PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Meta-Analysis of Neurobehavioral Outcomes in Very Preterm and/or Very Low Birth Weight Children

Cornelieke Sandrine Hanan Aarnoudse-Moens, Nynke Weisglas-Kuperus, Johannes Bernard van Goudoever and Jaap Oosterlaan

Pediatrics 2009;124;717-728; originally published online Jul 27, 2009;
DOI: 10.1542/peds.2008-2816

Meta-Analysis of Neurobehavioral Outcomes in Very Preterm and/or Very Low Birth Weight Children

CONTRIBUTORS: Cornelieke Sandrine Hanan Aarnoudse-Moens, MSc,[a] Nynke Weisglas-Kuperus, MD, PhD,[a] Johannes Bernard van Goudoever, MD, PhD,[a] and Jaap Oosterlaan, PhD[b]

[a]*Division of Neonatology, Department of Pediatrics, Erasmus University Medical Centre Sophia Children's Hospital, Rotterdam, Netherlands; and* [b]*Department of Clinical Neuropsychology, VU University Amsterdam, Amsterdam, Netherlands*

KEY WORDS

very preterm, very low birth weight, academic achievement, behavioral problems, executive function

ABBREVIATIONS

VLBW—very low birth weight
BW—birth weight
EF—executive function
GA—gestational age
CI—confidence interval
FSN—fail-safe *N*

www.pediatrics.org/cgi/doi/10.1542/peds.2008-2816

doi:10.1542/peds.2008-2816

Accepted for publication Feb 2, 2009

Address correspondence to Cornelieke Sandrine Hanan Aarnoudse-Moens, MSc, Erasmus University Medical Centre Sophia Children's Hospital Rotterdam, Division of Neonatology, Department of Pediatrics, Dr Molewaterplein 60, 3000 CB Rotterdam, Netherlands. E-mail: c.moens@erasmusmc.nl

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

FINANCIAL DISCLOSURE: *The authors have indicated they have no financial relationships relevant to this article to disclose.*

abstract

OBJECTIVE: Sequelae of academic underachievement, behavioral problems, and poor executive function (EF) have been extensively reported for very preterm (≤33 weeks' gestation) and/or very low birth weight (VLBW) (≤1500 g) children. Great variability in the published results, however, hinders the field in studying underlying dysfunctions and developing intervention strategies. We conducted a quantitative meta-analysis of studies published between 1998 and 2008 on academic achievement, behavioral functioning, and EF with the aim of providing aggregated measures of effect size for these outcome domains.

METHODS: Suitable for inclusion were 14 studies on academic achievement, 9 studies on behavioral problems, and 12 studies on EF, which compared a total of 4125 very preterm and/or VLBW children with 3197 term-born controls. Combined effect sizes for the 3 outcome domains were calculated in terms of Cohen's *d*. *Q*-test statistics were performed to test homogeneity among the obtained effect sizes. Pearson's correlation coefficients were calculated to examine the impact of mean birth weight and mean gestational age, as well as the influence of mean age at assessment on the effect sizes for academic achievement, behavioral problems, and EF.

RESULTS: Combined effect sizes show that very preterm and/or VLBW children score 0.60 SD lower on mathematics tests, 0.48 SD on reading tests, and 0.76 SD on spelling tests than term-born peers. Of all behavioral problems stacked, attention problems were most pronounced in very preterm and/or VLBW children, with teacher and parent ratings being 0.43 to 0.59 SD higher than for controls, respectively. Combined effect sizes for parent and teacher ratings of internalizing behavior problems were small (<0.28) and for externalizing behavior problems negligible (<0.09) and not significant. Combined effect sizes for EF revealed a decrement of 0.57 SD for verbal fluency, 0.36 SD for working memory, and 0.49 SD for cognitive flexibility in comparison to controls. Mean age at assessment was not correlated with the strength of the effect sizes. Mathematics and reading performance, parent ratings of internalizing problems, teacher ratings of externalizing behavior, and attention problems, showed strong and positive correlations with mean birth weight and mean gestational age (all r values > 0.51).

CONCLUSIONS: Very preterm and/or VLBW children have moderate-to-severe deficits in academic achievement, attention problems, and internalizing behavioral problems and poor EF, which are adverse outcomes that were strongly correlated to their immaturity at birth. During transition to young adulthood these children continue to lag behind term-born peers. *Pediatrics* 2009;124:717–728

Improvements in perinatal care have resulted in increased survival rates for children born very preterm (≤33 weeks' gestation) and/or with a very low birth weight (VLBW) (≤1500 g). The incidence of major disabilities such as cerebral palsy, mental retardation, deafness, or blindness is fairly low.[1] There is growing awareness that the majority of nondisabled survivors encounter more "subtle" problems such as academic underachievement,[2] behavioral problems,[3–5] and deficits in higher-order neurocognitive functions: the so-called executive functions (EFs),[6] which persist throughout childhood and young adulthood.[1,4,7] However, great variability exists in the published results because of small numbers of participants, high attrition rates, and substantial variations in methods and study design. We conducted a quantitative meta-analysis to integrate previous research on academic achievement, behavioral problems, and EF in very preterm and/or VLBW children to provide aggregated measures of effect size for these 3 outcome domains. Such an aggregation will facilitate the field to move forward to study underlying dysfunctions and develop intervention strategies.

