Using the PARCA-R to assess children's cognitive and language development at two years of age

Developmental assessment at two years of age is important for both clinical and research purposes, for which standardised measures are needed. Developmental tests frequently require significant time and resources to administer, making them challenging to implement on a large scale. In contrast, the PARCA-R (Parent Report of Children's Abilities – Revised) is a standardised parent questionnaire that can be used to assess children's cognitive and language development at 24 months of age. Here we describe the development of the PARCA-R and key information needed to use the questionnaire in clinical practice.

Samantha Johnson
PhD, AFBPsS
sjj19@le.ac.uk

Vasiliki Bountziouka
PhD

On behalf of the PARCA-R Study Group
(**TABLE 1**)

Keywords

assessment; cognitive development; language development; parent report; neurodevelopmental outcomes

Key points

Johnson S., Bountziouka V. Using the PARCA-R to assess children's cognitive and language development at two years of age. *Infant* 2020; 16(4): 159-63.
1. The PARCA-R can be used to assess neurodevelopmental outcomes at 23.5 to 27.5 months of age.
2. The questionnaire takes parents less than 15 minutes to complete.
3. It produces separate norm-referenced scores for non-verbal cognition and language development.
4. It can be used to identify developmental delay for all children in the general population.
5. The PARCA-R is free to use and all available resources can be accessed online at www.parca-r.info.

Neurodevelopmental assessment at two years of age

Routine screening for developmental delay is particularly important for populations of children known to be at high risk for developmental disorders. Indeed, developmental assessments are frequently carried out as part of routine clinical follow-up for preterm-born children and other neonatal intensive care unit graduates with significant neonatal morbidities. The classification of developmental status at two years of age is also widely regarded as a key outcome for parental counselling, benchmarking of neonatal services and quality audit,[1] and cognitive tests in early childhood are predictive of IQ in adulthood in very preterm populations.[2] In addition, developmental assessments at two years of age are increasingly considered as important outcomes in clinical trials of obstetric and perinatal interventions and are widely used as outcome measures in observational studies.

Neurodevelopment at two years of age is typically recorded as a composite outcome comprising assessments of vision, hearing, motor, communication and cognitive function. While routine clinical assessments may be sufficient for assessing sensory and motor outcomes, the use of a structured test is recommended for assessing cognitive and language development. For example, Wong and colleagues demonstrated that routine clinical assessments lack sensitivity for identifying very preterm-born children with

Samantha Johnson Professor of Child Development, Department of Health Sciences, University of Leicester

Vasiliki Bountziouka Research Fellow, Department of Health Sciences, University of Leicester

Peter Brocklehurst Professor of Women's Health, Birmingham Clinical Trials Unit, University of Birmingham

Louise Linsell Lead Medical Statistician, National Perinatal Epidemiology Unit Clinical Trials Unit, University of Oxford

Bradley Manktelow Associate Professor in Medical Statistics, Department of Health Sciences, University of Leicester

Neil Marlow Professor of Neonatal Medicine, Elizabeth Garrett Anderson Institute for Women's Health, University College London

Dieter Wolke Professor of Developmental Psychology and Individual Differences, Department of Psychology, University of Warwick

TABLE 1 The PARCA-R Study Group.

communication and cognitive impairments at two years of age; approximately 50% of children with at least mild communication impairment and 30% with at least mild cognitive impairment would have been classified as unimpaired on the basis of a clinical assessment alone relative to a developmental test.[3]

For both clinical and research purposes, an assessment carried out by a trained examiner using a standardised test, such as the Bayley Scales of Infant and Toddler Development III (Bayley-III), is typically considered the gold standard. However, the extensive resources required to administer such tests often prohibit their use, especially in large scale trials. Parent questionnaires can provide cost and time-efficient alternatives to standardised tests if they are known to have sound psychometric properties and good clinical validity. They can also be used to aid in remote developmental surveillance to address challenges associated with the provision of face-to-face assessments, such as in the context of COVID-19.

