



POPULATION STUDY ARTICLE

Early clinical and MRI biomarkers of cognitive and motor outcomes in very preterm born infants

Joanne M. George¹, Paul B. Colditz^{2,3}, Mark D. Chatfield¹, Simona Fiori⁴, Kerstin Pannek⁵, Jurgen Fripp⁵, Andrea Guzzetta⁴, Stephen E. Rose⁵, Robert S. Ware⁶ and Roslyn N. Boyd¹

BACKGROUND: This study aimed to identify which MRI and clinical assessments, alone or in combination, from (i) early (32 weeks postmenstrual age, PMA), (ii) term equivalent age (TEA) and (iii) 3 months corrected age (CA) are associated with motor or cognitive outcomes at 2 years CA in infants born <31 weeks gestation.

METHODS: Prospective cohort study of 98 infants who underwent early and TEA MRI ($n = 59$ males; median birth gestational age $28 + 5$ weeks). Hammersmith Neonatal Neurological Examination (HNNE), NICU Neonatal Neurobehavioural Scale and General Movements Assessment (GMs) were performed early and at TEA. Premie-Neuro was performed early and GMs, Test of Infant Motor Performance and visual assessment were performed at TEA and 3 months CA. Neurodevelopmental outcomes were determined using Bayley Scales of Infant and Toddler Development 3rd edition.

RESULTS: The best combined motor outcome model included 3-month GMs ($\beta = -11.41$; 95% CI = $-17.34, -5.49$), TEA MRI deep grey matter score ($\beta = -6.23$; 95% CI = $-9.47, -2.99$) and early HNNE reflexes ($\beta = 3.51$; 95% CI = $0.86, 6.16$). Combined cognitive model included 3-month GMs ($\beta = -10.01$; 95% CI = $-15.90, -4.12$) and TEA HNNE score ($\beta = 1.33$; 95% CI = $0.57, 2.08$).

CONCLUSION: Early neonatal neurological assessment improves associations with motor outcomes when combined with term MRI and 3-month GMs. Term neurological assessment combined with 3-month GMs improves associations with cognitive outcomes.

Pediatric Research _#####_; <https://doi.org/10.1038/s41390-021-01399-5>

IMPACT:

- We present associations between 32- and 40-week MRI, comprehensive clinical assessments and later 2-year motor and cognitive outcomes for children born <31 weeks gestation.
- MRI and clinical assessment of motor, neurological and neurobehavioural function earlier than term equivalent age in very preterm infants is safe and becoming more available in clinical settings. Most of these children are discharged from hospital before term age and so completing assessments prior to discharge can assist with follow up.
- MRI and neurological assessment prior to term equivalent age while the child is still in hospital can provide earlier identification of children at highest risk of adverse outcomes and guide follow-up screening and intervention services.

INTRODUCTION

Infants born very preterm are at risk of adverse motor, cognitive, language and behavioural outcomes which persist to school age and beyond.¹ They are four times more likely to experience motor delays,² six times more likely to have developmental coordination disorder³ and 40 times more likely to be later diagnosed with cerebral palsy compared to term born peers.⁴ Cognition is suggested to be affected more frequently than motor function with preterm born infants scoring on average 12 points lower on IQ testing than term born children.⁵ Early, accurate identification of preterm infants at risk of adverse outcomes is critical to enable structured surveillance, early commencement of targeted interventions and provision of family supports.

The search for early prognostic biomarkers has led to developments in neuroimaging techniques and clinical measures such as assessment of neurological, motor and neurobehavioural function. Currently the combination of magnetic resonance imaging (MRI) at term equivalent age (TEA) and a General Movements assessment (GMs) at 3 months corrected age (CA) has the best prediction for motor and cognitive outcomes following very preterm birth.^{6–10} Earlier MRI prior to term has been found to be safe and feasible.¹¹ Structural MRI at 30–32 weeks postmenstrual age (PMA) has demonstrated good predictive validity for later motor and cognitive outcomes¹² and associations with concurrent clinical measures both early in the neonatal period and at TEA.¹³ It is not yet known if earlier MRI and earlier clinical assessments

¹Queensland Cerebral Palsy and Rehabilitation Research Centre, Centre for Children's Health Research, Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia; ²University of Queensland Centre for Clinical Research, The University of Queensland, Brisbane, QLD, Australia; ³Perinatal Research Centre, Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia; ⁴Stella Maris Scientific Institute, Department of Developmental Neuroscience, The University of Pisa, Pisa, Italy; ⁵The Australian e-Health Research Centre, CSIRO, Brisbane, QLD, Australia and ⁶Menzies Health Institute Queensland, Griffith University, Brisbane, QLD, Australia
Correspondence: Joanne M. George (j.george2@uq.edu.au)

Received: 13 October 2020 Accepted: 22 January 2021

Published online: 24 February 2021

strengthen the associations with later outcomes when compared with MRI and clinical assessment at TEA and 3 months CA. Our prospective cohort study, 'PPREMO: Prediction of preterm motor outcome', aimed to develop a toolbox of structural and advanced diffusion MRI and clinical biomarkers that can predict outcomes prior to TEA and 3-month GMs.¹¹ The first step in the process is to examine the early variables individually and in combination to determine their associations with motor and cognitive outcomes at 2 years CA.

In a cohort of very preterm infants with early neonatal (32 weeks PMA) and TEA MRI and comprehensive clinical assessment, clinical assessment at 3 months CA and outcomes measured at 2 years CA, this study sought to determine (i) which combination of MRI and clinical measures were most strongly associated with motor outcomes at 2 years CA, and (ii) which combination of MRI and clinical measures were most strongly associated with cognitive outcomes at 2 years CA.

METHODS

Study design and participants

This prospective cohort study was conducted at the Royal Brisbane and Women's Hospital (RBWH), a specialist tertiary neonatal centre in Brisbane, Australia. Infants were eligible for inclusion between February 2013 and February 2016 if they were born <31 weeks gestational age, their parents/carers lived within a 200 km radius of the hospital and were English speaking. Infants were excluded if they had a known congenital or chromosomal abnormality likely to affect their neurodevelopmental outcome. Informed parental consent was obtained for all participants. This study was nested within a broader study, and sample size calculations are detailed in the study protocol.¹¹ Ethical approval was obtained from the RBWH Human Research Ethics Committee (HREC/12/QRBW/245) and The University of Queensland (2012001060). Trial registration: Australian New Zealand Clinical Trials Registry (ACTRN12613000280707).

