Australian Journal of Educational & Developmental Psychology. Vol 13, 2013, pp. 1-15

The Effect of Age-correction on IQ Scores Among School-aged Children Born Preterm[1]

Rachel M Roberts[1], Wing Man George[1], Carolyn Cole[2], Peter Marshall[3], Vanessa Ellison[2] & Helen Fabel[2]

[1] University of Adelaide, [2] Flinders Medical Centre, [3] Flinders University

ABSTRACT

This study examined the effect of age-correction on IQ scores among preterm school-aged children. Data from the Flinders Medical Centre Neonatal Unit Follow-up Program for 81 children aged five years and assessed with the WPPSI-III, and 177 children aged eight years and assessed with the WISC-IV, were analysed. Corrected IQ scores were significantly higher than not-corrected IQ scores (Full Scale IQ and all indices) for both the WPPSI-III and WISC-IV. The use of age-corrected IQ scores has the potential to exclude some children from support services.

INTRODUCTION

The International Statistical Classification of Diseases and Related Health Problems – 10th Revised Version (ICD-10) classifies prematurity by birth weight (of less than 2500 grams) or gestation period (of less than 37 completed weeks) (World Health Organisation, WHO, 2007). The classifications include: Extremely Low Birth Weight (ELBW, birth weight of 999 grams or less), Other Low Birth Weight (LBW, birth weight of 1000-2499 grams), Extreme Immaturity (gestation period of less than 28 completed weeks), and Other Preterm Infants (gestation period of 28 completed weeks or more but less than 37 completed weeks).

The Australian Institute of Health and Welfare reported that in 2007, 8.1% of live births in Australia were preterm, that is, 36 weeks or less of gestation, with 0.9% meeting the criteria for Extreme Immaturity (Laws & Sullivan, 2009). Also in 2007, 5.8% of Australian live born infants were of Low Birth Weight (1000-2499 grams) and 0.4% were of ELBW (Laws & Sullivan, 2009).

Developmental and Cognitive Functioning in Preterm Children

Children born preterm are at greater risk of developmental and cognitive impairments. Preterm children are more likely to exhibit deficits in executive, attentional, perceptual-motor, and visuo-spatial abilities than children who are born full-term (Anderson & Doyle, 2004; Johnson & Marlow, 2006). Infants of low birth weight were also found to be more likely to exhibit delays in speech and language, motor and visual-motor development, and intelligence (Tomchek & Lane, 1993). Further, previous research has suggested that preterm infants perform at a lower level on developmental tests when compared with their full-term peers (Eaves, Nuttall, Klonoff & Dunn, 1970; Ross, 1985).

[1] Contact:
Rachel M Roberts,
School of Psychology,
University of Adelaide,
Adelaide, SA, 5005,
Ph: 8 8313 5228,
Fax: 8303 3770,
rachel.roberts@adelaide.edu.au

A recent meta-analysis by Bhutta, Cleves, Casey, Craddock, and Anand (2002) investigated case-control studies that examined the cognitive development of preterm children after their fifth birthday. The authors found that overall, while preterm children produced IQ scores that were within the average range of intelligence, their performance was lower than that of their full-term peers. ELBW children have been found to be at particularly high risk for cognitive deficits, with higher rates of intellectual disability (i.e., IQ < 70), as well as cerebral palsy and visual disability (Hack et al., 1994).

Studies have shown that school-aged preterm children were inferior to their full-term peers in academic performance (e.g., reading, spelling and mathematics) (Aylward 2002; Taylor, Klein, & Hack, 2000). Further, preterm children have also been found to be more likely to exhibit academic underachievement and to receive additional assistance at school (Wocaldo & Rieger, 2006).

Such cognitive deficits and learning difficulties have been found to continue beyond childhood and into adolescence and adulthood. Longitudinal research has shown that when compared to their full-term counterparts, adolescents and young adults who were born preterm had lower IQs, performed less well academically, were more likely to repeat a grade due to unsatisfactory academic performance or to require special educational assistance, and were less likely to graduate from secondary school (Allin et al. 2008; Hack et al., 2002; Saigal, Lorraine, Streiner, Stoskopf, & Rosenbaum, 2000).

Psychological Testing at Neonatal Follow-Up

In light of the risks of developmental and cognitive deficits associated with preterm birth, it is important that the development of preterm children is monitored beyond birth and hospital discharge. Such monitoring via routine developmental and cognitive assessments should serve to: 1) inform clinical neonatal care and management; 2) detect any difficulties or impairments; 3) guide the assignment of interventions or services when necessary; and 4) provide insight into predictions of later functioning (Johnson & Marlow, 2006).

Age-Correction in Psychological Testing

Since the 1930s, there has been an increasingly common clinical practice of correcting the age of preterm infants and children for developmental follow-up in order to account for their prematurity (Lems, Hopkins & Samsom, 1993; Miller, Dubowitz, & Palmer, 1984; Rickards, Kitchen, Doyle, & Kelly, 1989; Siegel, 1983; Wilson & Craddock, 2004). This practice commonly involves deriving a corrected age by subtracting from chronological age the number of days or weeks the child is preterm (Wilson, 1987). For example, an 18-month old who was born 8 weeks preterm would be assessed at a corrected age of 16 months.

The theoretical basis behind this practice reflects a biological and maturation perspective, that early development continues to progress as a function of time since conception (rather than since birth), regardless of whether it is in an intrauterine or extrauterine environment (Gesell & Amatruda, 1947; Siegel, 1983). While preterm children are temporarily delayed in terms of development due to the immaturity of their central nervous system, they would eventually ‘catch-up’ during the first two or three years of life as their central nervous system matures. Then they would be no different developmentally from their full-term peers (Gesell & Amatruda, 1947; Palisano, 1986). In light of this, it has been argued that age-correction has important implications, in that the abilities of preterm infants and toddlers are less likely to be underestimated or under-predicted during developmental assessment. This approach then translates to a decreased likelihood of misdiagnoses (e.g., intellectual or developmental disability, and motor impairments) and subsequently, unnecessary stress on families and service providers (Wilson & Craddock, 2004). Indeed, a study has shown that preterm infants were neurologically less mature and produced significantly lower development quotients (DQs) than their full-term peers (Parmelee & Schulte, 1970). However, when their scores were corrected to account for being preterm, their DQs increased to a level comparable with the full-term infants.