Academic achievement includes mathematics, reading, and spelling, of which the literature suggests that the poorest performance of very preterm and/or VLBW children is observed in mathematics.[2] Behavioral problems in these children mainly manifest in an increased risk for attention-deficit/hyperactivity disorder[3] and internalizing behavioral problems such as withdrawn behavior,[6] although some studies have also found oppositional behavior.[8,9] A large body of evidence has shown that academic underachievement and behavioral problems arise from a deficit in EF,[10–13] a set of neurocognitive functions such as inhibitory control, working memory, cognitive flexibility, and planning.[14] EF, therefore, has attracted considerable interest, and in very preterm and/or VLBW children executive dysfunction has been reported, suggested to arise from disruptions of cortical and subcortical circuits connecting frontal, striatal, and thalamic regions.[6]

The primary aim of this study was to meta-analytically chart the outcome of very preterm and/or VLBW children in terms academic achievement, behavioral functioning, and EF. The second aim was to examine the relationship between age at assessment, birthweight (BW), and gestational age (GA) on the one hand and effect sizes for the indices of academic achievement, behavioral functioning, and EF on the other hand.

METHODS

Inclusion Criteria

The guidelines for reporting meta-analyses of observational studies published by Stroup et al[15] (2000) were taken into account in the design, performance, and report of this meta-analysis. We searched original articles using the search terms "child*," "low birth weight," "prematur*," "preterm," "outcome," "math*," "arithmetic," "reading," "spelling," "school," "academic," "behav*," "neurocogn*," and "executive function*." The studies were located in the PubMed, PsycINFO, and Web-of-Science computerized databases. The reference lists of published articles were used to identify other relevant articles on these topics.

The literature was reviewed to include studies that met the following inclusion criteria: (1) the study was published between 1998 and 2008, thereby demarcating the period of emerging research into EF; (2) the study concerned children born very preterm (≤33 weeks' gestation) and/or with VLBW (≤1500 g) to estimate the maximal impact of prematurity and VLBW; (3) a case-control design was used; (4) the mean age at assessment was at least 5 years, because at this age children start to receive formal education, which enables academic achievement to be charted; (5) the study reported data on academic achievement and/or behavioral problems and/or EF collected with standardized tests; and (6) there is a range of different tests and questionnaires available to measure academic achievement, behavioral functioning, and EF, and some tests or questionnaires may have been used in only 1 or 2 studies. Although meta-analytic procedures may be applied with few studies, the obtained results might then be unstable.[16] To control for this problem, a cutoff point was chosen of a minimum of 5 studies that used a particular test or questionnaire, if the study was to be included in the meta-analysis, and we included as our seventh criterion that these results were published in English-language peer-reviewed journals. Studies were excluded if they did not meet all of these inclusion criteria.

Academic Achievement

Fourteen studies met our inclusion criteria. Standardized academic-achievement tests that were used in these studies all had identical normative scales with age- and grade-based standard scores around a mean score of 100 (SD: 15) and included the Woodcock-Johnson Tests of Achievement[17] which measures reading and mathematics; the Wide Range Achievement Test,[18] which measures mathematics, reading, and spelling; the Wechsler Individual Achievement Test,[19] which measures mathematics, reading, and spelling; and the Woodcock Reading Mastery Test-Revised,[20] which measures reading. Details on the studies included are provided in Table 1.[21–34]

Behavioral Problems

Nine studies met the inclusion criteria. Standardized questionnaires that were used in these studies included

TABLE 1 Studies That Reported on Academic Achievement in Very Preterm and/or VLBW Children