Social demographic data were collected via a questionnaire and a scoring system, previously validated in cohorts of preterm infants, was applied to classify social risk status.^{14–16} Six aspects were considered, including family structure, education of primary caregiver, occupation of primary income earner, employment status of primary income earner, language spoken at home and maternal age. A score of 0 to 2 was applied for each of the six items for a maximum total score of 12. Higher scores indicate a higher level of social risk and a total score of 2 or above was defined as high social risk. Full details of items and scoring are available in Appendix 1.

MRI acquisition

Infants underwent brain MRI at 30–32 weeks PMA or when medically stable ('Early'), and again at 40–42 weeks PMA ('Term'). An MR compatible incubator equipped with a dedicated neonatal head coil (LMT Lammers Medical Technology, Lübeck, Germany) was used in conjunction with a 3T MRI Siemens Tim Trio (Erlangen, Germany) scanner. Ear protection attenuated noise, no sedation was used and MRI was performed during natural sleep. Coronal, axial and sagittal T2-weighted HASTE (TR/TE 2000/90 ms, flip angle 150°, field of view 200 × 160 mm, matrix 320 × 256, slice thickness 4 mm) were acquired. Axial T1 TSE (TR/TE 1490/90 ms, flip angle 150°, field of view 200 × 160 mm, matrix 256 × 180, slice thickness 2 mm) and an axial multi-echo T2 TSE (TR/TE1/TE2/TE3 10580/27/122/189 ms, flip angle 150°, field of view 144 × 180 mm, matrix 204 × 256, slice thickness 2 mm) were also acquired.

MRI scoring

An independent neurologist with training in brain MRI (SF) and blinded to history and clinical assessment findings applied a standardised scoring system to MRIs.^{12,17} The scoring system

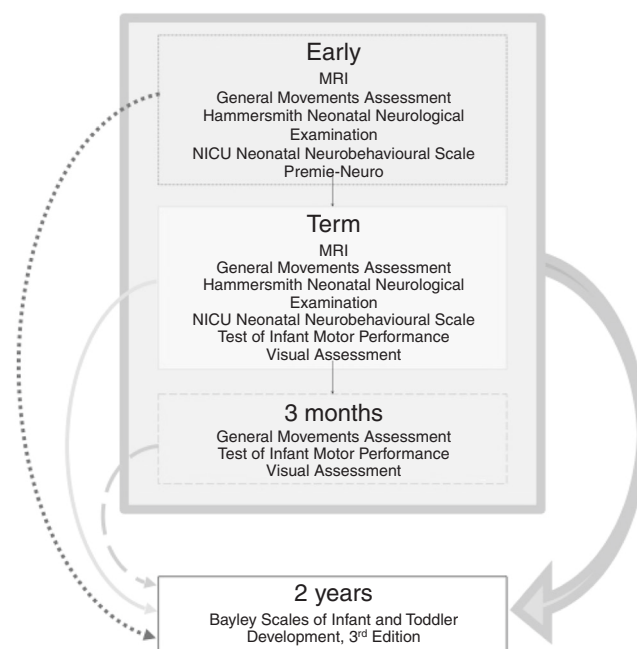


Fig. 1 Flowchart of assessment timepoints and measures. Early, 30–32 weeks postmenstrual age (PMA); Term, 40–42 weeks PMA; blue (dotted line) represents analyses examining relationships between Early measures and 2-year outcomes; yellow (solid line) represents analyses examining relationships between Term measures and 2-year outcomes; orange (dashed line) represents analyses examining relationships between 3-month measures and 2-year outcomes; green (bold lines) represents analyses examining relationships between all three timepoints combined and 2-year outcomes.

generates four subscales which can be summed to give a global total score; white matter, cortical grey matter, deep grey matter and cerebellum.^{12,17} Both T1 and T2 images were evaluated during scoring and T1 hyperintensities and T2 hypointensities were recorded and considered as signal abnormalities. Sagittal T2-weighted images were used to score the corpus callosum as it is clearly visualised as low signal intensity prior to myelination. Inter- and intrarater reproducibility of the scale have been demonstrated in two studies, one of which included SF as a rater.^{12,17}

Clinical measures

Clinical assessments were combined to remove duplicate items and reduce handling. They were completed within a week of MRI by an experienced assessor, blinded to birth history and brain imaging findings (J.G.). Full details of assessment tools and administration are available in the study protocol, including detailed summaries of their psychometric properties,¹¹ see also Fig. 1.

The GMs is a neuromotor assessment of observed spontaneous movements. Assessment at Early and Term timepoints are in the 'writhing period' (sensitivity 75–100%, specificity 40–48% for CP)^{18,19} and at 3 months CA is in the 'fidgety period' (sensitivity 98%, specificity 91% for CP).²⁰ Scoring was performed by two advanced GMs raters (J.G. and B.S.), with cases of non-agreement reviewed until consensus reached and advice sought from a third blinded rater.

The Hammersmith Neonatal Neurological Examination (HNNE) is a neurological assessment evaluating posture, tone, reflexes, spontaneous movements, orientation and behaviour.²¹ At the early assessment, all items except placing were administered. In the preterm period, sensitivity and specificity for predicting CP are 57–86% and 45–83%, respectively, increasing to 68–96% and 52–97% at TEA.²⁰ Inter-rater reliability between the clinical

assessor (J.G.) and an observer (P.C.) for the HNNE total optimality score was tested with intra-class correlation coefficients of 0.94 Early, and 0.99 at Term.

The NICU Neonatal Neurobehavioural Scale (NNNS) evaluates an infant's response to stimuli and handling, state regulation, motor performance, and neurological status.²² The NNNS at TEA has been shown to predict motor and cognitive outcomes at 18 months CA, motor outcomes at 24 months CA and cognitive outcomes at 4.5 years.^{23–25} The Premie-Neuro (PN) is a neurological examination which consists of three categories: neurological, movement and responsiveness.²⁶ It has been designed for use from 23 to 37 weeks PMA and has scoring based on expected performance at each week of PMA.²⁶

At TEA and 3 months CA, a visual assessment developed by Ricci et al.²⁷ was used to examine visual function. The test examines ocular motility, acuity and the ability to fix and follow and has demonstrated predictive validity for neurodevelopmental outcomes in preterm cohorts.²⁸ The Test of Infant Motor Performance (TIMP), performed at TEA and 3 months CA, is a standardised, discriminative and evaluative assessment of gross motor development.^{29,30} Construct validity enabling discrimination between infants at high and low risk of adverse motor outcomes has been demonstrated.²⁹

Outcomes

An experienced paediatric physiotherapist, blinded to all earlier clinical and MRI findings, assessed participants at 2 years CA using the Bayley Scales of Infant and Toddler Development 3rd edition (Bayley III).³¹ The Bayley III is a discriminative and norm-referenced tool. The motor and cognitive composite scores were utilised in this study. They have a mean of 100 and a SD of 15 and higher scores reflect better development.

