Early Human Development 195 (2024) 106068

Contents lists available at ScienceDirect

Early Human Development

journal homepage: www.elsevier.com/locate/earlhumdev

Sensitivity and specificity of the Neonatal Visual Assessment to predict motor and cognitive outcomes in infants born very preterm

Jessica W. Blazek [a], Paul B. Colditz [b,c], Andrea Guzzetta [d,e], Robert S. Ware [f], Mark D. Chatfield [g], Judith L. Hough [a,h,i], Roslyn N. Boyd [g], Joanne M. George [h,j,*]

[a] School of Allied Health, Australian Catholic University, Brisbane, Australia
[b] University of Queensland Centre for Clinical Research, The University of Queensland, Brisbane, Australia
[c] Perinatal Research Centre, Royal Brisbane and Women's Hospital, Brisbane, Australia
[d] Department of Developmental Neuroscience, IRCCS Stella Maris, Pisa, Italy
[e] Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
[f] Griffith Biostatistics Unit, Griffith University, Brisbane, Australia
[g] Queensland Cerebral Palsy and Rehabilitation Research Centre, Centre for Children's Health Research, Faculty of Medicine, The University of Queensland, Brisbane, Australia
[h] Child Health Research Centre, Faculty of Medicine, The University of Queensland, Brisbane, Australia
[i] Physiotherapy Department, Mater Hospital, Brisbane, Australia
[j] Physiotherapy Department, Queensland Children's Hospital, Children's Health Queensland Hospital and Health Service, Brisbane, Australia

ARTICLE INFO

Keywords:
Infant, premature
Vision screening
Neurodevelopment
Cohort study

ABSTRACT

Background: Very preterm infants are at increased risk of neurodevelopmental impairments. The Neonatal Visual Assessment (NVA) assesses visual function and outcomes and has been used to assess early neurodevelopmental outcomes. This study aimed to compare NVA results of very preterm and term-born infants and to calculate the sensitivity and specificity of the NVA at term equivalent age (TEA) and three months corrected age (CA) to predict motor and cognitive outcomes at 12 months CA in very preterm infants.

Methods: This prospective observational cohort study recruited infants born before 31 weeks gestation and a healthy term-born control group. The NVA was assessed at TEA and three months CA, and neurodevelopmental outcomes (Bayley Scales of Infant and Toddler Development, Third Edition; Neurosensory Motor Developmental Assessment; Alberta Infant Motor Scale) were performed at 12 months CA. The sensitivity and specificity of the NVA to predict outcomes were calculated based on a previously published optimality score.

Results: 248 preterm (54 % male) and 46 term-born infants (48 % male) were analysed. The mean NVA scores of preterm and term-born infants were significantly different at TEA (preterm 3.1±2.1; term-born 1.2±1.7, $p < 0.001$). The NVA had moderate sensitivity (59–78 %) and low specificity (25–27 %) at TEA, and low sensitivity (21–28 %) and high specificity (86–87 %) at three months CA for the prediction of preterm infants' outcomes at 12 months CA.

Conclusion: The NVA at TEA and three months CA was not a strong predictor of motor and cognitive impairments in this contemporary cohort of very preterm infants.

1. Introduction

Very preterm infants (born <32 weeks gestational age [GA]) are at an increased risk of developing a range of motor, cognitive and behavioural impairments compared to term-born infants due to adverse pre- and post-natal factors [1]. As the visual system is one of the first body systems to mature, impaired development of visual function may provide an early indication of the impact of perinatal brain damage in preterm infants [2,3]. It has been postulated that visual cortical function in the first months after birth may provide a sensitive measure of the functional effect of perinatal brain damage and may help to identify mild, yet functionally significant impairments in preterm infants [4]. Assessment of visual function in the first months of life therefore merits further exploration as a predictor of neurodevelopment. Ideally, both

* Corresponding author at: Child Health Research Centre, Faculty of Medicine, The University of Queensland, Brisbane, Australia.
E-mail address: j.george2@uq.edu.au (J.M. George).

https://doi.org/10.1016/j.earlhumdev.2024.106068
Received 2 May 2024; Received in revised form 13 June 2024; Accepted 14 June 2024
Available online 24 June 2024

mild and severe developmental impairments should be diagnosed soon after birth so the infant can be referred to early intervention while there is the greatest potential to optimise outcomes and prevent permanent disability [5].

The Neonatal Visual Assessment (NVA) is an assessment tool used to assess visual function and predict neurodevelopmental outcomes in infants from term age [6]. Current studies examining the NVA have been performed by one research group and have not specifically investigated very preterm infants born <32 weeks gestation [3,6–10], who are at greater risk of long-term developmental impairments [11].

The aims of this study were therefore to i) compare NVA results of very preterm and term-born control infants at term equivalent age (TEA), and ii) calculate the sensitivity and specificity of the NVA at TEA and three months corrected age (CA) to predict motor and cognitive outcomes at 12 months CA in a contemporary cohort of infants born very preterm.

2. Methods

2.1. Study design

This study of preterm infants born <31 weeks GA combines data from two prospective observational cohort studies conducted in Brisbane and Melbourne, Australia, between 2013 and 2019 [12]. Healthy term-born infants from three existing studies conducted at the same hospital over the same time period were a comparison group [12,13]. Ethical approval was obtained from the Human Research Ethics Committees at The Royal Brisbane and Women's Hospital (RBWH) (HREC/12/QRBW/245) and the University of Queensland (UQ) (2012001060) [12]. Additional term-born infant data was sourced from the RBWH (HREC/09/QRBW/296), Children's Health Queensland Hospital and Health Service (HREC/12/QCHQ/40) and UQ (2014001160) [13]. Studies were registered with the Australian New Zealand Clinical Trials Registry; PPREMO ACTRN12613000280707, PREBO ACTRN12619000155190 and PREMM ACTRN12612000335897 [13]. Written informed consent was obtained from the parents or guardians of participants in the study.

2.2. Preterm cohort

Preterm infants were recruited from the Neonatal Intensive Care Unit (NICU) at the RBWH between February 2013 and December 2019. Infants were eligible if they were born before 31 weeks gestation, lived within 200 km of the RBWH to allow for follow up appointments and home visits, and, due to a lack of funding for translators, were from English-speaking families. Infants were ineligible if they had diagnosed chromosomal or congenital abnormalities that could adversely affect neurodevelopmental outcomes [12]. To be included in the present analysis, infants must have had data available from the NVA at TEA and/or three months CA and at least one neurodevelopmental outcome measure at 12 months CA.

2.3. Term-born cohort

Term-born infants were recruited from the post-natal wards of the RBWH or as interested volunteers by word of mouth. Infants were eligible if they were born between 37 and 41 weeks gestation after an uncomplicated pregnancy and delivery, had a birth weight greater than the tenth percentile, were not treated in the NICU or special care unit after birth [12], and were assessed using the NVA at TEA.

2.4. Demographic and clinical characteristics

Several perinatal variables have been shown to impact neurodevelopmental outcomes in infants born preterm [12]. Details of each participant's birth history and neonatal course were collected from the medical discharge summaries, including premature rupture of membranes, chorioamnionitis, maternal antenatal corticosteroid use and social risk. Social risk was assessed using a questionnaire examining six areas including family structure, maternal age, education level, occupation, employment status and language spoken in the home, with each scoring between zero and two for a maximum total score of 12 [14–16]. Scores of two and above were classified as higher social risk [14,16].