Despite the seeming ‘common-sense’ of age-correction and the amount of clinical support and advocacy it has received over time, research in this area is mixed and inconclusive. On the one hand, some have found that while preterm children may initially exhibit some developmental delays, they tend to perform at a level comparable with their full-term peers by their second or third birthday. They argue that age-correction (up to the chronological age of two to three years) leads to more accurate estimations of current development (Allen & Alexander, 1990; Ouden, Rijken, Brand, Verloove-Vanhorick, & Ruys, 1991; Palisano 1986; Restiffe & Gherpelli, 2006; Ungerer & Sigman, 1983). Supporting this approach, Rickards et al. (1989) found that corrected developmental scores were better predictors of cognitive ability than not-corrected scores. In addition, more recent research by Ment et al. (2003) found that the cognitive ability of preterm children improved progressively with age (from three to eight years of age), which has been interpreted as support for age-correction beyond infanthood and early childhood (Wilson & Craddock, 2004).

On the other hand, the risks of developmental, cognitive, and academic difficulties associated with preterm birth that continue beyond infanthood and early childhood should not be overlooked. In fact, age-correction among the very preterm (i.e., less than 28 weeks gestation or LBW) has been suggested to overestimate developmental and cognitive functioning and hence lead to under-detection of any developmental difficulties and subsequent delay of needed intervention (Wilson & Craddock, 2004). Further, the validity of using corrected assessment scores as an indicator of current functioning has been called into question on a number of dimensions: measurement error involved in calculating gestational age (e.g., use of obstetric exams); the potential inaccuracy of assessing preterm children with measures that were developed for and standardised with full-term children; and the ongoing debate regarding whether preterm and full-term children develop at parallel rates or in the same pattern, and, as a consequence should they be compared at all (DiPietro & Allen, 1991; Miller et al., 1977; Kalmar, 1996)?

Limitations of Previous Research on Age-correction

Due to recent medical advances and improved perinatal care, which result in better chances of survival for preterm infants, preterm children born today are likely to be qualitatively different from those in past decades (Horbar et al., 2002; Wilson & Craddock, 2004). Since most of the research on age-correction was conducted between 15 and 30 years ago, research with more recent data is needed to improve generalisability. Additionally, a majority of previous research focused on age-correction in developmental assessments of preterm infants and toddlers (i.e., usually three years old or younger). In fact, a thorough review of previous literature and research yielded just one relevant study by Rickards, Kitchen, Doyle and Kelly (1989) who reported that in a sample of very low birth weight children assessed using the WPPSI at age five, corrected full scale IQ scores were 4.1 points lower than uncorrected scores.

This observed gap in the research, particularly with older children, warrants attention as cognitive assessments (e.g., IQ testing) are suggested to be more stable and reliable measures of current and future functioning than infant developmental tests (Halperin & McKay, 1998; McCall, Hogarty, & Hurlburt, 1972; Zigler, Balla, & Hodapp, 1984). Infant developmental tests are less reliable and less predictive of future functioning due to unexpected developmental spurts and delays commonly experienced in infanthood and early childhood, and such tests assess qualitatively different functions from these functions in cognitive tests (e.g., sensorimotor functions versus intelligence) (Bradley-Johnson, 2001; Humphreys & Davey, 1988; Sattler, 2001).

In light of these issues and the catch-up theory, that is, that preterm children eventually catch-up by the end of infanthood and hence age-correction is only necessary during the first two or three years of life, some have speculated that age-correction is not necessary for cognitive assessments during and beyond early childhood (Allen & Alexander, 1990; Ouden et al., 1991; Palisano 1986; Restiffe & Gherpelli, 2006; Ungerer & Sigman, 1983). However, this speculation has received little research to date and hence remains untested.

Age-correction and Intellectual Disability

In South Australia, one of the main criteria for diagnosing intellectual disability and determining eligibility to disability services is having an IQ of less than 70. This IQ must be derived from a standardised intellectual assessment (Hay, 2008). In light of this, age-correction in cognitive assessments could have significant implications for a child’s eligibility for support services. Specifically, the potential overestimation of abilities associated with age-correction could mean that children with genuine intellectual disabilities lose eligibility for funding and services. This potential clinical dilemma has received limited attention.

Objectives of Present Research

Despite the widespread practice of age-correction in the psychological testing of preterm infants and children, further research in this area is needed. As Wilson and Craddock (2004) point out, no one could be entirely confident in distinguishing preterm children with temporary developmental lags, as explained by the immaturity of their central nervous system, from those with genuine developmental deficits. Most of the previous research is outdated and has not explored age-correction in cognitive assessments of school-aged children. Additionally, the potential effect of age-correction on IQ scores and how this may influence the diagnosis of intellectual disability warrants investigation.

As such, the aim of the present research was to examine the effect of age-correction on IQ scores among preterm school-aged children, and to explore the clinical implications of age-correction on the diagnosis of intellectual disability.

Specifically, the research objectives were to:

1. Examine the difference between not-corrected and corrected IQ scores (Full Scale IQ and Indices) among preterm children aged five and eight.
2. Examine the relationship between IQ score differences (i.e., the difference between not-corrected and corrected scores) and level of preterm birth (i.e., birth weight and gestation period).
3. Examine how age-correction affects IQ scores in terms of meeting criteria for intellectual disability.

METHOD

Participants

At the Flinders Medical Centre (FMC) Neonatal Unit, a neonatal follow-up program is routinely offered to infants born very preterm or very low birth weight (i.e., less than 32 weeks gestation or less than 1500 grams birth weight). Participating infants are assessed using various measures of medical, health, developmental assessments, language, fine and gross motor skills, and psychological outcomes such as intellectual functioning, academic achievement and behaviour for up to the first eight years of their lives. For the present study, data were analysed from two cohorts of children who underwent intellectual assessment at ages five and eight. The younger cohort involved 81 children aged between 54 and 66 months (or from four and a half years to five and a half years, M = 60.56, SD = 2.40), and the older cohort involved 177 children aged between 90 and 102 months (or from seven and a half to eight and a half years, M = 96.94, SD = 1.94). Within the younger cohort, there were 46 males and 35 females, and within the older cohort, there were 88 males and 89 females. Birth weight of children ranged from 440 to 2080 grams (M = 1143.12, SD = 313.74) and gestational age ranged from 23 to 36 weeks (M = 28.41, SD = 2.14).

