



## Brain MRI before and at term equivalent age predicts motor and cognitive outcomes in very preterm infants

Alex M. Pagnozzi<sup>a,\*</sup> , Kerstin Pannek<sup>a,b</sup>, Roslyn N. Boyd<sup>c</sup>, Liza van Eijk<sup>a,d</sup> ,  
Joanne M. George<sup>c,e</sup> , Samudragupta Bora<sup>f</sup>, DanaKai Bradford<sup>a</sup>, Michael Fahey<sup>g</sup>,  
Michael Ditchfield<sup>h,i</sup>, Atul Malhotra<sup>j,k</sup>, Paul B. Colditz<sup>l</sup>, Jurgen Fripp<sup>a</sup>

<sup>a</sup> The Australian e-Health Research Centre, CSIRO, Brisbane, Australia

<sup>b</sup> The University of Queensland, School of Electrical Engineering and Computer Science, Brisbane, Australia

<sup>c</sup> Child Health Research Centre, Queensland Cerebral Palsy and Rehabilitation Research Centre, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia

<sup>d</sup> Department of Psychology, James Cook University, Townsville, Queensland, Australia

<sup>e</sup> Physiotherapy Department, Queensland Children's Hospital, Children's Health Queensland Hospital and Health Service, Brisbane, Australia

<sup>f</sup> Mothers, Babies and Women's Health Program, Mater Research Institute, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia

<sup>g</sup> Monash Health Paediatric Neurology Unit and Department of Paediatrics, School of Clinical Sciences, Monash University, Clayton, Victoria, Australia

<sup>h</sup> Monash Imaging, Monash Health, Melbourne, Victoria, Australia

<sup>i</sup> Department of Medicine, Monash University, Melbourne, Victoria, Australia

<sup>j</sup> Department of Paediatrics, Monash University, Melbourne, Victoria, Australia

<sup>k</sup> Monash Newborn, Monash Children's Hospital, Melbourne, Victoria, Australia

<sup>l</sup> Perinatal Research Centre, The University of Queensland Centre for Clinical Research, Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia

### ARTICLE INFO

#### Keywords:

Biomarkers  
Cerebral palsy  
Magnetic resonance imaging  
Neurodevelopmental impairment  
Prediction  
Preterm infants  
Structural

### ABSTRACT

Brain Magnetic Resonance Imaging (MRI) of high-risk infants in the neonatal period (from 26 weeks postmenstrual age to Term Equivalent Age (TEA)) is increasingly being used for the detection of brain injuries, and the early prognostication of adverse outcomes such as Cerebral Palsy (CP). While most imaging is performed around TEA in clinical practice for infants born preterm (<37 weeks of gestation), this would often require families to return to hospital for imaging. In this work, we extract structural biomarkers from MRI acquired both before and at TEA in a cohort of very preterm infants from the PPREMO and PREBO studies (n = 100), to determine if either time-point, or both combined, are predictive of both Bayley Scales of Infant and Toddler Development – Third Edition (Bayley-III) and the Neuro-sensory Motor Developmental Assessment (NSMDA) at 2 years. Using multivariable regression, moderately strong and statistically significant associations were found between brain structure on both early and TEA MRIs with 2-year outcomes (r = 0.39–0.55 for early MRI, r = 0.37–0.49 for Term MRI, r = 0.37–0.56 for early and TEA MRI combined). Importantly, brain biomarkers associated with early childhood outcomes from MRIs were identified, including white and grey matter volumes, deep grey matter and cerebellar volumes, and gyrification and surface area across the whole cortex. Early MRI showed the best prognostic accuracy along with combining timepoints, indicating the potential clinical benefit of Early MRI in predicting adverse outcomes.

### 1. Introduction

Preterm birth is the largest cause of death in newborn infants globally (Chang et al., 2013), and results in an increased risk of developmental delay (Pugliese et al., 2013). Increasingly, brain magnetic

resonance imaging (MRI) acquired during the neonatal period, from 26 weeks postmenstrual age (PMA) up to term equivalent age (TEA), is improving the ability to detect evidence of brain injury, and plays an important role in identification of infants at risk of adverse motor and cognitive outcomes (George et al., 2018; Woodward et al., 2006), such

\* Corresponding author. Australian e-Health Research Centre Level 7 STARS Building, 296 Herston Road Herston, QLD, 4029, Australia.

E-mail address: [Alex.Pagnozzi@csiro.au](mailto:Alex.Pagnozzi@csiro.au) (A.M. Pagnozzi).

<https://doi.org/10.1016/j.ynrp.2025.100262>

Received 29 November 2024; Received in revised form 4 February 2025; Accepted 9 April 2025

Available online 19 April 2025

2666-9560/Crown Copyright © 2025 Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

as intelligence quotient (Anderson et al., 2017). Identifying early prognostic markers of later development and outcomes is crucial, given the enhanced neuroplasticity at this age, early interventions starting in the infant period may have the potential to significantly reduce motor and cognitive impairments (Spittle et al., 2015).

On structural MRI, preterm infants exhibit altered cortical morphometrics (Dubois et al., 2019; Kim et al., 2016; Shimony et al., 2016), growth and morphological trajectories (Bouyssi-Kobar et al., 2016; Lefèvre et al., 2016) compared to term born infants. Specifically at TEA, a number of studies have found MRI-derived measures to be associated with neurodevelopmental outcomes at 2 years and beyond (Anderson, 2011), including quantitative structural measures (such as volume, cortical shape), microstructural measures using diffusion MRI (Parikh, 2016) and qualitative measures such as white matter abnormality (van't Hooft et al., 2015). In one study using the state-of-the-art developing Human Connectome Project (dHCP) structural pipeline (Makropoulos et al., 2018a) on TEA MRI data (Makropoulos et al., 2018b), anatomical volumes (for areas such as the thalamus) as well as increased cortical surface area and reduced cortical curvature were found to be associated with better 2-year Bayley-III motor and cognitive outcomes (Kline et al., 2020; Julia E. Kline et al., 2020). Furthermore, combining structural MRI measures with early assessments such as the General Movements Assessment (GMA) and Hammersmith Infant Neurological Examination (HINE) at 3–4 months corrected age (CA) have demonstrated stronger predictive accuracy for neurodevelopmental outcomes than TEA MRI alone (Setänen et al., 2014; Skiöld et al., 2013; Wang et al., 2021).

While much work on TEA imaging exists, it is unclear if performing imaging earlier, before TEA, provides similar or greater prognostic value. While early imaging (28–32 weeks PMA) has been performed in cohorts whose motor, cognitive and language trajectories were assessed up to 4.5 years of age, literature in this area is sparse. Using an MRI scoring system for white matter injury (WMI), the severe WMI classification from early MRI was associated with poor motor outcomes as measured by the MABC-2 at 4 years (Cayam-Rand et al., 2019). Using another semi-quantitative approach, WMI severity was strongly associated with neurodevelopmental impairments at 2 years (Spearman's rank correlation 0.88,  $p < 0.001$ ) (Martinez-Biarge et al., 2019). In contrast, quantitative MRI morphology (tissue volumes) at the early time point (30 weeks PMA) and at TEA were found to be independently predictive of motor and cognitive impairment at around 2 years of age (Gui et al., 2019), with an area under the curve of 0.8 and 0.85 for low motor/cognitive outcomes respectively (Moeskops et al., 2017). Early imaging has the advantage being conducted prior to discharge, when the infant is still in the hospital, as part of early standard of care, often without the need for sedation or general anaesthesia. This timing could aid in the early identification and prediction of adverse motor or neurodevelopmental outcomes in infants born preterm and facilitate referral to early interventions while the infant is still in the hospital.

We previously demonstrated that brain volumetrics from early structural MRI predict motor and cognitive outcomes at 2 years of age (Pagnozzi et al., 2023). In the current study, we compared these findings with structural MRI measures at TEA in predicting motor and cognitive outcomes at 2 years CA, and then investigate the potential benefit of utilising both time points for this prediction. For this, we utilise the internationally unique prospective Prediction of Preterm Motor Outcomes (PPREMO) and Prediction of Preterm Brain Outcomes (PREBO) studies of infants born <31 weeks gestational age (GA) which together form one of the largest cohorts worldwide with 3T MRI acquired at a median postmenstrual age of 32 weeks (PMA; 'early') as well as at TEA (George et al., 2015a). These cohorts also underwent neurobehavioural and motor assessment concurrent with imaging at early and TEA time-points, and neurodevelopmental assessments at 2 years corrected age. By determining which time point has the best prediction of later outcomes (early, or TEA), we can determine what assessments at what times within the neonatal window are best for predicting motor and cognitive outcomes.

## 2. Methods

### 2.1. Participants

This study examined data from two longitudinal prospective cohort studies of very preterm infants: the Prediction of Preterm Motor Outcomes (PPREMO) study (George et al., 2015b) and the Prediction of childhood Brain Outcomes (PREBO) study (NHMRC1084032). Both the PPREMO and PREBO cohorts recruited preterm infants born at <31 weeks' GA with no congenital or chromosomal abnormality, whose parents/caregivers were English-speaking and lived within a 200 km radius of the recruiting hospital. The total number of preterm and term-born infants recruited was  $n = 146$  for PPREMO and  $n = 187$  for PREBO. Infants were recruited at the Royal Brisbane and Women's Hospital (RBWH), Monash Children's Hospital (MCH), and Mater Mothers Hospital (MMH) for PPREMO between February 2013 and February 2016 (RBWH only), and for PREBO between February 2016 and December 2019 (RBWH, MCH, MMH). Some PREBO participants received follow up including imaging at the Queensland Children's Hospital (QCH).

Ethics approval was obtained from the Royal Brisbane and Women's Hospital Human Research Ethics Committee (HREC/12/QRBW/245), the Royal Children's Hospital (HREC/15/QRCH/7), and The University of Queensland (2012001060, 2015000290) with reciprocal approval from the CSIRO Health and Medical HREC (2019\_013\_RR). Both cohort studies were registered with the Australian New Zealand Clinical Trials Registry (PPREMO: ACTRN12613000280707; PREBO: ACTRN12615000591550). Informed written parental consent was obtained for each infant.

### 2.2. Image acquisition

Infants were scanned at 29–35 weeks PMA and again at TEA during natural sleep using a 3T MRI with a dedicated neonatal head coil in an MR compatible incubator or paediatric head coil with no incubator (Table 1). PREBO infants were placed on an immobilisation pillow in the incubator to minimise movement. Noise from the MRI scanner was attenuated using mini muffs (Natus Medical Inc., San Carlos, CA). All infants were monitored with pulse oximetry and electrocardiographic monitoring. No sedation or anaesthesia was used. The PPREMO study used multi-echo T2-weighted turbo spin-echo (TSE) volumes acquired in the axial plane, while in the later PREBO study at RBWH, QCH and MCH, we acquired three orthogonal T2-weighted images (in axial, coronal, and sagittal plane) to improve image quality in relation to motion. Imaging parameters are detailed in Table 1 below.

