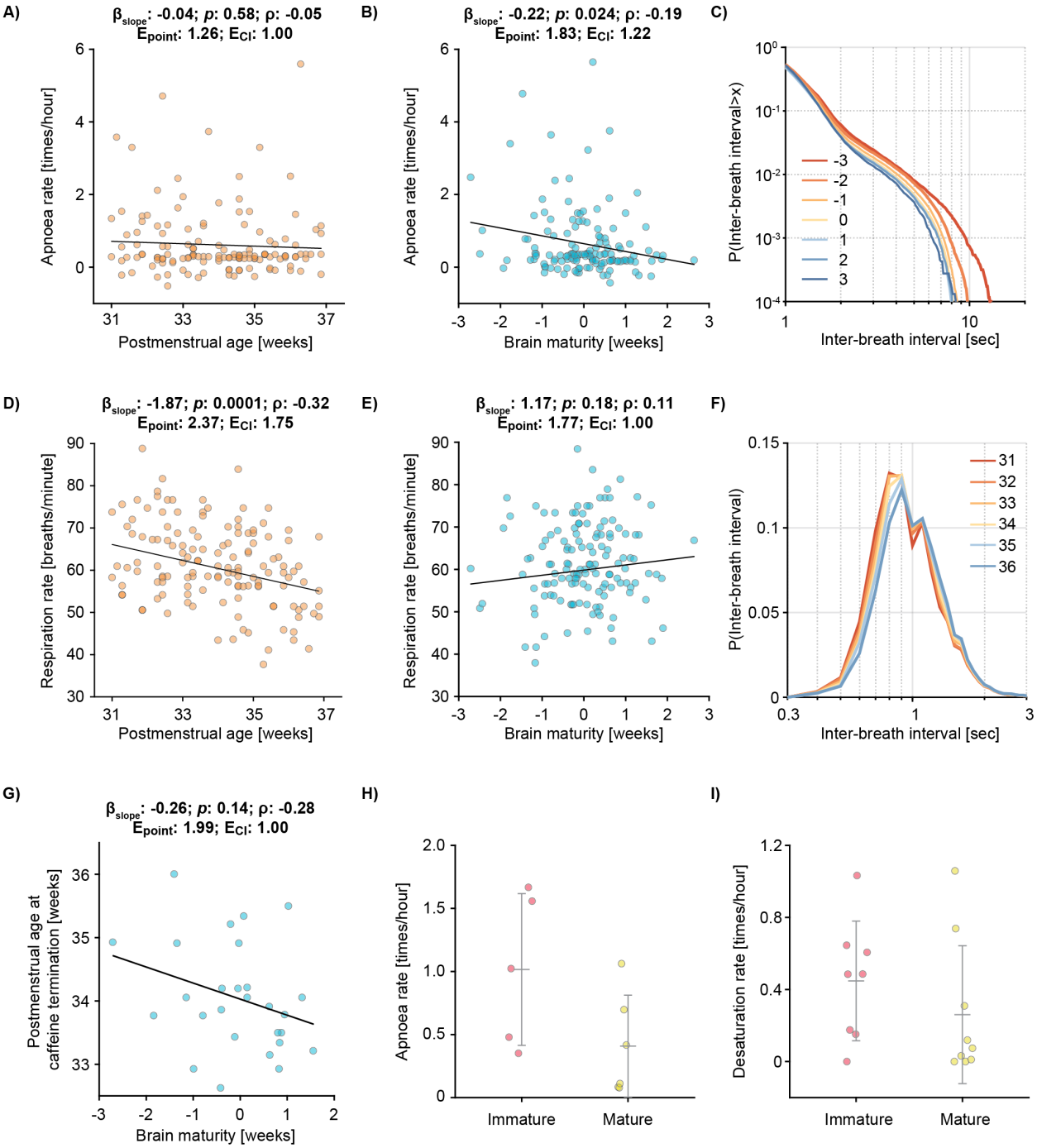

Figure 2. Apnoea rate and caffeine requirement are related to brain maturity.
The relationship between apnoea rate (defined as a pause in breathing of at least 15 seconds) and (A) postmenstrual age (PMA), and (B) brain maturity. Brain maturity is computed relative to the infant's PMA, with negative values indicating that the brain activity is immature relative to PMA. Each dot indicates an individual test occasion (A-F, 74 infants were included on 138 test occasions). Black line is the mean fit of