One such option is the PARCA-R, a parent questionnaire designed to assess children's cognitive and language development at 24 months of age. The questionnaire takes only 10-15 minutes to complete, less than five minutes to score and is available for use free of charge, making it a cost-efficient alternative to a formal developmental test. In 2017, the National Institute for Health and Care Excellence (NICE) identified the PARCA-R as the optimal developmental screening tool for use with preterm-born children and recommended its use at 24 months of age to identify developmental delay, early intellectual disability or language problems in children eligible for enhanced developmental surveillance (ie all children born <30 weeks of gestation, or those born at 30-36 weeks with additional risk factors for developmental delay).[4] The provision of this assessment at two years of age, including use of the PARCA-R, was identified as one of the four quality statements in the accompanying Quality Standard.[5]

In the following sections we describe the development and standardisation of PARCA-R, its psychometric properties and practical considerations for its application in clinical follow-up.

A brief history of the PARCA-R

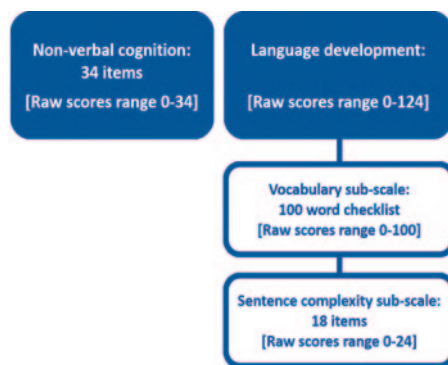The PARCA-R originally started life as the Parent Report of Children's Abilities

FIGURE 1 Composition of the PARCA-R and range of raw scores.

(PARCA), a 26-item parent-based assessment of children's non-verbal cognition developed as an outcome measure for the Twins Early Development Study.[6] This was appended with the UK short form version of the MacArthur Communicative Development Inventories, a parent questionnaire for assessing language development in children aged 16-30 months. Together, scores on these measures were significantly associated with scores on the Mental Development Index (MDI) of the Bayley Scales of Infant Development II (BSID-II) (r=0.64; p<0.0001). This demonstrated concurrent validity and led the authors to conclude that the measure could be used to provide valid estimates of young children's cognitive and language abilities.[6]

In 2002, in order to provide an outcome measure for use in the United Kingdom Oscillation Study (a randomised trial of high frequency ventilation for very preterm infants), the study investigators adapted the PARCA by adding eight items to assess cognitive abilities at a lower developmental level to improve its sensitivity for detecting severe developmental delay. Together with the language scale, the revised questionnaire became known as the PARCA-R.[7]

The PARCA-R comprises scales to assess non-verbal cognition and language development (**FIGURE 1**). The cognitive scale comprises 34 questions about the child's play (for example, can your child stack two small blocks or toys on top of each other?), for each of which parents indicate whether or not their child is able to perform the activity. Scores on these items are summed to produce a total raw score ranging from 0 to 34. The language scale comprises a 100-word expressive vocabulary checklist, on which parents indicate the number of words their child is able to say, and 18 items to assess grammatical development.

Item scores are summed to produce a total raw score for language development ranging from 0 to 124. Scores on the non-verbal cognitive scale and language scale can also be summed to produce a total Parent Report Composite (PRC) score ranging from 0 to 158. A full description of the PARCA-R can be seen in the *PARCA-R Technical and Interpretive Manual.*[8]

Using these raw scores, the validity and reliability of the PARCA-R were established through a series of studies in clinical populations, predominantly among children born preterm. For example, strong correlations were found between raw PRC scores and BSID-II MDI scores among two samples of children born very preterm (n=64, r=0.68;[7] n=164, r=0.77[9]), and between PRC scores and an averaged Bayley-III Cognitive and Language composite score among children born late and moderately preterm (n=219, r=0.66[10]), demonstrating good concurrent validity. At the sub-scale level, correlations of 0.48 and 0.81 were reported between PARCA-R cognitive and language raw scores and standardised scores on the corresponding scale of the Bayley-III, respectively, in children with sepsis in the neonatal period.[11] In one study, inter-rater reliability was reported as excellent with a correlation of 0.95 between PRC scores obtained for 52 children whose parents completed the questionnaire on two separate occasions approximately two weeks apart.[7]