### 2.3. Image processing

We utilised the same MRI processing pipeline as in our previous work (Pagnozzi et al., 2023), including conversion of MRI, "thick slice" neonatal MRIs to a high-resolution image and segmentation using the state-of-the-art dHCP structural pipeline (Makropoulos et al., 2018b). Super-resolution reconstructed images and dHCP segmentations were visually inspected and rated as 'failed' (no reconstruction obtained), 'poor' (major artefacts/accuracy errors), 'good' (minor artefacts/accuracy errors), or 'excellent' (no or negligible artefacts/accuracy errors). A quality rating of 'good' or 'excellent' was required for inclusion in further analysis. As the cohort included participants with intraventricular haemorrhage, in many cases of brain injury, specifically severely enlarged ventricles, resulted in an inaccurate segmentation. As a result, these participants where segmentation failed were excluded from the analysis. We extracted measures of cortical thickness (CT), sulcal depth (SD), surface area (SA) and gyrification index (GI) of each lobe of the cerebral cortex (frontal, parietal, temporal and occipital), as defined by the Gousias neonatal atlas (Gousias et al., 2013), and 8 tissue volumes (extracerebral CSF, cortical GM, WM, ventricles, cerebellum,

**Table 1**  
Scanner details and imaging parameters across the two studies and three sites.

Cohort	PPREMO	PREBO	
Site	RBWH	RBWH/ QCH	MCH
MRI scanner	Siemens Tim Trio	Siemens Skyra	Philips Ingenia
Head coil	8-channel neonatal head coil, Lammers LMT incubator (first timepoint only)		32-channel head coil, Phillips
T2 sequence	T2 TSE	3x T2 HASTE	3x T2 SSh TSE
TR	10,580 ms	2280 ms	2280 ms
TE	27/122/189 ms	117 ms	117 ms
Flip angle	150	120	120
Field of view	144 × 180mm	SAG: 162.6 × 200mm COR: 137.6 × 200mm AX: 137.6 × 200mm	SAG: 160 × 200mm COR: 160 × 200mm AX: 160 × 200mm
Matrix	204 × 256mm	SAG: 208 × 256mm COR: 176 × 256mm AX: 176 × 256mm	SAG: 198 × 252mm COR: 197 × 252mm AX: 198 × 252mm
Voxel size	0.7 × 0.7 mm	0.8 × 0.8 mm	0.8 × 0.8 mm
Slice thickness	2 mm	1.8 mm	1.8 mm
Scan time	5:40 min	SAG: 3:27 min COR: 3:11 min AX: 2:53 min	SAG: 2:05 min COR: 1:08 min AX: 1:54 min

AX, Axial; COR, Coronal; LMT, Lammers Medical Technology; MCH, Monash Children's Hospital; RBWH, Royal Brisbane & Women's Hospital; SAG, Sagittal; TE, Echo Time; TR, Repetition Time; TSE, Turbo Spin Echo.

DGM, brainstem, hippocampi/amygdala), resulting in a total of 24 variables. An example segmentation provided by the dHCP pipeline is illustrated in Fig. 1, for an early and TEA MRI.

#### 2.4. Concurrent neonatal assessments

Infants were assessed at Early and TEA timepoints, and additionally at 3 months CA, with the General Movements Assessment (GMA). This assessment can be conducted from birth to 20 weeks CA (Olsen et al., 2018) and is a reliable tool for identifying infants at risk of adverse neurodevelopmental outcomes, including CP (Hadders-Algra et al., 2024). In the GMA, movement was classified as normal, poor repertoire, cramped synchronised, or chaotic in the writhing period, and as normal, abnormal fidgety movements, or absent fidgety movements in the fidgety period. In addition, a measure of socioeconomic status (SES) and parenting factors were collected through a questionnaire completed by infants' primary care giver. This questionnaire provides a raw score from 0 to 12, with scores of 2 and above being considered high social risk (George et al., 2015a; Roberts et al., 2008) in line with other research in this population (see (George et al., 2015b), Supplementary Table 1).

#### 2.5. Neurodevelopmental outcomes at 24 months corrected

Neurodevelopmental outcomes at 2 years CA were measured with the Bayley Scales of Infant and Toddler Development – Third Edition (Bayley-III) standardised assessment (Bayley, 2006), with the motor, cognitive and language composite scores being used (Armstrong and Agazzi, 2010; Crais, 2010). Motor outcome was further measured using the standardised Neuro-Sensory Motor Developmental Assessment

(NSMDA) tool (Burns, 1992). These assessments were conducted by an experienced paediatric physiotherapist who was blinded to all earlier clinical and MRI findings. While the Bayley-III is a norm-referenced discriminate measure of developmental functioning of the infant (mean score 100, ±15 standard deviation based on Australian infant data (Anderson et al., 2010)) with higher scores indicating better development, the NSMDA is a categorical score of motor performance, with scores that are normal for developmental age (6–8), and delayed skill attainment with scores ranging from: minimal motor dysfunction (9–11), mild (12–14), moderate (15–19), severe (20–25), and profound disability (>25). Both the Bayley-III and NSMDA have been shown to predict functional outcomes beyond 3 years of age (Boswell et al., 2017; Lowe et al., 2023).

#### 2.6. Statistical analyses

Prior to associating MRI measures with 2-year outcomes and to minimise the risk of overfitting on our dataset, we investigated the collinearity of the extracted MRI measures using a correlation matrix. Highly collinear measures were then grouped if anatomically appropriate. Then the sparse, independent set of structural measures from MRIs taken at the early time point and TEA were associated separately and combined with 2-year outcomes using the Least Absolute Shrinkage and Selection Operator (LASSO) regression (Tibshirani, 1996). To further reduce the risk of model overfitting, the large number of structural measures, relative to the number of participants, was further reduced with model simplification (i.e. variable reduction) based on the L1 penalty. Key patient demographics and clinical variables were also included in all models as confounders, which include sex, GA at birth, PMA at MRI scans, socioeconomic status (raw score) and age at NSMDA assessment, as well as covariates for cohort (PPREMO/PREBO) to account for differences in the MRI acquisition and the demographics of the cohort itself, and MR image quality (which was visually classified as 'good', 'fair', 'poor' or 'unusable'). Age at Bayley-III assessment was not included into any model as the measure inherently accounts for this. Variable reduction is performed implicitly in LASSO with a penalty term (alpha), with higher values enforcing sparser (fewer non-zero) model coefficients. Models were constructed using 30-week data and TEA data separately, as well as combining MRI measures from both timepoints. In addition, models with only confounders (i.e. no exposure MRI measures) were also constructed to determine the potential benefit of adding MRI information. All MRI measures were standardised prior to LASSO regression.

For each time point, complete data was split into the identical training (75 %) and test (25 %) sets. Models were generated on the training set, with the optimal sparsity term (alpha) determined using mean-square error and 5-fold cross validation. The alpha of the best performing model was then applied to the entire training set, yielding an optimal training model (which included only the retained features from LASSO as well as all confounders). This optimal training was applied to the test set, producing a Pearson's r correlation between actual versus predicted 2-year outcome as well as a mean absolute error (MAE). This was repeated for each clinical score, (Bayley motor, cognitive, language, and NSMDA), with each using the same train and test splits. Additionally, outcome measures on the test set were dichotomised as 'normal' or 'poor outcome', which for Bayley-III was defined as 1 standard deviation below the mean (scores <85) and for NSMDA as any functional grade indicating disability (scores >11). From this the Area Under the Curve (AUC) was quantified, which measured the overall classification agreement between the predicted model and the dichotomised outcome.

To investigate potential changes in brain growth observed among preterm-born infants ex-utero compared to term born controls (Bouyssi-Kobar et al., 2016), frequently MRI measures retained by LASSO were compared between preterm infants at the 'early' and 'TEA' timepoints, and term-born infants (at 'Term').

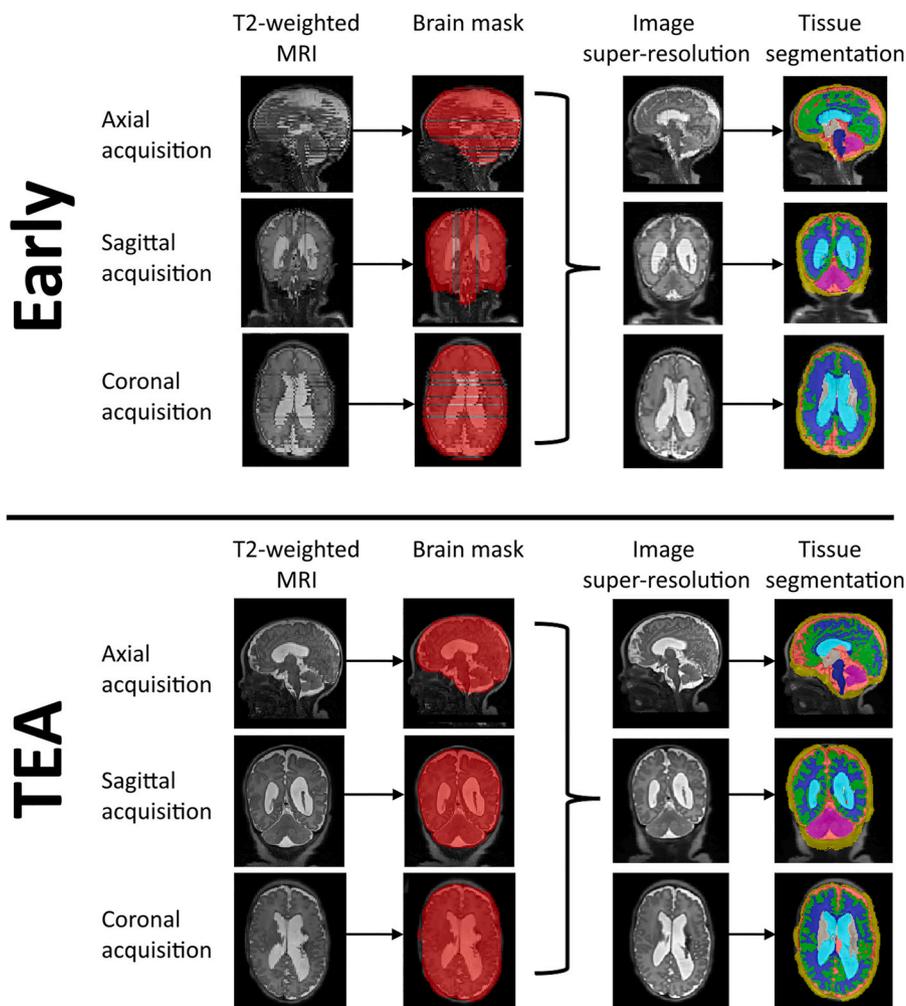


Fig. 1. For both an 'early' and 'TEA' MRI of the same PREBO participant, the original T2 MRIs are shown (first column) in the sagittal, coronal and axial planes, respectively. These images underwent brain masking (shown in red in the second column), with slices containing artefact removed. These three acquisitions are combined into a single high-resolution MRI (third column), and this was then segmented using the dHCP pipeline with the tissue labels illustrated in the fourth column. dHCP, Developing Human Connectome Project; MRI, Magnetic Resonance Images; TEA, Term Equivalent Age.