Children were assessed by a paediatrician for signs of cerebral palsy as per the definition by Rosenbaum et al.³² to allow description of the cohort. Any level of severity and distribution were included and recorded using the Gross Motor Function Classification System.³³

Statistical analysis

Pearson correlations and univariable regression analyses were performed to compare the relationship between individual variables from Early, Term and 3 months CA with 2-year outcomes. Clinical and MRI scores were used as dependent variables without categorisation. The 3-month GMs score was strongly associated with both motor and cognitive outcomes, so we calculated the associations between each of the variables and outcomes, adjusting for 3-month GMs score. Multivariable models of MRI and clinical variables at all three timepoints were constructed using forward stepwise selection of next variable. The variable with the smallest *p* value was added to the model if *p* < 0.01. For multivariable analyses we considered: (i) use of maximum possible data available for each model without imputing any data; (ii) as our dataset contained 10 sets of twins and 1 set of triplets, we conducted mixed effects linear regression models with family included as a random effect, (iii) GMs featured as an ordinal variable (0 = normal, 1 = poor repertoire/abnormal fidgety, 2 = cramped synchronised/absent fidgety). Results are presented as regression coefficients with 95% confidence intervals. Analysis was performed using Stata statistical software, v16 (StataCorp, College Station, TX).

RESULTS

Early MRI was available for 119 preterm infants, *N* = 92 and *N* = 98 of which had motor and cognitive outcome data for Bayley III. Full demographic, perinatal and clinical assessment data have been published previously¹³ and are summarised in Table 1. MRI scoring

Table 1. Characteristics of the study sample.

Birth and maternal data	<i>n</i> = 98
Gestational age at birth (weeks)	28.5 (26.7–29.4); range 23.1–30.9
Birthweight (g)	1065 (323); range 494–1886
Small for gestational age (birthweight < 10th percentile)	15 (15%)
Birth head circumference (cm)	25.6 (2.4), <i>n</i> = 94; range 20.5–30.5
Males	59 (60%)
Multiple births	32 (33%)
Premature rupture of membranes	21 (21%)
Caesarian section	72 (73%)
Chorioamnionitis	16 (16%)
Antenatal steroids	74 (76%)
Magnesium sulfate	51 (63%), <i>n</i> = 81
Higher social risk ^a	45 (46%)
Acquired medical factors	
Patent ductus arteriosus	48 (49%)
Any intraventricular haemorrhage	23 (23%)
Intraventricular haemorrhage grade III or IV	6 (6%)
Periventricular leukomalacia	4 (4%)
Hydrocephalus ^b	3 (3%)
Seizures treated with anticonvulsant therapy	0 (0%)
NEC diagnosed or suspected	4 (4%)
Confirmed sepsis	4 (4%)
Total parenteral nutrition (days)	11 (9–15); range 0–36
Postnatal corticosteroids	19 (19%)
Ventilation (days)	3 (0–15), <i>n</i> = 97; range 1–50
CPAP (days)	30 (7–47), <i>n</i> = 97; range 1–81
Oxygen therapy (hours)	56 (2–573), <i>n</i> = 90; range 1–3912
Bronchopulmonary dysplasia ^c	29 (30%)
PMA at Early MRI (weeks)	32.0 (31.3–33.6); range 29.4–35.3
PMA at Term MRI (weeks)	40.7 (40.0–41.4); range 38.4–46.6
Weight at Early MRI (g)	1467 (1240–1680); range 858–2715
Weight at Term MRI (g)	3000 (2700–3375); range 1900–5150
PMA at Early clinical assessment (weeks)	32.0 (31.1–33.4); range 29.4–35.3
PMA at Term clinical assessment (weeks)	40.6 (40.0–41.4); range 38.4–46.6
Neurodevelopmental outcomes	
Corrected Age at assessment (months)	24.1 (0.6); range 22.8–26.6
Bayley III motor composite score	95.6 (15.1), 95 (85–105), <i>n</i> = 91; range 46–124
Bayley III cognitive composite score	96.9 (14.4), 100 (91–107); range 55–140
Cerebral palsy	4 (4%)

Data are presented as mean (SD) and/or median (IQR); range min–max for continuous measures and *n* (%) for categorical measures.

'Early' refers to MRI or clinical assessment between 29 and 35 weeks PMA, 'Term' refers to MRI or clinical assessment between 40 and 42 weeks. PMA PMA postmenstrual age, NEC necrotising enterocolitis, CPAP continuous positive airway pressure, Bayley III Bayley Scales of Infant and Toddler Development 3rd edn.

^aSocial risk.

^bHydrocephalus.

^cBronchopulmonary dysplasia.

Table 2. Summary of MRI and clinical assessment scores.