Measures

Wechsler Preschool and Primary Scale of Intelligence – Third Edition (WPPSI-III). Cognitive ability was measured at approximately five years of age using the WPPSI-III (Wechsler, 2002). The WPPSI-III is a reliable, valid and well-standardised measure of intelligence among preschool and early primary children (Zhu & Weiss, 2005). The test provides US norms for ages between two and half years to seven years and three months. It contains seven core subtests and seven supplemental subtests, which assess four indices of intelligence: Verbal IQ, Performance IQ, Processing Speed Index (which could be derived for children aged four to seven years and three months), and General Language Composite. The General Language Composite was not routinely measured as part of the Follow-up Program cognitive assessments and hence such data were not available for analysis. Scores for these indices are summarised into a Full Scale composite score (i.e., Full Scale IQ), which has been considered as an indicator of overall cognitive ability (Wechsler, 2002).

While the WPPSI-III Australian Standardised version (i.e., with Australian norms) was published in 2004, it was not made available to the FMC Neonatal Unit during the time when the assessments took place and hence the US version was utilised for all participants.

Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV). Cognitive ability was measured at approximately eight years of age using the WISC-IV (Wechsler, 2003). The Follow-up Program initially assessed children using the WISC-IV Australian Adaptation version (2003), which has US norms and changes made to items to reflect Australian language and culture. From 2007, the program began to assess children with the WISC-IV Australian Standardised version (Wechsler, 2003), which has Australian norms and items. Both versions have excellent validity, reliability and standardisation samples (Wechsler, 2003; Zhu & Weiss, 2005). The test contains 10 core subtests and five supplemental subtests, which cover four indices of intelligence: Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed. Scores for these indices are summarised into a Full Scale composite score (i.e., Full Scale IQ), which is considered an indicator of general cognitive functioning (Wechsler 2003).

Design and Procedure

The present study performed retrospective analyses on quantitative data from two cohorts. Approval was obtained from Flinders Clinical Research Ethics Committee and the University of Adelaide Human Research Ethics Committee, Psychology Subcommittee. All data were derived from the FMC Neonatal database.

The present study reports data from cognitive assessments of preterm children as provided through the FMC Neonatal Follow-up Program. When preterm infants were recruited into the Follow-up Program, parents signed a form that gave their consent for the use of clinical information for publication in the scientific literature,. As part of the psychological component of the program, participating children who were born with a gestation period of less than 32 weeks or birth weight of less than 1500 grams were assessed at approximately eight years of age with the WISC-IV. Infants with gestational age of less than 30 weeks or birth weight of less than 1250 grams were also assessed at approximately five years of age with the WPPSI-III. Starting with the cohort of children born in 2002 or later, the criteria for assessment at eight years using the WISC-IV were changed to gestational age of less than 28 weeks or birth weight of less than 1000 grams.

The WPPSI-III data included assessments administered between May 2004 and June 2010, and the WISC-IV data included assessments administered between January 2005 and June 2010. The Neonatal Unit reported that during these periods, 47 children aged around five years of age and 69 children aged around eight years of age did not receive assessment or produce valid assessment outcomes on the WPPSI-III and WISC-IV due to administrative reasons (e.g., failure or refusal to attend assessments, cancellations, child moving or living interstate, and child no longer receiving care from FMC). It was further reported that at age five, seven children, and at age eight, eight children were not assessed or did not produce valid assessment outcomes due to having moderate to severe intellectual or developmental disabilities or conditions, auditory or visual impairments, and language impairments. It should also be noted that the present study included children with various perinatal conditions, such as intraventricular haemorrhage and intrauterine growth restriction, and developmental or learning disabilities, such as autism and attentional deficits.

The assessments were administered by psychologists and trainee psychologists at the FMC Neonatal Unit. The assessment results were recorded in the FMC Neonatal database, from which the researchers collected relevant information for the present study (e.g., age, gender, birth weight, gestation period, and WPPSI-III and WISC-IV results). All relevant data were then transferred to a separate database for statistical analysis.

Corrected scores were produced by using corrected age (i.e., chronological age subtracted by the number of weeks the child is preterm) when converting total raw scores to Scaled Scores for each subtest administered. For the subtests of Coding and Symbol Search, there are different items for children aged six to seven years and those aged eight to sixteen years. Therefore, it was not possible to produce corrected scores for these two subtests among those children who had a chronological age of eight years or older but a corrected age of seven years and eleven months or younger because of the differences in test items and hence the incompatibility of raw scores. For these children, chronological Scaled Scores were recorded as both not-corrected and corrected scores for these two subtests.

Statistical Analyses

The present study had adequate data (i.e., 81 records of WPPSI-III data and 177 records of WISC-IV data) to examine the first objective as only approximately 64 participants were needed to provide sufficient statistical power (0.80) in order to detect a medium effect (i.e., $d \geq 0.50$) at a significance level of $\alpha = .05$ (Cohen, 1992).

Initially, independent samples t-tests were conducted to detect any gender differences among variables, such as birth weight, gestational period, age at testing, and assessment results. Paired samples t-tests were then performed to compare not-corrected and corrected scores for WPPSI-III and WISC-IV (both Full Scale IQ and indices). Simple correlations were performed to explore the relationships between IQ score differences (i.e., difference between not-corrected and corrected IQ scores) and birth weight, and between IQ score differences and gestation period. Lastly, a chi-square test was conducted to compare not-corrected and corrected scores in terms of the proportions of children whose scores fell within the intellectual disability range.

RESULTS

Data Screening

There was no missing data in the scores for the WPPSI-III and WISC-IV Full Scale IQ. However, the Processing Speed Index for the WPPSI-III had three cases with missing scores. These cases were excluded from the specific analyses that involved this index. All other indices had no missing data.