the linear model (β_{slope}, p, ρ, E_{point}, and E_{CI} in the titles are the predictor's regression slope, and its significance, partial correlation coefficient, E-value point estimate, and E-value of the confidence interval respectively). (C) The cumulative probability distribution of inter-breath intervals split according to brain maturity. The distribution is shown on a double logarithmic scale to emphasise the prolonged inter-breath intervals (i.e., apnoeas). (D,E) The relationship between respiratory rate and (D) PMA and (E) brain maturity. (F) The probability density distributions of inter-breath intervals shown according to the infant's PMA. (G) The relationship between age at which caffeine treatment is discontinued (dependent on clinical decision) compared to brain maturity in the subset of 27 infants where brain activity was recorded in the two weeks before caffeine was discontinued. (H, I) Exploratory hypothesis-generating study in the subset of infants where vital signs were recorded longitudinally. Infants were grouped according to whether their brain activity was immature or mature relative to their PMA before they discontinued caffeine treatment. Error bars indicate mean and standard deviation. (H) Apnoea rate (n = 11 infants) and (I) desaturation rate (n = 17 infants) were computed in the 7 days after treatment was stopped. Statistical models in panels A and B are adjusted for data length of impedance pneumograph (i.e. the signal used to assess apnoea rate) and infection; panels D, E and G are adjusted for infection (see Supplementary Results for discussion of statistical models and Supplementary Figures S3-S10).

Guidelines for discontinuing caffeine treatment are based on PMA, but there is variation in when treatment is stopped[29], depending on clinical discretion and guided by the infant's physiological stability. Given that apnoea rate is related to brain maturity, we hypothesised that the age at which infants stopped caffeine treatment would be related to their brain maturity. In a subset of 27 infants who had their EEG recorded shortly before discontinuing caffeine treatment, we found that the age at which infants stopped caffeine treatment was correlated with their brain maturity (p:0.14; β [95% CI]:-0.26 [-0.61; 0.09]; ρ:-0.2839; E_{point}:1.99; E_{CI}:1.00; Figure 2**Figure 2**G) with infants with immature brain activity discontinuing caffeine at older ages (the decision to stop treatment was made by the clinical team who were unaware of the brain maturity or EEG of the infant; note all infants stop caffeine whilst in neonatal care, however, for most infants we did not record their EEG at the time at which they stopped caffeine treatment and so they were not included in this subset analysis). However, despite this association, some infants with relatively mature brain activity continued caffeine close to term-corrected age, and, *vice versa*, some infants with immature brain activity discontinued treatment at a much younger age. Using an infants' brain maturity will provide an objective metric to identify when to discontinue treatment.

Finally, in an exploratory hypothesis-generating study in a subset of infants who had longitudinal recordings of vital signs throughout their hospital stay (see Methods), we investigated whether there was a relationship between brain maturity (recorded before infants discontinued caffeine), and apnoea rate in the 7 days after they discontinued caffeine treatment. Infants with more mature brain activity (defined as brain maturity >0) had 60% fewer apnoeas and 42% fewer oxygen desaturations after they stopped caffeine treatment compared with those infants with immature brain activity (brain maturity <0) (Apnoea rate, n=11: infants with mature brain activity: 0.41±0.40 apnoeas/hour [mean±SD]; infants with immature brain activity: 1.02±0.60; Figure 2H; Desaturations n=17: infants with mature brain activity: 0.26±0.38 desaturations/hour; infants with immature brain activity: 0.45±0.33; Figure 2I). As this exploratory study was limited in sample size and conducted to determine if there was any preliminary evidence of a relationship between brain maturity and apnoea rate after caffeine discontinuation, no statistical analyses were performed. This exploratory study is also limited in that it was a

single-centre study comprising mainly Caucasian babies (Table 1), the study did not follow a pre-registered protocol and only examined the subset of infants that had longitudinal recordings at the time of stopping caffeine. These results should therefore be replicated in a prospective study before definitive conclusions can be drawn. Nevertheless, our findings provide promising initial evidence that physiological instability after discontinuation of caffeine is related to brain maturity.

The results presented here demonstrate that apnoea rate is related to brain maturity, not postmenstrual age, in infants from 31–36 weeks. We focused on this age range as this is the age at which doctors will consider whether or not infants require caffeine treatment and, unlike at younger ages, not all infants will experience apnoea. Future work should assess whether EEG can be used as a biomarker for caffeine treatment in younger infants – whilst all extremely preterm infants are likely to need caffeine at birth, it is plausible that brain maturity may be a useful biomarker to indicate the dose of caffeine an individual will require.

We identified a significant relationship between apnoea rate and brain maturity, nevertheless, there was some heterogeneity in responses. Within the statistical model we adjusted for data length and the presence of infection, and additionally mode of ventilation (see Supplementary Results). We did not adjust for other potential confounding factors, such as ethnicity and socioeconomic status, as information on these factors were not available for all infants. We also did not adjust for gestational age, as in our sample this was highly colinear with PMA (see Supplementary Results). Future work to investigate heterogeneity in apnoea rate and other potential confounds is required. This study included infants with a wide range of both gestational ages at birth (from 23-36 weeks) and postnatal ages (0-11 weeks). This enabled us to capture a broad spectrum of the infants who receive neonatal care and experience apnoea (which is important to assess the viability of biomarkers for use in the clinical setting). However, differences in the intra- and extra-uterine experiences of these infants (such as infection, drug exposure, experience of painful procedures, and medical conditions) may impact on brain development trajectories [30,31] and consequently brain maturity. The factors which have the largest impact on brain maturity are yet to be elucidated. Prospective studies, for example, with tightly controlled gestational and postmenstrual age groups are warranted to examine the impact of these factors on brain maturity and the relationship with apnoea rate.