	Early	Term	3 months CA
MRI Scores, median (IQR)	<i>n</i> = 98	<i>n</i> = 94	<i>n</i> = 96
White matter ↓	3 (2–4)	2 (1–3)	
Cortical grey matter ↓	0 (0–1)	0 (0–1)	
Deep grey matter ↓	0 (0–1)	0 (0–1)	
Cerebellum ↓	0 (0–0)	0 (0–1)	
Global ↓	4 (3–7)	3 (1–5)	
GMs, <i>n</i> (%)	<i>n</i> = 98	<i>n</i> = 92	<i>n</i> = 96
Normal	32 (33%)	29 (32%)	88 (92%)
Poor repertoire/abnormal fidgety	60 (61%)	56 (61%)	2 (2%)
Cramped synchronised/absent fidgety	6 (6%)	7 (8%)	6 (6%)
HNNE, mean (SD)	<i>n</i> = 98	<i>n</i> = 95	
Posture & tone ↑	3.8 (1.8), <i>n</i> = 90	6.9 (1.6)	
Tone patterns ↑	3.9 (0.8), <i>n</i> = 90	3.6 (0.8)	
Reflexes ↑	2.4 (1.0), <i>n</i> = 92	4.1 (1.1)	
Spontaneous movements ↑	1.0 (0.8), <i>n</i> = 89	2.3 (0.8)	
Abnormal signs ↑	2.0 (0.6)	2.5 (0.5)	
Orientation & behaviour ↑	2.9 (1.5)	5.2 (1.2)	
HNNE total score↑	16.1 (3.5), <i>n</i> = 88	24.7 (3.6)	
NNNS, mean (SD)	<i>n</i> = 98	<i>n</i> = 93	
Quality of movement ↑	3.4 (0.5)	4.4 (0.6)	
Regulation ↑	4.1 (0.6), <i>n</i> = 97	4.9 (0.6)	
Nonoptimal reflexes ↓	6.9 (1.5)	6.8 (2.8)	
Stress/abstinence↓	0.2 (0.1)	0.2 (0.1)	
Arousal ↑↓	3.1 (0.6)	4.3 (0.6)	
Hypertonicity ↓	0.1 (0.4)	0.2 (0.6)	
Hypotonicity ↓	1.6 (1.1)	0.5 (0.7)	
Asymmetric reflexes ↓	0.9 (1.0)	0.7 (0.9)	
Excitability ↓	3.0 (1.7)	3.6 (1.8)	
Lethargy ↓	8.6 (2.0)	4.9 (2.1)	
Premie-neuro, mean (SD)	<i>n</i> = 98		
Factor 1 Neurological ↑	31.6 (4.2)		
Factor 2 Movement ↑	34.3 (4.6), <i>n</i> = 97		
Factor 3 Responsiveness ↑	31.2 (3.5), <i>n</i> = 90		
Total score ↑	97.4 (7.6), <i>n</i> = 90		
TIMP, mean (SD)		<i>n</i> = 94	<i>n</i> = 94
z-score ↑		−0.6 (0.7)	−0.4 (0.9)
Visual Score, mean (SD)		<i>n</i> = 92	<i>n</i> = 94
Total score ↑		16.4 (6.0)	24.7 (2.6)

Data are presented as mean (SD) or median (IQR) for continuous measures, and *n* (%) for categorical measures. The total number of assessments at each timepoint are indicated; *n* is further specified for any items or subscales that had missing data.

‘Early’ refers to MRI or clinical assessment between 29 and 35 weeks PMA, ‘Term’ refers to MRI or clinical assessment between 40 and 42 weeks PMA, CA corrected age, GMs General Movements Assessment, HNNE Hammer-smith Neonatal Neurological Examination, NNNs NICU Neonatal Neurobehavioral Scale, TIMP Test of Infant Motor Performance, SD standard deviation, IQR interquartile range, ↑ higher scores better, ↓ lower scores better.

and clinical assessment results are summarised in Table 2. Differences in perinatal data for participants with and without a 2-year outcome are summarised in Supplementary Table 1. Those children who failed to return at 2 years had a slightly greater birthweight (1225 vs 1065 g) and underwent Term MRI at a slightly younger PMA (40 + 0 vs 40 + 4) which are unlikely of clinical significance. They also had a lower rate of receiving antenatal steroids (43% vs 76%).

Motor outcome

Univariable analyses between MRI/clinical scores and 2-year Bayley III motor outcome. Univariable associations and correlation results

are presented in Supplementary Table 2. The 3-month GMs score demonstrated the strongest associations with motor outcome (regression coefficient $\beta = -13.2$; 95% confidence interval CI = -18.5 to -7.9 , $p < 0.001$). Early MRI global score and Term MRI white matter, deep grey matter, cerebellar and global scores were all associated with motor outcome. Early GMs, Early HNNE reflexes and abnormal signs, Term HNNE total, and the domain of NNNs regulation at both Early and Term timepoints were associated with motor outcome. The overall TIMP z-score and the visual assessment score at 3 months CA were both significantly associated with motor outcome.

Associations after adjustment for 3-month GMs score. Results are presented in Supplementary Table 2. After adjustment for 3-month GMs, Term MRI deep grey matter score had the strongest association with motor outcome ($\beta = -3.8$; 95% confidence interval CI = -6.7 to -1.0 , $p < 0.01$). Term MRI cerebellar and global total scores, and Early GMs and early HNNE reflexes score were significantly associated with motor outcome.

Multivariable models for 2-year Bayley III motor outcome. Multivariable model results are presented in Table 3. When motor outcome multivariable models were developed for each timepoint individually the results are as follows:

- 32 weeks PMA model included early HNNE reflexes subscale score ($\beta = 4.30$; 95% CI = 1.47 to 7.12; $p = 0.003$) and the early HNNE abnormal signs subscale score ($\beta = 6.15$; 95% CI = 1.64 to 10.65; $p = 0.007$).
- TEA model included TEA deep grey matter MRI score ($\beta = -5.45$; 95% CI = -8.25 to -2.65 ; $p < 0.001$) and the TEA NNNs regulation subscale ($\beta = 6.28$; 95% CI = 1.90 to 10.66; $p = 0.005$).
- 3 months CA included 3-month CA GMs score ($\beta = -13.19$; 95% CI = -18.52 to 7.86; $p < 0.001$).

When the three timepoints were combined, the best motor outcome multivariable model included 3-month CA GMs score ($\beta = -11.41$; 95% CI = -17.34 to -5.49 ; $p < 0.001$), TEA deep grey matter MRI score ($\beta = -6.23$; 95% CI = -9.47 to -2.99 ; $p < 0.001$) and the early HNNE reflexes subscale score ($\beta = 3.51$; 95% CI = 0.86 to 6.16; $p = 0.009$). The model was constructed from data of 83 participants who had data available for all three variables.

Cognitive outcome

Univariable analyses between MRI/clinical scores and 2-year Bayley III cognitive outcome. Univariable associations and correlation results are presented in Supplementary Table 3. Term HNNE score demonstrated the strongest association with cognitive outcome ($\beta = 1.6$; 95% CI = 0.8 to 2.4; $p = 0.000047$) followed closely by 3-month GMs ($\beta = -12.2$; 95% CI = -18.2 to -6.3 ; $p = 0.000054$). Early MRI cerebellar score, Term MRI cortical grey matter, deep grey matter, cerebellar and global score were associated with cognitive outcome. Early HNNE reflexes, NNNs stress, NNNs arousal and the neurological subscale of the Premie-Neuro were associated with cognitive outcome. Term HNNE posture and tone, tone patterns, reflexes, spontaneous movements and total score, term NNNs regulation, TIMP and visual assessment, as well as 3-month TIMP and visual assessment were associated with 2-year cognitive outcome.