Skewness and kurtosis values were examined to assess normality of the scores for the WPPSI-III Full Scale IQ, WISC-IV Full Scale IQ and all indices. All skewness values were close to zero (i.e., >-1.00 and <1.00) and all kurtosis values were acceptable considering the size of the sample (Tabachnick & Fidell, 2000). Histograms were also inspected and the shape of distribution for the scores appeared reasonably normal.

Preliminary Analyses

An independent samples t-test showed that the children in the WPPSI-III cohort (M = 27.98, SD = 1.68) had a significantly lower mean gestational age than those in the WISC-IV cohort [M = 28.61, SD = 2.30, $t(207) = -2.49$, $p = .013$]. In addition, independent samples t-tests were conducted to assess gender differences for a number of variables. As presented in Table 1, there were no significant gender differences except for males having a higher birth weight, which had a small effect size ($\eta^2 = .025$).

Given that there were no significant gender differences in WPPSI-III and WISC-IV Full Scale IQ scores, statistical analyses were performed using the total data (i.e., genders combined).

Difference between not-corrected and corrected scores.

Paired samples t-tests were conducted and it was found that there was a significant difference between not-corrected and corrected scores for WPPSI-III Full Scale IQ [$t(80) = -16.28$, $p < .001$] and WISC-IV Full Scale IQ [$t(176) = -13.70$, $p < .001$]. The corrected scores were significantly higher than the not-corrected scores, for both the WPPSI-III (i.e., mean difference of 5.20 IQ points) and WISC-IV Full Scale IQ (i.e., mean difference of 1.75 IQ points). See Table 2 for means and standard deviations.

Table 1: Means (and standard deviations) for variables by gender

	Male	Female	
	M (*SD*)	*M* (*SD*)	*t*(*df*)
Gestational age (weeks)	28.49 (2.12)	28.33 (2.17)	.578 (256)
Birth weight (grams)	1190.90 (303.80)	1091.49 (318.37)	2.57 (256)*
WPPSI-III Full Scale IQ not-corrected	96.83 (12.72)	96.57 (13.99)	.09 (79)
WPPSI-III Full Scale IQ corrected	101.98 (13.48)	101.83 (14.75)	.05 (79)
WISC-IV Full Scale IQ not-corrected	89.58 (14.95)	91.70 (14.17)	-.97 (175)
WISC-IV Full Scale IQ corrected	91.27 (14.83)	93.49 (14.42)	-1.01 (175)
Age at testing WPPSI-III (months)	60.35 (2.56)	60.83 (2.19)	-.89 (79)
Age at testing WISC-IV (months)	96.85 (2.02)	97.02 (1.87)	-.58 (175)

*$p < .05$

As can be seen in Table 2, a series of paired samples t-tests showed significant differences between not-corrected and corrected scores for all WPPSI-III indices: Verbal [$t(80) = -14.91$, $p < .001$], Performance [$t(80) = -14.29$, $p < .001$], and Processing Speed [$t(77) = -12.19$, $p < .001$]. Similarly, results revealed significant differences between not-corrected and corrected scores for all WISC-IV indices: Verbal [$t(176) = -10.53$, $p < .001$], Perceptual Reasoning [$t(176) = -11.45$, $p < .001$], Working Memory [$t(176) = -10.07$, $p < .001$], and Processing Speed [$t(176) = -3.20$, $p = .002$]. The corrected scores were significantly higher than the not-corrected scores for all indices of both the WPPSI-III and WISC-IV.

Relationship between IQ score differences and birth weight and gestational age.

The relationship between differences in Full Scale IQ scores (i.e., the difference between not-corrected and corrected scores) and birth weight and gestational age was investigated using Pearson product-moment correlation coefficients. There was no significant correlation between IQ score differences and birth weight for either WPPSI-III ($r = -.175$, $N = 81$, $p = .117$) or WISC-IV scores ($r < .001$, $N = 177$, $p = .997$). While there was also no significant correlation between IQ score differences and gestational age for WISC-IV scores ($r = -.059$, $N = 177$, $p = .435$), a significant negative correlation was found between these two variables for WPPSI-III scores ($r = -.289$, $N = 81$, $p = .009$), albeit a relatively weak relationship.

Table 2: Means (and standard deviations) for not-corrected and corrected Full Scale IQ scores and indices

	Not-corrected	Corrected
WPPSI-III Full Scale IQ	96.72 (13.20)	101.91 (13.95)
WISC-IV Full Scale IQ	90.64 (14.56)	92.34 (14.63)
WPPSI-III indices		
Verbal	97.88 (11.50)	102.05 (12.02)
Performance	97.82 (13.94)	102.33 (14.69)
Processing Speed	92.76 (15.42)	97.45 (15.47)
WISC-IV indices		
Verbal	90.44 (12.53)	91.79 (12.52)
Perceptual Reasoning	94.11 (14.92)	95.91 (14.78)
Working Memory	92.69 (13.96)	94.29 (14.03)
Processing Speed	93.24 (15.05)	93.62 (15.01)

Number of cases in Wechsler classifications of cognitive functioning.

Table 3 presents a summary of how many cases were categorised into each of Wechsler's classifications of cognitive functioning (Wechsler, 2003; 2004) when examining not-corrected versus corrected scores. Also see Figure 1 and 2.

Proportions of children with intellectual disability range IQ scores.

There was no WPPSI-III Full Scale score that fell within the range of Intellectual Disability (i.e., IQ < 70). As shown in Table 4, 5.6% of WISC-IV Full Scale scores fell within this category when not-corrected scores were used versus 4.5% when corrected scores were used. However, a chi-square test showed that this difference in proportion was not statistically significant [$\chi^2(1, N = 354) = .059, p = .809$].

DISCUSSION

There has been limited research on age-correction in cognitive assessments of school-aged children. Therefore, the aim of the present study was to examine the effect of age-correction on IQ scores in children aged five and eight years who were assessed on the WPPSI-III and WISC-IV. The relationship between IQ score differences (i.e., the difference between not-corrected and corrected scores) and prematurity, and the influence of age-correction on IQ scores in terms of meeting criteria for intellectual disability were explored.

