Performance of the Parent Report of Children's Abilities–Revised (PARCA-R) versus the Bayley Scales of Infant Development III

Andrew J Martin,[1] Brian A Darlow,[2] Alison Salt,[3] Wendy Hague,[1] Lucille Sebastian,[1] Nicki McNeill,[2] William Tarnow-Mordi[1,4]

[1]NHMRC Clinical Trials Centre, University of Sydney, Sydney, Australia
[2]Department of Paediatrics, University of Otago, Christchurch, New Zealand
[3]Neurodisability Service, Great Ormond Street Hospital for Children, Institute of Child Health, London, UK
[4]Westmead International Network for Neonatal Education and Research, Westmead Hospital, Westmead, Australia

Correspondence to
Andrew J Martin, National Health and Medical Research Council Clinical Trials Centre, University of Sydney, Sydney, NSW 1450, Australia; andrew.martin@ctc.usyd.edu.au

ABSTRACT

Background The Parent Report of Children's Abilities-Revised (PARCA-R) assesses cognitive and language development at 24 months. It was validated against the Mental Development Index of the Bayley Scales of Infant Development II (BSID II), but this has now been superseded by BSID III.

Objective To compare the PARCA-R against the BSID III.

Methods PARCA-R and BSID III assessments scheduled at 24 months of age (corrected for prematurity) were completed in 204 infants with suspected or proven neonatal sepsis in the International Neonatal Immunotherapy Study. Associations between the scales were measured and the predictive accuracy of the PARCA-R for moderate cognitive delay and moderate language delay was assessed using Receiver Operating Characteristic (ROC) analysis.

Results Median birthweight was 911 g, median gestational age at birth was 27 weeks and 100 (49.0%) were girls. 4.4% and 8.4% met standard BSID III criteria for cognitive delay and language delay, respectively. These rates increased to 19.6% and 12.6% when an independent sample of normal term infants were used as the reference group suggesting standard BSID III reference norms may tend to underestimate delay. The Spearman correlation between PARCA-R and BSID scales were 0.43 for cognition and 0.71 for language. The PARCA-R successfully predicted cases of cognitive delay and language delay with the area under the ROC curves ranging from 0.83 to 0.97 depending on reference norms used.

Conclusions The results support the PARCA-R as a practical tool for the identification of appreciable cognitive and language delay at 24 months among critically ill premature and extremely low birthweight neonates.

What is already known on this subject

- The Parent Report of Children's Abilities-Revised (PARCA-R) is a parent-administered developmental questionnaire that is a practical and cheaper alternative to tests administered by trained assessors for outcome assessment in neonatal trials.
- The PARCA-R was validated against the Bayley Scales of Infant Development II (BSID II), but this has now been superseded by BSID III.
- The aim was to compare the PARCA-R as an indicator of cognitive and language development against the latest version of the BSID (BSID III)

What this study adds

- A sample of 204 infants previously treated for suspected or proven neonatal sepsis was assessed with the The Parent Report of Children's Abilities-Revised (PARCA-R) and Bayley Scales of Infant Development III (BSID III).
- The PARCA-R language and cognitive scales correlated with those from the BSID III, and were able to predict cases of appreciable delay accurately.
- The results support the use of the PARCA-R for the assessment of 24 month outcomes in neonatal clinical trials.

BACKGROUND

Parent-completed assessments of children's abilities represent a more practical and cost-effective method for collecting longer-term outcome data in neonatal clinical trials,[1] as well as in the monitoring of high-risk populations (eg, preterm infants),[2] than formal neurodevelopmental appraisals performed by trained assessors. The Parent Report of Children's Abilities–Revised (PARCA–R) is an example of a parent-completed questionnaire developed to evaluate language and cognitive ability at 24 months[3 4] that has been successfully used in a number of neonatal clinical trials[5 6] and has been previously validated against the Mental Development Index of the Bayley Scales of Infant Development II (BSID II).[1 3 4] The BSID II was a widely used and well regarded developmental assessment designed to be administered by a trained examiner; however, it has recently been superseded by a new version (BSID III)[7 8] that has scales that are not directly comparable with its predecessor. For example, the BSID III produces separate language and cognitive skills scores, whereas the BSID II combined these skills into the Mental Development Index. The aim of the present study was to compare the PARCA-R as an indicator of development in high-risk infants against the latest version of the BSID (BSID III), given that instrument's potential to gain acceptance as a new criterion approach to developmental assessment.