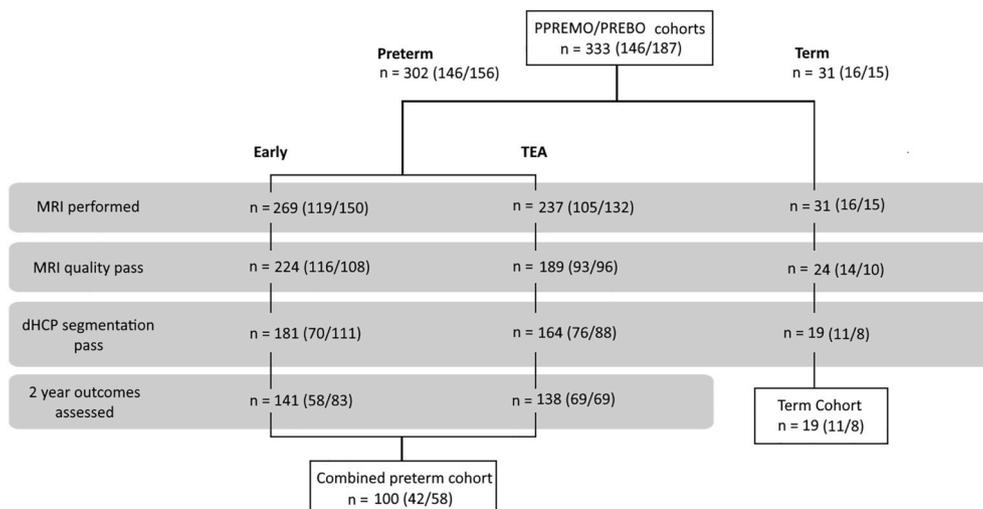


Fig. 2. Diagram illustrating data available for statistical analysis, with reasons for exclusion. Number in brackets shows participants from each cohort (PPREMO/PREBO). dHCP, Developing Human Connectome Project; MRI, Magnetic Resonance Images; TEA, Term Equivalent Age.

**Table 2**

Baseline characteristics of the very preterm cohort with MRI at both early and TEA timepoints with 2-year follow-up assessments. Characteristics of the term-born infants (scanned only at TEA) used to compare structural biomarkers with the preterm cohort is also given below.

	Preterm Cohort (n = 100)		Term-born (n = 19)
MRI PMA early, median weeks + days (range)	32 <sup>+4</sup> (29 <sup>+2</sup> - 36 <sup>+2</sup> )		NA
MRI PMA TEA, median weeks + days (range)	41 <sup>+6</sup> (38 <sup>+3</sup> - 45 <sup>+2</sup> )		41 <sup>+6</sup> (39 <sup>+3</sup> - 44 <sup>+5</sup> )
Gestational age, median weeks + days (range)	28 <sup>+2</sup> (24 <sup>+3</sup> - 30 <sup>+6</sup> )		39 <sup>+6</sup> (38 <sup>+4</sup> - 41 <sup>+0</sup> )
Male, n (%)	44 (44 %)		7 (37 %)
Socioeconomic status, mean (% at risk, >2)	1.71 (25 %)		0.21 (0 %)
GMs classification at 30w and TEA			
Normal	42	41	–
Poor Repertoire - writhing period	53	46	–
Cramped Synchronised - writhing period	3	6	–
Chaotic - writhing period	0	0	–
Missing	2	7	–
Bayley motor composite at 2 years CA (SD)	96 (23)		NA
Bayley cognitive composite at 2 years CA (SD)	98 (16)		NA
Bayley language composite at 2 years CA (SD)	92 (22)		NA
NSMDA functional grade at 2 years CA			
Normal for developmental age (6–8)	67		NA
Minimal motor dysfunction (9–11)	22		NA
Mild motor dysfunction (12–14)	10		NA
Moderate motor dysfunction (15–19)	0		NA
Severe motor dysfunction (20–25)	1		NA
Profound disability (>25)	0		NA
Observed brain injury			
Hydrocephalus, n (%)	1 (1 %)		NA
Periventricular leukomalacia, n (%)	1 (1 %)		NA
Intraventricular haemorrhage, n (%)	30 (30 %)		NA
Grade I, n	19		NA
Grade II, n	7		NA
Grade III, n	1		NA
Grade IV, n	3		NA
Clinical characteristics			
Clinical Chorioamnionitis, n (%)	16 (16 %)		NA
Confirmed sepsis, n (%)	1 (1 %)		NA
Days until discharge, days, median (SD)	74 (34)		NA
Clinical measures of care			
Surgery, n (%)	13 (13 %)		NA
Antenatal corticosteroids, n (%)	68 (68 %)		NA
Days of ETT ventilation, days, median (SD)	3 (10.1)		NA
Days of CPAP, days, median (SD)	16 (13.1)		NA
Hours of oxygen therapy, hours, median (SD)	54 (390.2)		NA
Days of TPN, days, median (SD)	11 (5.2)		NA
Hours of phototherapy, hours, median (SD)	51 (53.7)		NA

CA, corrected age; CPAP, Continuous Positive Airway Pressure; ETT, Endotracheal Tube; GMs, General Movements assessment; NSMDA, Neuro-Sensory Motor Developmental Assessment; PMA, Postmenstrual age; SD, Standard deviation; TEA, Term Equivalent Age; TPN, Total Parenteral Nutrition.

### 3. Results

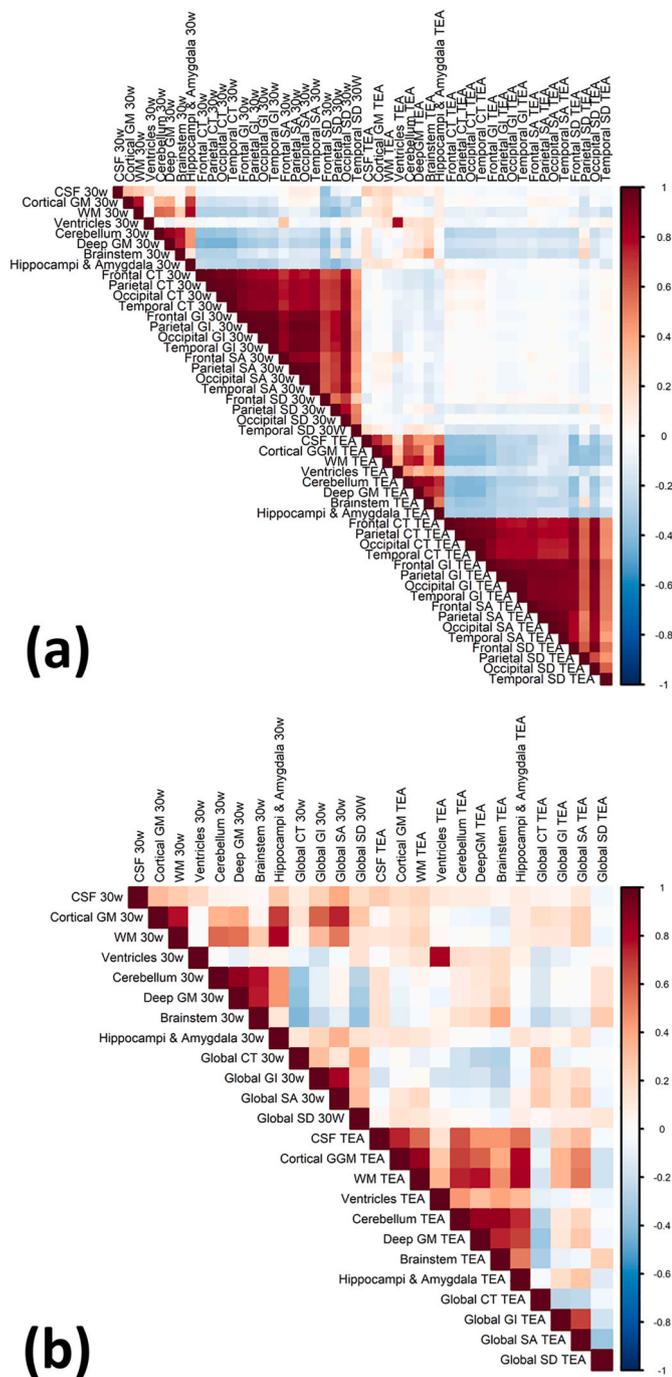
#### 3.1. Participants

Of the total number of participants in the PPREMO/PREBO cohorts (N = 333), the total number of participants available for analyses was n = 100 at both time points (see flowchart Fig. 2). This is accounting for poor MRI data quality (n = 62 early, 96 TEA) and/or segmentation errors (n = 47 early, 48 TEA) at either time-point, or missing assessment data at 2-years (n = 93), with the breakdown of data loss at each timepoint illustrated in Fig. 2. All of these participants formed part of the cohort of the our previous paper (n = 139) (Pagnozzi et al., 2023). Of these n = 100 infants, n = 42 were part of the PPREMO study and 58 were part of the PREBO (RBWH) study. The overall demographics, clinical assessments and measures of care of the preterm infants are detailed in Table 2 below. While there was only one case of confirmed sepsis in this cohort (1 %), thirteen infants underwent surgery (13 %), and there was a high incidence of chorioamnionitis during pregnancy (16 %) as well as intraventricular haemorrhage (30 %) as observed on cranial ultrasound. The demographics of the included participants in the analysis and the excluded participants are in Supplementary Table 1, with the excluded cohort having a significantly greater GA (1 week median difference, p = 0.04) and an increased incidence of

hydrocephalus (n = 9, p = 0.03), sepsis (n = 11, p < 0.01) and endotracheal tube ventilation (p = 0.04).

#### 3.2. Correlation of structural measures

All 24 structural measures (8 vol, 16 cortical shape) at each of the two timepoints were put into a correlation matrix (Fig. 3a). Independent variables with a Pearson's correlation above 0.9 are highly correlated and can be reduced or combined. We observed that all cortical measures were highly correlated with each other, independently for both the 'early' (labelled 30w) and 'TEA' timepoints. To reduce model complexity and for the feature set to remain anatomically consistent, all cortical measures were grouped from a frontal, temporal, occipital and parietal lobule average score to a global average. This reduced the number of structural measures from 48 to 24 (12 at each time point), and this reduced set of structural features showed markedly reduced collinearity (Fig. 3b). Furthermore, to test the importance of cortical shape, volume-only models were also constructed and validated on the same training and testing sets.