The diagnostic utility of PARCA-R for identifying children with developmental delay as defined by scores on a gold standard test was also established within these studies. For example, the optimal cut-point for identifying moderate to severe delay (corresponding with BSID-II MDI <70) in infants born very preterm (<32 weeks' gestation) was a PRC score <44 (sensitivity 85%; specificity 87%).[9] PRC scores <38 were also found to have sensitivity of 82% and specificity of 90% for identifying children with severe developmental delay (BSID-II MDI <55).[12] Similar predictive values were found for the association between PARCA-R cognitive and language scores and delayed cognitive and language development on the Bayley-III.[8,11]

The need for standardised scores

Although the PARCA-R was found to have good validity, reliability and diagnostic utility, its use was limited by a lack of standardised or norm-referenced scores. Published cut-off scores for identifying

developmental delay were typically derived from small samples (n <500), resulting in cut-points that varied significantly between different populations and with wide confidence intervals for predictive values. This limited its use only to those clinical populations in which the test was validated. Moreover, as cut-off points were developed to identify only moderate and/or severe delay, it was not possible to identify children with more subtle delays, or to classify advanced development. Ultimately, standardised scores were needed to allow clinicians and researchers to precisely quantify a child's development relative to the norm, identify delays of any severity, and enable the PARCA-R to be used to assess the development of *all* children in the general population.

Developing PARCA-R standardised scores

In 2018, thanks to funding from Action Medical Research, we embarked on a project to standardise the PARCA-R. The main aims of the project were to:

- use existing data to identify a large standardisation sample that was representative of the UK population
- develop standardised scores that have a normative mean of 100 and SD of 15, in keeping with other psychometric tests
- test the validity of the new standardised scores.

To produce a standardisation sample, we first obtained anonymised PARCA-R data for children assessed between 23.5 and 27.5 months of age from three existing sources:[8]

- 6,196 children born at 35-42 weeks of gestation in the INFANT trial (a randomised trial of computerised interpretation of fetal heart rate during labour)
- 186 children born at 32-36 weeks of gestation in LAMBS (a cohort study of birth at late and moderately preterm gestations)
- 20 children born <32 weeks' gestation

from the PANDA Study (a feasibility study of using parent questionnaires for clinical follow-up).

To determine whether this was representative of the UK population, we compared the standardisation sample with census data on a range of key characteristics known to affect developmental outcomes. Shown in **TABLE 2**, the standardisation sample was representative of the UK population in terms of sex, deprivation and multiple births. However, due to the availability of existing data, there were slightly fewer very preterm births in the standardisation sample than would be expected in the general population, and it was not possible to assess representativeness in terms of ethnicity as this was not recorded for almost 14% of children in the standardisation sample.

Using PARCA-R data for children in the standardisation sample, we then applied an extension of the Lamda-Mu-Sigma (LMS) method to develop standardised scores.[13] The LMS method is commonly used to estimate percentiles and produced age-related normalised reference ranges for use in clinical practice (for example, growth centiles). Using this method, standard scores with a normative mean of 100 (SD 15) were produced for the cognitive and language scale. For each scale, standard scores were produced separately for males and females in four age bands:

- 23 months 16 days to 24 months 15 days
- 24 months 16 days to 25 months 15 days
- 25 months 16 days to 26 months 15 days
- 26 months 16 days to 27 months 15 days.

Norms tables were produced for converting raw scores to standard scores, similar to those used in other developmental tests.

External validity

In order to test external validity, we applied the standard scores to anonymised PARCA-R data from a different sample of children similar to that of the standardisation sample. If the PARCA-R has good external validity, then the standard scores in the external validation sample should be similar, on average, to those derived from the standardisation sample, and thus close to what would be expected in the general population. Using data from 709 term-born children recruited to the control group in LAMBS, a geographical population-based birth cohort study, we applied the equations derived from the LMS models to these data. As shown in **TABLE 3**,