Finally, further work is needed to translate the results presented here into a clinically useable tool: EEG is already used in neonatal units, however, technology to incorporate brain age assessment within an EEG device is needed to readily assess this metric in the neonatal unit. Additionally, brain maturity should be validated as a biomarker for caffeine requirement in a randomised clinical trial. Such a trial would compare stopping caffeine based on the brain age of an infant with a control arm where caffeine is stopped based on current standard practice (i.e. based on clinical judgement and postmenstrual age criteria). Before such a trial can be conducted a prospective study evaluating the relationship between brain age and apnoea rate after stopping caffeine treatment is

needed. This study could be used to build a model to identify the optimal brain age at which to stop caffeine treatment in premature infants.

In summary, we demonstrate that apnoea rate is related to brain maturity, not PMA, in infants from 31-36 weeks PMA. We also provide promising initial evidence that brain maturity is related to apnoea rate and oxygen desaturations in the week after caffeine is stopped. There is a need for personalised medicine in neonatology; this study demonstrates that brain maturity is a key candidate biomarker for tailored caffeine treatment.

References

1. Ohuma, E. O. *et al. Lancet* **402**, 1261–1271 (2023).
2. Finer, N. N. *et al. Pediatrics* **117**, S47–S51 (2006).
3. Pillekamp, F. *et al. Neonatology* **91**, 155–161 (2007).
4. Janvier, A. *et al. J. Perinatol.* **24**, 763–768 (2004).
5. Williamson, M. *et al. Front. Pediatr.* **9**, 1–9 (2021).
6. Henderson-Smart, D. J. *Aust Paediatr J* **17**, 273–276 (1981).
7. Moschino, L. *et al. ERJ Open Res.* **6**, 1–19 (2020).
8. Care of Preterm or Low Birthweight Infants Group. *eClinicalMedicine* **63**, 1–13 (accessed 2023).
9. Puia-Dumitrescu, M. *et al. J Pediatr* **211**, 27–32 (2019).
10. Schmidt, B. *et al. N. Engl. J. Med.* **354**, 2112–2121 (2006).
11. Haddad, W. *et al. J. Neonatal. Perinatal. Med.* **8**, 41–45 (2015).
12. Schmidt, B. *J. Pediatr.* **259**, 1–5 (2023).
13. Zandvoort, C. S. *et al. Clin. Neurophysiol.* **157**, 61–72 (2024).
14. Ansari, A. *et al. Clin. Neurophysiol.* **163**, 226–235 (2024).
15. André, M. *et al. Clin. Neurophysiol.* **40**, 59–124 (2010).
16. Van Der Vaart, M. *et al. Cereb. Cortex* **32**, 3799–3815 (2022).
17. Colonnese, M. T. *et al. Neuron* **67**, 480–498 (2010).
18. Chipaux, M. *et al. PLoS One* **8**, e79028 (2013).
19. Molloy, M. F. *et al. Neuroimage* **253**, 119101 (2022).
20. Lenroot, R. K. *et al. Dev. Psychopathol.* **20**, 1161–1175 (2008).
21. Cornelissen, L. *et al. Br. J. Anaesth.* **120**, 1274–1286 (2018).
22. Davidson, A. J. *et al. Paediatr. Anaesth.* **18**, 60–70 (2008).
23. Henderson-Smart, D. J. *et al. N Engl J Med* **308**, 353–57 (1983).
24. Adjei, T. *et al. BMJ Open Respir. Res.* **8**, 1–9 (2021).
25. Elder, D. E. *et al. J. Paediatr. Child Health* **49**, E388–E396 (2013).
26. Erickson, G. *et al. J. Perinatol.* **41**, 2111–2123 (2021).
27. Heldt, G. P. *J. Appl. Physiol.* **65**, 441–444 (1988).
28. Trachsel, D. *et al. Paediatr. Anaesth.* **32**, 108–117 (2022).
29. Ji, D. *et al. J. Perinatol.* **40**, 288–293 (2020).
30. Grunau, R. E. *Rambam Maimonides Med J* **4**, 1–13 (2013).
31. Schiller, R. M. *et al. J. Clin. Pharmacol.* **58**, S140–S150 (2018).

Methods

Study design and participants

This was a prospective cohort study conducted at the Newborn Care Unit of the John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK between 2019 and 2023. Clinically stable preterm infants aged between 31- and 36-weeks postmenstrual age (PMA) at the time of test occasion were eligible to be included in the study (from an age of 31 weeks and 0 days up to 36 weeks and 6 days). Infants were excluded if they had grade III or IV intraventricular haemorrhage, hypoxic ischaemic encephalopathy, congenital malformations, were on mechanical ventilation, receiving opioid analgesics at the time of study, or if there was a history of maternal substance misuse during pregnancy. These conditions are known to impact brain activity and may interfere with the interpretation of EEG by acting as confounders. Previous studies on brain age models utilised similar inclusion and exclusion criteria[13,14].