Children in the WPPSI-III cohort had a lower gestational age than those in the WISC-IV cohort. This was anticipated because the Neonatal Follow-up Program only offered assessments on the WPPSI-III to children with less than 30 weeks gestation (or birth weight of less than 1250 grams), whereas the majority of the children assessed on the WISC-IV had gestational ages of less than 32 weeks (or birth weight of less than 1500 grams). Interestingly, there was no difference between the cohorts in terms of birth weight despite their differing inclusion criterion. In addition, male participants were heavier at birth than female participants. This finding is consistent with previous research on gender differences in perinatal factors (Bertino et al., 2009; Lehre, Lehre, Laake, & Danbolt, 2009).

Table 3: Number (and percentage) of cases per Wechsler classification of cognitive functioning between not-corrected and corrected scores

	WPPSI-III			WISC-IV		
	Not-corrected	Corrected	Difference	Not-corrected	Corrected	Difference
Wechsler Classifications (IQ)						
Intellectual Disability (<70)	0	0	0	10 (5.6)	8 (4.5)	2 (1.1)
Borderline (70-79)	9 (11.1)	8 (9.9)	1 (1.2)	28 (15.8)	24 (13.6)	4 (2.2)
Low Average (80-89)	14 (17.3)	8 (9.9)	6 (7.4)	49 (27.7)	51 (28.8)	2 (1.1)
Average (90-109)	46 (56.8)	35 (43.2)	11 (13.6)	76 (42.9)	75 (42.4)	1 (0.5)
High Average (110-119)	9 (11.1)	24 (29.6)	15 (18.5)	11 (6.2)	16 (9.0)	5 (2.8)
Superior (120-129)	2 (2.5)	5 (6.2)	3 (3.7)	3 (1.7)	3 (1.7)	0
Very Superior (≥130)	1 (1.2)	1 (1.2)	0	0	0	0

Table 4: *Number (and percentage) of WISC-IV cases with IQ less than 70*

IQ	Not-corrected	Corrected
< 70	10 (5.6)	8 (4.5)
≥ 70	167 (94.4)	169 (95.5)

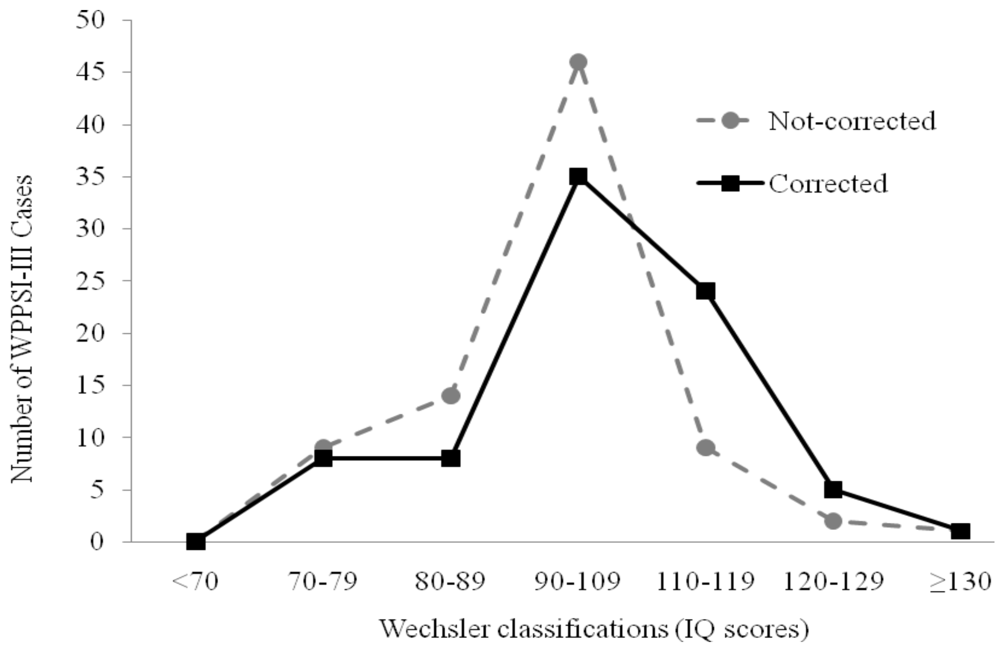

Figure 1: Number of WPPSI-III cases by Wechsler classifications (IQ scores).

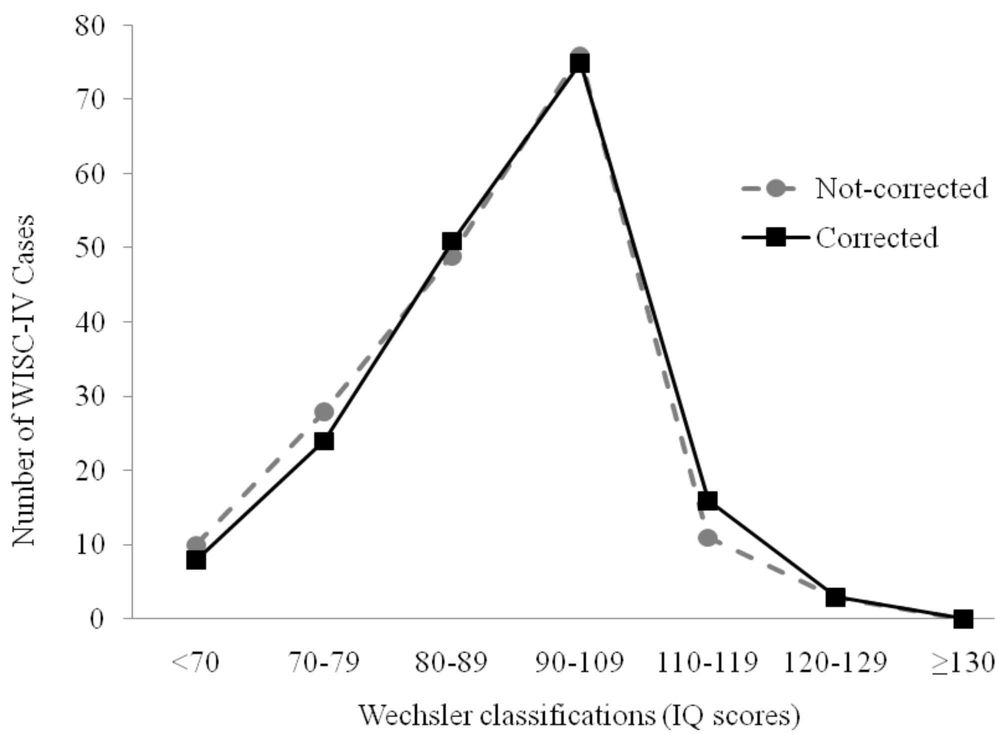

Figure 2: Number of WISC-IV cases by Wechsler classifications (IQ scores).