Associations after adjustment for 3-month GMs score. A very similar pattern of associations between the clinical assessment scores and cognitive outcome was seen after adjusting for 3-month GMs. Only Term NNNs regulation and the TIMP z-score were no longer significant. Fewer MRI scores remained associated with the outcome, with only the Term MRI deep grey matter score maintaining significance ($\beta = -3.5$; 95% CI = -6.4 to -0.7 ; $p = 0.02$).

Table 3. Multivariable models for motor outcome on the Bayley III.

Timepoint	Variables included	Bayley III Motor Composite Score		
		β	95% CI	<i>p</i> value
32 weeks PMA	HNNE reflexes subscale	4.30	1.47 to 7.12	0.003
	HNNE abnormal signs subscale	6.15	1.64 to 10.65	0.007
TEA	Intercept = 74.81 (the model on $n = 86$ reduced the unexplained variation from SD = 14.3 to SD = 13)			
	MRI deep grey matter score	-5.45	-8.25 to -2.65	<0.001
	NNNS regulation subscale	6.28	1.90 to 10.66	0.005
	Intercept = 68.63 (the model on $n = 86$ reduced the unexplained variation from SD = 14.7 to SD = 12.9)			
3 months CA	GMs	-13.19	-18.52 to -7.86	<0.001
	Intercept = 98.72 (the model on $n = 90$ reduced the unexplained variation from SD = 14.4 to SD = 12.7)			
All 3 timepoints combined	3-month GMs	-11.41	-17.34 to -5.49	<0.001
	TEA MRI deep grey matter score	-6.23	-9.47 to -2.99	<0.001
	32 week PMA HNNE reflexes subscale	3.51	0.86 to 6.16	0.009
	Intercept = 93.50 (the model on $n = 83$ reduced the unexplained variation from SD = 14.5 to SD = 11.7)			

Bayley III Bayley Scales of Infant and Toddler Development 3rd edition. CA corrected age, CI confidence interval, GMs general movement assessment, HNNE Hammersmith Neonatal Neurological Examination, MRI magnetic resonance imaging, PMA postmenstrual age, TEA term equivalent age, β regression coefficient, n varies slightly between models as an effect of each subject having data available for the variables included in the model.

Table 4. Multivariable models for cognitive outcome on the Bayley III.

Timepoint	Variables included	Bayley III Cognitive Composite Score		
		β	95% CI	<i>p</i> value
32 weeks PMA	HNNE reflexes subscale	4.14	1.43 to 6.84	0.003
	Intercept = 85.91 (the model on $n = 92$ reduced the unexplained variation from SD = 14.7 to SD = 13.9)			
TEA	HNNE total score	1.15	0.35 to 1.94	0.005
	MRI deep grey matter score	-4.01	-6.84 to -1.17	0.006
	TEA GMs	-6.24	-10.85 to -1.63	0.008
	Intercept = 74.85 (the model on $n = 90$ reduced the unexplained variation from SD = 14.7 to SD = 13.0)			
3 months CA	3-month GMs	-9.39	-15.54 to -3.23	0.003
	TIMP total z-score	4.64	1.16 to 8.11	0.009
	Intercept = 98.45 (the model on $n = 94$ reduced the unexplained variation from SD = 15.4 to SD = 13.6)			
All 3 timepoints combined	3-month GMs	-10.01	-15.90 to -4.12	0.001
	TEA HNNE total score	1.33	0.57 to 2.08	0.001
	Intercept = 64.45 (the model on $n = 95$ reduced the unexplained variation from SD = 15.3 to SD = 13.4)			

Bayley III Bayley Scales of Infant and Toddler Development 3rd edition. CA corrected age, CI confidence interval, GMs General Movement Assessment, HNNE Hammersmith Neonatal Neurological Examination, MRI magnetic resonance imaging, PMA postmenstrual age, TEA term equivalent age, TIMP total z-score, β regression coefficient; n varies slightly between models as an effect of each subject having data available for the variables included in the model.

Multivariable models for Bayley III cognitive outcome. When cognitive outcome multivariable models were developed for each timepoint individually the results are as follows:

- 32 weeks PMA model included early HNNE reflexes subscale score ($\beta = 4.14$; 95% CI = 1.43 to 6.84; $p = 0.003$).
- TEA model included TEA HNNE total score ($\beta = 1.15$; 95% CI = 0.35 to 1.94; $p = 0.005$), TEA deep grey matter MRI score ($\beta = -4.01$; 95% CI = -6.84 to -1.17; $p = 0.006$) and the TEA GMs score ($\beta = -6.24$; 95% CI = -10.85 to -1.63; $p = 0.008$).
- 3 months CA included 3-month CA GMs score ($\beta = -9.39$; 95% CI = -15.54 to -3.23; $p = 0.003$) and the 3-month TIMP total z-score ($\beta = 4.64$; 95% CI = 1.16 to 8.11; $p = 0.009$).

When the three timepoints were combined, the best overall model for cognitive outcome included the 3-month CA GMs score ($\beta = -10.01$; 95% CI = -15.90 to -4.12; $p = 0.001$) and TEA HNNE score ($\beta = 1.33$; 95% CI = 0.57 to 2.08; $p = 0.001$). Results are

presented in Table 4. The model was constructed from data of 95 participants who had TEA HNNE, 3-month GMs and 2-year cognitive data available.

DISCUSSION

This study analysed a large cohort ($n = 98$) of very preterm born infants using Early MRI (32 weeks PMA) and TEA with concurrent clinical measures and clinical assessment at 3 months CA to determine associations with motor and cognitive outcomes at 2 years CA. Our study identified that 3-month GMs, Term MRI deep grey matter score and the Early HNNE reflexes score was the strongest model associated with motor outcome when data from all timepoints was included. The TEA HNNE total score and 3-month GMs were most strongly associated with cognitive outcome when data from all timepoints were included. Assessment before TEA strengthens associations for motor outcomes, but less so for cognitive outcomes. Clinical measures alone can be

used for associations with cognitive outcomes, a very important finding given the limited availability of MRI in the majority of clinical settings.

This study evaluated whether the addition of data from earlier in the neonatal period, to the traditional TEA and 3 months CA assessment timepoints, could improve associations with outcomes. Our data suggests that addition of neurological assessment data (HNNE) from the earlier timepoint improves associations with motor outcome. Addition of neurological assessment at TEA, improves associations for cognitive outcome. The clinical implications are that there is value in adding earlier neurological testing, to clinical management of very preterm infants, for the purposes of screening and planning further monitoring and surveillance.