The study was approved by the National Research Ethics Service (references: 19/LO/1085, 12/SC/0447). Eligible families were given verbal and written information about the study, and written parental consent was signed before inclusion in the study. The study conformed to the standards set by the Declaration of Helsinki and Good Clinical Practice.

An initial total of 76 infants were studied on 141 separate test occasions; 53 infants were included in one to three test occasions, 23 infants took part in the ongoing *Breathing and Brain Development* study[32] where their EEG is recorded approximately once a week and their vital signs recorded continuously whilst on the Newborn Care Unit (included on average on 3±1 [mean ± SD] occasions). During each test occasion (which aimed to last ~2 hours), simultaneous electroencephalography (EEG) and vital signs (respiration, heart rate and oxygen saturation) were recorded. Three test occasions were excluded due to very short recordings of inter-breath intervals (see ‘computing respiratory and apnoea rate’), leaving a total of 74 infants studied on 138 test occasions included in the analysis. Infant demographic information is provided in Table 1. Only one infant had suspected necrotising enterocolitis, and this was treated medically, while 10 infants had chronic lung disease at the time of the EEG recordings.

354 **Table 1. Infant demographics.**

355

Factors	
Age	
Gestational age at birth (weeks)	31.0 (23.6-36.6)
Postmenstrual age at recording (weeks)	34.0 (31.0-36.9)
Postnatal age at recording (weeks)	3.6 (0.0-11.4)
Birthweight (g)	1,573 (630-4,525)
Sex	
Females	25 (33.8)
Males	49 (66.2)
Mode of delivery	
Normal vaginal delivery	13 (17.6)
Vaginal breech	2 (2.7)
Vaginal assisted (ventouse/forceps/kiwi)	4 (5.4)
Elective Caesarean section	8 (10.8)
Emergency Caesarean section	47 (63.5)
Apgar scores	
Apgar at 1 minute	6.9 (1-10)
Apgar at 5 minutes	8.9 (2-10)
Apgar at 10 minutes	9.6 (6-10)
Ethnicity	
White	26 (35.1)
Asian	5 (6.8)
Black	1 (1.4)
Mixed	6 (8.1)
Other	1 (1.4)
Information not recorded	35 (47.3)
Resuscitation at birth	
Yes	55 (74.3)
No	19 (25.7)
Respiratory distress syndrome	
Ongoing condition	47 (34.1)
Past treated	57 (41.3)
No	34 (24.6)
Persistent pulmonary hypertension	
Past treated	4 (2.9)
No	134 (97.1)

Patent ductus arteriosus	
Ongoing condition	9 (6.5)
Past treated	7 (5.1)
No	122 (88.4)
Jaundice	
Ongoing condition	6 (4.3)
Past treated	78 (56.5)
No	54 (39.1)
Infection	
None	61 (44.2)
Suspected	17 (12.3)
Treated	60 (43.5)
Level of care	
Intensive Therapy Unit	6 (4.3)
High Dependency Unit	65 (47.1)
Low Dependency Unit	59 (42.8)
Postnatal care	5 (3.6)
Information not recorded	3 (2.2)
Respiratory support	
High flow therapy	35 (25.4)
Low flow therapy	14 (10.1)
Spontaneous ventilation	88 (63.8)
Information not recorded	1 (0.7)
Medications received on the day of EEG recording	
Vitamins and minerals	112 (81.2)
Probiotics	59 (42.8)
Breastmilk fortifier	56 40.6)
Caffeine	71 (51.4)
Antibiotics	6 (4.3)
Gaviscon	6 (4.3)
Paracetamol	2 (1.4)
Budesonide	6 (4.3)
Nystatin	1 (0.7)
Diuretics	1 (0.7)
No medication	16 (11.6)

356 Reported values are mean (range) or number (%). All demographic details are computed for every recording except
357 ethnicity, need for resuscitation at birth, gestational age (which are provided per infant). Antibiotics include:
358 benzylpenicillin, gentamicin, coamoxiclav, cefotaxime, chloramphenicol, and amoxicillin; vitamins and minerals include:
359 abidec, calcium, phosphate, vitamin D, folic acid, and iron; diuretics include: chlorothiazide and spironolactone.

360

361 ***Data acquisition***

362 *Vital signs recordings*

363 Each infant had continuous vital signs monitoring as the standard of care. Vital signs were
364 monitored using Phillips IntelliVue MX800 or MX750 monitors and data were continuously
365 downloaded from the vital signs monitor using an electronic data capture software
366 (iXtrend, ixitos, Germany). For infants in the *Breathing and Brain Development* study, vital
367 signs were downloaded continuously throughout the infant's hospital stay; the recordings
368 from the day of the EEG test occasion are included in the main analysis (i.e., up to 24
369 hours of recording). For other infants, vital signs were recorded during EEG data
370 acquisition for an average of 1.5 hours (median, range: 0.4-3.3 hours).