The Effect of Age-correction on IQ Scores

Findings from the present study suggest that the practice of age-correction increases IQ scores among school-aged children. Specifically, when corrected age was used, there was an increase of approximately five IQ points for the WPPSI-III, and approximately one to two IQ points for the WISC-IV. Similarly, all Index scores for the WPPSI-III and WISC-IV increased when corrected age was used. These findings are consistent with previous research on age-correction in developmental assessments of infants and toddlers, which found age-correction to increase DQ scores (e.g., Ouden et al., 1991; Parmelee & Schulte, 1970) and to increase IQ scores at age 5 (Rickards et al., 1989).

It was noteworthy to find that there was a greater increase of IQ points for the WPPSI-III as compared to the WISC-IV. This finding was expected because the children in the WPPSI-III cohort were on average more preterm (i.e., shorter gestation) and hence required more correction. Secondly, the age norms for the WPPSI-III are presented in two-month bands (e.g., from 5 years 0 months to 5 years 2 months), whereas the age norms for the WISC-IV are presented in three-month bands (e.g., 8 years 0 months to 8 years 3 months). As such, age-correction in the WPPSI-III involved reference to one to two (but more frequently two) preceeding age norm bands, whereas the WISC-IV involved reference to only one (and very rarely, two) preceding age norm bands. Therefore, age-correction in the WPPSI-III was more likely to produce greater increase of IQ scores than in the WISC-IV. This finding may be interpreted as support for the 'catch up theory', as presented in the study by Ment et al. (2003), that the cognitive ability of preterm children improves progressively with age.

Since there is limited research on the effect of age-correction on IQ scores within this age group, the present findings should be interpreted cautiously. The significant increase in IQ scores within both cohorts could suggest that age-correction should be implemented beyond infanthood and early childhood (i.e., beyond two to three years of age). Furthermore, the greater increase in IQ scores within the WPPSI cohort, as compared to the WISC cohort could imply that the effect of age-correction is more prominent in the younger children. As a result, it could be argued that age-correction should be routinely implemented at five years of age, and, with some caution, at eight years of age.

IQ Score Differences and Preterm Birth

The present study did not find a relationship between IQ score differences (i.e., the difference between not-corrected and corrected scores) and birth weight within the WPPSI-III and WISC-IV cohort. This finding was expected because birth weight did not influence the amount of correction which determined the differences between scores. On the other hand, there was a weak relationship between score differences and gestational age within the WPPSI-III cohort. This showed that prematurity was related to the amount of correction. This was anticipated as it was assumed that lower gestational age would translate to more correction, and hence greater differences between not-corrected and corrected scores. Interestingly, there was no such relationship within the WISC-IV cohort. On average, the WISC-IV cohort was less preterm than the WPPSI-III cohort, and it is possible that this could explain the absence of a relationship.

The Effect of Age-correction on Intellectual Disability Eligibility

There were no children in the WPPSI-III cohort who had IQ scores that fell within the range of Intellectual Disability (i.e., IQ < 70). This could be because the children with cognitive abilities within the Intellectual Disability range were too developmentally delayed to be assessed at that age. Further, it has been suggested that some of the subtests in the WPPSI-III do not have adequate test floors, that is, the subtest cannot take on a lower value and hence all children with lower cognitive abilities produce similar scores. This could overestimate cognitive ability (Gordon, 2004). In contrast, there were ten children in the WISC-IV cohort who had IQ scores that fell within the range of Intellectual Disability. After age-correction, two of these children were no longer considered to have an intellectual disability. Rather, their corrected IQ scores fell within the Borderline range of cognitive functioning. This finding suggests that the use of corrected scores could overestimate the cognitive abilities of a small percentage of children who may have genuine intellectual disabilities, and hence affect their eligibility for access to funding and services.

While the present findings are consistent with research by Bhutta et al. (2002) with most of the children falling within the range of "Average or above IQ", a notable proportion of children were within either the "Low Average" or "Borderline" range. When corrected age was used, a number of these children were placed within a higher IQ range. This could have important implications because lower cognitive abilities (e.g., as represented by an IQ below the Average range) are associated with academic difficulties and needing extra assistance at school (Hamilton, 2006). If corrected scores were considered, the cognitive ability of these children would be overestimated and their difficulties and extra needs remain undetected.

Limitations and Implications for Future Research

The present findings should be interpreted with consideration of the following limitations. First, the data were not longitudinal and hence it was not possible to compare IQ score differences between the WPPSI-III and WISC-IV cohorts. Such comparisons within a longitudinal study could provide further insight into the 'catch-up theory' proposed by Ment et al. (2003). Also, examination of IQ score differences over time may shed more light on the predictive ability of corrected scores (versus not-corrected scores). Future research could examine the differences between not-corrected and corrected scores at two or more age points. If Ment et al.'s suggestions are valid, the difference between scores should decrease as age increases.

Second, corrected scores were not produced for the subtests of Coding and Symbol Search among the children who had a chronological age of eight years or older but a corrected age of seven years and eleven months or younger. This was due to the presence of different items for younger and older age bands, which led to incompatibility of raw scores. As such, it is possible that the effect of age-correction among these children would have been greater if corrected scores were produced for these subtest scores. Future research could address this by examining children whose chronological age and corrected age fall within the same age band in terms of item administration for the subtests of Coding and Symbol Search.