2.5. The Neonatal Visual Assessment (NVA)

The NVA comprises assessment of nine aspects of visual function including spontaneous ocular motility, ocular movements with a target, fixation, horizontal tracking, vertical tracking, arc tracking, tracking coloured stimulus, stripes discrimination and attention at distance [7]. When performed at TEA, its sensitivity and specificity to predict global neurodevelopmental outcomes at 12 months CA in preterm infants have been reported as 92 % and 74 %, respectively [10]. The NVA has strong clinical utility as it can be performed in under ten minutes, does not require extensive assessor training, and utilises minimal equipment. At TEA and three months CA, a physiotherapist blinded to all perinatal characteristics and neuroimaging findings conducted the NVA at the homes of the enrolled infants, at outpatient appointments or in hospital if they were inpatients at the time. The examiners were trained to conduct the assessment by authors of this paper (AG, PC), who had been taught by one of the original authors of the NVA (Daniela Ricci) [7]. An infant's performance in each test item was scored zero (indicating good performance) or one (indicating poor performance) based on the optimality scoring from Ricci and colleagues [10], resulting in a total global NVA score ranging from zero to nine.

2.6. Neurodevelopmental outcome measures at 12 months CA

As no single measure has been shown to provide conclusive data on motor and cognitive outcomes in preterm infants [17], a combination of the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III), the Neurosensory Motor Developmental Assessment (NSMDA) and the Alberta Infant Motor Scale (AIMS) were performed at 12 months CA. A physiotherapist blinded to the infants' clinical history performed these assessments during a follow-up appointment at the RBWH.

The Bayley-III is a discriminative, norm-referenced tool used to assess global developmental outcomes in children aged one to 42 months [18,19]. It consists of cognitive, motor and language domains, which are assessed via the interaction between the infant and the examiner in a standardised series of play tasks [18]. Composite scores are derived and range from 40 to 160 [20]. Although the Bayley-III is the most widely used standardised assessment tool internationally for neurodevelopmental outcomes in preterm infant follow up, there are concerns regarding overestimation of cognitive and motor performance [21], and poor prediction of later cognitive [22,23] and motor performance [24]. For this reason, a further two clinical outcome assessments were included, the NSMDA and the AIMS.

The NSMDA is a criterion-referenced tool used to evaluate six domains of development, including gross and fine motor function, neurological status, patterns of movement, postural development and motor responses to sensory input [25]. Each item is scored from one (definitely abnormal) to four (above average) [26]. Motor performance is further classified as normal (6 to 8), minimal (9 to 11), mild (12 to 14), moderate (15 to 19), severe (20 to 24) or profound (25 to 30) dysfunction based on functional grades for each domain [26]. When performed at four months of age, the NSMDA has a reported sensitivity of 80 % and specificity of 56 % to predict motor developmental outcomes at 24 months [27]. Moderate and severe dysfunction on the NSMDA at eight months, two years and four years of age has been shown to correlate with later motor impairments in children without cerebral palsy (CP) or other neurological conditions at 11 to 13 years of age (eight months

sensitivity 80 %, specificity 29 %; two years sensitivity 48 %, specificity 82 %; four years sensitivity 65 %, specificity 80 %) [25].

The AIMS assesses gross motor development in infants from birth to 18 months of age [28]. This norm-referenced tool consists of 58 items which involve observing the infant in prone, supine, sitting and standing [28]. A total score is recorded and can be converted to centile ranks and age-equivalent growth scores [29]. Scores below the tenth centile at four months and below the fifth centile at eight months are considered valid and reliable indicators of motor developmental delay [30]. The AIMS has a reported sensitivity of 86 % and specificity of 93 % to predict motor outcomes in infants born preterm at 18 months CA [28].

2.7. Statistical analysis

To summarise NVA scores, the mean, standard deviation (SD), ranges and centiles were calculated. An independent samples *t*-test was used to analyse the difference in preterm and term-born infants' NVA scores at TEA. The level of significance was set at $p < 0.05$. The Bayley-III cognitive and motor composite scores, NSMDA functional classification and AIMS were dichotomised using cut-off points from previously published studies ($\leq$ 85 for Bayley-III cognitive and motor composite scores [31]; $\geq$ 12 on the NSMDA [32]; and $\leq$ the 5th centile for the AIMS [28]). Global NVA scores were dichotomised based on the optimality scoring from Ricci and colleagues (scores of zero and one were considered normal and scores of $\geq$ two were considered abnormal) [10]. Sensitivity and specificity were calculated to evaluate the ability of the NVA to predict neurodevelopmental outcomes at 12 months CA. The percentage correctly classified and 95 % CIs were calculated for these estimates based on the binomial distribution. Area under the receiver operating characteristic (ROC) curve analyses of NVA scores at TEA and three months CA were undertaken. The cut-off points that most accurately predict neurodevelopmental outcomes at 12 months CA were determined according to i) the greatest combination of sensitivity and specificity [determined as the largest sum of sensitivity and specificity], and ii) and the best balance of sensitivity and specificity. To further evaluate the predictive accuracy of the NVA, a sensitivity analysis was performed using our term cohort to generate our own optimality score (Appendix A). For each NVA test item, our preterm cohort was scored a zero if their performance fell within the 90th centile of our term cohort, and one if their performance fell outside the 90th centile [10]. The individual NVA test items were summed to provide a total global score ranging from zero to nine, where scores of zero and one were considered normal and scores of $\geq$ two were considered abnormal [10]. Data analysis was performed using Stata statistical software v14 (StataCorp, College Station, TX, USA).

3. Results

Two hundred and forty-eight eligible preterm infants had data available from the NVA at TEA or three months CA. Of these infants, 199 had at least one 12-month CA outcome assessment performed. At 12 months CA, the neurodevelopmental outcomes of 197 preterm infants were assessed using the Bayley-III and NSMDA, and 199 infants were assessed using the AIMS. Forty-six term-born infants had NVA data available at TEA and were included in the term-born cohort. The characteristics of both preterm and term-born cohorts are presented in Table 1. The term-born cohort had a significantly higher mean birth weight ($p < 0.001$) and birth head circumference ($p < 0.001$), lower rate of caesarean sections ($p < 0.001$), and 15 % were classified as having higher social risk compared to 52 % of the preterm cohort ($p < 0.001$). Table B.1 compares the characteristics of preterm infants who were assessed using the NVA at TEA with or without an outcome at 12 months CA. The infants who did not have an outcome assessment performed had a significantly higher gestational age at birth ($p < 0.05$), birth weight ($p < 0.05$), and proportion of infants with chorioamnionitis ($p < 0.001$) and higher social risk ($p < 0.05$) compared to those who had an outcome

Table 1
Characteristics of preterm and term-born study samples. Data are presented as mean (SD) or median (IQR) for continuous measures and frequency (percentage) for categorical measures.