Original article

METHODS

The children in this study comprised a sample of participants in the International Neonatal Immunotherapy Study (INIS). INIS was a double-blind, randomised, controlled trial of human pooled intravenous immunoglobulin versus placebo as adjunctive treatment for proven or suspected neonatal sepsis.[6] The 9 countries (UK, Australia, Argentina, New Zealand, Serbia, Greece, Denmark, Belgium, Ireland) participating in INIS enrolled 3493 babies with a median birthweight of 1000 g and median gestational age of 28 weeks.

The primary outcome for INIS was death or major disability at 2 years, adjusted for gestational age and assessed by a parent questionnaire that included the PARCA-R. A sample of 858 infants at selected sites in Australia and New Zealand were also planned to be assessed using the BSID, of which 690 (80%) did complete a BSID assessment and 747 (87%) did complete a PARCA-R assessment. The higher completion rate for the PARCA-R suggests there were no particular problems with parent acceptability, and the feasibility of administration by phone to parents, in the few cases where necessary, underscores the instrument's broad applicability to parents of different abilities.

During the INIS follow-up period, the BSID III replaced the BSID II as the standard neurodevelopmental assessment tool used in a number of participating hospitals. Of the 690 infants assessed with the BSID, 484 received the BSID II and 206 received the BSID III. An assessment with PARCA-R was completed in 204/206 (99%) infants that received the BSID III.

The parent questionnaire (comprising the PARCA-R) was mailed to parents for completion approximately 4 weeks before the child reached 24 months of age (corrected for any prematurity).The BSID III was administered by a certified psychologist, or other trained assessor, at the time of the scheduled 24 month INIS follow-up visit.

The PARCA–R assesses non-verbal cognition using 34 items and linguistic skill using a 100-word checklist assessing vocabulary plus 18 items assessing sentence construction complexity. Scored responses to items within each domain are summed to produce a non-verbal cognition scale and a linguistic skill scale. An item that is left blank does not contribute to the scale score (ie, a blank is interpreted as a negative response). These scales may be combined to form an overall Parent Report Composite score with a range of 0–158.

The BSID III items fall into the developmental areas of cognition, language and motor skills. BSID III can also be used to derive scores for social-emotional as well as adaptive behaviour. For each developmental aspect, the BSID III yields a score that is age-adjusted and standardised against a normative population to have a mean of 100 and an SD of 15. Data obtained from the cognition and language scales from the BSID III were included in the current analysis.

The strength of association between the PARCA-R and BSID III cognitive and language scales were quantified using Spearman correlation coefficients. A receiver-operator characteristics (ROC) curve was constructed to evaluate the ability of the PARCA-R to predict moderate or worse delay on the BSID III cognition and language composite scales. The area under the ROC curve (AUC) was calculated to provide a summary measure of predictive performance.

Moderate or worse delay in cognitive ability and in language ability was defined as a BSID III score on the relevant component scale that was at least 2 SD below the norm of 100 (ie, a score <70). A secondary analysis was also undertaken using the published scores from an Australian normative cohort of infants

Table 1 PARCA-R and BSID III scale scores

	PARCA-R			BSID III	
	Non–verbal cognition subscale	Linguistic skills scale	Parent report composite	Cognitive composite	Language composite
N	204	204	204	204	190*
Mean	24.6	48.3	72.9	94.8	94.3
SD	5.3	34.0	37.4	15.5	16.6
Median	26	42	67	95	94
IQR	22–28	19–72	43–98	85–105	86–106
Min-Max	0 to 34	0 to 123	0 to 150	55 to 145	47 to 141

*Missing for 14 infants.
PARCA-R, Parent Report of Children's Abilities–Revised; BSID III, Bayley Scales of Infant Development III

born at term to define delay.[9] A BSID III Cognitive Composite Score <80.3 and a BSID III Language Composite Score <78.6, each equated to 2 SDs below the mean for the Australian normative cohort.