Studies	Participants	GA, Mean (SD), wk	BW, Mean (SD), g	Age, Mean (SD), y	Type of Test	Academic Achievement Test Scores		
						Mathematics, Mean (SD)	Reading, Mean (SD)	Spelling, Mean (SD)
Chaudhari et al[21] (2004)	78 VLBW	NA	NA	12.0	WRAT	80.4 (15.1)	NA	NA
	90 NC	NA	NA			87.8 (15.8)[a]	NA	NA
Anderson and Doyle[22] (2003)	250 ELBW	26.7 (1.9)	884.0 (162.0)	8.7 (0.3)	WRAT	89.2 (14.3)	96.6 (16.0)	94.4 (12.6)
	217 NC	39.3 (1.4)	3407.0 (443.0)	8.9 (0.4)		98.0 (13.4)	103.3 (14.7)	100.0 (13.3)
Grunau et al[23] (2002)	74 VPT	26.0	718.8	9.0 (8.4–12.5)	WRAT	90.3 (11.0)	94.5 (16.5)	NA
	30 NC	40.0	3540.0	9.3 (9.0–10.0)		99.9 (10.5)	107.0 (14.1)	NA
Grunau et al[24] (2004)	53 ELBW	25.8	719.0	17.3 (16.3–19.7)	WRAT	91.4 (13.6)[a]	103.9 (10.2)[a]	100.2 (13.5)[a]
	31 NC	40.0	3506.0	17.8 (16.5–19.0)		106.3 (14.5)[a]	110.6 (10.2)[a]	105.33 (12.2)[a]
Hack et al[25] (2002)	242 VLBW	29.7 (0.2)	1179.0 (219.0)	20.0	WJ-TOA	89.0 (13.2)[a]	95.8 (19.5)[a]	NA
	233 NC	NA	3279.0 (584.0)			95.18 (14.4)[a]	102.7 (21.0)[a]	
Litt et al[26] (2005)	31 at <750 g	27.7 (2.1)	964.6 (149.6)	11.2 (1.2)	WJ-TOA	100.6 (14.4)[a]	101.8 (11.7)[a]	NA
	41 at 750–1499 g	NA	3390.8 (623.6)	11.1 (1.3)				
	52 NC			11.2 (1.1)	WIAT	105.3 (10.3)[a]	105.3 (12.8)[a]	NA
Kilbride et al[27] (2004)	25 ELBW	26.0 (1.6)	702.0 (76.0)	5.0 (.3)	WRAT	74.0 (15.0)	81.0 (13.0)	69.0 (18.0)
	25 NC	38.8 (1.5)	3215.0 (509.0)			81.0 (17.0)	87.0 (9.0)	84.0 (18.0)
Rickards et al[28] (2001)	120 VLBW	29.3 (2.0)	1167.0 (215.0)	14.0	WRAT	89.0 (13.8)	96.8 (14.4)	93.7 (16.2)
	41 NC	39.9 (1.0)	3417.0 (432.0)			95.9 (13.6)	100.4 (12.7)	98.6 (13.8)
Saigal et al[29] (2000)	150 VPT	27.0	833.0 (126.0)	14.0 (1.6)	WRAT	75.0 (18.0)	85.0 (21.0)	83.0 (20.0)
	124 NC	NA	3395.0 (483.0)	14.4 (1.3)		92.0 (15.0)	101.0 (15.0)	101.0 (15.0)
Short et al[30] (2003)	75 VLBW	30.0 (2.0)	1256.0 (176.0)	8.0	WJ-TOA	98.9 (17.5)[b]	100.3 (18.0)[b]	NA
	99 NC	40.0 (1.0)	3451.0 (547.0)			109.3 (17.0)[b]	105.1 (18.0)[b]	NA
Taylor et al[31] (2006)	219 ELBW	26.4 (0.2)	810.0 (124.0)	8.7 (.6)	WJ-TOA	88.2 (15.6)	88.6 (17.7)	88.2 (19.1)
	176 NC	NA	3300.0 (513.0)	9.2 (.8)		98.1 (13.6)	95.7 (13.7)	95.2 (12.7)
Taylor et al[32] (2000)	65 at <750 g	25.7 (1.8)	665.6 (68.2)	11.0 (1.1)	WJ-TOA	87.6 (24.2)	93.9 (18.9)[a]	NA
	54 at 750–1499 g	29.4 (2.4)	1173.2 (217.1)	11.1 (1.3)				NA
	49 NC	40.0	3360 (660.0)	11.2 (1.2)		103.2 (12.7)[a]	105.6 (14.8)[a]	NA
Downie et al[33] (2005)	39 VPT	25.8 (1.4)	815.0 (149.0)	11.5 (1.3)	WRMT-R	NA	94.8 (9.1)	97.7 (11.4)
	11 NC	40.6 (1.4)	3842.0 (697.0)	12.1 (1.1)			102.5 (8.4)	107.6 (7.4)
Gross et al[34] (2001)	118 VPT	28.3	1164.6	10.1	WIAT	94.8 (9.0)	90.5 (10.3)	91.3 (10.6)
	119 NC	NA	NA	10.1		96.2 (9.9)	99.6 (11.3)	100.5 (10.4)

NA indicates not available; NC, normal control; ELBW, extremely low birth weight; VPT, very preterm; WJ-TOA, Woodcock-Johnson Tests of Achievement; WRAT, Wide Range Achievement Test; WIAT, Wechsler Individual Achievement Test; WRMT-R, Woodcock Reading Mastery Test-Revised.

[a] Means and SDs are weighted.

[b] Mean subtest scores are averaged.

Achenbach's Child Behavior Checklist and Teachers Report Form.[35] For the purposes of the meta-analysis we clustered participants' behavioral problems following the taxonomy developed by Achenbach,[35] which distinguishes the broadband scales internalizing behavioral problems (eg, anxiety or depression) and externalizing behavioral problems (eg, oppositional behavior). In addition, we examined the narrow-band scale attention problems, because very preterm and/or VLBW children have been reported to show these symptoms in particular.[3] In case of missing data, authors were contacted.[5,28,32,36–38] Some authors were not able to provide missing data[5,28] or could not be reached.[39] These studies, therefore, were not included in the meta-analysis. Details on the 9 studies included are provided in Table 2.[5,24,28,32,36–40]

Executive Function

Twelve publications met the inclusion criteria. EF tests that were used in these studies included the Controlled Word Association Test,[41,42] Animal Naming Test,[43] Digit Span,[44,45] and the Trail-Making Test.[46] The Controlled Word Association Test and Animal Naming Test measure letter and semantic fluency, respectively, which are both components of verbal fluency. Verbal fluency is the ability to quickly generate as many different solutions for a particular (verbal) problem as possible[42] and also involves heavy linguistic requirements. Both tests were used in each of the studies on verbal fluency and are identical in test administration, response mode, and scoring,[42] and for the purposes of this meta-analysis, a mean verbal fluency score was calculated for each study. Digit Span is a test of working memory, in which series of digits are read aloud to the child.[47] Digits forward requires repetition of series of digits in the same order, whereas digits backward requires repetition of series of digits in reverse order.[47] The total number of correctly repeated series on digits forward and backward served as an index for working memory. The Trail-Making Test is a test that measures cognitive

TABLE 2 Studies That Reported on Behavioral Problems in Very Preterm and/or VLBW Children