Characteristic		Standardisation sample	General population
Sex, n (%)		6,402	757,686
	Males	3,321 (51.9)	386,833 (51.1)
	Females	3,081 (48.1)	370,853 (48.9)
Ethnic background, n (%)		6,402	3,789,571
	White	5,009 (78.2)	2,956,304 (78)
	Non-white	508 (7.9)	833,267 (22)
	Missing	885 (13.8)	–
Deprivation, quintile, n (%)		6,402	755,118
	1st quintile (most deprived)	1,651 (25.8)	195,221 (25.9)
	2nd quintile	1,284 (20.1)	162,041 (21.5)
	3rd quintile	1,081 (16.9)	141,071 (18.7)
	4th quintile	1,217 (19.0)	130,374 (17.3)
	5th quintile (least deprived)	1,078 (16.8)	126,117 (16.7)
	Missing	91 (1.4)	–
Multiple birth, n (%)		6,402	772,814
	Singleton	6,234 (97.4)	748,342 (96.8)
	Multiple birth	168 (2.6)	24,472 (3.2)
Gestational age, n (%)		6,402	772,814
	Term (≥37 weeks)	6,039 (94.3)	717,277 (92.8)
	Late & moderately preterm (32-36 weeks)	343 (5.4)	45,896 (5.9)
	Very preterm (<32 weeks)	20 (0.3)	9,613 (1.2)

TABLE 2 Socio-demographic and birth characteristics of children in the standardisation sample and in the UK general population. Deprivation was defined according to the Index of Multiple Deprivation, 2010. Distribution of sex, ethnicity and deprivation obtained from the Office for National Statistics; Distribution of preterm and multiple births obtained from the Office for National Statistics for England and Wales, and from the Information Services Division for Scotland.

the mean (SD) standard scores in the external validation sample approximated the normative mean (100; SD 15) on both scales, demonstrating good external validity. Moreover, the proportions of children with developmental delay were close to the proportions expected in the general population (**FIGURE 2A**). For example, 16% of children in the external validation sample had mild to severe cognitive delay and 16% had mild to severe language delay, both of which are comparable with the 16% expected within the normal distribution. Similarly, 2% and 3% had at least moderate to severe cognitive or language delay respectively, similar to the 2.5% expected in the normal distribution. Small differences between the two samples were expected as children in the external validation sample were all born at term and more were living in less deprived areas than children in the standardisation sample.

Clinical validity

In order to test clinical validity, we applied the standard scores to anonymised PARCA-R data from two samples of children known to be at high risk for developmental disorders. If the PARCA-R has good clinical validity, then the standard scores in the clinical validation samples should be lower, on average, than those expected in the general population. Indeed this was the case. Using the norms tables, the standard scores were applied to PARCA-R data for 692 children born very

preterm (<32 weeks' gestation; PANDA Study) and 764 children with suspected or confirmed sepsis during the neonatal period (UK arm of the INIS Trial). As shown in **TABLE 3**, mean standard scores for very preterm children were 6-9 points lower than the normative mean, and for children with neonatal sepsis they were 11-14 points lower. Furthermore, the proportions of children with developmental delay were substantially greater than in the general population for both very preterm children (**FIGURE 2B**) and those with neonatal sepsis (**FIGURE 2C**), demonstrating clinical validity.

Using the PARCA-R

The methods and results of the standardisation study, including the investigations of external and clinical validity, were published in 2019.[13] We have since produced a test manual,[8] which includes all the information needed to score and interpret the questionnaire and use it in clinical practice. The PARCA-R is available for use non-commercially; the questionnaire, test manual and associated resources can be downloaded for free from the PARCA-R website (www.parca-r.info). The website also includes an electronic version of the questionnaire that can be completed by parents or professionals to generate raw scores, and an online score calculator that can be used to generate the standard scores. The output from the score calculator can either be saved as a PDF or printed for the child's medical notes. The

questionnaire can also be scored by hand in less than five minutes using the instructions and norms tables provided in the manual. To correct for prematurity when assessing children born preterm, standard scores can be derived by hand using the appropriate norms table in the manual for the child's corrected age at assessment, rather than his or her chronological age, or by entering the child's expected date of delivery, rather than date of birth, in the online score calculator.

As the PARCA-R is now standardised and has demonstrated clinical validity, it can be used with confidence to determine a child's developmental level relative to the norm and to identify delay by applying conventional SD-banded cut-offs, as follows:

- mild delay: standard scores 70-84 (corresponding to standardised scores -2 SD to < -1 SD)
- moderate delay: standard scores 55-69 (corresponding to standardised scores -3 SD to < -2 SD)
- severe delay: standard scores ≤54 (corresponding to standardised scores < -3 SD).