**Fig. 3.** Correlation matrix of (a) all structural measures, and (b) manually reduced feature set using global measures of cortical shape. CSF, Cerebrospinal Fluid; CT, Cortical Thickness; GI, Gyrfication Index; GM, Grey Matter; SA, Surface Area; SD, Sulcal Depth; WM, White Matter; TEA, Term Equivalent Age.

**3.3. Structural brain measures impacting 2-year Bayley-III composite scores and NSMDA**

For each of the 12 models (4 outcome assessments, each using 2 MRI time points individually and combined), the features were found to be associated with changes in 2-year outcomes and hence were retained from LASSO regression. The coefficients for the ‘early’ and ‘TEA’ models are provided in Table 3. With the exception of the NSMDA models, most structural measures were retained from LASSO. In all cases, confounders were specifically retained by LASSO. Table 3 outlines the coefficient of each retained structural measure, and at which time point, and the

model fit provided by the pseudo R2 for the complete model and for models without MRI features (i.e. only clinical variables). In all models the pseudo R2 of models with MRI features was increased compared to the clinical variables only models, indicating the MRI measures independently explained variance in the outcomes. Structural biomarkers that are associated with better 2-year outcomes will show a positive association for the Bayley-III composite scores, and negative associations for NSMDA. This is because while higher Bayley-III scores indicate better outcomes, higher NSMDA scores indicate greater dysfunction.

Cortical grey matter was positively associated with better outcomes in most models (5 of 6), and conversely larger ventricles were consistently associated with poorer outcomes across 5 of the 6 models. White matter volume was generally positively associated with better outcomes (in 4 of 6 models), and similarly cerebellum volume was positively associated with better outcomes in 4 of 6 models. The remaining structural measures were more mixed in terms of association with outcome; however we note that gyrfication index was positively associated with higher Bayley motor, cognitive and language composite scores, whilst sulcal depth was positively associated with improved Bayley scores and NSMDA grade. Extracerebral CSF (in all 6 models) and ventricular CSF (in 5 of 6 models) was negatively associated with Bayley outcomes, which was also the only finding consistent with the univariate analysis that did not adjust for confounders (Supplementary Table 2, Supplementary Fig. 1).

Higher social risk was negatively associated with all Bayley composite scores (average reduction –3.50 per SES risk level, accounting for MRI measures). Higher GA at birth was associated with better outcomes (score increase 0.35 per week). PMA at early MRI was negatively associated with Bayley outcomes (score decrease 5.34 per week), however the PMA at Term MRI was not strongly associated with 2-year outcomes (score decrease of 0.08 per week). As shown previously (Pagnozzi et al., 2023), females performed better than males across all models (average increase 4.29), and PREBO participants had higher Bayley outcomes at 2 compared to PPREMO participants (average increase 6.24). There was little association between Bayley outcomes and the included confounders; GMAs classification at the early (0.02) and TEA timepoint (0.04), and image quality (0.03).

**3.4. Prediction model validation and classification performance**

The trained TEA LASSO models were used to predict assessments scores for the unseen test data (25 %). Pearson’s r correlations between actual and predicted assessment scores are provided in Table 4. Moderate to strong correlations were observed for all assessments, with all but two statistically significant in the test set when accounting for multiple comparisons (Bonferroni correction  $p < 0.05/12$  tests = 0.004). All models using structural measures from the early MRI, and both early and TEA MRI combined, were significant in the test set, with the combined model showing the greatest predictive accuracy in all cases except the Bayley-III language score. When looking at measures of volume alone, excluding measures of cortical shape which are more computationally difficult to obtain, predictive performance was overall slightly reduced on the test set (Supplementary Table 3).

In addition, the sensitivity, specificity and AUC metrics for classifying mild developmental delay from Bayley composite scores (<1 SD below mean, scores <85) (Celik et al., 2020) and disability from the NSMDA functional grade (scores >11) are shown in Table 5. The scatter plots of the model predicted score versus the actual score are illustrated in Fig. 4.

**3.5. Comparing ex-utero brain development of preterm-born infants to term-born infants**

For the frequently retained structural measures obtained from LASSO, the measures were compared between the preterm infants at the ‘early’ and ‘TEA’ timepoints (n = 100), as well as the term-born infants

**Table 3**

Morphometric biomarkers and their association with 2-year Bayley-III composite scores and NSMDA functional grade. The MRI (exposure) variables are shown above the included confounders. Model fit (pseudo  $R^2$ ) with and without the MRI exposure variables is presented at the bottom of the table. Exposure variables dropped by LASSO are designated ‘-’.

Feature	Bayley motor composite score		Bayley cognitive composite score		Bayley language composite score		NSMDA functional grade	
	Early coefficient	Term coefficient	Early coefficient	Term coefficient	Early coefficient	Term coefficient	Early coefficient	Term coefficient
Extracerebral CSF	-10.4	-5.41	-1.85	-0.91	-6.73	-4.72	0.29	3.20
Cortical grey matter	16.6	13.7	-	3.16	10.7	5.43	-0.16	-
White matter	-	10.1	2.82	4.05	-	4.47	-0.92	-
Ventricles	-1.48	-3.45	-1.27	-5.19	-1.60	-	-	0.09
Cerebellum	-	5.38	1.41	-1.05	10.9	3.47	-	-
Deep grey matter	8.51	0.53	8.35	-	5.68	5.89	-0.28	-0.26
Brainstem	11.8	5.86	-	4.11	-9.23	7.86	-	-0.76
Hippocampus/ amygdala	-7.07	1.04	5.90	-	-10.3	-0.51	-	-0.12
Global CT	-0.28	-4.86	-2.07	-3.14	-0.85	-2.51	-	0.25
Global GI	-0.01	9.59	0.13	-2.17	1.61	2.96	-	-1.03
Global SA	-	-0.07	1.05	0.007	9.70	-6.42	-	-
Global SD	0.22	10.3	1.24	-0.58	1.83	-0.99	-0.29	-0.83
Sex (REF: Male)	0.85	1.99	2.88	7.69	9.96	2.35	0.44	0.54
GA at birth	0.64	0.35	0.15	0.12	0.49	0.35	-0.007	0.009
PMA at MRI	-6.42	-0.47	-2.47	-0.13	-7.12	0.35	0.59	-0.02
SES	-2.11	-6.23	-3.59	-2.39	-2.54	-4.12	0.30	0.28
GMs classification	0.05	0.05	-0.02	-0.005	0.04	0.08	0.001	0.04
Cohort (REF: PPREMO)	11.3	13.4	0.83	0.66	6.11	5.16	-0.008	-0.56
Image quality	0.05	0.01	0.08	0.01	0.03	0.02	0.003	0.03
<b>Pseudo R2</b>	0.813	0.799	0.811	0.799	0.807	0.709	0.919	0.722
<b>Pseudo R2 (no MRI)</b>	0.215	0.203	0.230	0.217	0.260	0.176	0.142	0.258

CSF; CT, Cortical Thickness; GA, Gestation Age; GI, Gyrfication Index; MRI; Magnetic Resonance Imaging; PMA, Postmenstrual age; SA, Surface Area; SD, Sulcal Depth; SES, Socio-economic status.

**Table 4**

Correlations between the best performing models and the test set assessment scores unseen by the model.

	Early MRI		Term MRI		Combined	
	Pearson's r	MAE	Pearson's r	MAE	Pearson's r	MAE
Bayley motor composite	0.480* (p = 0.0004)	14.14	0.496* (p = 0.0002)	18.67	<b>0.561** (p &lt; 0.0001)</b>	<b>16.96</b>
Bayley cognitive composite	<b>0.551** (p &lt; 0.0001)</b>	<b>10.87</b>	0.456* (p = 0.0008)	12.11	0.492* (p = 0.0004)	14.99
Bayley language composite	0.399 (p = 0.004)	16.41	0.373 (p = 0.007)	16.92	0.373 (p = 0.004)	15.11
NSMDA functional grade	0.511* (p = 0.0001)	1.84	0.435* (p = 0.0016)	2.22	<b>0.534** (p &lt; 0.0001)</b>	<b>2.04</b>

Bold indicates the timepoints with best statistically significant prediction accuracy.

\*p < 0.004, \*\*p < 0.0001. MRI, Magnetic Resonance Images; NSMDA: Neuro-sensory Motor Developmental Assessment.

**Table 5**

Area Under the Curve (AUC), sensitivity and specificity metrics indicating agreement between the best performing models and the test set assessment scores unseen by the model.

	Early MRI			Term MRI			Combined		
	AUC	Sens (%)	Spec (%)	AUC	Sens (%)	Spec (%)	AUC	Sens (%)	Spec (%)
Bayley motor composite	0.84	90	80	0.82	78	100	0.88	85	100
Bayley cognitive composite	0.82	72	86	0.58	95	50	0.84	100	78
Bayley language composite	0.87	83	86	0.78	78	85	0.75	95	50
NSMDA functional grade	0.95	100	91	0.91	100	71	0.98	100	95

Bolded values are the time-point with the highest AUC metric for each outcome.

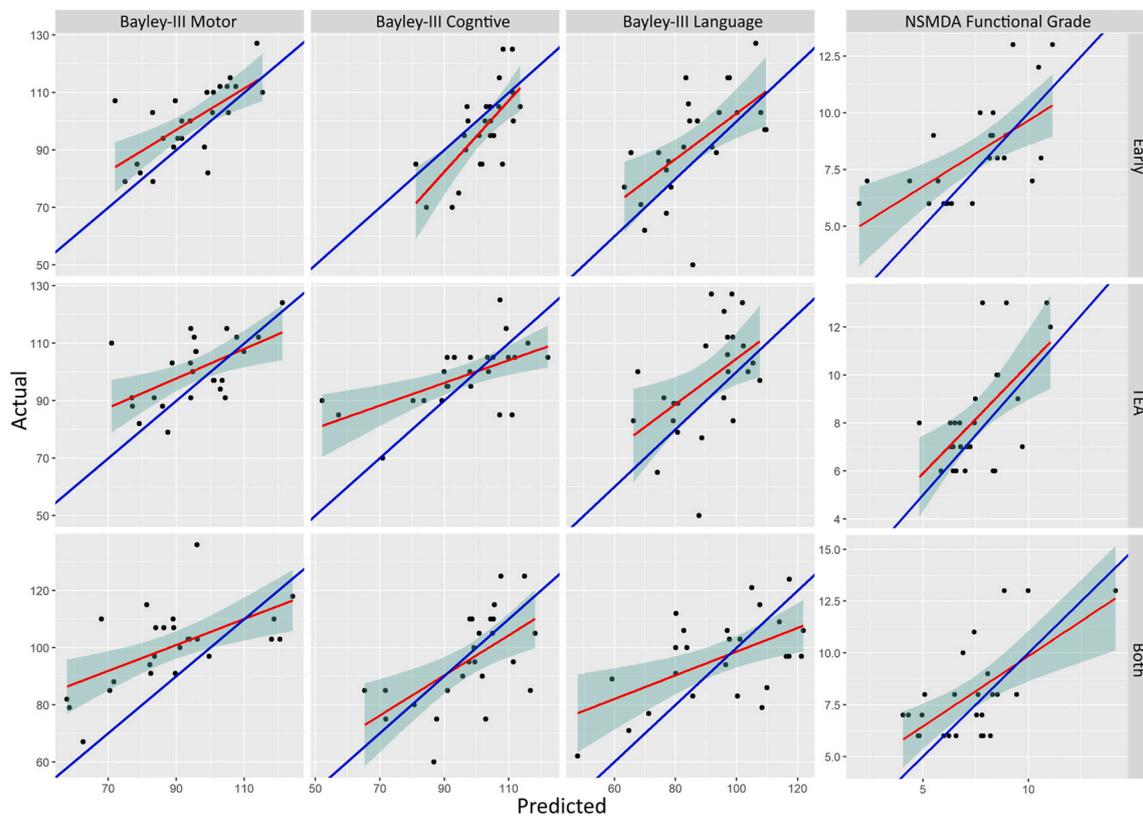
AUC, Area Under the Curve; MRI, Magnetic Resonance Images; Sens, Sensitivity; Spec, Specificity; NSMDA: Neuro-sensory Motor Developmental Assessment.