Studies	Participants	GA, Mean (SD), wk	BW, Mean (SD), g	Age, Mean (SD), y	Questionnaire	Behavioral Problems		
						Externalizing Problems	Internalizing Problems	Attention Problems
Greenley et al[40]	48 at <750 g	25.8 (1.8)	660.3 (72.8)	11.2 (1.5)	CBCL	VLBW = 47.9 (9.7)	VLBW = 49.9 (10.1)	NA
(2007)	46 at 750–1499 g	29.4 (2.4)	1169.0 (215.1)	11.1 (1.3)		NC = 48.1 (11.5)	NC = 49.6 (11.6)	NA
	51 NC			11.2 (1.3)	TRF	VLBW = 51.2 (8.8)	VLBW = 52.3 (8.6)	NA
						NC = 50.4 (8.0)	NC = 51.2 (8.7)	NA
Farooqi et al[36]	83 EPT	24.6 (0.7)	765.0 (111.0)	10.9 (.8)	CBCL	ELBW = 50.9 (12.8)	ELBW = 57.5 (16.8)	ELBW = 61.7 (16.8)
(2007)	86 NC	39.2 (2.7)	3520.0 (601.0)	11.6 (.8)		NC = 49.2 (13.0)	NC = 48.8 (8.8)	NC = 51.5 (13.9)
					TRF	ELBW = 52.5 (12.7)	ELBW = 55.4 (10.4)	ELBW = 56.8 (11.6)
						NC = 50.0 (9.9)	NC = 50.0 (9,9)	NC = 50.0 (9.9)
Grunau et al[24]	53 ELBW	25.8	719.0	17.3 (16.3–19.7)	CBCL	ELBW = 50.1 (11.2)	ELBW = 53.7 (11.3)	ELBW = 57.8 (8.2)
(2004)	31 NC	40.0	3506.0	17.8 (16.5–19.0)		NC = 44.0 (8.9)	NC = 46.9 (14.5)	NC = 51.6 (3.1)
Weindrich et al[38]	29 VLBW	30.7 (2.0)	1212.0 (185.0)	10.9 (.1)	CBCL	VLBW = 51.6 (8.2)	VLBW = 53.5 (11.6)	VLBW = 58.1 (9.8)
(2003)	112 NC	39.9 (1.1)	3344.0 (382.0)	10.9 (.2)		NC = 51.3 (10.4)	NC = 52.6 (8.8)	NC = 54.6 (6.7)
						VLBW = 49.7 (8.1)	VLBW = 54.9 (11.7)	VLBW = 54.7 (5.6)
						NC = 51.1 (9.7)	NC = 51.1 (9.5)	NC = 53.4 (5.4)
Saigal et al[5]	141 ELBW	27.0 (2.4)	838.0 (123.0)	14.1 (1.5)	CBCL	NA	NA	NA
(2003)	122 NC	NA	3391.0 (48.0)	14.4 (1.2)				NA
Rickards et al[28]	120 VLBW	29.3 (2.0)	1167.0 (215.0)	14.0	CBCL	NA	NA	NA
(2001)	41 NC	39.9 (1.0)	3417.0 (432.0)	14.0				NA
Nadeau et al[37]	61 EPT	27.4 (1.1)	1024.3 (204.2)	7.0	CBCL	EPT = 50.9 (8.8)	EPT = 52.4 (10.0)	EPT = 57.7 (8.7)
(2001)	44 NC	39.8 (1.6)	3453.4 (497.8)		TRF	NC = 53.3 (9.7)	NC = 53.3 (10.6)	NC = 56.1 (8.3)
						EPT = 50.5 (8.4)	EPT = 54.4 (9.2)	EPT = 55.3 (7.4)
						NC = 50.9 (9.7)	NC = 53.3 (10.3)	NC = 53.1 (5.1)
Taylor et al[32]	60 < 750 g	25.7 (1.8)	665.6 (68.2)	11.0 (1.1)	CBCL	VLBW = 48.4 (9.8)	VLBW = 49.6 (9.9)	VLBW = 56.9 (8.4)
(2000)	55 750–1499 g	29.4 (2.4)	1173.2 (217.1)	11.0 (1.3)	TRF	NC = 46.6 (11.1)	NC = 48.0 (11.4)	NC = 52.4 (5.1)
	49 NC	40.0	3360 (660.0)	11.0 (1.2)		VLBW = 51.5 (8.9)	VLBW = 52.7 (8.6)	VLBW = 56.0 (7.2)
						NC = 50.2 (8.0)	NC = 51.0 (8.0)	NC = 53.7 (5.3)
Stjernqvist and	61 EPT	27.1 (1.1)	1042 (252.0)	10.5 (0.6)	CBCL	NA	NA	NA
Svenningsen[39]	61 NC	40.1 (1.4)	3648 (533.0)	10.6 (0.6)		NA	NA	NA
(1999)								

EPT indicates extremely preterm; NA, not available; NC, normal control; ELBW, extremely low birth weight.

flexibility[48] and involves switching between mental sets.[42] In part A of this test, the child needs to draw lines to connect consecutively numbered circles; in part B, the child has to connect consecutively numbered circles and lettered circles while alternating between the 2 sequences.[42] The score on part B of the Trail-Making Test served as an index for cognitive flexibility.

If data of 2 measurements pertaining to a partially overlapping sample had been reported,[49] results of the first measurement were included in our meta-analysis to avoid retest effects that would confound our results. Studies were excluded if they did not report scores for either the Controlled Word Association Test and/or the Animal Naming Test, separately.[50,51] Details on the studies included are provided in Table 3.[28,49,50,52–60]

Statistical Analyses

Meta-analysis was conducted by using the computer program Comprehensive Meta-analysis.[61] For studies that reported results for subgroups of very preterm and/or VLBW children or controls, we calculated a weighted group mean and weighted SD by multiplying each subgroup mean and SD, respectively, by its sample size, adding the subtotals, and dividing the obtained sum by the total sample size.[24,25,32–34,51] Most dependent measures were not standardized. Hence, the variability metric for the dependent measures differed both between studies and between groups within studies (very preterm and/or VLBW children and controls). Therefore, we calculated effect sizes and 95% confidence intervals (CIs) in terms of Cohen's *d* for each study separately.