	Preterm infants n = 248	Term-born infants n = 46	p-value[†]
Birth and Maternal Data			
Gestational age at birth (weeks), median (IQR)	28.3 (26.9 to 30.6)	39.6 (38.9 to 40.4), n = 44	**<0.001**
Birth weight (g), mean (SD)	1122.8 (308.0)	3486.0 (273.2), n = 44	**<0.001**
Birth head circumference (cm), mean (SD)	26.0 (2.3), n = 242	34.8 (1.1), n = 38	**<0.001**
Males	134 (54 %)	21 (48 %), n = 44	0.47
Multiple births	75 (30 %)	0 (0 %), n = 42	0.06
Premature rupture of membranes	57 (23 %)	4 (10 %), n = 40	0.06
Caesarean section	168 (68 %)	16 (38 %), n = 42	**<0.001**
Chorioamnionitis	39 (16 %), n = 247		
Antenatal corticosteroids	172 (69 %)		
Magnesium sulphate	129 (60 %), n = 215		
Higher social risk[‡]	122 (52 %), n = 234	6 (15 %), n = 39	**<0.001**
Acquired Medical Factors			
Patent ductus arteriosus	89 (36 %)		
Any IVH	72 (29 %)		
IVH grade III or IV	22 (9 %)		
Periventricular leukomalacia	8 (3 %), n = 241		
Hydrocephalus	10 (4 %), n = 241		
Seizures treated with anticonvulsant therapy	1 (0.4 %), n = 241		
NEC diagnosed or suspected	11 (5 %), n = 241		
Confirmed sepsis	11 (5 %), n = 241		
Total parenteral nutrition (days), median (IQR)	11 (8 to 14), n = 240		
Postnatal corticosteroids	35 (15 %), n = 241		
Ventilation (days), median (IQR)	2 (0 to 7), n = 240		
CPAP (days), median (IQR)	28 (7 to 47), n = 240		
Oxygen therapy (hours), median (IQR)	74 (4 to 771), n = 231		
Bronchopulmonary dysplasia[§]	102 (42 %), n = 241		
Presence of ROP at 3 months CA	101 (42 %), n = 240		
ROP stage 1 at 3 months CA	53 (52 %), n = 101		
ROP stage 2 at 3 months CA	41 (41 %), n = 101		
ROP stage 3 at 3 months CA	7 (7 %), n = 101		
ROP zone 1 at 3 months CA	1 (1 %), n = 101		
ROP zone 2 at 3 months CA	54 (53 %), n = 101		
ROP zone 3 at 3 months CA	46 (46 %), n = 101		
PMA (weeks) at TEA clinical assessment, median (IQR)	40.9 (40 to 41.6), n = 241	41.2 (40.7 to 42.1), n = 44	0.67
CA (weeks) at 3-month clinical assessment, Median (IQR)	13.3 (13 to 13.7), n = 184		

SD standard deviation; IQR interquartile range; IVH intraventricular haemorrhage; NEC necrotizing enterocolitis; CPAP continuous positive airway pressure; ROP retinopathy of prematurity; CA corrected age; PMA postmenstrual age; TEA term equivalent age.

[†] Independent samples *t*-test, significance set at $p < 0.05$.

‡ Social risk was assessed using a questionnaire examining 6 areas with each scoring 0–2 for a maximum social score of 12: including family structure, maternal age, education level, occupation, employment status and language spoken in the home. Scores of 2 and above were classified as indicating higher social risk.

§ Defined as oxygen requirement at 36 weeks postmenstrual age.

performed. The infants without an outcome performed also had a significantly lower proportion of intraventricular haemorrhage [IVH] Grade III or IV ($p < 0.05$) and bronchopulmonary dysplasia ($p < 0.05$), and required significantly fewer days of continuous positive airway pressure ($p < 0.05$).

3.1. The NVA and neurodevelopmental outcomes

Preterm and term-born global NVA scores at TEA are presented in Fig. 1. Preterm infants had significantly higher average NVA scores at TEA (3.1±2.1) compared to term-born infants (1.2±1.7), $p < 0.001$. When NVA scores were dichotomised using the optimality scoring from Ricci and colleagues, 26 % of preterm infants had a normal NVA global score at TEA and 85 % had a normal NVA score at three months CA [10]. Dichotomised neurodevelopmental outcome scores at 12 months CA are summarised in Table 2. The percentage of preterm-born infants with adverse outcomes was 9 % (NSMDA), 10 % (Bayley-III cognitive subscale), 11 % (AIMS) and 22 % (Bayley-III motor subscale) [Table 2]. Of the infants who had the NVA performed at TEA or three months CA, the number who obtained sub-optimal results in one, two, three or all four of the 12 month outcome assessments were 24, 13, 5 and 11, respectively.

3.2. Diagnostic accuracy of the NVA

The NVA had moderate sensitivity (59 to 78 %) and low specificity (25 to 27 %) for the prediction of motor and cognitive outcomes when performed at TEA (Table 3). The NVA only correctly classified a low proportion (28 to 38 %) of preterm infants as having neurodevelopmental delay. When performed at three months CA, the NVA had

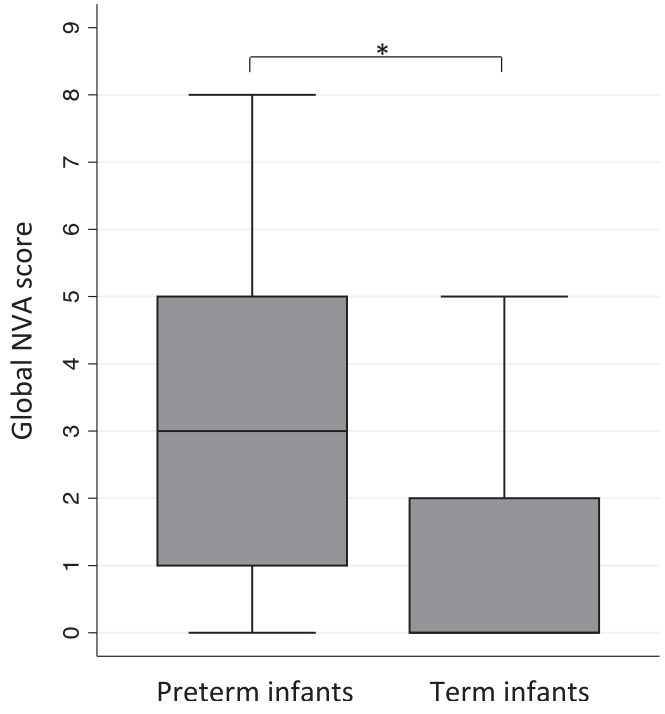

Fig. 1. Box plot of global Neonatal Visual Assessment scores at term equivalent age for preterm and term-born infants. Legend: * indicates $p < 0.001$; NVA Neonatal Visual Assessment.

Table 2
Motor and neurodevelopmental outcomes of preterm-born infants at 12 months corrected age. Data are presented as mean (SD) or median (IQR); range min-max for continuous measures and n (%) for categorical measures.

Age at 12-month assessment (weeks)	Outcome	Score	Cut-off point	Number (%) of preterm-born infants with adverse outcomes at 12 months CA
53.0 (52.1 to 53.9); range 48.0 to 63.9	Bayley-III cognitive composite $n = 197$	102.7 (11.6), 100 (95 to 110); range 60 to 140	85	19 (10 %)†
	Bayley-III motor composite n = 197	96.4 (14.5), 97 (88 to 107); range 46 to 136	85	45 (22 %)†
	NSMDA functional classification n = 197	8.0 (3.0), 7 (6 to 9); range 6 to 23	12	18 (9 %)‡
	AIMS $n = 200$	51.0 (7.5), 52 (51 to 54); range 14 to 58	5th centile	22 (11 %)†

Legend: CA corrected age; Bayley-III Bayley Scales of Infant and Toddler Development 3rd Edition; NSMDA Neurosensory Motor Developmental Assessment; AIMS Alberta Infant Motor Scale.

† Refers to the number of infants with scores less than or equal to the cut-off point.

‡ Refers to the number of infants with scores greater than or equal to the cut-off point.

Table 3
Sensitivity, specificity and correct classification of the Neonatal Visual Assessment of preterm-born infants at TEA and three months CA for predicting neurodevelopmental outcomes at 12 months CA.