(n = 19) (Fig. 5). In most cases, the significant difference we see at the ‘early’ timepoint largely disappears at TEA, with the preterm infants showing no significant difference from the term-born group. This is not true for cortical grey matter however, which shows a slight but significant reduction in volume at term compared to term-born infants (7.5 %). While this difference is not reflected in cortical thickness, we note that the range of the cortical thickness measured was biologically congruent (1.3 mm at 28 weeks of PMA to 1.8 mm at term equivalent) (Liu et al., 2021). Also due to the presence of ventriculomegaly in the preterm cohort, lateral ventricle volumes are significantly larger at TEA compared to the term group. We also observed differences in volume

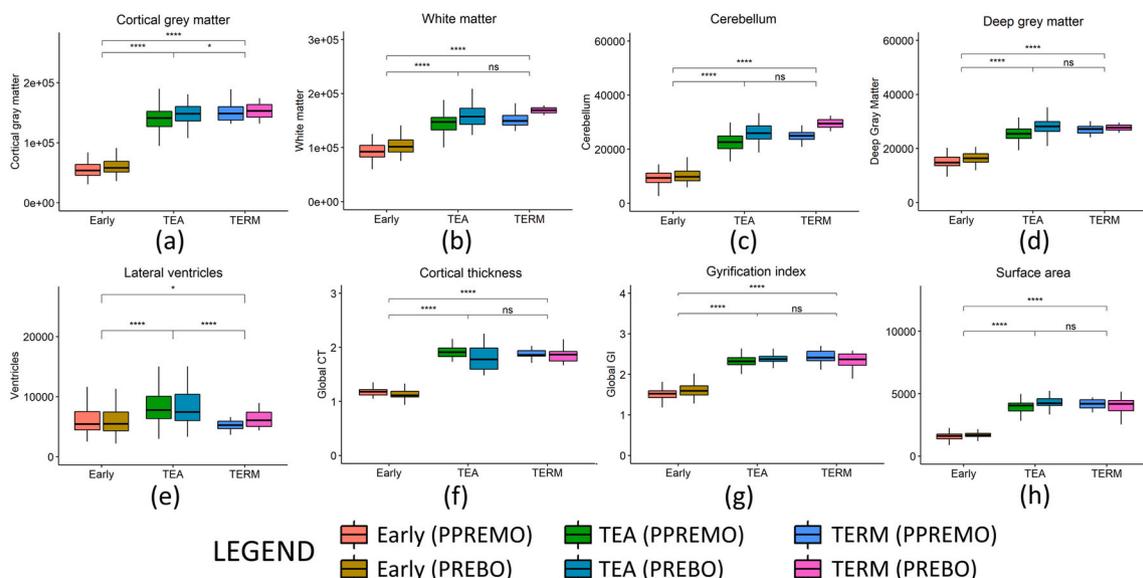
between the PPREMO and PREBO studies, with PREBO reporting larger cortical GM (2.5 and 4.3 % at Early and TEA time points respectively), WM (7.8 and 11.8 %), cerebellar (2.0 and 13.4 %) and deep GM (8.5 and 10.5 %), which likely reflect the differences in image acquisition (Table 1).

#### 4. Discussion

Measures from structural MRI taken at <32 weeks PMA (‘early’) and at TEA are associated with motor, cognitive and language ability at two-years CA in a very preterm cohort. Twelve LASSO models were



**Fig. 4.** Test set correlations between the predicted outcome from the best performing LASSO model and the actual Bayley III assessment scores. Columns illustrate; Bayley-III motor composite score, first column; cognitive composite score, second column; language composite score, third column; and NSMDA (fourth column) using measures from the early MRI (first row), TEA MRI (second row) and both time points combined (third row). The regression line is shown in red, and the  $y = x$  line is shown in blue. LASSO, Least Absolute Shrinkage and Selection Operator; MRI, Magnetic Resonance Images; NSMDA, Neuro-sensory Motor Developmental Assessment; TEA, Term Equivalent Age.



**Fig. 5.** Box and whisker plots of several extracted structural measures comparing the preterm cohort acquired early and at TEA, and the term-born cohort acquired at TEA (TERM). PPREMO and PREBO cohorts are plotted separately, however statistics are only compared between ‘early’, ‘TEA’ and ‘TERM’ groups. CT, Cortical thickness; GI, Gyrfication index; SA, Surface area; TEA, Term Equivalent Age.

generated, one for each of the four 2-year outcomes, and one for each timepoint (early, TEA, and both combined). By combining brain morphometrics with key clinical confounders using LASSO, ten of the twelve investigated models demonstrated moderate associations with 2-year

functional outcomes in an independent test set (25 % of available data) unseen by the model. This study demonstrates that automated structural biomarkers from TEA MRI is predictive of 2-year outcomes (AUC 58–82 %), albeit showed lower prognostic accuracy than

biomarkers from the early MRI (AUC 82–87 %), which is similar to our previous findings (Pagnozzi et al., 2023). Furthermore, there is some independence between the time-points as both timepoints combined led to the highest predictive performance for motor outcomes (Table 4). LASSO regression revealed volumes of the CSF, cortical GM, WM and cerebellum were consistently associated with outcomes at 2 years CA (Table 3), which expands on the list of structural biomarkers found in previous work on this cohort that utilised MRI scoring methods at both early and TEA time-points (George et al., 2021).

This indicates the potential value of quantitative analysis in clinical practice, specifically at the early time-point, for the prediction of long-term motor and cognitive outcomes (Hinojosa-Rodríguez et al., 2017). In terms of implications for clinical practice, this research suggests that acquiring early MRI in the first weeks of life, in addition to the TEA MRI (>38 weeks) typically performed in clinical practice, may more fully characterise brain structure and the trajectory of brain development (AUC = 84 %) (Inder et al., 2021).

While most structural measures were retained at both time points, we observed a consistent volumetric effect of increased GM and WM volume, and decreased extracerebral CSF and ventricle volume, to be associated with improved 2-year outcomes. This is consistent with findings in the literature comparing preterm and term-born infants at TEA (Kelly et al., 2020; Thompson et al., 2007), and reflects that underlying cerebral GM/WM dysmaturation leads to poorer outcomes (Inder et al., 2005). Similarly we found that greater deep grey matter and cerebellar volume, specifically at the TEA time-point, is positively associated with improved outcomes, which has previously been shown to predict neurodevelopmental outcomes at 4 years of age (Young et al., 2015). In terms of cortical morphology, gyrification index was found to be associated with better Bayley-III scores which has been shown previously (Julia E Kline et al., 2020), while cortical thickness at TEA was negatively associated with better Bayley-III scores. This latter result has been shown in previous morphological/microstructural studies, finding greater cortical thickness in frontal, insular and anterior parietal cortices in preterm infants compared to term-born infants at TEA (Dimitrova et al., 2021). At both time-points, surface area was positively associated with neurodevelopmental outcomes at 2-years of age which has been observed previously (Julia E Kline et al., 2020; Julia E. Kline et al., 2020). Despite grouping of structural measures, both gyrification index and sulcal depth as well as surface area reflect some of the same anatomical features, suggesting there may still be some collinearity in the models and further variable reduction may improve model performance. Interestingly, while LASSO aims to reduce the number of structural biomarkers, in all models most structural markers were retained, indicating a more global impact of brain structure on developmental outcome (Thompson et al., 2005). This could, in part be due to our efforts to retain independent structural features through grouping cortical features.

As we found in our previous work (Pagnozzi et al., 2023), higher social risk was negatively associated with poorer motor, cognitive and language outcomes, and was one of the biggest drivers we found of 2-year outcomes. Social risk is known to be a strong factor impacting childhood development (Dalmaijer et al., 2023). Further, we observed GMAs classification to be weakly, positively associated with Bayley composite scores, in contrast to our previous work (average reduction –1.01 per early GMAs score). The authors note that 3-month GMAs have been found to be the most predictive of later outcomes (John et al., 2022). Interestingly while PMA at Term MRI was not strongly associated with 2-year outcomes (score decrease of 0.08 per week), PMA at early MRI was negatively associated with Bayley outcomes (score decrease 5.34 per week) i.e. a later “early” MRI was associated with lower Bayley outcomes. The authors speculate that this could be due to babies with more complex clinical needs having their early MRI delayed until they were clinically stable enough to undergo MRI. These babies are also more likely to have poorer outcomes; potentially explaining the observed association between PMA at early MRI and 2-year outcomes.

Alternatively, this may be due to collinearity between brain size and development at this time point and PMA at this time point, the latter of which was manually retained in all LASSO models. Brain structure of the very preterm infants at TEA did not significantly differ from the term-born controls (Fig. 5), with the exception of cortical grey matter and the lateral ventricles which are a reflection of underlying brain injury (Agut et al., 2020). This does not support the observed volumetric differences between preterm-born infants ex-utero compared to term born controls (Bouyssi-Kobar et al., 2016), however we note the limited statistical power in being able to detect potential differences due to the small term-born cohort (n = 19).

Further, using these structural biomarkers we were able to classify mild/moderate/severe motor and cognitive developmental delay with an AUC of 0.84 with early MRI, 0.82 with TEA MRI, and 0.88 combined (Table 5). This result is slightly improved on our previous deep-learning approach to detect mild motor delay on a smaller sample from the same cohort (AUC 0.72, n = 77) (Saha et al., 2020), and similar to a structural MRI study utilising TEA MRI (AUC 0.78 for low cognitive outcome, and an AUC 0.80 for low motor outcome) (Moeskops et al., 2017). These findings are consistent with similar research utilising early MRI for the identification of white matter injury (AUC 0.75 for mild motor delay, and AUC 0.8 for mild cognitive delay), which was observed at a later timepoint (4.5 years) suggesting this developmental delay continues through early childhood (Cayam-Rand et al., 2019). We show a similar classification accuracy of mild language delay (AUC 0.80) to mild motor and cognitive delay, which is lower than previous research (AUC 0.85) (He et al., 2021) which instead utilised a multi-modal MRI (structural, diffusion and functional MRI) to leverage both the structural and functional connectome for prediction. When looking at diagnostic accuracy, obtained sensitivity (72–100 %) and specificity (50–100 %) are on par with previous findings using TEA MRI to predict 18–24 month CA outcomes (sensitivity 33–100 %, specificity 30–97 %) (Banihani et al., 2021). Overall, we found that both early and TEA MRI was able to classify normal vs. mild/moderate/severe motor, cognitive and language delay. This was observed looking at both timepoints independently, as well as combined which led to slightly improved classification performance (Table 5).