Cohen's *d* is defined by the difference between 2 means divided by the pooled SD for those means.[62] Combined effect sizes for each of the dependent variables of the 3 outcome domains were computed by weighting the domain-specific effect sizes according to the studies' sample sizes. Cohen's guidelines were followed to indicate the strength of the combined effect sizes, with 0.20, 0.50, and 0.80 referring to small, medium, and large effect sizes, respectively.[62]

Q-test statistics[63] were performed to test homogeneity among the studies' effect sizes (ie, whether findings are consistent among studies) and among combined effect sizes for the various indices of academic achievement, behavioral problems, and EF.

Pearson's correlation coefficients (*r*) were calculated to test the impact of

TABLE 3 Studies That Reported on EF in Very Preterm and/or VLBW Children

Studies	Participants	GA, Mean (SD), wk	BW, Mean (SD), g	Age, Mean (SD), y	Type of Test	EF Domain	Test Scores, Mean (SD)
Narberhaus et al[53] (2008)	52 VPT	29.7 (2.0)	1273.0 (337.7)	14.2 (1.7)	COWAT	Phonetic fluency	28.0 (7.9)
	50 NC	39.6 (1.5)	3421.0 (428.0)	14.3 (2.2)			33.2 (10.4)
Narberhaus et al[53] (2008)	52 VPT	29.7 (2.0)	1273.0 (337.7)	14.2 (1.7)	TMT-B	Cognitive flexibility	54.4 (26.7)
	50 NC	39.6 (1.5)	3421.0 (428.0)	14.3 (2.2)			41.2 (21.3)
Narberhaus et al[53] (2008)	52 VPT	29.7 (2.0)	1273.0 (337.7)	14.2 (1.7)	Digit Span	Working memory	9.5 (3.4)
	50 NC	39.6 (1.5)	3421.0 (428.0)	14.3 (2.2)			11.2 (2.6)
Narberhaus et al[53] (2008)	52 VPT	29.7 (2.0)	1273.0 (337.7)	14.2 (1.7)	ANT	Category fluency	19.0 (5.1)
	50 NC	39.6 (1.5)	3421.0 (428.0)	14.3 (2.2)			21.5 (4.2)
Nosarti et al[54] (2007)	61 VPT	29.5 (1.8)	1296.0 (295.8)	22.3 (1.1)	COWAT	Phonetic fluency	39.3 (13.0)
	64 NC	NA	NA	23.2 (1.5)			50.8 (13.5)
Nosarti et al[54] (2007)	61 VPT	29.5 (1.8)	1296.0 (295.8)	22.3 (1.1)	ANT	Category fluency	43.7 (13.2)
	64 NC	NA	NA	23.2 (1.5)			50.5 (12.6)
Nosarti et al[54] (2007)	61 VPT	29.5 (1.8)	1296.0 (295.8)	22.3 (1.1)	TMT-B	Cognitive flexibility	66.4 (24.5)
	64 NC	NA	NA	23.2 (1.5)			56.6 (19.0)
Allin et al[49] (2008)	94 VPT	NA	NA	15.5 (0.7)	ANT	Category fluency	19.9 (5.3)
	44 NC			15.0 (0.7)			19.3 (4.6)
Allin et al[49] (2008)	94 VPT	NA	NA	15.5 (0.7)	COWAT	Phonetic fluency	28.7 (9.0)
	44 NC			15.0 (0.7)			32.9 (8.9)
Shum et al[52] (2008)	45 VPT	26.4 (1.9)	838.2 (151.7)	8.3 (0.9)	TMT-B	Cognitive flexibility	84.7 (43.7)
	49 NC	39.9 (1.5)	3577.8 (516.5)	8.2 (0.9)			63.3 (42.1)
Caldú et al[55] (2006)	25 VPT	29.5 (2.5)	NA	13.4 (1.9)	COWAT	Phonetic fluency	27.1 (8.4)
	25 NC	39.9 (1.4)		13.9 (2.5)			32.1 (11.8)
Caldú et al[55] (2006)	25 VPT	29.5 (2.5)	NA	13.4 (1.9)	ANT	Category fluency	16.4 (4.1)
	25 NC	39.9 (1.4)		13.9 (2.5)			21.3 (4.1)
Giménez et al[56] (2006)	30 VPT	29.1 (2.0)	1107.8 (240.3)	14.3 (2.0)	COWAT	Phonetic fluency	28.0 (8.5)
	30 NC	NA	NA	14.1 (2.0)			32.6 (8.8)
Giménez et al[56] (2006)	30 VPT	29.1 (2.0)	1107.8 (240.3)	14.3 (2.0)	ANT	Category fluency	16.7 (3.6)
	30 NC	NA	NA	14.1 (2.0)			21.2 (4.7)
Kulseng et al[57] (2006)	54 VLBW	28.9 (2.7)	1178.0 (234.0)	14.1 (0.3)	TMT-B	Cognitive flexibility	46.7 (22.0)
	83 NC	39.6 (1.2)	3690.0 (458.0)	14.2 (0.3)			31.9 (18.6)
Anderson and Doyle[58] (2004)	298 ELBW/VPT	26.7 (1.9)	884.0 (162.0)	8.7 (0.3)	Digit Span	Working memory	8.5 (2.8)
	223 NC	39.3 (1.4)	3407.0 (443.0)	8.9 (0.4)			9.5 (2.9)
Foulder-Hughes and Cooke[59] (2003)	280 VPT	29.8 (23.0–32.0)	1467.0 (424.0)	7.5	Digit Span	Working memory	8.6 (2.7)
	210 NC	—	—	7.5			10.0 (3.0)
Rushe et al[50] (2001)	75 VPT	29.6 (1.8)	1299.0 (284.0)	14.9 (0.4)	TMT-B	Cognitive flexibility	75.0 (24.5)
	53 NC	NA	NA	14.9 (0.6)			69.2 (25.2)
Rickards et al[28] (2001)	120 VLBW	29.3 (2.0)	1167.0 (215.0)	14.0	Digit Span	Working memory	9.9 (3.6)
	41 NC	39.9 (1.0)	3417.0 (432.0)				9.8 (3.6)
Rushe et al[50] (2001)	75 VPT	29.6 (1.8)	1299.0 (284.0)	14.9 (0.4)	Digit Span	Working memory	13.6 (2.9)
	53 NC	NA	NA	14.9 (0.6)			14.2 (4.1)
Olsen et al[60] (1998)	42 VPT	31.0	1410.0	8.0	Digit Span	Working memory	9.3 (2.1)
	42 NC	39.0	3323.0				9.9 (2.6)