Ethics approval was obtained for all Australian participants from the research and ethics committee of The University of Sydney and the ethics committees of all participating hospitals. Ethics approval was obtained for New Zealand participants from the Ministry of Health's Multi-regional Ethics Committee.

RESULTS

The median birthweight of the 204 infants in the study was 911 g (IQR: 718–1163), the median gestational age at birth was 27 (IQR: 25–30) weeks and 100 (49.0%) were girls. At the follow-up assessment, 28 (13.7%) had a confirmed or suspected hearing loss, 79 (38.7%) had chest symptoms, 15 (7.4%) had experienced seizures and 25 (12.3%) had a confirmed or suspected diagnosis of cerebral palsy.

The median (corrected) age of subjects in months at the time of the assessments was 24.3 (IQR: 23.6–25.6, min-max: 11.7–44.8) for the PARCA-R and 24.6 (IQR: 24.0–25.4, min-max: 12.0–38.5) for the BSID III. Over half of the infants were assessed within ±1 month of the 24 month target with the PARCA-R (55.8%), and the BSID III (57.4%). The interval between the administration of the two assessments was less than 1 month for 70.1% of the sample, and less than 2 months for 82% of the sample. Complete PARCA-R data were available for 186/204 (91%) infants, and no infant had more than five missing item responses.
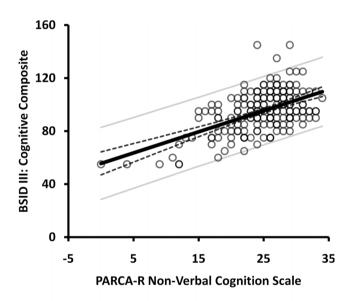
Table 1 provides descriptive statistics on the PARCA-R and BSID III scales. The Spearman correlation between PARCA-R and BSID scales was moderate (0.43) for the cognition component and greater (0.71) for the language component (table 2). These correlations were stronger (ie, 0.48 and 0.81, respectively) for the subset of infants that had both assessments completed within ±1 month of the target. Scatter plots depicting the relationship between the cognitive and language components of the PARCA-R versus the BSID III are shown in figure 1. The points appeared relatively symmetrically distributed above and below the fitted regression line. There was nevertheless some suggestion of floor effects in the PARCA-R language scale in figure 1, and all three infants with extremely positive scores BSID cognitive scale (ie, ≥99th percentile) appeared as outlying cases.

Table 2 Correlations between PARCA-R and BSID III scales*

	PARCA-R		BSID III	
	Non–verbal cognition subscale	**Linguistic skills scale**	**Cognitive composite**	**Language composite**
PARCA-R				
Non-verbal cognition scale	1			
Linguistic skills scale	0.60 (0.66†)	1		
BSID III				
Cognitive composite	0.43 (0.48†)	0.51 (0.55†)	1	
Language composite	0.48 (0.57†)	0.71 (0.81†)	0.76 (0.73†)	1

*Spearman correlations shown. All were highly statistically significant with p<0.0001.
†Estimates for subgroup of 94 patients who completed both assessments within 1 month of target and each other shown in square brackets.
PARCA-R, the Parent Report of Children's Abilities–Revised; BSID III, Bayley Scales of Infant Development III

The 95% prediction limit (ie, the CI for an individually predicted value) was approximately ±26 points for the PARCA-R cognitive scale and approximately ±24 points for the PARCA-R language scale. The average width of the 95% CI for the regression line was ±2.4 for both the cognitive scale and the language scale; with precision being notably lower in regions of the scales occupied by few cases (eg, for infants with very low PARCA-R cognition scores).

Using the standard normative scoring for the BSID, 9 infants (4.4%, 95% CI 1.6% to 7.2%) met the criteria for at least moderate cognitive delay (BSID cognitive composite score <70), and 16 (8.4%, 95% CI 4.5% to 12.4%) met the criteria for at least moderate language delay (BSID language composite score<70). The PARCA–R cognitive component identified the cases of cognitive delay accurately achieving an AUC of 0.96 (95% CI 0.90 to 1.00; a cut-point of ≤19 on the cognitive component had a sensitivity of 0.89 and specificity of 0.89). The PARCA-R language component likewise identified the cases of language delay accurately achieving an AUC of 0.97 (95% CI 0.94 to 0.99; a cut-point of ≤23 on the language component had a sensitivity of 0.75 and specificity of 0.79).