Third, the present study included children with various perinatal conditions (e.g., intraventricular haemorrhage and intrauterine growth restriction) and developmental or learning disabilities (e.g., autism and attentional deficits). These conditions could have obscured the findings because it is uncertain whether the lower cognitive abilities detected could be attributed to preterm birth or comorbidity. Such comorbidity is common among preterm infants and children (Johnson & Marlow, 2006; Msall & Tremont, 2002). Therefore, the present study intentionally employed an inclusion criterion that captured all children who were assessed with the WPPSI-III or WISC-IV as opposed to only those without any conditions or disabilities. This criterion allowed the study to examine most of the preterm children recruited by the Neonatal Follow-up Program and hence increased the study's generalisability.

Lastly, it was beyond the scope of the present study to address the question of whether age-correction is an appropriate or valid procedure, and at what age it should cease. It could be argued that the findings suggest that age-correction should be implemented beyond infanthood and early childhood. However, it should be noted that the differences in IQ scores, while statistically significant, were small, particularly at age eight. In order to investigate more extensively the effects of age-correction (both statistically and clinically), future research could examine longitudinal data that extend from early childhood to young adulthood. At this stage, our recommendation is that when eligibility for services and support for school age children is being considered uncorrected scores be used. This would ensure children are not excluded from services of potential benefit to them.

Conclusion

The current findings provide some insight into the effect of age-correction on IQ scores among preterm school-aged children who were assessed at age five and eight years. Specifically, the study compared not-corrected IQ scores with corrected IQ scores. Does age-correction affect IQ scores in terms of meeting criteria for intellectual disability? The results revealed that age-correction increased IQ scores for both age cohorts, but especially among those assessed at five years of age. The result was that two out of ten children who originally had IQs within the Intellectual Disability range were moved into a higher range of cognitive functioning (i.e., Borderline range) after age-correction was implemented. In sum, the present study has demonstrated that age-correction produces significant increases in IQ scores among preterm school-aged children, potentially excluding them from some support services.

REFERENCES

Allen, M. C., & Alexander, G. R. (1990). Gross motor milestones in preterm infants: Correction for degree of prematurity. *Journal of Pediatrics, 116*, 955-959.

Allin, M., Walshe, M., Fern, A., Nosarti, C., Cuddy, M., Rifkin, L., Murray, R., Rushe, T., & Wyatt, J. (2008). Cognitive maturation in preterm and term born adolescents. *Journal of Neurology, Neurosurgery and Psychiatry, 79*, 381-386.

Anderson, P. J., & Doyle, L. W. (2004). Executive functioning in school-aged children who were born very premature or with extremely low birth weight in the 1990's. *Pediatrics, 114*, 50-57.

Aylward, G. P. (2002). Cognitive and neuropsychological outcomes: More than IQ scores. *Mental Retardation and Developmental Disabilities Research Reviews, 8*, 234-240.

Bertino, E., Coscia, A., Boni, L., Rossi, C., Martano, C., Giuliani, F., Fabris, C., Spada, E., Zolin, A., & Milani, S. (2009). Weight growth velocity of very low birth weight infants: Role of gender, gestational age and major morbidities. *Early Human Development, 85*, 339-347.

Bhutta, A. T., Cleves, M. A., Casey, P. H., Craddock, M. M., & Anand, K. J. S. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *Journal of the American Medical Association, 288*, 728-737.

Bradley-Johnson, S. (2001). Cognitive assessment for the youngest children: A critical review of tests. *Journal of Psychoeducational Assessment, 19*, 19-44.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*, 155-159.

DiPietro, J., & Allen, M. C. (1991). Estimation of gestational age: Implications for developmental research. *Child Development, 62*, 1184-1199.

Eaves, L. C., Nuttall, J. C., Klonoff, H., & Dunn, H. G. (1970). Developmental and psychological test scores in children of low birth weight. *Pediatrics, 45*, 9-20.

Gesell, A., & Amatruda, C. S. (1947). *Developmental Diagnosis*. New York: Hoeber.

Gordon, B. (2004). Test review: The Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III). *Canadian Journal of School Psychology, 19*, 205-220.

Hack, M., Flannery, D. J., Schluchter, M., Cartar, L., Borawski, E., & Klein, N. (2002). Outcomes in young adulthood for very-low-birth-weight infants. *The New England Journal of Medicine, 346,* 149-157.

Hack, M., Taylor, H. G., Klein, N., Eiben, R., Schatschneider, C., & Mercuri-Minich, N. (1994). School-age outcomes in children with birth weights under 750 grams. *The New England Journal of Medicine, 331*, 753-759.

Halperin, J. M., & McKay, K. E. (1998). Psychological testing for child and adolescent psychiatrists: A review of the past ten years. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 575-584.

Hamilton, S. (2006). Screening for developmental delay: Reliable, easy-to-use tools. *Applied Evidence, 55*, 415-422.

Hay, R. (2008). *Eligibility and Access Guideline*. South Australia: Disability SA.

Horbar, J. D., Badger, G. J., Carpenter, J. H., Fanaroff, A. A., Kilpatrick, S., LaCorte, M., Phibbs, R., & Soll, R. F. (2002). Trends in mortality and morbidity for very low birth weight infants, 1991-1999. *Pediatrics*, *110*, 143-151.

Humphreys, L. G., & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence, 12*, 183-197.

Johnson, S., & Marlow, N. (2006). Developmental screen or developmental testing? *Early Human Development, 82*, 173-183.

Kalmar, M. (1996). The course of intellectual development in preterm and fullterm children: An eight-year longitudinal study. *International Journal of Behavioral Development, 19*, 491-516.

Laws, P., & Sullivan, E. A. (2009). *Australia's Mothers and Babies 2007.* Perinatal Statistics Series no. 23. Cat. No. PER 48. Sydney: AIHW National Perinatal Statistics Unit.

Lehre, A., Lehre, K. P., Laake, P., & Danbolt, N. C. (2009). Greater intrasex phenotype variability in males than in females is a fundamental aspect of the gender differences in humans. *Developmental Psychobiology, 51*, 198-206.