VPT indicates very preterm; ELBW, extremely low birth weight; ANT, Animal Naming Test; COWAT, Controlled Word Association Test; NA, not available; NC, normal control; TMT-B, Trail-Making Test part B.

mean BW, mean GA, and mean age at assessment on the strength of the studies' effect sizes for all indices of academic achievement, behavioral problems, and EF. Cohen's guidelines were followed to indicate the strength of the correlation coefficients, with 0.10, 0.30, and 0.50 referring to small, medium, and large coefficients, respectively.[64]

A major concern in conducting meta-analyses is the existence of publication bias. Publication bias results from studies reporting nonsignificant results that have failed to be published and, therefore, are not included in a meta-analysis. If these studies had been included, they would nullify observed effects.[16] We examined the potential for publication bias by using 2 methods. First, we computed Rosenthal's fail-safe *N*[16] (FSN) (ie, the number of studies that would be required to nullify the observed effect) for each combined effect size, separately. A FSN is often considered robust if it is more than $5k + 10$ (k = number of studies in the meta-analysis).[16] Second, we correlated sample sizes to the effect sizes. A negative correlation between sample sizes and effect sizes indicates that small studies with significant results may be published more often than small studies with nonsignificant results, which has recently been shown to exist in 80% of the meta-analyses.[65]

TABLE 4 Sample Sizes, Number of Studies, Combined Effect Sizes in Terms of Cohen's *d*, 95% CIs, Heterogeneity Statistics, Correlations With Sample Sizes, and FSNs for Outcome Measures

	Sample Sizes	No. of Studies	*d*	95% CI	*P*	*Q*	*P*	FSN	*r*
Academic achievement									
Mathematics	2753	13	−0.60	−0.74, −0.46	<.001	34.59	<.001	705	0.03
Reading	2639	13	−0.48	−0.60, −0.34	<.001	26.21	.01	417	0.31
Spelling	1251	8	−0.76	−1.13, −0.40	<.001	80.76	<.001	355	0.22
Behavioral problems									
CBCL internalizing	930	6	−0.20	−0.48, 0.08	.16	17.63	<.001	18	−0.16
TRF internalizing	920	5	−0.28	−0.45, −0.12	<.01	4.32	.37	10	−0.54
CBCL externalizing	930	6	−0.09	0.05, 0.22	.22	8.64	.13	3	0.26
TRF externalizing	920	5	−0.08	−0.24, 0.07	.30	2.46	.65	0	−0.87
CBCL attention	930	5	−0.59	−0.74, −0.44	<.001	6.95	.14	67	−0.13
TRF attention	920	4	−0.43	−0.61, −0.25	<.001	2.76	.43	17	−0.74
EF									
Verbal fluency	475	5	−0.57	−0.82, −0.32	<.001	6.70	.15	41	0.81
Working memory	1580	7	−0.36	−0.47, −0.20	<.001	9.09	.17	56	0.33
Cognitive flexibility	586	5	−0.49	−0.66, −0.33	<.001	4.03	.41	39	−0.06

Negative effect sizes indicate underperformance on academic achievement and EF tests and higher ratings of behavioral problems for very preterm and/or VLBW children in comparison to controls. CBCL indicates Child Behavior Checklist; TRF, Teachers Report Form.

RESULTS

Table 4 depicts the sample sizes, number of studies, combined effect sizes in terms of Cohen's *d*, 95% CIs, *Q*-test statistics, FSNs, and correlations with sample sizes for effect sizes pertaining to academic achievement, behavioral problems, and EF.

Academic Achievement

Mathematics, reading, and spelling were significantly poorer in very preterm and/or VLBW children. Combined effect sizes were −0.48 for reading, −0.60 for mathematics, and −0.76 for spelling. The combined effect sizes for mathematics and spelling were medium to close to large and did not differ significantly ($Q_1 = 2.41$; $P = .12$). The combined effect size for reading, however, was significantly lower than the combined effect sizes for mathematics ($Q_1 = 5.73$; $P = 0.02$) and spelling ($Q_1 = 12.47$; $P < .001$). Within each of the indices for academic achievement, strength of the studies' effect sizes varied significantly between studies (*P* values < .01). FSNs ranged from 355 to 705, and small-to-medium, albeit nonsignificant, correlation coefficients were observed between sample sizes and indices for academic achievement (all *P* values > .32), indicating that there was no evidence for publication bias.