Using the Australian normative cohort as the reference group to define moderate or worse cognitive and language delay, 40 infants in our cohort (19.6%, 95% CI 14.2% to 25.1%) met the criteria for cognitive delay (ie, a BSID cognitive composite <80.3), and 24 (12.6%, 95% CI 7.9% to 17.4%) met the criteria for language delay (ie, a BSID cognitive composite<78.6). The PARCA–R cognitive component identified the cases of cognitive delay with an AUC of 0.83 (95% CI 0.77 to 0.90; a cut-point of ≤10 on the cognitive component had a sensitivity of 0.94 and specificity of 0.93). The PARCA-R language component identified the cases of language delay with an AUC of 0.91 (95% CI 0.86 to 0.94; a cut-point of ≤17 on the language component had a sensitivity of 0.88 and specificity of 0.87).

DISCUSSION

This study of 204 infants from Australia and New Zealand previously diagnosed with proven or suspected neonatal sepsis estimated moderate correlations within and between the scales of the PARCA–R and BSID III instruments at 2 years of age. The correlation between the cognitive and language scales was 0.48 for the PARCA–R and 0.76 for the BSDI III, indicating that, while cognitive and language skills are clearly distinct aspects of development, they were not independent among our sample of infants at risk of delay.

The PARCA-R scales were not perfect surrogates for their counterpart scales from the BSID III; with correlations reaching 0.48 and 0.81 for the cognitive component and language component, respectively, among infants that had the two assessments performed within ±1 month of the target. Predictions of individual children's BSID III scores for cognition and language from the PARCA-R were relatively imprecise, as depicted

Figure 1 Legend: The solid black lines are the fitted regression lines. The dashed back lines are the 95% CIs for the regression lines. The solid grey lines are the 95% prediction limits for individual children.

graphically by the scatter about the fitted regression lines in figure 1; however, an expectation for complete agreement between developmental assessments, even between two administrations of the BSID III, is unrealistic owing to the constraints imposed by the random measurement error inherent in the assessment of latent constructs such as cognitive function. Nevertheless, a convincing association between the scales of the PARCA-R and BSID III was found, and this was modelled with considerable precision for the vast majority of cases as demonstrated by the narrow 95% CIs for the regression lines (figure 1). The precision of fit was lower for the few infants at the upper and lower extremes of ability, with possible floor effects observed with the PARCA-R language scale for seven infants scoring below the first percentile on the BSID III language scale (ie, score ≤60), and an underestimate of cognitive function by the PARCA-R for three infants scoring at or above the 99th percentile on the BSID cognitive scale (ie, score ≥124). These potential limitations, if confirmed, would have negligible implications for the value of the PARCA-R as a practical tool for highlighting instances of appreciable developmental delay in neonatal clinical trials. Further work to fully evaluate PARCA-R scores as continuous variable substitutes for BSID III measures of development would be of value.

The PARCA-R performed well in predicting cases of moderate or worse language delay and moderate or worse cognitive delay. The AUC for the prediction of language delay was high (ie, over 0.9) irrespective of whether cases were defined using standard BSID III normative criteria (ie, BSID III language composite <70) or the Australian normative data (ie, BSID III language composite <78.6). The AUC for the prediction of cognitive delay was high (ie, 0.83) when the Australian normative data were used to define cases (ie, BSID III cognitive composite <80.3) and near perfect (ie, 0.96) when cases were defined relative to standard BSID III normative criteria (ie, BSID III cognitive composite <70). This latter result needs to be interpreted with caution, however, as the number of cases identified using the standard normative criteria was very small (N=9; 4.4%). This estimate was in fact much lower than that previously obtained (14.9%) using the preceding version of the BSID (ie, BSID II) in an earlier cohort of 476 infants from INIS.[1]

When an equal importance was applied to maximising sensitivity and specificity, and depending on the normative criteria applied, the PARCA-R would miss between 6% to 11% of cases of cognitive delay, and 12% to 25% of cases of language delay. However, the optimal choice a cut-point will depend upon the relative value placed upon constructing a test with high sensitivity versus constructing a test with high specificity (see discussion in Martin et al[1]).