Outcome at 12 months	Age at assessment	Sensitivity (%), 95 % CI	Specificity (%), 95 % CI	Correctly classified (%)
Bayley-III cognitive	TEA	78 (52 to 94)	27 (20 to 34)	32
	3 months CA	21 (6 to 46)	86 (80 to 91)	80
Bayley-III motor	TEA	76 (60 to 88)	27 (20 to 35)	38
	3 months CA	21 (20 to 35)	87 (81 to 92)	72
NSMDA functional classification	TEA	59 (33 to 82)	25 (18 to 32)	28
	3 months CA	28 (10 to 54)	87 (81 to 91)	81
AIMS	TEA	73 (50 to 89)	26 (20 to 34)	32
	3 months CA	27 (11 to 50)	87 (81 to 91)	80

Legend: TEA term equivalent age; CA corrected age; CI confidence interval; Bayley-III Bayley Scales of Infant and Toddler Development 3rd Edition; NSMDA Neurosensory Motor Developmental Assessment; AIMS Alberta Infant Motor Scale.

low sensitivity (21 to 28 %), high specificity (86 to 87 %), and correctly classified a high proportion (72 to 81 %) of infants as having neurodevelopmental delay (Table 3).

3.3. Area under the curve analyses

The results of ROC analyses to determine if using an alternative NVA cut-off point (rather than scores of zero or one are considered normal and scores of ≥ two are considered abnormal) would result in higher predictive accuracies are displayed in Table 4. At TEA, the cut-off point of ≥ three yielded the best combinations of sensitivity and specificity for all outcomes (sensitivity 67 to 68 %; specificity 42 to 44 %) except for

Table 4
Receiver operating characteristic analysis of the Neonatal Visual Assessment to determine which cut-off point is able to most accurately predict neurodevelopmental outcomes at 12 months CA in preterm infants.

Outcome performed at 12 months CA	Age at assessment	Cut-off point with best combination of sensitivity and specificity[†]	Sensitivity (%)	Specificity (%)	Cut-off point with most balanced sensitivity and specificity[†]	Sensitivity (%)	Specificity (%)	Area under the ROC curve (95 % CI)
Bayley-III cognitive	TEA	3	67	42	3	67	42	0.5 (0.4 to 0.6)
	3 months CA	1	32	76	1	32	76	0.5 (0.4 to 0.7)
Bayley-III motor	TEA	3	68	44	3	68	44	0.5 (0.4 to 0.6)
	3 months CA	1	32	78	1	32	78	0.6 (0.5 to 0.6)
NSMDA functional classification	TEA	8	6	98	4	41	55	0.4 (0.3 to 0.6)
	3 months CA	1	50	78	1	50	78	0.6 (0.5 to 0.8)
AIMS	TEA	3	68	42	4	50	57	0.5 (0.4 to 0.7)
	3 months CA	1	45	78	1	45	78	0.6 (0.5 to 0.7)

Legend: ROC receiver operating characteristic; NVA Neonatal Visual Assessment; CA corrected age; CI confidence interval; Bayley-III Bayley Scales of Infant and Toddler Development 3rd Edition; TEA term equivalent age; NSMDA Neurosensory Motor Developmental Assessment; AIMS Alberta Infant Motor Scale.
[†] Refers to scores greater than or equal to the cut-off point.

the NSMDA, in which the cut-off point of ≥ eight was best (sensitivity 6 %; specificity 98 %). The cut-off point with the most balanced sensitivity and specificity at TEA was ≥ three for Bayley-III cognitive and motor outcomes (sensitivity 67 to 68 %; specificity 42 to 44 %), and ≥ four for NSMDA and AIMS outcomes (sensitivity 41 to 50 %; specificity 55 to 57 %). At three months CA, the NVA cut-off point with the greatest combination and the best balance of sensitivity and specificity was ≥ one for all outcomes (sensitivity 32 to 50 %; specificity 76 to 78 %).

3.4. Sensitivity analysis

As our results mainly revealed low to moderate sensitivity and specificity at TEA and three months CA, a sensitivity analysis was performed to determine if the predictive accuracy could be improved by using our term cohort to generate our own optimality score. Details of this analysis are presented in Appendix A. Fig. 1.A displays the original NVA scoring sheet published by Ricci and colleagues [10] and Fig. 2.A displays the NVA scoring sheet based on our optimality score. Using our optimality score at TEA, the NVA had high sensitivity (94 to 100 %) but low specificity (13 to 15 %) to predict neurodevelopmental outcomes at 12 months CA. At 3 months CA, the sensitivity was low (30 to 45 %) and specificity was high (74 to 76 %) for all outcomes.

4. Discussion

Preterm infants assessed at TEA have poorer NVA scores compared to their term-born peers assessed shortly after birth. No predictive relationship between the NVA and neurodevelopmental outcomes at 12 months CA was identified in this cohort of very preterm infants. Our findings of significantly poorer visual assessment scores in preterm infants compared to term-born infants contrast with previous studies, which suggested that visual function in low-risk preterm infants is similar to, if not better than, term-born infants assessed at the same post-term age [9,33–36]. Previous studies have proposed maturation of the cortically-mediated aspects of visual development is dependent on the post-menstrual age at assessment rather than the length of extrauterine life, and subcortical aspects of visual function are accelerated by premature exposure to the extrauterine environment and early visuomotor experiences [8,9]. Diagnostic accuracy of the NVA was lower than that of a previous study [10], which reported a sensitivity of 92 % and specificity 74 % to predict neurodevelopmental outcomes at 12 months CA.

Differences between the current cohort and that of previous studies may be the reason for the variances observed. Our preterm cohort included younger infants (range 23.1 to 30.9 weeks GA) compared to earlier cohorts (ranges 25 to 30.9 weeks GA [8] and 25 to 30 weeks GA [9]). This may account for our cohort's greater immaturity in NVA responses at TEA and poorer correlation with later outcomes. Our current preterm cohort ($n = 248$) had a greater proportion of infants with retinopathy of prematurity (ROP) stage 2 (41 %) during the neonatal course compared to earlier studies which reported 28 % [8]. The number of infants with ROP has increased over the past decade due to enhanced imaging systems and treatment modalities [37]. These advancements have also resulted in fewer impacts of ROP on functional vision over time [37]. Our preterm cohort also had a significantly higher social risk (52 %) compared to the term-born cohort (15 %). As social risk factors such as maternal age and marital status have been associated with higher rates of preterm birth [38], this may explain preterm infants' lower NVA scores at TEA.

The difference in diagnostic accuracy compared to earlier studies may also be explained by the difference in prevalence of brain injuries between the cohorts. Our cohort had 9 % IVH Grade III or IV compared with 4.8 % in Ricci's cohort [10]. Cystic periventricular leukomalacia (PVL) is considered a major cause of visual impairment in children born preterm [39,40] and children with PVL are at greater risk of also being diagnosed with CP compared to those without PVL [39]. Our cohort had a small proportion of infants with PVL ($n = 8$, 3 %), which is similar to normative data for Australian and New Zealand preterm infants ($n = 33$, born <32 weeks of gestation; 1.2 %) [40]. These results contrast to the previous study investigating the diagnostic accuracy of the NVA [10], in which a significantly larger proportion of their preterm cohort had PVL ($n = 20$, 13.8 %). The previously published cohort also had a high number of infants diagnosed with CP ($n = 23$, 15.8 %) [10]. This data suggests that in a contemporary cohort of preterm infants with lower rates of major brain injuries the associations between visual scores and neurodevelopmental outcomes at 12 months CA may be less strong than in cohorts with higher rates of brain injuries.