It is important to note that, as the PARCA-R was standardised on children aged 23.5 to 27.5 months, the questionnaire should be completed at this age (chronological or corrected) in order to derive the standard scores. Although previous studies primarily used the raw PRC scores for identifying children with delay, the use of PRC standard scores may
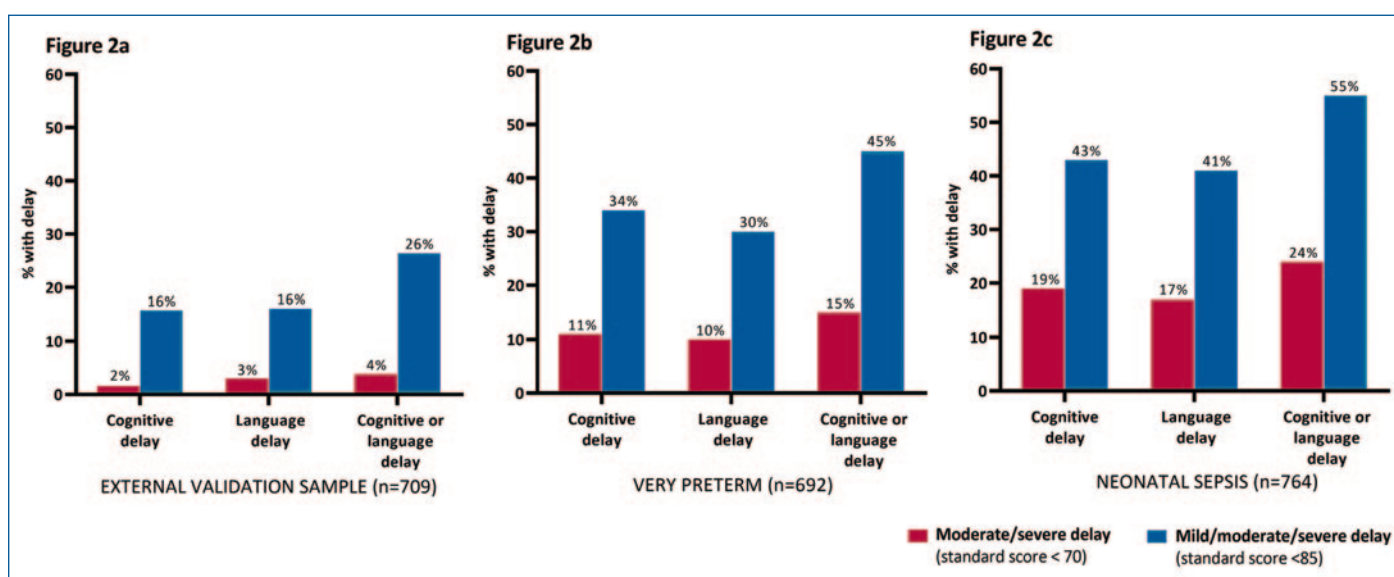

FIGURE 2 Proportion of children with developmental delay in the external and clinical validation samples. Figure 2a – external validation sample; Figure 2b – children born very preterm (<32 weeks' gestation); Figure 2c – children with suspected or confirmed sepsis in the neonatal period. Expected proportion of children with moderate/severe delay in the general population: 2.5%; Expected proportion of children with mild/moderate/severe delay in the general population: 16%.

mask children who have delays in individual domains. Therefore, although standard scores for the PRC scale were developed to enable comparison of data with previous studies, standard scores for the separate cognitive and language scales should be used to assess development and classify delay for both clinical and research purposes.

As a result of the PARCA-R being used by research groups in pan-European studies, the questionnaire (or parts thereof) is currently translated into 14 languages other than English, all of which are freely available from the PARCA-R website. However, standard scores have only been developed for the UK population and it is not yet known how these will perform in other populations. For non-English speaking families in the UK, for whom an appropriate translation is not yet available, we recommend that the English version be used to enable derivation of the standard scores. In the external validation sample, there was no significant difference in standard scores between children who did and did not come from homes in which English was the first language. This suggests that the PARCA-R is valid for use with families who do not speak English as a first language, but assistance may need to be provided for the family to complete the questionnaire.