The disparity in estimates of prevalence for moderate or worse cognitive delay between the versions of the BSID may be a consequence of a miscalibration of the normative scoring for the BSID III cognitive composite scale. We obtained a more consistent estimate to our previous study (19.6% compared to 14.9%) when the Australian normative data were applied to the BSID III data. This result suggests that the standard BSID III reference norms may tend to underestimate cognitive disability and is consistent with other recent studies that compared the BSID III against the previous BSID II and found a tendency for the BSID III cognitive component to produce systematically greater scores than the BSID II Mental Development Index.[9–13]

In conclusion, the results of this study provide additional support for the PARCA-R as a practical tool for the identification of appreciable cognitive and language delay at 24 months among critically ill, premature and extremely low birthweight neonates. They also suggest that the BSID III reference norms may tend to underestimate cognitive delay.

REFERENCES

1 Martin AJ, Darlow BA, Salt A, et al. Identification of infants with major cognitive delay using parental report. Dev Med Child Neurol 2012;54:254–9.
2 Johnson S, Wolke D, Marlow N. Outcome monitoring in preterm populations—measures and methods. Zeitschrift Fur Psychologie-J Psychol 2008;216:135–46.
3 Johnson S, Marlow N, Wolke D, et al. Validation of a parent report measure of cognitive development in very preterm infants. Dev Med Child Neurol 2004;46:389–97.
4 Johnson S, Wolke D, Marlow N. Developmental assessment of preterm infants at 2 years: validity of parent reports. Dev Med Child Neurol 2008;50:58–62.
5 Marlow N, Greenough A, Peacock JL, et al. Randomised trial of high frequency oscillatory ventilation or conventional ventilation in babies of gestational age 28 weeks or less: respiratory and neurological outcomes at 2 years. Arch Dis Child Fetal Neonatal Ed 2006;91:F320–6.
6 Brocklehurst P, Farrell B, King A, et al. Treatment of neonatal sepsis with intravenous immune globulin. N Engl J Med 2011;365:1201–11.
7 Albers CA, Grieve AJ. Bayley scales of infant and toddler development, third edition. J Psychoeducational Assess 2007;25:180–90.
8 Bayley N. Bayley scales of infant and toddler development. 3rd edn. San Antonio: Harcourt Assessment, 2006.
9 Anderson PJ, De Luca CR, Hutchinson E, et al. Underestimation of developmental delay by the new Bayley-III Scale. Arch Pediatr Adolesc Med 2010;164:352–6.
10 Lowe JR, Erickson SJ, Schrader R, et al. Comparison of the Bayley II Mental Developmental Index and the Bayley III cognitive scale: are we measuring the same thing? Acta Paediatr 2012;101:E55–8.
11 Reuner G, Fields AC, Wittke A, et al. Comparison of the developmental tests Bayley-III and Bayley-II in 7-month-old infants born preterm. Eur J Pediatr 2013;172:393–400.
12 Acton BV, Biggs WSG, Creighton DE, et al. Overestimating Neurodevelopment using the Bayley-III after early Complex Cardiac Surgery. Pediatrics 2011;128:E794–800.
13 Moore T, Johnson S, Haider S, et al. Relationship between test scores using the second and third editions of the Bayley Scales in extremely preterm children. J Pediatr 2012;160:553–8.

ADC

Performance of the Parent Report of Children's Abilities−Revised (PARCA-R) versus the Bayley Scales of Infant Development III

Andrew J Martin, Brian A Darlow, Alison Salt, Wendy Hague, Lucille Sebastian, Nicki McNeill and William Tarnow-Mordi

---

Updated information and services can be found at:

http://adc.bmj.com/content/early/2013/09/12/archdischild-2012-30328 8

These include:

References	This article cites 12 articles, 1 of which you can access for free at: **http://adc.bmj.com/content/early/2013/09/12/archdischild-2012-30328 8#BIBL**
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections	Articles on similar topics can be found in the following collections Child and adolescent psychiatry (paedatrics) (681) Developmental paediatrics (127) Immunology (including allergy) (1999)

Notes