Methodological differences may account for the lower diagnostic accuracy of the NVA compared to the previously published study [10]. The same nine test items of the NVA were performed in our study at TEA and three months CA for preterm infants. The earlier study completed these same nine NVA test items at TEA but performed a slightly different behavioural visual assessment which included assessment of only four items (ability to fix and follow, acuity, attention at distance and visual fields) at three and 12 months CA [10]. This previously published study from 2011 also used the Griffiths Mental Development Scales as the main outcome assessment, which has been shown to more accurately estimate neurodevelopmental impairment compared to the Bayley-III [41]. The Griffiths Developmental Quotient includes a motor component and their study had a higher proportion of infants with CP [41] than the current study. An experienced clinician (AG) who was taught by the original NVA research group trained one of our assessors (PC) to conduct the assessment. This newly trained assessor then taught the other examiners how to complete the NVA. Subtle variations in the administration of the NVA and interpretation of scores may account for the differences in findings.

When performed at TEA, the NVA had higher sensitivities but lower specificities for all outcomes compared to when it was performed at three months CA. Generally, a balance of sensitivity and specificity is recommended in order to identify the infants who are at risk of neurodevelopmental impairments and require early intervention, and to prevent the misallocation of resources to infants without impairments [11]. At TEA, the NVA had a high rate of false positives where a significant proportion of infants who may not have a neurodevelopmental impairment tested positive, making them subject to further testing and early intervention when it may not have been required if this test was to be used in clinical practice. At three months CA, there was a high rate of false negatives where many preterm infants who may have had impairments tested negative on the NVA and may not have been referred to early intervention when it may in fact have been appropriate. Area under the ROC curve analyses also suggested that higher cut-off points at TEA produced better combinations of sensitivity and specificity, as well as more balanced sensitivity and specificity. Using these higher cut-off points would identify a large number of infants who are at risk of adverse outcomes but is likely not sensitive enough to identify the preterm infants who may have mild developmental impairments. The authors acknowledge that the Bayley-III assessment at 12 months CA may not identify all who may progress to having an impairment and so these results need to be interpreted with caution.

Our study had a considerably larger number of preterm infants who scored one SD below the mean on the Bayley-III motor subscale ($n = 44$, 22 %) than the Bayley-III cognitive subscale ($n = 19$, 10 %, see Table 2). Previous studies have demonstrated that the Bayley-III underestimates developmental delay compared to the Bayley Scales of Infant Development, Second Edition [21,42–45] and differences between test scores are most obvious with respect to motor performance [46]. In typically-developing Australian cohorts, 17 % of infants are expected to score one SD below the mean [18], and 22 % (n = 44) of the current preterm cohort scored below this on the Bayley-III motor subscale. The means and SDs obtained by our cohort on the Bayley-III were consistent with those of other prospective studies of preterm infants [46,47].

Strengths of this study were the large, contemporary sample of very preterm infants, the inclusion of a term-born cohort, and evaluating the NVA's sensitivity and specificity with three standardised and clinically useful outcome measures. A limitation of this study was not collecting NVA results for term control infants at three months CA which could have allowed for more detailed analysis of visual system maturation compared to preterm infants. Another limitation was only assessing neurodevelopmental outcomes at one time point (12 months CA). A Bayley-III assessment at 12 months CA does not exclude the possibility of a developmental impairment as there may be a fall in neurodevelopmental abilities as measured by the Bayley-III or other assessments at two years or beyond. As a single assessor completed the outcome assessments for all of the participants at 12 months CA, it is possible that findings from one assessment may have influenced the others. The research team is currently collecting data for six-year outcomes in these cohorts, which will be used to further evaluate the true predictive validity of the NVA into childhood [48]. Future studies investigating the associations between key risk factors of preterm birth and neurodevelopmental outcomes may also allow for earlier detection and treatment of developmental delays.

5. Conclusion

Preterm infants born <31 weeks GA had significantly lower visual assessment scores at TEA compared to term-born infants. When performed at TEA and three months CA, the NVA was not a strong predictor of motor and cognitive impairments at 12 months CA in this contemporary cohort of very preterm infants. Despite its clinical utility and ease of administration, the NVA performed at TEA is likely to overestimate the number of children who will have adverse neurodevelopmental outcomes at 12 months CA and underestimate those who will have impairments when performed at three months CA. Additional assessments performed in the neonatal period should be used to identify very preterm infants who have or are at risk of acquiring neurodevelopmental impairments.

Funding

This study was supported by grants from the Cerebral Palsy Alliance Research Foundation (IRG1413), the Financial Markets Foundation for Children (2014-074), the Queensland Government (Smart State; Health Practitioner Research Stimulus Grant) and the National Health and Medical Research Council of Australia (NHMRC1084032). The authors received funding from UQ (UQ Research Scholarship [JMG]), and the National Health and Medical Research council (Research Fellowship NHMRC1105038 [RNB]), Investigator Grant NHMRC1195602 [RNB], an NHMRC Clinical CRE (NHMRC1116442 CRE Australasian CP Clinical Trials Network [RNB]) and from the Italian Ministry of Health (RC Linea 1 RC-2024 and 5 X 1000 Health Research voluntary contributions [AG]). No funder was involved in the study design, data collection, data analysis, manuscript preparation and/or publication decisions.

CRediT authorship contribution statement

Jessica W. Blazek: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis. **Paul B. Colditz:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Andrea Guzzetta:** Writing – review & editing, Methodology, Formal analysis. **Robert S. Ware:** Writing – review & editing, Methodology, Formal analysis. **Mark D. Chatfield:** Writing – review & editing, Software. **Judith L. Hough:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Formal analysis. **Roslyn N. Boyd:** Writing – review & editing, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Joanne M. George:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

None declared.

Acknowledgements

We would like to thank the infants and their families for their participation in this study, the research nurses who undertook recruitment and study management, Christine Finn for completing the 12-month clinical assessments and Dr. Marcella Danks for imparting her knowledge about this research topic.

Appendix A. Sensitivity analysis using our term cohort data to generate our own optimality score

Background and Aim: In light of the Neonatal Visual Assessment (NVA) displaying poor prediction of neurodevelopmental outcomes at 12 months corrected age (CA) in our cohort of very preterm infants, we aimed to determine if generating an optimality score using our own term control

data would improve the NVA's predictive accuracy.

Methods: NVA scores for our preterm cohort were converted to a zero to nine scale based on the optimality concept in a similar manner to Ricci and colleagues [10]. For each NVA test item, an infant who performed within the 90th centile of our term cohort data was scored zero, and one if their performance fell <10th centile resulting in a global NVA score from zero to nine. Fig. 1.A displays NVA proforma published by Ricci [10]; Fig. 2.A displays converted NVA proforma based on our term control data. Sensitivity and specificity of the NVA performed at TEA and three months CA were calculated based on our optimality score to predict neurodevelopmental outcomes at 12 months CA (Bayley Scales of Infant and Toddler Development, Third Edition; Neurosensory Motor Developmental Assessment; Alberta Infant Motor Scale).