Our data support previous studies of 3-month GMs to be strongly associated with both motor^{6,20,34} and cognitive^{8,9} outcomes and that the combination of TEA MRI and 3-month GMs has the greatest predictive accuracy for motor outcomes for preterm infants.²⁰ For the first time, we have demonstrated that addition of the HNNE reflexes score at 32 weeks PMA strengthens the associations with later motor outcomes.

It is an interesting and novel finding from this study that the early HNNE reflexes score demonstrates such strong associations with both motor and cognitive outcomes. The reflexes subscale at TEA is also strongly associated with cognitive outcome, and it is the elements of posture and tone, tone patterns and spontaneous movements which are significantly associated with cognitive outcome at TEA but not early on, which strengthens the total HNNE score at TEA. The HNNE reflex subscale assesses primitive reflexes of sucking, hand and toe grasp, Moro reflex, deep tendon reflexes and pacing. Other elements of the HNNE such as posture and tone, movements and behaviour are assessing functions that likely evolve over the first months of life in line with cortical development, becoming more discriminative by TEA and therefore more strongly associated with cognitive outcomes. Our data suggest that neonatal reflexes both early and at term are an important marker of outcomes and further analyses to tease out predictive accuracy and which of the reflex items contribute the most valuable information is warranted.

Diagnostic accuracy statistics are required to determine the true predictive accuracy of individual and combined variables for neurodevelopmental outcomes. This manuscript presents data from the first half of our preterm cohort. We have subsequently completed a total of $n = 262$ early MRI and clinical assessment, data collection is still in progress for 2-year outcomes and will be complete in February 2022. With the full dataset, we will undertake diagnostic accuracy analyses including sensitivity, specificity, positive predictive value, negative predictive value and accuracy for each of the individual variables and the combined models. The work completed and presented here is an important first step which has identified which of the measures, individual variables and combined models should be investigated further and tested with diagnostic accuracy analyses, both in our larger cohort when data collection is complete, and to guide others working the field as to which measures and variables warrant further investigation. In addition, future work will need to evaluate the predictive contributions of the measures and combined models, independent of other known risk factors such as GA at birth, presence of intraventricular haemorrhage or bronchopulmonary dysplasia, etc.

We have identified which measures and variables from a single timepoint could be selected by clinical or research teams, as well as which combination of variables over time can be combined to assist with risk stratification for follow-up screening and intervention requirements. Further work will further assist in understanding if the associations found are clinically meaningful as despite having statistically significant associations, the actual correlations were relatively low. GMs at all timepoints appear to

make a valuable contribution. The HNNE certainly warrants further investigation, having been found in this work to contribute to associations with both motor and cognitive outcomes. The NNNS is complex to administer and score, and is unlikely to confer clinical benefit over and above simple clinical tools such as the GMs and HNNE. The Premie-Neuro has good clinical utility and the neurological subscale was associated with cognitive outcomes. Given the HNNE was so much more strongly associated with cognitive outcomes, it is unlikely that it would be justifiable to select the Premie-Neuro as a clinical tool over the HNNE which also evaluates neurological function. The TIMP warrants further investigation to determine if the associations, especially at 3 months CA, are clinically meaningful and truly predictive, and whether it adds value beyond the administration of the 3-month GMs. The assessment of visual function is very simple and quick to administer and if found to be equivalent to the TIMP in terms of predictive accuracy, may be more efficient to administer in preterm follow-up services. Both early and TEA MRI variables were significantly associated with neurodevelopmental outcomes. Diagnostic accuracy analyses will be able to definitively answer the question as to whether early MRI really adds value to predicting long-term outcome given the expense and challenges of doing this compared with TEA MRI. While our study purposely completed multiple different assessments to compare their relative associations with outcomes, ultimately the long-term aim of this project is to determine the minimum assessment schedule that offers the best cost-benefit balance for both families and clinical services.

In the landscape previously of a 'wait and see' approach to outcomes from prematurity, our data and that of others provide strong evidence for assessments prior to 3 months CA. In the cases for those children with the most serious adverse outcomes of cerebral palsy, clinicians are encouraged by recent International Clinical Practice Guidelines for the early detection of Cerebral Palsy, to be using these tools and providing clear early diagnosis and prognosis to families as early as 3–4 months CA.⁶ Early interventions are being offered earlier than previously possible, starting from 3 to 6 months CA.^{35,36} It could be argued then that little is added by assessment earlier than term, and the resource cost may not be justified. Many neuroprotective, neurorestorative and rehabilitative treatment options are however being developed for application shortly after birth and while the infant is still in the neonatal intensive care unit. These include treatments with creatine, melatonin, amnion cells and developmental care. In that context, both clinical assessment and MRI, early in the neonatal period, are critical to enable selection of infants for eligibility and participation in these research trials. Our data provide evidence for tools for very early identification of participants who may warrant and benefit from new early interventions aimed at improving outcomes. Items from the early timepoint with significant associations with motor outcome include the MRI global score, GMs, HNNE reflexes, HNNE abnormal signs and NNNS regulation. Early MRI cerebellar score, HNNE reflexes, NNNS stress, NNNS arousal and the neurological subscale of the Premie-Neuro were significantly associated with cognitive outcome. This very early timepoint of assessment therefore offers a critical new window for early interventions to be implemented and evaluated.

Our finding that TEA HNNE total score was as strongly associated with cognitive outcome as the 3-month GMs is an important finding. The HNNE is a simple, clinically accessible tool that can be used anytime from birth until 2 months post term age in very preterm born infants.²¹ Together with Early HNNE reflexes being strongly associated with both motor and cognitive outcomes, and Early HNNE abnormal signs being associated with motor outcomes, inclusion of this tool in clinical care of these high-risk infants is warranted. Cognitive outcome is often overlooked with a tendency to focus on motor outcomes of preterm

survivors. A systematic review and meta-analysis of >64,000 preterm infants found that while risk for adverse cognitive outcomes was highest for infants with the lowest gestational ages, children with gestational ages at birth of 28–34 weeks performed almost as poorly as those born <28 weeks.¹ Traditional follow-up and early intervention programs that focus predominantly on extremely or very preterm infants are therefore likely missing many children at risk of adverse cognitive outcomes. Targeted early interventions need to incorporate and evaluate cognitive development strategies.³⁵ The fact that the HNNE can be completed in the neonatal intensive care unit prior to discharge can assist clinicians in planning future screening, enabling targeted surveillance and service requirements for the infant and family before they return home.