Lems, W., Hopkins, B., & Samsom, J. F. (1993). Mental and motor development in preterm infants: The issue of corrected age. *Early Human Development, 34,* 113-123.

McCall, R. B., Hogarty, P. S., & Hurlburt, N. (1972). Transitions in infant sensorimotor development and the prediction of childhood IQ. *American Psychologist, 27*, 728-748.

Ment, L. R., Vohr, B., Allen, W., Katz, K. H., Schneider, K. C., Westerveld, M., Duncan, C. C., Makuch, R. W. (2003). Change in cognitive function over time in very low-birth-weight infants. *Journal of the American Medical Association, 286*, 705-711.

Miller, G., Dubowitz, L. M. S., & Palmer, P. (1984). Follow-up of pre-term infants: Is correction of the developmental quotient for prematurity helpful? *Early Human Development, 9*, 137-144.

Msall, E. M., & Tremont, M. R. (2002). Measuring functional outcomes after prematurity: Developmental impact of very low birth weight and extremely low birth weight status on childhood disability. *Mental Retardation and Developmental Disabilities Research Reviews, 8*, 258-272.

Ouden, L. D., Rijken, M., Brand, R., Verloove-Vanhorick, S. P., & Ruys, J. H. (1991). Is it correct to correct? Developmental milestones in 555 "normal" preterm infants compared with term infants. *Journal of Pediatrics, 118*, 399-404.

Palisano, R. J. (1986). Uses of chronological and adjusted ages to compare motor development of healthy preterm and fullterm infants. *Developmental Medicine and Child Neurology, 28*, 180-187.

Parmelee, A. H., & Schulte, F. J. (1970). Developmental testing of pre-term and small-for-date infants. *Pediatrics, 45*, 21-28.

Restiffe, A. P., & Gherpelli, J. L. D. (2006). Comparison of chronological and corrected ages in the gross motor assessment of low-risk preterm infants during the first of year of life. *Arq Neuropsiquiatr, 64*, 418-425.

Rickards, A. L., Kitchen, W. H., Doyle, L. W., & Kelly, E. A. (1989). Correction of developmental and intelligence test scores for premature birth. *Australian Paediatric Journal, 25*, 127-129.

Ross, G. (1985). Use of the Bayley Scales to characterize abilities of premature infants. *Child Development, 56*, 835-842.

Saigal, S., Lorraine, A. H., Streiner, D. L., Stoskopf, B. L., & Rosenbaum, P. L. (2000). School difficulties at adolescence in a regional cohort of children who were extremely low birth weight. *Pediatrics, 105*, 325-331.

Sansavini, A., Guarini, A., Alessandroni, R., Faldella, G., Giovanelli, G., Salvioli, G. (2007). Are early grammatical and phonological working memory abilities affected by preterm birth? *Journal of Communication Disorders, 40*, 239-256.

Sattler, J. M. (2001). *Assessment of Children: Cognitive Applications*. (4th ed.). San Diego: Sattler Publisher Inc.

Siegel, L. S. (1983). Correction for prematurity and its consequences for the assessment of the very low birth weight infant. *Child Development*, *54*, 1176-1188.

Tabachnick, B. G., & Fidell, L. S. (2000). *Using Multivariate Statistics*. (4th ed.). Boston, MA: Allyn and Bacon.

Taylor, H. G., Klein, N., & Hack, M. (2000). School-age consequences of birth weight less than 750 g: A review and update. *Developmental Neuropsychology*, *17*, 289-321.

Tomchek, S. D., & Lane, S. J. (1993). Full-term low birth weight infants: Etiology and developmental implications. *Physical and Occupational Therapy in Pediatrics*, *13*, 43-65.

Ungerer , J. A., & Sigman, M. (1983). Developmental lags in preterm infants from one to three years of age. *Child Development*, *54*, 1217-1228.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence*. (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children*. (4th ed.). San Antonio, TX: Psychological Corporation.

Wilson, W. M. (1987). Age adjustment in psychological assessment of children born prematurely. *Journal of Pediatric Psychology*, *12*, 445-450.

Wilson, S. L., & Craddock, M. M. (2004). Review: Accounting for prematurity in developmental assessment and the use of age-adjusted scores. *Journal of Pediatric Psychology*, *29*, 641-649.

Wocaldo, C., & Rieger, I. (2006). Educational and therapeutic resource dependency at early school-age in children who were born very preterm. *Early Human Development*, *82*, 29-37.

World Health Organisation (WHO). (2007). *The International Statistical Classification of Diseases and Related Health Problems 10th Revision Version for 2007* (ICD-10). Geneva: Author.

Zhu, J., & Weiss, L. (2005). The Wechsler Scales. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 297-324). New York: Guilford Press.

Zigler, E., Balla, D., & Hodapp, R. M. (1984). On the definition and classification of mental retardation. *American Journal of Mental Deficiency*, *89*, 215-230.

Biographical Notes:

Dr Rachel Roberts M Psych (Clinical), PhD is a Senior Lecturer in the School of Psychology, University of Adelaide. Her interests include child clinical and neuropsychology.

Wing Man George M.Psych (Clinical) completed her Masters degree at the University of Adelaide and is currently working as a psychologist at Behavioural Intervention Service, Child and Adolescent Mental Health Services.

Carolyn Cole M. Psych (Clinical) is a psychologist at Flinders Medical Centre, Adelaide. As a member of the Neo-Natal Follow-up Program, Carolyn assesses the cognitive outcomes for school age children who are born preterm. Carolyn is also involved in the assessment and diagnosis of developmental and learning disorders as a member of the Children's Assessment Team.

A/Prof Peter Marshall AM MBBS FRACP is the Director of Neonatal Unit at the Flinders Medical Centre Adelaide. Peter's interests include the neuro-developmental outcomes of high-risk newborn infants and related events in the perinatal period; benchmarking and audit of clinical care, patient costing, and clinical information systems.

Dr Vanessa Ellison FRACP is a Neonatologist and Lecturer in Paediatrics at Flinders Medical Centre, Adelaide. She is a member of the Neonatal Follow-up team.

Helen Fabel RN, RM is the Co-ordinator of the Neonatal Follow-up Program, Flinders Medical Centre, Adelaide.