NEONATAL VISUAL ASSESSMENT *(Mod. from Ricci et al. 2007)*

Baby's name ..... Today's date ..... / ..... / .....
(day) (month) (year)

		score 3	score 2	score 1	score 0
Spontaneous ocular motility *Note spontaneous ocular movements before presenting a target*		Mainly conjugate	Occasional strabismus nystagmus R L R L	Intermittent strabismus nystagmus R L R L	Continuous strabismus nystagmus R L R L
Ocular movements with a target *Note spontaneous ocular movements while presenting a target*		Mainly conjugate	Occasional strabismus nystagmus R L R L	Intermittent strabismus nystagmus R L R L	Continuous strabismus nystagmus R L R L
Fixation *With the target in front of the infant at 25 cm, note the ability of the infant to fix on the target*		Stable (>3 sec)		Unstable (<3 sec)	Absent
Horizontal tracking *With the target at 25 cm and starting in the midline move it slowly to both left and right. Note the infant's eye movement*		Complete R L	Incomplete R L	Brief R L	Absent
Vertical tracking *With the target at 25 cm and starting in the midline move it slowly upwards and downwards. Note the infant's eye movement*		Complete U D	Incomplete U D	Brief U D	Absent
Arc tracking *With the target at 25 cm move it slowly tracing an arc. Note the infant's eye movement*		Complete	Incomplete	Brief	Absent
Tracking colored stimulus *Note the infant's eye movements in response to the target movements, starting from the midline towards lateral*		Present			Absent
Stripes discrimination *Note the infant's ability to fixate on a series of targets of decreasing stripe widths held at a distance of 38 cm starting with the widest stripe. Note the narrowest stripe width on which the infant fixates*		3.2 c/g	2.4 c/g 1.6 c/g	1.3 c/g 0.86 c/g	0.64 c/g 0.43 c/g 0.32 c/g
Attention at distance *After eliciting central fixation move the target slowly away and a few cm laterally from the infant. Note the maximal distance at which focusing is maintained*		>70 cm Cm	50-69 cm Cm	30-49 cm Cm	<30 cm Cm .....

Fig. 1.A. Original Neonatal Visual Assessment scoring sheet published Ricci et al. [10]. For each NVA test item, an infant who performed within the 90th centile of the term cohort was scored zero (coloured red in the Figure above) and scored one (coloured white in the Figure above) if their performance fell outside the 90th centile. A global NVA score from zero to nine was recorded, where scores of zero and one were considered normal and scores of greater than one were considered abnormal [10].

NEONATAL VISUAL ASSESSMENT *(Mod. from Ricci et al. 2007)*

Baby's name ..... Today's date ..... / ..... / .....
(day) (month) (year)

		score 3	score 2	score 1	score 0
Spontaneous ocular motility *Note spontaneous ocular movements before presenting a target*		Mainly conjugate	Occasional strabismus nystagmus R L R L	Intermittent strabismus nystagmus R L R L	Continuous strabismus nystagmus R L R L
Ocular movements with a target *Note spontaneous ocular movements while presenting a target*		Mainly conjugate	Occasional strabismus nystagmus R L R L	Intermittent strabismus nystagmus R L R L	Continuous strabismus nystagmus R L R L
Fixation *With the target in front of the infant at 25 cm, note the ability of the infant to fix on the target*		Stable (>3 sec)		Unstable (<3 sec)	Absent
Horizontal tracking *With the target at 25 cm and starting in the midline move it slowly to both left and right. Note the infant's eye movement*		Complete R L	Incomplete R L	Brief R L	Absent
Vertical tracking *With the target at 25 cm and starting in the midline move it slowly upwards and downwards. Note the infant's eye movement*		Complete U D	Incomplete U D	Brief U D	Absent
Arc tracking *With the target at 25 cm move it slowly tracing an arc. Note the infant's eye movement*		Complete	Incomplete	Brief	Absent
Tracking colored stimulus *Note the infant's eye movements in response to the target movements, starting from the midline towards lateral*		Present			Absent
Stripes discrimination *Note the infant's ability to fixate on a series of targets of decreasing stripe widths held at a distance of 38 cm starting with the widest stripe. Note the narrowest stripe width on which the infant fixates*		3.2 c/g	2.4 c/g 1.6 c/g	1.3 c/g 0.86 c/g	0.64 c/g 0.43 c/g 0.32 c/g
Attention at distance *After eliciting central fixation move the target slowly away and a few cm laterally from the infant. Note the maximal distance at which focusing is maintained*		>70 cm Cm	50-69 cm Cm	30-49 cm Cm	<30 cm Cm .....

Fig. 2.A. Neonatal Visual Assessment scoring sheet based on the optimality score concept and using our term control cohort to derive the 90th and 10th percentile categories. For each NVA test item, an infant who performed within the 90th centile of our term cohort was scored zero (coloured red in the Figure above) and scored one (coloured white in the Figure above) if their performance fell outside the 90th centile. A global NVA score from zero to nine was recorded, where scores of zero and one were considered normal and scores of greater than one were considered abnormal.

Results: Using our cohort's optimality score, the sensitivity of the NVA at TEA improved from 59–78 % to 94–100 % for all outcomes but specificity worsened from 25–27 % to 13–15 %. At three months CA, the sensitivity improved from 21–28 % to 30–45 % for all outcomes and specificity worsened from 86–87 % to 74–76 %.

Conclusion: Results of our sensitivity analysis improved the NVA's sensitivity and slightly reduced specificity compared to our previous results. However the NVA was still not a strong predictor of motor and cognitive impairments at 12 months CA.

Appendix B

Table B.1
Characteristics of the preterm study samples who had the Neonatal Visual Assessment performed with and without an outcome at 12 months CA. Data are presented as mean (SD) or median (IQR) for continuous measures and frequency (percentage) for categorical measures.

	Preterm infants with NVA and at least one outcome performed n = 199	Preterm infants with NVA but no outcome performed n = 49	p-value[†]
Birth and Maternal Data			
Gestational age at birth (weeks), median (IQR)	28.4 (26.9 to 29.9)	29.3 (28.1 to 30.3)	**<0.05**
Birth weight (g), mean (SD)	1102.2 (306.2)	1206.3 (304.1)	**<0.05**
Birth head circumference (cm), mean (SD)	25.9 (2.3), n = 193	26.4 (2.2)	0.16
Males	110 (55 %)	24 (49 %)	0.53
Multiple births	62 (31 %)	13 (27 %)	0.51
Premature rupture of membranes	47 (24 %)	10 (20 %)	0.53
Caesarean section	137 (69 %)	31 (63 %)	0.42
Chorioamnionitis	33 (17 %), n = 198	6 (12 %)	**<0.001**
Antenatal corticosteroids	139 (70 %)	33 (67 %)	0.68
Magnesium sulphate	102 (58 %), n = 175	27 (68 %), n = 40	0.24
Higher social risk[‡]	96 (48 %)	26 (74 %), n = 35	**<0.05**
Acquired Medical Factors			
Patent ductus arteriosus	75 (38 %)	14 (29 %)	0.25
Any IVH	63 (32 %)	9 (18 %)	0.06
IVH grade III or IV	21 (11 %)	1 (2 %)	**<0.05**
Periventricular leukomalacia[§]	8 (4 %), n = 196	0 (0 %), n = 45	0.18
Hydrocephalus	9 (5 %), n = 196	1 (2 %), n = 45	0.37
Seizures treated with anticonvulsant therapy	1 (1 %), n = 196	0 (0 %), n = 45	0.34
NEC diagnosed or suspected	10 (5 %), n = 196	1 (2 %), n = 45	0.39
Confirmed sepsis	9 (5 %), n = 196	2 (4 %), n = 45	0.77
Total parenteral nutrition (days), median (IQR)	11 (8 to 14), n = 196	10 (7.5 to 14), n = 44	0.78
Postnatal corticosteroids	30 (15 %), n = 196	5 (11 %), n = 45	0.49
Ventilation (days), median (IQR)	2 (0 to 7.5), n = 196	0.5 (0 to 6), n = 44	0.10
CPAP (days), median (IQR)	30 (7 to 48), n = 197	13 (3 to 36), n = 46	**<0.05**
Oxygen therapy (hours), median (IQR)	112 (5 to 840), n = 186	20 (1 to 319), n = 45	0.18
Bronchopulmonary dysplasia[¶]	90 (46 %), n = 196	12 (27 %), n = 45	**<0.05**
Presence of ROP	87 (44 %), n = 196	14 (32 %), n = 44	0.15
ROP stage 1	45 (52 %), n = 87	8 (57 %), n = 14	
ROP stage 2	35 (40 %), n = 87	6 (43 %), n = 14	
ROP stage 3	7 (8 %), n = 87	0 (0 %), n = 14	
ROP zone 1	1 (1 %), n = 87	0 (0 %), n = 14	
ROP zone 2	43 (50 %), n = 87	11 (79 %), n = 14	
ROP zone 3	43 (50 %), n = 87	3 (21 %), n = 14	
PMA (weeks) at TEA clinical assessment, median (IQR)	40.7 (40.1 to 41.4), n = 196	40.3 (40.0 to 41.7), n = 45	0.93
CA (weeks) at 3-month clinical assessment, median (IQR)	13.4 (13.1 to 13.8), n = 148	13.1 (12.5 to 13.4), n = 36	0.26