Our cohort had 4 children out of 98 with an outcome of CP and it is tempting to consider this a low-risk cohort. This prevalence of 4% is however aligned with other contemporary cohorts with documented reductions in CP prevalence compared to earlier cohorts.^{37,38} A prevalence rate of 3.8% in extremely low birth-weight infants (<1000 g) and 3.6% in very low birthweight infants (1000–1499 g) has been reported from data of 20 population-based CP registers of infants born in 2003.³⁷ Outcomes on the Bayley III motor composite score in the present study were a mean of 96 and SD 15, similar to the standardised Mean of 100 and SD 15.³¹ There has however been criticism of the Bayley III in underestimating adverse motor outcomes.^{39,40} Data of healthy term born infants in Australia, assessed at 2 years, reported a mean motor composite score of 118 and a standard deviation of 17.³⁹ Our preterm cohort therefore scored lower by 1 SD on average when compared to typically developing children in Australia and is in line with other cohorts of very preterm infants assessed at 2 years CA.³⁹ These data confirm that our cohort is representative of contemporary populations of very preterm born infants.

Strengths of the current study include the large sample of Early MRI data with concurrent clinical data of infants born very preterm in a contemporaneous study cohort with masked analysis of both clinical and MRI assessments and good participant retention at 2 years CA (>80%). MRI was acquired at 3T and we chose a clinically accessible scoring system for structural MRI rather than more complex volumetric or diffusion analyses. Limitations exist for multivariable modelling. Different statistical methodologies produce slightly different models, and it is a great challenge to decide which methodology truly produces the 'correct' or 'best' model. There is a tendency to give abundant attention to the measures that are included in the final model, and disregard all others. To mitigate this, we have presented the results in multiple ways (univariable, adjusted for 3-month GMs and full multivariable models). The variables that narrowly missed out inclusion in the full models, all feature in the univariable analyses. We did not correct for multiple comparisons, however set our methodology for the multivariable model building to use $p < 0.01$ for selection of addition to the model.

CONCLUSION

Our study findings support previous work highlighting the value of term MRI and 3-month GMs in identifying motor outcomes in very preterm born infants. The clinical implications of our study findings suggest that early and term neurological assessment, in particular the HNNE, adds value in understanding motor and cognitive outcomes. These findings lay the groundwork for our PPREMO toolbox and future work will examine predictive accuracy of each of the biomarkers and also include our advanced diffusion imaging biomarkers. Until diagnostic accuracy analyses definitively compare the predictive accuracy of early MRI compared with TEA, TEA MRI remains recommended over early MRI for prediction of longer term outcomes.

DATA AVAILABILITY

Deidentified individual participant data will not be made available.

ACKNOWLEDGEMENTS

We thank the families of preterm and term infants who participated in this study, the medical and nursing staff of the neonatal unit and radiology and administrative staff in the department of medical imaging. Dr. Melissa Lai undertook clinical study management and escorted infants to MRI, research nurses Donna Hovey, Kellie McGrory and Kylie Smart conducted recruitment, study management and infant MRI, Kym Morris performed clinical assessment and Bernadette Shannon independently scored GMs. We also thank Chris Finn who conducted the 2-year outcome assessments. Grants from the Cerebral Palsy Alliance Research Foundation (IRG1413), the Financial Markets Foundation for Children (2014-074), and the Queensland Government (Smart State; Health Practitioner Stimulus Grant) support this project. The authors received funding from The University of Queensland (University of Queensland Research Scholarship [to J.M.G.], the Queensland Government (Smart State PhD Top-Up Scholarship [to J.M.G.]) and the National Health and Medical Research Council (Research Fellowship 1105038 [to R.N.B.]; Centre of Research Excellence: Australasian Cerebral Palsy Clinical Trials Network). No funder was involved in study design, data collection, data analysis, manuscript preparation and/or publication decisions.

AUTHOR CONTRIBUTIONS

J.M.G.: substantial contributions to conception and design, acquisition of data, analysis and interpretation of data; drafting the article and revising it critically for important intellectual content; and final approval of the version to be published. P.B.C. and K.P.: substantial contributions to conception and design, acquisition of data and interpretation of data; revising the article critically for important intellectual content; and final approval of the version to be published. M.D.C.: substantial contributions to analysis and interpretation of data, revising the article critically for important intellectual content; and final approval of the version to be published. S.F. and J.F.: substantial contributions to acquisition of data, revising the article critically for important intellectual content; and final approval of the version to be published. A.G.: substantial contributions to conception and design, acquisition of data, and final approval of the version to be published. S.E.R.: substantial contributions to conception and design, acquisition of data, and revising manuscript critically for important intellectual content; and final approval of the version to be published. R.S.W. and R.N.B.: substantial contributions to conception and design, analysis and interpretation of data; revising the article critically for important intellectual content; and final approval of the version to be published.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41390-021-01399-5>.

Competing interests: The authors declare no competing interests.

Consent statement: Informed parental consent was obtained for all participants.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Allotey, J. et al. Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64 061 children. *BJOG* **125**, 16–25 (2018).
- Williams, J., Lee, K. J. & Anderson, P. J. Prevalence of motor-skill impairment in preterm children who do not develop cerebral palsy: a systematic review. *Dev. Med. Child Neurol.* **52**, 232–237 (2010).
- Edwards, J. et al. Developmental coordination disorder in school-aged children born very preterm and/or at very low birth weight: a systematic review. *J. Dev. Behav. Pediatr.* **32**, 678–687 (2011).
- Report of the Australian Cerebral Palsy Register Birth years 1995–2012. Australian Cerebral Palsy Register, 2018.
- Kerr-Wilson, C. O., Mackay, D. F., Smith, G. C. & Pell, J. P. Meta-analysis of the association between preterm delivery and intelligence. *J. Public Health (Oxf.)* **34**, 209–216 (2012).
- Novak, I. et al. 2017 early, accurate diagnosis and early intervention in cerebral palsy: advances in diagnosis and treatment. *JAMA Pediatr.* **171**, 897–907 (2017).