There are several limitations to the present study, the first being the variability in the structural sequences used due to recruitment spanning two separate projects over 7 years. There was variation introduced by sequence specific preprocessing for the T2-TSE and T2 HASTE sequences to conform with processing of the dHCP structural pipeline. As a result, both image quality and sequence (cohort) were included as confounders in the LASSO analysis to account for this difference. Further, we note that both sequences have a smaller slice thickness (<2 mm) than a standard clinical T2 HASTE sequence. As such the pipeline remains to be validated on clinical data, and additionally to leverage the T1w sequence to better define tissue boundaries and to explore myelination through the T1w/T2w ratio (Soun et al., 2017). However, we found the T1w sequences to be of poorer image quality in our cohort, largely due to motion artefact caused by the long duration of the T1w scan (~5 min), and hence these were not included in the analysis. Additionally, there are several methodological improvements that could be performed to improve robustness to more severe image artefact or brain injury. More advanced image super-resolution approaches now exist leveraging deep learning to interpolate accurate anatomical detail for images with slice thickness as high as 6.5 mm (Delannoy et al., 2020; Li et al., 2021). Such approaches can utilise variable slice selection to avoid sampling from motion affected slices (Sui et al., 2021), which unlike the current approach could accommodate multiple adjacent slices impacted by motion. Furthermore, deep learning segmentation approaches, such as iBEAT V2.0 (Wang et al., 2023) have shown remarkable performance compared to other neonatal pipelines. Such an approach may prove effective for MRIs with severe brain injury, which typically failed with the dHCP structural pipeline due to its disparity to the age-matched healthy atlas, and thus were largely excluded from the current

analysis. Finally, there are logistical challenges in performing early MRI in very preterm cohorts in most clinical settings, requiring more careful monitoring of the infant and dedicated hardware (neonatal head coil, incubator).

Future work will incorporate measures of microstructure from diffusion MRI in addition to structural MRI as part of the prediction models (Dubois et al., 2021), to enhance the prediction of long-term outcomes. Associations between diffusion tensor imaging (DTI), yielding measures of fractional anisotropy and mean diffusivity, and neurodevelopmental outcomes at 3 years have been found (Kidowaki et al., 2017). More advanced diffusion models such as fixel-based analysis (Raffelt et al., 2017) provided advantages over simpler models such as DTI in their ability to resolve crossing fibres, and demonstrated stronger associations with motor and cognitive outcomes compared to DTI on the same cohort (Pannek et al., 2020). The authors also note that a limitation of the present study was the included cohort had a relatively low incidence of adverse outcomes (n = 15 infants had a Bayley motor composite score <1 SD below the mean, only n = 1 had severe motor dysfunction according to the NSMDA). As such, in future work the prediction models presented here will be validated on an external cohort to determine their generalisability to the wider preterm population. We also intend to apply these prediction models beyond 2 years CA, to reveal the utility of each imaging time-point in predicting later neurodevelopment (beyond 3 years CA), when intellectual, learning and behavioural delay can be fully determined (Arcangeli et al., 2012). This cohort is currently being followed up at 6 years of age with both advanced neuroimaging and concurrent clinical assessment of motor, neurological and neurobehavioural function as part of the PREBO-6 study (George et al., 2020). At this later time-point, a diagnosis of CP, ASD and school readiness outcomes can be reliably obtained, and can elucidate the long-term predictive value of early and TEA neonatal MRI.

## 5. Conclusions

We have leveraged state-of-the-art structural image analysis on our large, multi-centre, regional cohort of 100 very preterm infants, who underwent 3T MRI imaging at early and Term time points. Using LASSO regression and a model validation strategy to minimise overfitting, we found strong and significant association with motor, cognitive and language function at 2 years CA. We revealed consistent volume differences in the white and grey matter, deep grey matter and cerebellum, as well as surface area and gyrification of the cortex, associated with improved 2-year outcomes. In addition to key patient demographics and clinical variables, including socio-economic status, GMs assessment classification, sex and GA at birth, these biomarkers revealed both time-points are independently predictive of outcomes, suggesting either could be used in clinical practice. Early MRI showed the best prognostic accuracy compared to TEA MRI, indicating that imaging preterm infants before TEA not only help ensure at-risk infants are scanned prior to discharge from hospital, but that brain structure at this time may better identify at-risk infants. These early biomarkers are key to provide targeted interventions to at-risk preterm infants much earlier, when they are most effective.

## CRedit authorship contribution statement

**Alex M. Pagnozzi:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Data curation. **Kerstin Pannek:** Writing – review & editing, Supervision, Methodology, Data curation, Conceptualization. **Roslyn N. Boyd:** Writing – review & editing, Project administration. **Liza van Eijk:** Software, Methodology, Investigation, Data curation. **Joanne George:** Writing – review & editing, Project administration, Funding acquisition, Data curation, Conceptualization. **Samudragupta Bora:** Writing – review & editing, Supervision, Project administration,

Funding acquisition. **DanaKai Bradford:** Writing – review & editing, Supervision. **Michael Fahey:** Writing – review & editing, Data curation. **Michael Ditchfield:** Writing – review & editing, Data curation. **Atul Malhotra:** Writing – review & editing, Data curation. **Paul B. Colditz:** Writing – review & editing, Project administration, Funding acquisition, Data curation, Conceptualization. **Jurgen Fripp:** Writing – review & editing, Supervision, Project administration.

## Funding

The following support has been provided by the Australian National Health and Medical Research Council (NHMRC): Project grant (PREBO NHMRC 1084032) and Investigator Fellowship (RB: NHMRC 1195602) and an NHMRC Clinical CRE (NHMRC 1116442 CRE Australasian CP Clinical Trials Network). PPREMO was funded through grants from the Cerebral Palsy Alliance Research Foundation (IRG1413), the Financial Markets Foundation for Children (2014–074), the Queensland Government (Smart State; Health Practitioner Stimulus Grant) and the Royal Brisbane and Women's Hospital.

## Declaration of competing interest

The authors declare that they have no conflict of interest.

## Acknowledgements

We would like to acknowledge the families of preterm infants who participated in this study, the medical and nursing staff of the neonatal unit, radiology and administrative staff in the department of medical imaging and the Physiotherapists and Psychologists who assessed 24-month outcomes, and Julie Trinder for her management of the imaging, demographic, and clinical data for this cohort.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ynirp.2025.100262>.

## Data availability

The authors do not have permission to share data.

## References

- Agut, T., Alarcon, A., Cabañas, F., Bartocci, M., Martinez-Biarge, M., Horsch, S., 2020. Preterm white matter injury: ultrasound diagnosis and classification. *Pediatr. Res.* 87, 37–49. <https://doi.org/10.1038/s41390-020-0781-1>.
- Anderson, B.J., 2011. Plasticity of gray matter volume: the cellular and synaptic plasticity that underlies volumetric change. *Dev. Psychobiol.* 53, 456–465. <https://doi.org/10.1002/dev.20563>.
- Anderson, P.J., De Luca, C.R., Hutchinson, E., Roberts, G., Doyle, L.W., Victorian Infant Collaborative Group, 2010. Underestimation of developmental delay by the new Bayley-III Scale. *Arch. Pediatr. Adolesc. Med.* 164, 352–356. <https://doi.org/10.1001/archpediatrics.2010.20>.
- Anderson, P.J., Treyvaud, K., Neil, J.J., Cheong, J.L.Y., Hunt, R.W., Thompson, D.K., Lee, K.J., Doyle, L.W., Inder, T.E., 2017. Associations of newborn brain magnetic resonance imaging with long-term neurodevelopmental impairments in very preterm children. *J. Pediatr.* 187, 58–65.e1. <https://doi.org/10.1016/j.jpeds.2017.04.059>.
- Arcangeli, T., Thilaganathan, B., Hooper, R., Khan, K.S., Bhide, A., 2012. Neurodevelopmental delay in small babies at term: a systematic review. *Ultrasound Obstet. Gynecol.* 40, 267–275. <https://doi.org/10.1002/uog.11112>.
- Armstrong, K.H., Agazzi, H.C., 2010. The bayley-III cognitive scale. *Bayley-III Clinical Use and Interpretation* 29–45. <https://doi.org/10.1016/B978-0-12-374177-6.10002-9>.
- Banihani, R., Seesahai, J., Asztalos, E., Terrien Church, P., 2021. Neuroimaging at term equivalent age: is there value for the preterm infant? A narrative summary. *Children* 8, 227. <https://doi.org/10.3390/children8030227>.
- Bayley, N., 2006. *Bayley Scales of Infant and Toddler Development*. Pearson. PsychCorp.
- Boswell, L., Weck, M., Santella, M., Patrick, C., Russow, A., Deregnier, R., 2017. Neurosensory motor developmental assessment at 18–24 months predicts quality of life at 3–1/2 to 5 years. *Dev. Med. Child Neurol.* 59, 61–62. <https://doi.org/10.1111/dmcn.9313511>.