SD standard deviation; IQR interquartile range; IVH intraventricular haemorrhage; NEC necrotizing enterocolitis; CPAP continuous positive airway pressure; ROP retinopathy of prematurity; CA corrected age; PMA postmenstrual age; TEA term equivalent age.

[†] Independent samples t-test, significance set at $p < 0.05$.

[‡] Social risk was assessed using a questionnaire examining 6 areas with each scoring 0–2 for a maximum social score of 12: including family structure, maternal age, education level, occupation, employment status and language spoken in the home. Scores of 2 and above were classified as indicating higher social risk.

[§] n = 6 had cystic PVL, n = 7 had PVL in the motor pathway, n = 5 also had IVH grade III or IV.

[¶] Defined as oxygen requirement at 36 weeks postmenstrual age.

References

[1] Australian Institute of Health and Welfare, Australia's Mothers and Babies 2016 - in Brief, Australian Institute of Health and Welfare, Canberra, 2018.

[2] J. Atkinson, O. Braddick, S. Anker, et al., Cortical vision, MRI and developmental outcome in preterm infants, Arch. Dis. Child Fetal Neonatal Ed. 93 (4) (2008) 292–297, https://doi.org/10.1136/adc.2007.116988.

[3] D.M. Romeo, D. Ricci, F. Serrao, et al., Visual function assessment in late-preterm newborns, Early Hum. Dev. 88 (5) (2012) 301–305, https://doi.org/10.1016/j.earlhumdev.2011.08.024.

[4] J. Atkinson, O. Braddick, Visual attention in the first years: typical development and developmental disorders, Developmental Medicine & Child Neurology. 54 (7) (2012) 589–595, https://doi.org/10.1111/j.1469-8749.2012.04294.x.

[5] M. Hadders-Algra, Challenges and limitations in early intervention, Developmental Medicine & Child Neurology. 53 (s4) (2011) 52–55, https://doi.org/10.1111/j.1469-8749.2011.04064.x.

[6] D. Ricci, D.M. Romeo, F. Serrao, et al., Early assessment of visual function in preterm infants: how early is early? Early Hum. Dev. 86 (1) (2010) 29–33, https://doi.org/10.1016/j.earlhumdev.2009.11.004.

[7] D. Ricci, L. Cesarini, M. Groppo, et al., Early assessment of visual function in full term newborns, Early Hum. Dev. 84 (2) (2008) 107–113, https://doi.org/10.1016/j.earlhumdev.2007.03.010.

[8] D. Ricci, L. Cesarini, D.M.M. Romeo, et al., Visual function at 35 and 40 weeks' postmenstrual age in low-risk preterm infants, Pediatrics 122 (6) (2008) e1193, https://doi.org/10.1542/peds.2008-1888.

[9] D. Ricci, L. Cesarini, F. Gallini, et al., Cortical visual function in preterm infants in the first year, J. Pediatr. 156 (4) (2010) 550–555, https://doi.org/10.1016/j.jpeds.2009.10.042.

[10] D. Ricci, D.M. Romeo, F. Gallini, 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 (3) (2011) 177–182, https://doi.org/10.1016/j.earlhumdev.2010.12.003.

[11] A.J. Spittle, K.J. Lee, M. Spencer-Smith, L.E. Lorefice, P.J. Anderson, L.W. Doyle, Accuracy of two motor assessments during the first year of life in preterm infants for predicting motor outcome at preschool age, PloS One 10 (5) (2015) e0125854, https://doi.org/10.1371/journal.pone.0125854.

[12] J.M. George, R.N. Boyd, P.B. Colditz, et al., PPREMO: a prospective cohort study of preterm infant brain structure and function to predict neurodevelopmental outcome, BMC Pediatr. 15 (2015) 123, https://doi.org/10.1186/s12887-015-0439-z.

[13] M.M. Lai, G. D'Acunto, A. Guzzetta, et al., PREMM: preterm early massage by the mother: protocol of a randomised controlled trial of massage therapy in very preterm infants, BMC Pediatr. 16 (1) (2016) 146, https://doi.org/10.1186/s12887-016-0678-7.

[14] M. Hack, N. Breslau, D. Aram, B. Weissman, N. Klein, E. Borawski-Clark, The effect of very low birth weight and social risk on neurocognitive abilities at school age, J. Dev. Behav. Pediatr. 13 (6) (1992) 412–420, https://doi.org/10.1097/00004703-199212000-00005.

[15] M. Hack, A.A. Fanaroff, Outcomes of children of extremely low birthweight and gestational age in the 1990s, Semin. Neonatol. 5 (2) (2000) 89–106, https://doi.org/10.1053/siny.1999.0001.

[16] G. Roberts, K. Howard, A.J. Spittle, N.C. Brown, P.J. Anderson, L.W. Doyle, Rates of early intervention services in very preterm children with developmental disabilities at age 2 years, J. Paediatr. Child Health 44 (5) (2008) 276–280, https://doi.org/10.1111/j.1440-1754.2007.01251.x.

[17] J.W. Van Hus, M. Jeukens-Visser, K. Koldewijn, et al., Comparing two motor assessment tools to evaluate neurobehavioral intervention effects in infants with very low birth weight at 1 year, Phys. Ther. 93 (11) (2013) 1475–1483, https://doi.org/10.2522/ptj.20120460.

[18] C. Albers, A. Grieve, Test Review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development Third Edition. San Antonio, TX: Harcourt Assessment, Journal of Psychoeducational Assessment 25 (2007) 180–190, https://doi.org/10.1177/0734282906297199.

[19] S. Ranjitkar, I. Kvestad, T.A. Strand, et al., Acceptability and reliability of the Bayley scales of infant and toddler development-III among children in Bhaktapur, Nepal. Frontiers in Psychology. 9 (1265) (2018), https://doi.org/10.3389/fpsyg.2018.01265.

[20] D.E. Ballot, T. Ramdin, D. Rakotsoane, et al., Use of the Bayley scales of infant and toddler development, third edition, to assess developmental outcome in infants and young children in an urban setting in South Africa, International Scholarly Research Notices. 2017 (2017) 1631760, https://doi.org/10.1155/2017/1631760.

[21] Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW, Group TVIC, Underestimation of developmental delay by the new Bayley-III scale, Arch. Pediatr. Adolesc. Med. 164 (4) (2010) 352–356, https://doi.org/10.1001/archpediatrics.2010.20.

[22] J. Månsson, K. Stjernqvist, F. Serenius, U. Ådén, K. Källén, Agreement between Bayley-III measurements and WISC-IV measurements in typically developing children, J. Psychoeduc. Assess. 37 (5) (2018) 603–616.