371 Heart rate, oxygen saturation and respiratory rate (calculated by the monitor) were
372 downloaded onto the laptop at a sampling rate of 0.97 Hz; the electrocardiograph (ECG,
373 to measure heart rate) at 250 Hz and recorded with three electrodes placed on the infant's
374 chest; the impedance pneumography (IP, to measure respiration) at 62.5 Hz and
375 recorded from the chest electrodes, and the photoplethysmography (PPG, to measure
376 oxygen saturation and pulse) at 125 Hz from a probe placed on the infant's foot or hand.

377

378 *EEG recordings*

379 EEG was recorded from DC to 400 Hz using a SynAmps RT 64-channel headbox and
380 amplifiers and CURRYscan7 neuroimaging suite (Compumedics Neuroscan) at a
381 sampling rate of 2 kHz. Eight electrodes were placed at Cz, CPz, C3, C4, Oz, FCz, T3,
382 and T4, according to the modified international 10-20 system. The ground electrode was
383 placed at FPz and the reference electrode at Fz. Infant scalp preparation was done with
384 NuPrep gel (D.O. Weaver and Co., Aurora, USA) to achieve impedances of 5-10 kΩ. A
385 conductive paste (Elefix EEG paste, Nihon Kohden, Tokyo, Japan) was used to attach
386 disposable Ag/AgCl cup EEG electrodes (Ambu Neuroline, Ballerup, Denmark) to the
387 infant's scalp. At the beginning of each recording, approximately 10 background
388 annotations were time-locked on the EEG and vital signs recordings by manually pushing
389 a button connected to a bespoke triggering device (the PiNe box)[33] which simultaneously
390 annotates the EEG and vital signs recordings. This enabled synchronisation of the vital
391 signs and EEG recordings. Background annotations were performed by a researcher
392 during a period when the infant was quietly resting in their cot/incubator. Resting-state
393 EEG was recorded for an average of 1.7 ± 0.7 hours (mean ± SD).

394

395 *Tactile and visual stimuli*

396 A tendon hammer was used to lightly tap the infant's foot to elicit touch responses. At
397 least 10 gentle touch stimuli, with inter-stimulus interval of approximately 10 seconds

398 (longer if the infant was unsettled) were applied. The tendon hammer was modified with
399 a built-in force transducer (Brüel & Kjær, Denmark) that sends a trigger pulse at the point
400 of stimulation[34]. For visual-evoked responses, the stimulus was a flash of light presented
401 using a Lifelines Photic Stimulator (intensity level 4, approximately 514 lumens). The
402 Photic Stimulator was placed approximately 30 cm away from the infants, in line with their
403 direct field of vision and a train of 10 light flashes was delivered at least 10 seconds apart.
404 The tactile and visual stimuli (order pseudo-randomised) were performed either at the
405 start or end of the EEG recording. The stimuli were annotated simultaneously on the EEG
406 and vital signs recordings using our custom-made time-locking device (the PiNe box)[33].

407

408 *Recording of infant demographics and clinical details*

409 Relevant demographic and clinical information were obtained from the medical notes of
410 each infant and updated at each test occasion where relevant. These included the infant's
411 age at birth and sex, and at the time of the test occasion, the postmenstrual age, current
412 mode of ventilation and presence of infection. Infection was defined based on the
413 neonatal guidelines for infection treatment at the John Radcliffe Hospital as either an
414 elevated C-reactive protein of >10mg/L, abnormal white blood cell count (>20 x 10^9/L or
415 <5 x 10^9/L), a positive blood culture; or a symptomatic infant with predisposing factors for
416 sepsis receiving antibiotics at the time of EEG recording.

417

418 ***Data analysis***

419 All analyses were performed in MATLAB (ver. 2022b; MathWorks Inc., Natick, USA). In
420 brief, we first computed respiratory and apnoea rates from the impedance pneumography
421 (IP) signals, and brain maturity from 20 minutes of resting state EEG activity and sensory-
422 evoked (visual and tactile) potentials. Next, we constructed linear (mixed-effects) models
423 to test if the respiratory outcomes – respiratory and apnoea rates – were associated with
424 either PMA or brain maturity. We also explored whether the PMA at which caffeine
425 treatment is discontinued relates to brain maturity. In a final exploratory analysis to
426 explore if brain maturity could be utilised as a predictive biomarker for stopping caffeine
427 treatment, we investigated whether the rate of episodes of physiological instabilities
428 (apnoea and oxygen desaturation rate) in the week after caffeine was stopped is related
429 to brain maturity. Each of these analysis steps are outlined in more detail below.