- Bouyssi-Kobar, M., du Plessis, A.J., McCarter, R., Brossard-Racine, M., Murnick, J., Tinkleman, L., Robertson, R.L., Limperopoulos, C., 2016. Third trimester brain growth in preterm infants compared with in utero healthy fetuses. *Pediatrics* 138. <https://doi.org/10.1542/peds.2016-1640>.
- Burns, Y.R., 1992. *NSMDA: Physiotherapy Assessment for Infants & Young Children*. CopyRight Publishing, Brisbane.
- Cayam-Rand, D., Guo, T., Grunau, R.E., Benavente-Fernández, I., Synnes, A., Chau, V., Branson, H., Latal, B., McQuillen, P., Miller, S.P., 2019. Predicting developmental outcomes in preterm infants. *Neurology* 93, e1231–e1240. <https://doi.org/10.1212/WNL.00000000000008172>.
- Celik, P., Ayranci Sucakli, İ., Yakut, H.İ., 2020. Which Bayley-III cut-off values should be used in different developmental levels? *Turk. J. Med. Sci.* 50, 764–770. <https://doi.org/10.3906/sag-1910-69>.
- Chang, H.H., Larson, J., Blencowe, H., Spong, C.Y., Howson, C.P., Cairns-Smith, S., Lackritz, E.M., Lee, S.K., Mason, E., Serazin, A.C., Walani, S., Simpson, J.L., Lawn, J. E., 2013. Preventing preterm births: analysis of trends and potential reductions with interventions in 39 countries with very high human development index. *Lancet* (London, England) 381, 223–234. [https://doi.org/10.1016/S0140-6736\(12\)61856-X](https://doi.org/10.1016/S0140-6736(12)61856-X).
- Crais, E.R., 2010. The bayley-III language scale. Bayley-III Clinical Use and Interpretation 47–75. <https://doi.org/10.1016/B978-0-12-374177-6.10003-0>.
- Dalmajjer, E.S., Gibbons, S.G., Bignardi, G., Anwyl-Irvine, A.L., Siugzdaite, R., Smith, T. A., Uh, S., Johnson, A., Astle, D.E., 2023. Direct and indirect links between children's socio-economic status and education: pathways via mental health, attitude, and cognition. *Curr. Psychol.* 42, 9637–9651. <https://doi.org/10.1007/s12144-021-02232-2>.
- Delannoy, Q., Pham, C.-H., Cazorla, C., Tor-Díez, C., Dollé, G., Meunier, H., Bednarek, N., Fablet, R., Passat, N., Rousseau, F., 2020. SegSRGAN: super-resolution and segmentation using generative adversarial networks — application to neonatal brain MRI. *Comput. Biol. Med.* 120, 103755. <https://doi.org/10.1016/j.compbiomed.2020.103755>.
- Dimitrova, R., Pietsch, M., Ciarrusta, J., Fitzgibbon, S.P., Williams, L.Z.J., Christiaens, D., Cordero-Grande, L., Bataille, D., Makropoulos, A., Schuh, A., Price, A.N., Hutter, J., Teixeira, R.P., Hughes, E., Chew, A., Falconer, S., Carney, O., Egloff, A., Tournier, J.-D., McAlonan, G., Rutherford, M.A., Counsell, S.J., Robinson, E.C., Hajnal, J.V., Rueckert, D., Edwards, A.D., O'Muircheartaigh, J., 2021. Preterm birth alters the development of cortical microstructure and morphology at term-equivalent age. *Neuroimage* 243, 118488. <https://doi.org/10.1016/j.neuroimage.2021.118488>.
- Dubois, J., Alison, M., Counsell, S.J., Hertz-Pannier, L., Hüppi, P.S., Benders, M.J.N.L., 2021. MRI of the neonatal brain: a review of methodological challenges and neuroscientific advances. *J. Magn. Reson. Imag.* 53, 1318–1343. <https://doi.org/10.1002/jmri.27192>.
- Dubois, J., Lefèvre, J., Angleys, H., Leroy, F., Fischer, C., Lebenberg, J., Dehaene-Lambertz, G., Borradori-Tolsa, C., Lazeyras, F., Hertz-Pannier, L., Mangin, J.-F., Hüppi, P.S., Germanaud, D., 2019. The dynamics of cortical folding waves and prematurity-related deviations revealed by spatial and spectral analysis of gyrification. *Neuroimage* 185, 934–946. <https://doi.org/10.1016/j.neuroimage.2018.03.005>.
- George, J.M., Boyd, R.N., Colditz, P.B., Rose, S.E., Pannek, K., Fripp, J., Lingwood, B.E., Lai, M.M., Kong, A.H., Ware, R.S., Coulthard, A., Finn, C.M., Bandaranayake, S.E., 2015a. PPREMO: a prospective cohort study of preterm infant brain structure and function to predict neurodevelopmental outcome. *BMC Pediatr.* 15, 123. <https://doi.org/10.1186/s12887-015-0439-z>.
- George, J.M., Boyd, R.N., Colditz, P.B., Rose, S.E., Pannek, K., Fripp, J., Lingwood, B.E., Lai, M.M., Kong, A.H., Ware, R.S., Coulthard, A., Finn, C.M., Bandaranayake, S.E., 2015b. PPREMO: a prospective cohort study of preterm infant brain structure and function to predict neurodevelopmental outcome. *BMC Pediatr.* 15, 123. <https://doi.org/10.1186/s12887-015-0439-z>.
- George, J.M., Colditz, P.B., Chatfield, M.D., Fiori, S., Pannek, K., Fripp, J., Guzzetta, A., Rose, S.E., Ware, R.S., Boyd, R.N., 2021. Early clinical and MRI biomarkers of cognitive and motor outcomes in very preterm born infants. *Pediatr. Res.* 90, 1243–1250. <https://doi.org/10.1038/S41390-021-01399-5>.
- George, J.M., Pagnozzi, A.M., Bora, S., Boyd, R.N., Colditz, P.B., Rose, S.E., Ware, R.S., Pannek, K., Bursle, J.E., Fripp, J., Barlow, K., Iyer, K., Leishman, S.J., Jendra, R.L., 2020. Prediction of childhood brain outcomes in infants born preterm using neonatal MRI and concurrent clinical biomarkers (PREBO-6): study protocol for a prospective cohort study. *BMJ Open* 10, e036480. <https://doi.org/10.1136/bmjopen-2019-036480>.
- George, J.M., Pannek, Kerstin, Rose, S.E., Ware, R.S., Colditz, P.B., Boyd, R.N., 2018. Diagnostic accuracy of early magnetic resonance imaging to determine motor outcomes in infants born preterm: a systematic review and meta-analysis. *Dev. Med. Child Neurol.* 60, 134–146. <https://doi.org/10.1111/dmcn.13611>.
- Gousias, I.S., Hammers, A., Counsell, S.J., Srinivasan, L., Rutherford, M.A., Heckemann, R.A., Hajnal, J.V., Rueckert, D., Edwards, A.D., 2013. Magnetic resonance imaging of the newborn brain: automatic segmentation of brain images into 50 anatomical regions. *PLoS One* 8, e59990. <https://doi.org/10.1371/journal.pone.0059990>.
- Gui, L., Loukas, S., Lazeyras, F., Hüppi, P.S., Meskaldji, D.E., Borradori Tolsa, C., 2019. Longitudinal study of neonatal brain tissue volumes in preterm infants and their ability to predict neurodevelopmental outcome. *Neuroimage* 185, 728–741. <https://doi.org/10.1016/j.neuroimage.2018.06.034>.
- Hadders-Algra, M., Tacke, U., Pietz, J., Rupp, A., Philippi, H., 2024. Predictive value of the general movements assessment and standardized infant NeuroDevelopmental assessment in infants at high risk of neurodevelopmental disorders. *Dev. Med. Child Neurol.* 66, 1361–1368. <https://doi.org/10.1111/dmcn.15901>.
- He, L., Li, H., Chen, M., Wang, J., Altaye, M., Dillman, J.R., Parikh, N.A., 2021. Deep multimodal learning from MRI and clinical data for early prediction of neurodevelopmental deficits in very preterm infants. *Front. Neurosci.* 15.
- Hinojosa-Rodríguez, M., Harmony, T., Carrillo-Prado, C., Van Horn, J.D., Irimia, A., Torgerson, C., Jacokes, Z., 2017. Clinical neuroimaging in the preterm infant: diagnosis and prognosis. *Neuroimage: Clinical* 16, 355–368. <https://doi.org/10.1016/j.NICL.2017.08.015>.
- Inder, T.E., Vries, L.S. de, Ferriero, D.M., Grant, P.E., Ment, L.R., Miller, S.P., Volpe, J.J., 2021. Neuroimaging of the preterm brain: review and recommendations. *J. Pediatr.* 237, 276–287.e4. <https://doi.org/10.1016/j.jpeds.2021.06.014>.
- Inder, T.E., Warfield, S.K., Wang, H., Hüppi, P.S., Volpe, J.J., 2005. Abnormal cerebral structure is present at term in premature infants. *Pediatrics* 115, 286–294. <https://doi.org/10.1542/peds.2004-0326>.
- John, H.B., Oommen, S.P., Swathi, T.O., Kumar, M., Stoen, R., Adde, L., 2022. Preterm general movements in prediction of neurodevelopmental disability and cerebral palsy at two years: a prospective cohort study. *Indian Pediatr.* 59, 769–773.
- Kelly, C.E., Thompson, D.K., Spittle, A.J., Chen, J., Seal, M.L., Anderson, P.J., Doyle, L. W., Cheong, J.L., 2020. Regional brain volumes, microstructure and neurodevelopment in moderate-to-late preterm children. *Arch. Dis. Child. Fetal Neonatal Ed.* 105, 593–599. <https://doi.org/10.1136/archdischild-2019-317941>.
- Kidowaki, S., Morimoto, M., Yamada, K., Sakai, K., Zuiki, M., Maeda, H., Yamashita, S., Morita, T., Hasegawa, T., Chiyonobu, T., Tokuda, S., Hosoi, H., 2017. Longitudinal change in white matter in preterm infants without magnetic resonance imaging abnormalities: assessment of serial diffusion tensor imaging and their relationship to neurodevelopmental outcomes. *Brain and Development* 39, 40–47. <https://doi.org/10.1016/j.braindev.2016.07.007>.
- Kim, H., Lepage, C., Maheshwary, R., Jeon, S., Evans, A.C., Hess, C.P., Barkovich, A.J., Xu, D., 2016. NEOCIVET: towards accurate morphometry of neonatal gyrification and clinical applications in preterm newborns. *Neuroimage* 138, 28–42. <https://doi.org/10.1016/j.neuroimage.2016.05.034>.
- Kline, Julia E., Illapani, V.S.P., He, L., Altaye, M., Logan, J.W., Parikh, N.A., 2020. Early cortical maturation predicts neurodevelopment in very preterm infants. *Archives of disease in childhood. Fetal and neonatal edition* 105, 460–465. <https://doi.org/10.1136/archdischild-2019-317466>.
- Kline, Julia E., Sita Priyanka Illapani, V., He, L., Parikh, N.A., 2020. Automated brain morphometric biomarkers from MRI at term predict motor development in very preterm infants. *Neuroimage: Clinical* 28, 102475. <https://doi.org/10.1016/j.nicl.2020.102475>.
- Lefèvre, J., Germanaud, D., Dubois, J., Rousseau, F., de Macedo Santos, I., Angleys, H., Mangin, J.-F., Hüppi, P.S., Girard, N., De Guio, F., 2016. Are developmental trajectories of cortical folding comparable between cross-sectional datasets of fetuses and preterm newborns? *Cerebr. Cortex* 26, 3023–3035. <https://doi.org/10.1093/cercor/bhv123>.
- Li, Z., Yu, J., Wang, Y., Zhou, H., Yang, H., Qiao, Z., 2021. DeepVolume: brain structure and spatial connection-aware network for brain MRI super-resolution. *IEEE Trans. Cybern.* 51, 3441–3454. <https://doi.org/10.1109/TCYB.2019.2933633>.
- Liu, M., Lepage, C., Kim, S.Y., Jeon, S., Kim, S.H., Simon, J.P., Tanaka, N., Yuan, S., Islam, T., Peng, B., Arutyunyan, K., Surento, W., Kim, J., Jahanshad, N., Styner, M. A., Toga, A.W., Barkovich, A.J., Xu, D., Evans, A.C., Kim, H., 2021. Robust cortical thickness morphometry of neonatal brain and systematic evaluation using multi-site MRI datasets. *Front. Neurosci.* 15.
- Lowe, J., Bann, C.M., Dempsey, A.G., Fuller, J., Taylor, H.G., Gustafson, K.E., Watson, V. E., Vohr, B.R., Das, A., Shankaran, S., Yolton, K., Ball, M.B., Hintz, S.R., Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network, 2023. Do bayley-III composite scores at 18–22 Months corrected age predict full-scale IQ at 6–7 years in children born extremely preterm? *J. Pediatr.* 263, 113700. <https://doi.org/10.1016/j.jpeds.2023.113700>.
- Makropoulos, A., Robinson, E.C., Schuh, A., Wright, R., Fitzgibbon, S., Bozek, J., Counsell, S.J., Steinweg, J., Vecchiato, K., Passerat-Palmbach, J., Lenz, G., Mortari, F., Tenev, T., Duff, E.P., Bastiani, M., Cordero-Grande, L., Hughes, E., Tusor, N., Tournier, J.-D., Hutter, J., Price, A.N., Teixeira, R.P.A.G., Murgasova, M., Victor, S., Kelly, C., Rutherford, M.A., Smith, S.M., Edwards, A.D., Hajnal, J.V., Jenkinson, M., Rueckert, D., 2018a. The developing human connectome project: a minimal processing pipeline for neonatal cortical surface reconstruction. *Neuroimage* 173, 88–112. <https://doi.org/10.1016/j.neuroimage.2018.01.054>.
- Makropoulos, A., Robinson, E.C., Schuh, A., Wright, R., Fitzgibbon, S., Bozek, J., Counsell, S.J., Steinweg, J., Vecchiato, K., Passerat-Palmbach, J., Lenz, G., Mortari, F., Tenev, T., Duff, E.P., Bastiani, M., Cordero-Grande, L., Hughes, E., Tusor, N., Tournier, J.-D., Hutter, J., Price, A.N., Teixeira, R.P.A.G., Murgasova, M., Victor, S., Kelly, C., Rutherford, M.A., Smith, S.M., Edwards, A.D., Hajnal, J.V., Jenkinson, M., Rueckert, D., 2018b. The developing human connectome project: a minimal processing pipeline for neonatal cortical surface reconstruction. *Neuroimage* 173, 88–112. <https://doi.org/10.1016/j.neuroimage.2018.01.054>.
- Martinez-Biarge, M., Groenendaal, F., Kersbergen, K.J., Benders, M.J.N.L., Foti, F., van Haastert, I.C., Cowan, F.M., de Vries, L.S., 2019. Neurodevelopmental outcomes in preterm infants with white matter injury using a new MRI classification. *Neonatology* 116, 227–235. <https://doi.org/10.1159/000499346>.
- Moeskos, P., Işgum, I., Keunen, K., Claessens, N.H.P., van Haastert, I.C., Groenendaal, F., de Vries, L.S., Viergever, M.A., Benders, M.J.N.L., 2017. Prediction of cognitive and motor outcome of preterm infants based on automatic quantitative descriptors from neonatal MR brain images. *Sci. Rep.* 7, 2163. <https://doi.org/10.1038/s41598-017-02307-w>.
- Olsen, J.E., Allinson, L.G., Doyle, L.W., Brown, N.C., Lee, K.J., Eeles, A.L., Cheong, J.L.Y., Spittle, A.J., 2018. Preterm and term-equivalent age general movements and 1-year neurodevelopmental outcomes for infants born before 30 weeks' gestation. *Dev. Med. Child Neurol.* 60, 47–53. <https://doi.org/10.1111/dmcn.13558>.