[23] J. Månsson, K. Källén, E. Eklöf, F. Serenius, U. Ådén, K. Stjernqvist, The ability of Bayley-III scores to predict later intelligence in children born extremely preterm, Acta Paediatr. 110 (11) (2021) 3030–3039.

[24] A.J. Spittle, M.M. Spencer-Smith, A.L. Eeles, K.J. Lee, L.E. Lorefice, P.J. Anderson, et al., Does the Bayley-III motor scale at 2 years predict motor outcome at 4 years in very preterm children? Developmental Medicine & Child Neurology. 55 (5) (2012) 448–452.

[25] M. Danks, M.F. Maideen, Y.R. Burns, et al., The long-term predictive validity of early motor development in "apparently normal" ELBW survivors, Early Hum. Dev. 88 (8) (2012) 637–641, https://doi.org/10.1016/j.earlhumdev.2012.01.010.

[26] Y.R. Burns, R.M. Ensbey, M.A. Norrie, The neuro-sensory motor developmental assessment part 1: development and administration of the test, Aust. J. Physiother. 35 (3) (1989) 141–149, https://doi.org/10.1016/s0004-9514(14)60503-1.

[27] Y.R. Burns, R.M. Ensbey, M.A. Norrie, The neuro-sensory motor developmental assessment part II: predictive and concurrent validity, Aust. J. Physiother. 35 (3) (1989) 151–157, https://doi.org/10.1016/s0004-9514(14)60504-3.

[28] J. Darrah, M. Piper, M.J. Watt, Assessment of gross motor skills of at-risk infants: predictive validity of the Alberta infant motor scale, Developmental Medicine & Child Neurology. 40 (7) (2008 Nov 12) 485–491.

[29] S.K. Campbell, T.H. Kolobe, B.D. Wright, J.M. Linacre, Validity of the test of infant motor performance for prediction of 6-, 9- and 12-month scores on the Alberta infant motor scale, Dev. Med. Child Neurol. 44 (4) (2002) 263–272, https://doi.org/10.1017/s0012162201002043.

[30] J. Darrah, D. Bartlett, Maguire TO, W.R. Avison, T. Lacaze-Masmonteil, Have infant gross motor abilities changed in 20 years? A re-evaluation of the Alberta infant motor scale normative values, Dev. Med. Child Neurol. 56 (9) (2014) 877–881, https://doi.org/10.1111/dmcn.12452.

[31] C. Peyton, M.D. Schreiber, M.E. Msall, The test of infant motor performance at 3 months predicts language, cognitive, and motor outcomes in infants born preterm at 2 years of age, Developmental Medicine & Child Neurology. 60 (12) (2018) 1239–1243, https://doi.org/10.1111/dmcn.13736.

[32] O. Kaya-Kara, M. Kerem Günel, S. Yigit, Correlation of the Bayley Scales of Infant-Toddler Development - 3rd Edition and Neuro-Sensory Motor Assessment in preterm infants during the first year of life, The Turkish Journal of Pediatrics 61 (2019) 399–406, https://doi.org/10.24953/turkjped.2019.03.012.

[33] A.R. O'Connor, A.R. Fielder, Visual outcomes and perinatal adversity, Semin. Fetal Neonatal Med. 12 (5) (2007) 408–414, https://doi.org/10.1016/j.siny.2007.07.001.

[34] A.R. O'Connor, C.M. Wilson, A.R. Fielder, Ophthalmological problems associated with preterm birth, Eye (Lond.) 21 (10) (2007) 1254–1260, https://doi.org/10.1038/sj.eye.6702838.

[35] J. van Hof-van Duin, D.J. Heersema, F. Groenendaal, W. Baerts, W.P.F. Fetter, Visual field and grating acuity development in low-risk preterm infants during the first 2 1/2 years after term, Behav. Brain Res. 49 (1) (1992) 115–122, https://doi.org/10.1016/s0166-4328(05)80201-3.

[36] E. Mercuri, A. Guzzetta, S. Laroche, D. Ricci, I. van Haastert, A. Simpson, et al., Neurologic examination of preterm infants at term age: comparison with term infants, J. Pediatr. 142 (6) (2003) 647–655, https://doi.org/10.1067/mpd.2003.215.

[37] A.S.H. Tsai, L. Acaba-Berrocal, M. Sobhy, E. Cole, S. Ostmo, K. Jonas, et al., Current Management of Retinopathy of prematurity, Curr. Treat. Options Pediatr. 8 (3) (2022 Jul 12) 246–261.

[38] K. Beeckman, S. van De Putte, K. Putman, F. Louckx, Predictive social factors in relation to preterm birth in a metropolitan region, Acta Obstet. Gynecol. Scand. 88 (7) (2009) 787–792, https://doi.org/10.1080/00016340902974007.

[39] J.W. Gotardo, NdFV Volkmer, G.P. Stangler, A.D. Dornelles, BbdA Bohrer, C. G. Carvalho, Impact of peri-intraventricular haemorrhage and periventricular leukomalacia in the neurodevelopment of preterms: a systematic review and meta-analysis, PloS One 14 (10) (2019) e0223427, https://doi.org/10.1371/journal.pone.0223427.

[40] S.S.W. Chow, P. Creighton, G.M. Chambers, K. Lui, Report of the Australian and New Zealand Neonatal Network 2019, Australian and New Zealand Neonatal Network, Sydney, 2021.

[41] O. Picciolini, C. Squarza, C. Fontana, M.L. Giannì, I. Cortinovis, S. Gangi, et al., Neurodevelopmental outcome of extremely low birth weight infants at 24 months corrected age: a comparison between Griffiths and Bayley scales, BMC Pediatr. 15 (1) (2015 Sep 30).

[42] J.R. Lowe, S.J. Erickson, R. Schrader, A.F. Duncan, Comparison of the Bayley II mental developmental index and the Bayley III cognitive scale: are we measuring the same thing? Acta Paediatr. 101 (2) (2012) e55–e58, https://doi.org/10.1111/j.1651-2227.2011.02517.x.

[43] T. Moore, S. Johnson, S. Haider, E. Hennessy, N. Marlow, Relationship between test scores using the second and third editions of the Bayley scales in extremely preterm children, J. Pediatr. 160 (4) (2012) 553–558, https://doi.org/10.1016/j.jpeds.2011.09.047.

[44] M.E. Msall, The Bayley-III scale underestimates developmental delay in extremely premature and extremely low birth weight infants, J. Pediatr. 157 (5) (2010) 863–864, https://doi.org/10.1016/j.jpeds.2010.09.022.

[45] G. Reuner, A.C. Fields, A. Wittke, M. Löpprich, J. Pietz, Comparison of the developmental tests Bayley-III and Bayley-II in 7-month-old infants born preterm, Eur. J. Pediatr. 172 (3) (2013) 393–400, https://doi.org/10.1007/s00431-012-1902-6.

[46] F. Veiga de Góes, M.D.B.B. Méio, R. Reis de Mello, D. Morsch, Evaluation of neurodevelopment of preterm infants using Bayley III scale, Revista Brasileira de Saúde Materno Infantil. 15 (1) (2015) 47–55, https://doi.org/10.1590/S1519-38292015000100004.

[47] M.M. Greene, K. Patra, M.N. Nelson, J.M. Silvestri, Evaluating preterm infants with the Bayley-III: patterns and correlates of development, Res. Dev. Disabil. 33 (6) (2012) 1948–1956, https://doi.org/10.1016/j.ridd.2012.05.024.

[48] J.M. George, A.M. Pagnozzi, S. Bora, et al., 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 (5) (2020) e036480, https://doi.org/10.1136/bmjopen-2019-036480.