430

431 *Computing respiratory and apnoea rate*

432 To obtain respiratory and apnoea rates, IP signals were processed using an algorithm
433 validated for the identification of inter-breath intervals and apnoeas in infants[24] . Briefly,
434 the algorithm first filters the signal to reduce the noise introduced by movements and

435 cardiac activity. An adaptive amplitude threshold (set at 0.4 times the standard deviation
436 of the signal over the preceding 15 breaths) is used to identify individual breaths. Finally,
437 a support vector machine algorithm was applied to all potential episodes of apnoea
438 (defined as inter-breath intervals longer than 15 seconds) to remove periods of low
439 amplitude erroneously identified as apnoeas due to noise or shallow breathing. This
440 model was trained and validated on IP signals acquired during a clinical trial[35]. The data
441 from the clinical trial was completely independent from our current data sample. Long
442 inter-breath intervals (IBI) marked as noise/shallow breathing were discarded from further
443 analysis. We then computed respiratory and apnoea rates estimated from all available
444 IBIs on the day of the EEG recording, which was 8.5 ± 8.7 hours (mean ± SD over
445 recordings). We computed respiratory rates (expressed in times per minute) and apnoea
446 rates (in times per hour) estimated from all available IBIs on the day of the EEG recording.
447 We excluded three recordings which had fewer than 100 IBIs available. Apnoea rate was
448 calculated as the number of apnoeas (IBIs greater than 15 seconds) per hour. Respiratory
449 rate was computed as the median number of breaths per minute (i.e., 60 divided by
450 median IBI of the recording).

451

452 *Estimating brain maturity from EEG activity*

453 To obtain brain maturity, we first computed brain age using two brain age
454 models[13,14].These two models estimate the functional maturation of the brain using
455 resting state[14] and sensory-evoked EEG activity[13] respectively. These models aim to
456 capture different developmental aspects. The resting state model is an index for the
457 maturation of ongoing brain activity and the sensory-evoked model for the maturation of
458 sensory pathways (Supplementary Figure S11). Both are machine learning models which
459 are trained to estimate the age of the infant from the EEG activity, with the aim of such
460 models being to identify infants who deviate from normal neurodevelopment. Thus, brain
461 age is correlated with PMA, and large differences between PMA and brain age may be
462 indicative of abnormal outcome[13,14]. Brain maturity was defined as the mean brain age of
463 the two models minus PMA (note, brain maturity has been referred to in the literature as
464 the brain age gap or brain age delta). Figure S10 compares brain age and
465 respiration/apnoea rate.

466

467 Before estimating brain maturity, EEG data were first pre-processed using the Brainstorm
468 (version 3)[36] and EEGLAB (version 2022.1)[37] toolboxes. EEG time series were filtered
469 with pass-band edges at 0.1 and 30 Hz by consecutively high-pass (Hamming windowed-
470 sinc FIR filter with a cut-off frequency at 0.05 Hz) and low-pass filtering the signals
471 (Hamming windowed-sinc FIR filter with a cut-off frequency at 33.75 Hz).

472

To calculate the infant's brain age from their sensory-evoked responses (visual and tactile) we used the model developed by Zandvoort and colleagues (2024)[13]. Briefly, this model capitalises on the PMA-related changes in the morphology of sensory-evoked potentials during the preterm period. The sensory-evoked data were first epoched from -0.5 to 1.0 seconds around the stimulus onsets for channels Oz and Cz. Epochs were baseline normalised by subtracting the mean amplitude of the time window between -0.5 and 0 sec. Epochs were then visually inspected and those with excessive amplitudes or artefacts excluded (an average of 10% of epochs were excluded). Individual test occasions with fewer than five epochs of sensory-evoked data remaining after artefact rejection were excluded from sensory-evoked brain age estimation (9 out of 138 recordings) and brain age was estimated using the resting state model alone for these recordings. Visual and tactile evoked potentials were fitted to neurodynamic response functions based on the characteristic waveforms of the evoked stimulus responses[13] (Figures S12-S13 show the average evoked potentials). Zandvoort et al. (2024) extracted these characteristic waveforms using principal component analysis. The visual and tactile responses at channels Oz and Cz respectively can be captured by four visual and two tactile neurodynamic response functions[13]. Before fitting, evoked responses were Woody filtered to the neurodynamic response functions to improve their temporal alignment allowing for individual infant differences in the latency of the evoked response (maximum jitter: 50 ms; number of iterations: 1)[38]. Neurodynamic response functions were then fitted to the individual's evoked responses using a linear regression. The regression's slope indicates the magnitude of the evoked response. The slopes were obtained for the six neurodynamic response functions and forwarded as input for the sensory brain-age model, which consisted of a support vector regression with a linear kernel function using MATLAB's fitrsvm function (version 2022b; MathWorks). The output of the model is the predicted sensory brain age for the infant (for further details see Zandvoort et al., 2024[13]). Zandvoort et al. (2024) validated their model by applying it to an independent held-out dataset. Through leave-one-infant-out cross-validation, their mean absolute error (MAE; between the PMA and brain age) is 1.55 weeks which is comparable to the MAE of 1.41 weeks as found in the dataset to train the model. They also show that the MAE of 1.55 weeks for the held-out dataset is significantly better than a MAE from a null model estimating the mean PMA of the dataset.