- Pagnozzi, A.M., van Eijk, L., Pannek, K., Boyd, R.N., Saha, S., George, J., Bora, S., Bradford, D., Fahey, M., Ditchfield, M., Malhotra, A., Liley, H., Colditz, P.B., Rose, S., Fripp, J., 2023. Early brain morphometrics from neonatal MRI predict motor and cognitive outcomes at 2-years corrected age in very preterm infants. *Neuroimage* 267, 119815. <https://doi.org/10.1016/j.neuroimage.2022.119815>.
- Pannek, K., George, J.M., Boyd, R.N., Colditz, P.B., Rose, S.E., Fripp, J., 2020. Brain microstructure and morphology of very preterm-born infants at term equivalent age: associations with motor and cognitive outcomes at 1 and 2 years. *Neuroimage* 221, 117163. <https://doi.org/10.1016/J.NEUROIMAGE.2020.117163>.
- Parikh, N.A., 2016. Advanced neuroimaging and its role in predicting neurodevelopmental outcomes in very preterm infants. *Semin. Perinatol.* 40, 530–541. <https://doi.org/10.1053/J.SEMPERI.2016.09.005>.
- Pugliese, M., Rossi, C., Guidotti, I., Gallo, C., Della Casa, E., Bertocelli, N., Coccolini, E., Ferrari, F., 2013. Preterm birth and developmental problems in infancy and preschool age Part II: cognitive, neuropsychological and behavioural outcomes 1653–1657. <https://doi.org/10.3109/14767058.2013.794205>.
- Raffelt, D.A., Tournier, J.-D., Smith, R.E., Vaughan, D.N., Jackson, G., Ridgway, G.R., Connelly, A., 2017. Investigating white matter fibre density and morphology using fixel-based analysis. *Neuroimage* 144, 58–73. <https://doi.org/10.1016/j.neuroimage.2016.09.029>.
- Roberts, G., Howard, K., Spittle, A.J., Brown, N.C., Anderson, P.J., Doyle, L.W., 2008. Rates of early intervention services in very preterm children with developmental disabilities at age 2 years. *J. Paediatr. Child Health* 44, 276–280. <https://doi.org/10.1111/J.1440-1754.2007.01251.X>.
- Saha, S., Pagnozzi, A., Bourgeat, P., George, J.M., Bradford, D., Colditz, P.B., Boyd, R.N., Rose, S.E., Fripp, J., Pannek, K., 2020. Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model. *Neuroimage* 215, 116807. <https://doi.org/10.1016/J.NEUROIMAGE.2020.116807>.
- Setänen, S., Lahti, K., Lehtonen, L., Parkkola, R., Maunu, J., Saarinen, K., Haataja, L., 2014. Neurological examination combined with brain MRI or cranial US improves prediction of neurological outcome in preterm infants. *Early Hum. Dev.* 90, 851–856. <https://doi.org/10.1016/j.earlhumdev.2014.09.007>.
- Shimony, J.S., Smyser, C.D., Wideman, G., Alexopoulos, D., Hill, J., Harwell, J., Dierker, D., Van Essen, D.C., Inder, T.E., Neil, J.J., 2016. Comparison of cortical folding measures for evaluation of developing human brain. *Neuroimage* 125, 780–790. <https://doi.org/10.1016/j.neuroimage.2015.11.001>.
- Skiöld, B., Eriksson, C., Eliasson, A.-C., Adén, U., Vollmer, B., 2013. General movements and magnetic resonance imaging in the prediction of neuromotor outcome in children born extremely preterm. *Early Hum. Dev.* 89, 467–472. <https://doi.org/10.1016/j.earlhumdev.2013.03.014>.
- Soun, J.E., Liu, M.Z., Cauley, K.A., Grinband, J., 2017. Evaluation of neonatal brain myelination using the T1- and T2-weighted MRI ratio. *J. Magn. Reson. Imag.* 46, 690–696. <https://doi.org/10.1002/jmri.25570>.
- Spittle, A., Orton, J., Anderson, P.J., Boyd, R., Doyle, L.W., 2015. Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants. *Cochrane Database Syst. Rev.* 2015. <https://doi.org/10.1002/14651858.CD005495.PUB4>.
- Sui, Y., Afacan, O., Gholipour, A., Warfield, S.K., 2021. Fast and high-resolution neonatal brain MRI through super-resolution reconstruction from acquisitions with variable slice selection direction. *Front. Neurosci.* 15. <https://doi.org/10.3389/fnins.2021.636268>.
- Thompson, D.K., Warfield, S.K., Carlin, J.B., Pavlovic, M., Wang, H.X., Bear, M., Kean, M. J., Doyle, L.W., Egan, G.F., Inder, T.E., 2007. Perinatal risk factors altering regional brain structure in the preterm infant. *Brain* 130, 667–677. <https://doi.org/10.1093/brain/awl277>.
- Thompson, P.M., Sowell, E.R., Gogtay, N., Giedd, J.N., Vidal, C.N., Hayashi, K.M., Leow, A., Nicolson, R., Rapoport, J.L., Toga, A.W., 2005. Structural MRI and brain development. In: *International Review of Neurobiology, Neuroimaging, Part B*. Academic Press, pp. 285–323. [https://doi.org/10.1016/S0074-7742\(05\)67009-2](https://doi.org/10.1016/S0074-7742(05)67009-2).
- Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *J. Roy. Stat. Soc. B* 58, 267–288.
- van't Hoof, J., van der Lee, J.H., Opmeer, B.C., Aarnoudse-Moens, C.S.H., Leenders, A.G. E., Mol, B.W.J., de Haan, T.R., 2015. Predicting developmental outcomes in premature infants by term equivalent MRI: systematic review and meta-analysis. *Syst. Rev.* 4, 1–10.
- Wang, J., Shen, X., Hu, X., Yang, H., Yin, H., Zhu, X., Gao, H., Wu, Y., Meng, F., 2021. Early detection relationship of cerebral palsy markers using brain structure and general movements in infants born. *Early Hum. Dev.* 163, 105452. <https://doi.org/10.1016/J.EARLHUMDEV.2021.105452>.
- Wang, L., Wu, Z., Chen, L., Sun, Y., Lin, W., Li, G., 2023. iBEAT V2.0: a multisite-applicable, deep learning-based pipeline for infant cerebral cortical surface reconstruction. *Nat. Protoc.* 18, 1488–1509. <https://doi.org/10.1038/s41596-023-00806-x>.
- Woodward, L.J., Anderson, P.J., Austin, N.C., Howard, K., Inder, T.E., 2006. Neonatal MRI to predict neurodevelopmental outcomes in preterm infants. *N. Engl. J. Med.* 355, 685–694. <https://doi.org/10.1056/NEJMoa053792>.
- Young, J.M., Powell, T.L., Morgan, B.R., Card, D., Lee, W., Smith, M.L., Sled, J.G., Taylor, M.J., 2015. Deep grey matter growth predicts neurodevelopmental outcomes in very preterm children. *Neuroimage* 111, 360–368. <https://doi.org/10.1016/j.neuroimage.2015.02.030>.