Behavioral Problems

Parents and teachers did not differ significantly in their ratings of internalizing behavioral problems ($Q_1 = 0.02$; $P = .88$), externalizing behavioral problems ($Q_1 = 0.007$; $P = .93$), and attention problems ($Q_1 = 1.95$; $P = .16$).

Significant (*P* values < .001) and close-to-medium combined effect sizes were found for parent and teacher ratings of attention problems: −0.59 and −0.43, respectively. Small combined effect sizes were found for parent and teacher ratings of internalizing behavioral problems, which were −0.20 ($P < .01$) and −0.28 ($P = .16$), respectively, and for externalizing behavioral problems, which were −0.08 and −0.09 and not significant (*P* values > .22). Parent and teacher ratings for attention problems were significantly larger than parent and teacher ratings of externalizing and internalizing behavioral problems ($Q_1 > 12.09$; $P < .001$). Within parent and teacher ratings, combined effect sizes for attention problems, internalizing behavioral problems, and externalizing behavioral problems did not differ significantly ($Q_1 < 3.03$; *P* values > .08). Except for parent ratings of internalizing behavioral problems, findings were consistent across studies.

FSNs for parent and teacher ratings of internalizing behavioral problems were 18 and 10, respectively, for parent and teacher ratings of externalizing behavioral problems were 3 and 0, respectively, and for parent and teacher ratings of attention problems were 67 and 17, respectively. Nonsignificant, small correlations were observed between sample sizes and parent ratings of internalizing and externalizing behavior problems and attention problems (all *P* values > .61). Nonsignificant, albeit large and negative, correlations were observed between sample sizes and teacher ratings of internalizing and externalizing behavior problems and attention problems (all *P* values > .08). The results point to possible publication bias in studies on teacher ratings of problem behavior.

Executive Function

Verbal fluency (Controlled Word Association Test and Animal Naming Test), working memory (Digit Span), and cognitive flexibility (Trail-Making Test part B) were significantly poorer in children born very preterm and/or with VLBW than in controls. The combined

TABLE 5 Pearson's Correlation Coefficients Between Outcome Measures and Age at Assessment, BW, and GA

	N	Age		BW[a]		GA[a]	
		r	*P*	*r*	*P*	*r*	*P*
Academic achievement							
Mathematics	11	−0.19	.55	0.60[b]	.02[b]	0.51[b]	.05[b]
Reading	13	0.09	.77	0.70[b]	.01[b]	0.65[b]	.01[b]
Spelling	8	−0.16	.72	0.43[c]	.17[c]	0.42[c]	.18[c]
Behavioral problems							
CBCL internalizing	6	−0.56[c]	.33[c]	0.71[b]	.06	0.82[b]	.03[b]
TRF internalizing	5	−0.54	.35	0.18	.39	0.25	.34
CBCL externalizing	6	−0.37[c]	.54[c]	0.56	.13	0.47	.18
TRF externalizing	5	−0.06	.93	0.98[b]	.002[b]	0.93[b]	.01[b]
CBCL attention	5	−0.47[c]	.53[c]	0.71[b]	.09[b]	0.45	.23
TRF attention	4	−0.31	.70	0.91[b]	.05[b]	0.94[b]	.03[b]
EF							
Verbal fluency	5	−0.04	.95	NA[d]	NA[d]	NA[d]	NA[d]
Working memory	7	0.33	.47	0.43[c]	.24[c]	0.03	.48
Cognitive flexibility	5	0.17	.79	0.24	.35	0.19	.38

CBCL indicates Child Behavior Checklist; *N*, number of studies; NA, not available; TRF, Teachers Report Form.

[a] Given the hypothesis that a decrease in BW and GA is associated with higher combined effect sizes, and the fact that the small number of studies included for some indices might reduce statistical power, 1-tailed tests of significance were conducted.

[b] Significant and trend correlations.

[c] Results after omission of 1 extreme effect size.

[d] Correlation coefficients for verbal fluency were not calculated because the values for GA for the pertinent studies ranged from 29.0 to 30.0 weeks and the values for BW ranged from 1107.0 to 1296.0 g; therefore, findings might be unreliable because of restriction of range.

effect sizes were small to medium and were −0.36 for working memory, −0.49 for cognitive flexibility, and −0.57 for verbal fluency (all P values $< .001$). Differences between the combined effect sizes for these indices of EF were not significant ($Q_2 = 6.33$; $P = .10$). Within these indices of EF, effect sizes did not vary significantly between studies (all P values $> .15$). FSNs ranged from 39 to 56. Correlations observed between sample sizes and effect sizes for EF ranged from small ($r = -0.06$) to large ($r = 0.81$) but were not significant (all P values $>$ .10). There was no clear evidence for publication bias.

Age at Assessment

Table 5 displays Pearson's correlation coefficients for the relationship between mean age at assessment and the studies' effect sizes for academic achievement, behavioral problems, and EF. All correlation coefficients for the relationship between effect sizes for academic achievement and mean age at assessment (5.0–20.0 years) and EF and mean age at assessment (7.5–22.3 years) were small and not significant (all r values less than −0.19; all P values $> .55$). After exclusion of 1 extreme effect size,[45] which would confound the results, correlations between parent and teacher ratings of internalizing, externalizing, and attention problems, and mean age at assessment (5.9–17.3 years), ranged from small to large but were not significant (all r values less than −0.56; all P values $> .33$).