To calculate the infants resting state brain age, we used the model developed by Ansari and colleagues (2024)[14]. Data were first downsampled to 64 Hz and the first 20 minutes of every recording was selected for analysis as this should warrant reliable brain age estimates[14]. EEG recordings were started when the infant was settled in their cot/incubator and no overt artefacts were present (data length: 105 ± 41 minutes [31, 193]; mean ± standard deviation [range]). Since our studies comprised resting-state and stimulus-evoked recordings, we selected a 20-minute window in which no visual or tactile

stimuli were applied. This enabled us to standardise the recording length used for analysis across infants. However, in line with Ansari et al. (2024), we show that brain age predictions estimated over 20 minutes are comparable to predictions from the full recordings (Figure S14A) and that using the full recording to estimate the brain age does not alter the relationship between apnoea rate and brain maturity (Figure S14B). For one recording, the total data duration was less than 20 minutes, meaning that we did not have sufficient data to calculate a reliable resting state brain age. For this recording, we only used the brain-age estimate of the sensory model. For all recordings, the bipolar derivative between channels C3 and C4 was used for brain age estimation. The data were input into the deep neural network, the output of the model is the predicted resting state brain age for the recording (for further details see Ansari et al. 2024[14]). Briefly, this algorithm segments the data into 30-second epochs and estimates a brain age for each of them for ten iterations (the model consisted of a 10-learner ensemble method). To obtain a single brain age value, the median over all ten ensembles was taken for every 30-second epoch, after which a median over all 30-second epochs resulted in a single value. To validate the output, Ansari et al. (2024) also used an independent test set and found a MAE of 1.03 weeks for the held-out dataset. Their model also yielded a significantly lower MAE than a null model estimating the mean PMA of the dataset. Moreover, for model training, Ansari et al. (2024) used data of infants with normal neurodevelopmental outcome at 24-month follow up (assessed with BSID-II). Their model was validated on EEG data that were labelled as normal by a neurophysiologist. Both brain age models do not take into account different sleep states. Whilst it is known that sleep states alter EEG activity in infants[15], derived brain ages are marginally affected by it according to earlier versions of the model[39].

We obtained a single brain age prediction by averaging the brain ages of both models. Since the models may estimate brain age with systematic prediction errors relative to PMA[40], we corrected for this by first fitting a linear model between the combined brain age and PMA. We next predicted PMA from this model and defined the model bias as the difference between the PMA and predicted PMA (meaning that brain maturity and PMA are completely uncorrelated). Next, brain age was corrected by subtracting the bias from the brain age (relative to PMA). This age-bias correction is a standard approach in the brain age field[40,41]. This also ensured that brain maturity and PMA are uncorrelated. However, without this bias correction, similar results are obtained (Supplementary Figure S15).

Probability density functions for the inter-breath intervals

To study how the IBI dynamics changed with PMA and brain maturity, we created probability density functions. To do so, we first created a frequency distribution from the

IBIs of each recording a resolution of 0.1 seconds. IBIs longer than 50 seconds were removed from further analysis. Distributions of each recording were normalised so that the area under the curve was equal to 1, allowing for inter-recording comparisons. We next transformed these recording-specific probability density functions into functions specific for a PMA/brain maturity. For this, a weight of 1 or lower was assigned to the recording-specific probabilities, depending on the infant's PMA/brain maturity. Weights were determined using a Gaussian window with a full width at half maximum of 13 days. Weightings for differences of more than 14 days were set to 0. Distributions that were specific for a certain PMA/brain maturity were computed as the weighted median of all distributions together with their weights[42]. We created probability density functions between 31 and 36 weeks for the PMA and between -3 and 3 weeks for the brain maturity.

Statistical analysis

To compare PMA and brain maturity with apnoea rate and respiration rate, we used linear mixed-effects models (computing four models for each of apnoea rate and respiration rate against PMA and brain maturity separately). Previous literature suggested that experiencing apnoeas is inversely associated with gestational age at birth[23,26]. The course of this relationship is however unknown which is why we used linear models as a starting point. Nevertheless, we tested the assumption of linearity (see Supplementary material and Table S1) using the mfp package in R. Moreover, it is not known how apnoea rate varies with brain maturity. A linear fit was appropriate for the data collected here so as not to overfit the data; however, whether a linearly association underlies the brain maturity/apnoea relationship needs to be studied in future research. Data length of the IBI recordings and infection ('no', 'suspected', or 'treated') were included as fixed factors and infant was included as a random factor (to account for the repeated test occasions in the same infant). Data length and infection were included as confounding factors. P-values (for the slopes of PMA and brain maturity) and partial correlation coefficients and their 95% confidence intervals are reported from these models. The normality assumption was verified with Q-Q plots. Further discussion of the fixed factors to include within the model, and assessment of the results with alternative factors is included in the Supplementary Material. To test for the robustness of the associations and to estimate the strength of unmeasured confounders we report E-values[43]. Reporting E-values is the most common method for observational medical studies[44]. We provide E-values for the observed association estimate and the confidence interval limit that is closest to null, computed via https://www.evalue-calculator.com. For visualisation (Figure 2), we used the plotAdjustedResponse function in MATLAB to adjust for the fixed factors (calculating a separate model without the random effect of infant to do so). Regression lines are visualised by calculating the line-of-best-fit on the adjusted data.