7. Anderson, P. J. et al. TE 2017 associations of newborn brain magnetic resonance imaging with long-term neurodevelopmental impairments in very preterm children. *J. Pediatr.* **187**, 58–65 (2017).
8. Einspieler, C., Bos, A. F., Libertus, M. E. & Marschik, P. B. The general movement assessment helps us to identify preterm infants at risk for cognitive dysfunction. *Front. Psychol.* **7**, 406 (2016).
9. Spittle, A. J. et al. General movements in very preterm children and neurodevelopment at 2 and 4 years. *Pediatrics* **132**, e452–e458 (2013).
10. Spittle, A. J., Boyd, R. N., Inder, T. E. & Doyle, L. W. Predicting motor development in very preterm infants at 12 months' corrected age: the role of qualitative magnetic resonance imaging and general movements assessments. *Pediatrics* **123**, 512–517 (2009).
11. George, J. M. et al. PPREMO: a prospective cohort study of preterm infant brain structure and function to predict neurodevelopmental outcome. *BMC Pediatr.* **15**, 123 (2015).
12. George, J. M. et al. Validation of an MRI brain injury and growth scoring system in very preterm infants scanned at 29- to 35-week postmenstrual age. *AJNR Am. J. Neuroradiol.* **38**, 1435–1442 (2017).
13. George, J. M. et al. Relationship between very early brain structure and neuro-motor, neurological and neurobehavioral function in infants born <31weeks gestational age. *Early Hum. Dev.* **117**, 74–82 (2018).
14. Hack, M. et al. The effect of very low birth weight and social risk on neurocognitive abilities at school age. *J. Dev. Behav. Pediatr.* **13**, 412–420 (1992).
15. Roberts, G. et al. Rates of early intervention services in very preterm children with developmental disabilities at age 2 years. *J. Paediatr. Child Health* **44**, 276–280 (2008).
16. Roberts, G., Lim, J., Doyle, L. W. & Anderson, P. J. High rates of school readiness difficulties at 5 years of age in very preterm infants compared with term controls. *J. Dev. Behav. Pediatr.* **32**, 117–124 (2011).
17. Kidokoro, H., Neil, J. J. & Inder, T. E. New MR imaging assessment tool to define brain abnormalities in very preterm infants at term. *AJNR Am. J. Neuroradiol.* **34**, 2208–2214 (2013).
18. Einspieler, C. *Prechtl's Method on the Qualitative Assessment of General Movements in Preterm, Term and Young Infants* (Mac Keith Press, 2004).
19. Darsaklis, V., Snider, L. M., Majnemer, A. & Mazer, B. Predictive validity of Prechtl's method on the qualitative assessment of general movements: a systematic review of the evidence. *Dev. Med. Child Neurol.* **53**, 896–906 (2011).
20. Bosanquet, M., Copeland, L., Ware, R. & Boyd, R. A systematic review of tests to predict cerebral palsy in young children. *Dev. Med. Child Neurol.* **55**, 418–426 (2013).
21. Dubowitz, L., Mercuri, E. & Dubowitz, V. An optimality score for the neurologic examination of the term newborn. *J. Pediatr.* **133**, 406–416 (1998).
22. Lester, B. M. & Tronick, E. Z. History and description of the Neonatal Intensive Care Unit Network Neurobehavioral Scale. *Pediatrics* **113**, 634–640 (2004).
23. El-Dib, M., Massaro, A. N., Glass, P. & Aly, H. Neurobehavioral assessment as a predictor of neurodevelopmental outcome in preterm infants. *J. Perinatol.* **32**, 299–303 (2012).
24. Stephens, B. E. et al. Neurobehavioral assessment predicts motor outcome in preterm infants. *J. Pediatr.* **156**, 366–371 (2010).
25. Liu, J. et al. Neonatal neurobehavior predicts medical and behavioral outcome. *Pediatrics* **125**, e90–e98 (2010).
26. Daily, D. K. & Ellison, P. H. The premie-neuro: a clinical neurologic examination of premature infants. *Neonat. Netw.* **24**, 15–22 (2005).
27. Ricci, D. et al. Application of a neonatal assessment of visual function in a population of low risk full-term newborn. *Early Hum. Dev.* **84**, 277–280 (2008).
28. Ricci, D. et al. Early visual assessment in preterm infants with and without brain lesions: correlation with visual and neurodevelopmental outcome at 12 months. *Early Hum. Dev.* **87**, 177–182 (2011).
29. Campbell, S. K. & Hedeker, D. Validity of the test of infant motor performance for discriminating among infants with varying risk for poor motor outcome. *J. Pediatr.* **139**, 546–551 (2001).
30. Campbell, S. K., Kolobe, T. H., Osten, E. T., Lenke, M. & Girolami, G. L. Construct validity of the test of infant motor performance. *Phys. Ther.* **75**, 585–596 (1995).
31. Bayley, N. *Bayley Scales of Infant and Toddler Development* 3rd edn (Harcourt Assessment, 2006).
32. Rosenbaum, P. et al. A report: the definition and classification of cerebral palsy April 2006. *Dev. Med. Child Neurol. Suppl.* **109**, 8–14 (2007).
33. Wood, E. & Rosenbaum, P. The gross motor function classification system for cerebral palsy: a study of reliability and stability over time. *Dev. Med. Child Neurol.* **42**, 292–296 (2000).
34. Skjold, B., Eriksson, C., Eliasson, A. C., Aden, U. & Vollmer, B. General movements and magnetic resonance imaging in the prediction of neuromotor outcome in children born extremely preterm. *Early Hum. Dev.* **89**, 467–472 (2013).
35. Morgan, C., Novak, I., Dale, R. C., Guzzetta, A. & Badawi, N. GAME (Goals - Activity - Motor Enrichment): protocol of a single blind randomised controlled trial of motor training, parent education and environmental enrichment for infants at high risk of cerebral palsy. *BMC Neurol.* **14**, 203 (2014).
36. Boyd, R. N. et al. REACH: study protocol of a randomised trial of rehabilitation very early in congenital hemiplegia. *BMJ Open* **7**, e017204 (2017).
37. Sellier, E. et al. Decreasing prevalence in cerebral palsy: a multi-site European population-based study, 1980 to 2003. *Dev. Med. Child Neurol.* **58**, 85–92 (2016).
38. Reid, S. M. et al. Temporal trends in cerebral palsy by impairment severity and birth gestation. *Dev. Med. Child Neurol.* **58**, 25–35 (2016).
39. Anderson, P. J., De Luca, C. R., Hutchinson, E., Roberts, G. & Doyle, L. W. Under-estimation of developmental delay by the new Bayley-III Scale. *Arch. Pediatr. Adolesc. Med.* **164**, 352 (2010).
40. Anderson, P. J. & Burnett, A. Assessing developmental delay in early childhood – concerns with the Bayley-III scales. *Clin. Neuropsychol.* **31**, 371–381 (2017).