In line with recommendations for the developmental follow-up of children born preterm published in 2017,[4] the PARCA-R is increasingly being used with success in clinical assessments at two years of age. For example, the team at Glasgow's Royal Hospital for Children introduced the PARCA-R into their neurodevelopmental outpatient service in 2018. Since its introduction, 100% of families attending the clinic at two years corrected age have completed the PARCA-R questionnaire,[14] demonstrating its acceptability to parents and its feasibility for use in routine clinical care.

Moreover, now that the PARCA-R has been standardised and produces norm-

	Non-verbal cognition Mean (SD)	Language development Mean (SD)
External validation sample (n=709)	101 (16)	100 (16)
Very preterm sample (n=692)	91 (17)	94 (17)
Neonatal sepsis sample (n=764)	86 (21)	89 (19)
Normative mean: 100; SD: 15.		

TABLE 3 PARCA-R scores in the external validation sample and clinical validation samples.

referenced scores, it can be used to detect even small differences in development between individuals or groups of children, making it an ideal continuous outcome measure for clinical trials and epidemiological studies. This is especially the case for large scale trials in which the resources required to carry out examiner-administered tests are prohibitive. Moreover, the questionnaire can now be used to assess the development of all children in the general population, not just those born preterm, and to identify both advanced development and delays of any severity. Notably, the PARCA-R standardisation sample of 6,402 children far exceeds that of other gold standard tests. For example, the Bayley-III was standardised on only 200 children in the age range corresponding with the PARCA-R.

Conclusion

In summary, the PARCA-R provides a cost-efficient, standardised outcome measure with robust psychometric properties that can be used to assess children's cognitive and language development and quantify delay at two years of age, for both clinical and research purposes.

For questions about the PARCA-R, please email the study team at parca-r@le.ac.uk.

References

1. **British Association of Perinatal Medicine**. Report of a BAPM/RCPCH Working Group: classification of health status at 2 years as a perinatal outcome. 2008. London: BAPM.
2. **Breeman LD, Jaekel J, Baumann N, et al.** Preterm cognitive function into adulthood. Pediatrics 2015;136:315-423.
3. **Wong HS, Cowan FM, Modi N, on behalf of the**
Medicines for Neonates Investigator Group. Arch Dis Child Fetal Neonatal Ed 2017;0:F1-F6.
4. **Wong HS, Cowan FM, Modi N.** Validity of neurodevelopmental outcomes of children born very preterm assessed during routine clinical follow-up in England. Arch Dis Child Fetal Neonatal Ed 2018;105:F479-84.
5. **National Institute for Health and Care Excellence.** Developmental follow-up of children and young people born preterm. Quality Standard. QS169. 2018: NICE.
6. **Saudino K, Dale PS, Oliver B, et al.** The validity of parent-based assessment of the cognitive abilities of 2-year-olds. Brit J Dev Psychol 1998;16:349-63.
7. **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.
8. **Johnson S, Bountziouka V, Linsell L, et al.** PARCA-R Technical and Interpretive Manual. 2019: University of Leicester.
9. **Johnson S, Marlow N, Wolke D, et al.** Developmental assessment of preterm infants at 2 years: validity of parent reports. Dev Med Child Neurol 2008;50:58-62.
10. **Blaggan S, Guy A, Boyle S, et al.** A parent questionnaire for developmental screening in infants born late and moderately preterm. Pediatrics 2014;134:e55-62.
11. **Martin AJ, Darlow BA, Salt A, et al.** Performance of the Parent Report of Children's Abilities-Revised (PARCA-R) versus the Bayley Scales of Infant Development III. Arch Dis Child 2013;98:955-58.
12. **Martin AJ, Darlow BA, Salt A, et al.** Identification of infants with major cognitive delay using parental report. Dev Med Child Neurol 2012;54:254-59.
13. **Johnson S, Bountziouka V, Brocklehurst P, et al.** Standardisation of the Parent Report of Children's Abilities-Revised (PARCA-R): a norm-referenced assessment of cognitive and language development at age 2 years. Lancet Child Adolesc Health 2019;3:705-12.
14. **Leven L, Grant A, Russell S.** Neurodevelopmental follow-up: multidisciplinary team working to standardise the delivery, uptake and recording of two-year assessments. In: National Neonatal Audit Programme 2019 Annual Report on 2018 data. Royal College of Paediatrics and Child Health 2019.