To test if apnoea rate is better associated with brain maturity than PMA, we first bootstrapped the partial correlation coefficient between apnoea rate and PMA using 10,000 repetitions. This allowed us to generate a 95% confidence interval for the association between apnoea rate and PMA. To test if the partial correlation coefficient between brain maturity and apnoea rate lies outside this interval, the significance value was defined as the number of bootstrapped partial correlation coefficients that were equal to or lower than the true partial correlation coefficient between brain maturity and apnoea rate. Significance was tested one-tailed with an alpha level of 0.05.

For the association between brain maturity and age at caffeine termination, we focused on a subset of 27 infants who had their brain maturity assessed within two weeks before they stopped caffeine treatment (note that infants were all receiving caffeine during the EEG recording). Note, the age at which caffeine was stopped was dependent on clinical decision and independent of the research study. A linear regression was used with infection status as fixed factors (data length was not included as a fixed factor as all EEG was analysed using a fixed data length and respiratory data was not analysed for this question). For visualisation, we used the plotAdjustedResponse function adjusting for the fixed effect of infection status only and the regression line was calculated as the line-of-best-fit of the adjusted data.

In the exploratory analysis, we compared brain maturity before infants stopped caffeine treatment with apnoea rate and oxygen desaturation rate in the week after they stopped caffeine treatment. In a subset of infants (n=17) with longitudinal recordings of vital signs throughout their hospital stay, infants were divided into two groups depending on their brain maturity – either mature if their brain maturity >0 weeks or immature if their brain maturity <0 weeks. Brain maturity was calculated from the EEG recording taken in the two weeks before they stopped caffeine, and used a single EEG recording for calculation. Apnoeas and desaturations were identified from the recordings in the 7 days after caffeine was stopped (the age at which caffeine was stopped was dependent on clinical decision and independent of the research study). Desaturation events were defined as 10-second periods in which the oxygen saturation was less than 80%[35]. The type of vital signs monitoring was dependent on clinical decision and so 6 out of 17 infants did not have their IP signal measured during this week (only oxygen saturation and pulse were monitored using a PPG), and they were consequently excluded from the apnoea rate analysis. Due to the limited sample sizes, and as this was an exploratory hypothesis-generating study, we did not apply any statistics and reported outcomes as mean and standard deviation.

Additional References

32. Hartley, C. https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/breathing-and-brain-development-version-10/ (accessed 2024).
33. Worley, A. *et al. PLoS One* **18**, e0288488 (2023).
34. Worley, A. *et al. J. Neurosci. Methods* **205**, 252–257 (2012).
35. Hartley, C. *et al. Lancet* **392**, 2595–2605 (2018).
36. Tadel, F. *et al. Comput. Intell. Neurosci.* **879716**, 1–13 (2011).
37. Delorme, A. *et al. J. Neurosci. Methods* **134**, 9–21 (2004).
38. Woody, C. D. *Med. Biol. Eng.* **5**, 539–554 (1967).
39. Pillay, K. *et al. Sci. reports* **10**, 1–14 (2020).
40. Smith, S. M. *et al. Neuroimage* **200**, 528–539 (2019).
41. Kalc, P. *et al. Hum Brain Mapp* **45**, 1–11 (2024).
42. Haase, S. *MATLAB Central File Exchange* https://www.mathworks.com/matlabcentral/fileexchange/23077-weighted-median (accessed 2024).
43. Vanderweele, T. J. *et al. Ann Intern Med* **167**, 268–274 (2017).
44. Latour, C. D. *et al. Am. J. Epidemiol.* **00**, 1–12 (2024).

Acknowledgements

This work is funded by the Wellcome Trust and Royal Society through a Sir Henry Dale Fellowship awarded to CH (grant number: 213486/Z/18/Z). FU is funded by the Commonwealth Scholarship Commission. We would like to thank the babies and their parents who took part in this research.

Data and code availability

Example resting state data as well as processed data (including brain maturity, apnoea rates and predictor variables for each individual) and the codes for the analysis are available at: https://github.com/CoenZandvoort/Brain_maturity_respiration. The brain age models can be found at: https://github.com/amirans65/brainagemodel and https://gitlab.com/paediatric_neuroimaging/sensory-brain-age-model. The algorithm for inter-breath interval detection is online at: https://gitlab.com/paediatric_neuroimaging/identify_ibi_from_ip. Due to ethical requirements, the raw data that support the findings of this study are available upon request to the corresponding author (caroline.hartley@paediatrics.ox.ac.uk).