BW and GA

Table 5 displays Pearson's correlation coefficients for the relationship between mean BW and mean GA and the studies' effect sizes for academic achievement, behavioral problems, and EF. Mean BW (702–1265 g) and mean GA (25.8–30.0 weeks) were strongly and positively correlated with studies' effect sizes for mathematics and reading (all r values > 0.51; all P values $< .05$). After exclusion of 1 extreme effect size,[34] correlations between mean BW (702.0–1176.0 g), mean GA (25.8–29.3 weeks), and spelling were small and not significant (r values < 0.43; P values $> .17$).

Mean GA (24.6–30.7 weeks) was strongly and positively correlated with parent ratings of internalizing behavior problems and teacher ratings of externalizing behavioral problems and attention problems (all r values > 0.82; all P values $< .03$). Mean BW (765.0–1212.0 g) was strongly and positively correlated with teacher ratings of externalizing behavioral problems and attention problems (r values > 0.91; P values $< .05$). There was a trend toward a significant association between mean BW (719.0–1212.0 g) and parent ratings of internalizing behavioral problems (r values $= 0.71$; P values $\geq .06$) and attention problems (r values $= 0.71$; P values $\geq .09$). Mean BW (719.0–1212.0 g) was not correlated with effect sizes for teacher ratings of internalizing problems or parent ratings of externalizing problems, and mean GA (24.6–30.7 weeks) was not correlated with effect sizes for teacher ratings of internalizing behavioral problems or parent ratings of externalizing behavioral and attention problems (all r values < 0.56; all P values $> .13$).

Correlation coefficients for verbal fluency were not calculated, because the obtained results might be unreliable because of restriction of range for BW and GA. After exclusion of 1 extreme effect size,[28] which would confound the results, mean BW (838.3–1467.0 g) and mean GA (26.4–31.0 weeks) were not significantly correlated with effect sizes for working memory (r values < 0.43; P values $> .24$). Mean BW (838.3–1299.0 g) and mean GA (26.4–29.7 weeks) were not correlated with effect sizes for cognitive flexibility (all r values < 0.24; all P values $> .35$).

DISCUSSION

This meta-analysis provides sound evidence for the presence of major difficulties in academic achievement, symptoms of inattention, internalizing behavioral problems, and poor EF in very preterm and/or VLBW children in comparison to controls. The results show that very preterm and/or VLBW children were 0.48 to 0.76 SD behind their term-born peers in reading, mathematics, and spelling, which translates into a 7.2- to 11.4-point decrement for these key academic-achievement areas. Spelling was found to be just as compromised as mathematics; differences between both combined effect sizes were not significant. Previous research has suggested that mathematics was the most pronounced academic-achievement deficit,[2,29] thereby overlooking the major spelling difficulties of very preterm and/or VLBW children.

Attention problems were most pronounced in very preterm and/or VLBW children, with teacher and parent ratings being 0.43 to 0.59 SD, respectively, higher than for controls. Teachers also reported significantly more internalizing behavior problems for these children than for peers. It should be noted, however, that the results for teacher-reported problem behavior should be interpreted cautiously, because there was some evidence for publication bias. Parents and teachers did not differ significantly in their ratings of behavioral problems for very preterm and/or VLBW children. This finding does not, however, imply a high level of agreement at the individual level between informants. Our results indicate that internalizing problems (ie, withdrawn behavior and symptoms of depression) do occur in these children but that these symptoms are not as prominent as symptoms of inattention. Our meta-analysis did not find significantly increased parent and teacher ratings of externalizing problems (ie, delinquent and risk-taking behaviors) in very preterm and/or VLBW children in comparison to their term-born peers, although in a previously conducted meta-analysis by Bhutta et al[3] it was found that 69% of the studies included reported a high prevalence of externalizing behavioral problems. Unclear is, however, whether the authors subsumed attention problems under externalizing behavioral problems. In addition, Bhutta et al[3] conducted a narrative review on behavior and did not take a quantitative meta-analytic approach, which precludes comparison of their results with our findings.

This meta-analytic study was the first to aggregate studies on the neurocognitive domain EF. Although EF covers a variety of capabilities, the majority of studies into very preterm and/or VLBW children have focused on verbal fluency, working memory, and cognitive flexibility, thereby allowing meta-analytic aggregation of findings. Our results show that very preterm and/or VLBW children score 0.36 to 0.57 SD lower than their term-born peers on these measures, differences that translate into small-to-medium effect sizes. These findings indicate that very preterm and/or VLBW children display difficulties in holding information in mind, switching between mental sets, and generating as many different solutions for a particular problem as possible. These EFs have been strongly related to academic achievement and/or behavioral functioning[10–12,66] and might form an explanation of the problems that very preterm and/or VLBW children face in these domains of functioning. However, other well-established EFs of importance for academic and behavioral functioning, such as inhibitory control, which has been considered to be the underlying symptoms of inattention,[11] have only scarcely been assessed in these children. Therefore, in the search toward the understanding of academic underachievement and behavioral problems in very preterm and/or VLBW children, insight into other EF domains may be of great merit.

Smaller and more premature infants were found to be more prone to poor academic achievement and internalizing and externalizing behavior problems than more mature and heavier peers. Despite the small number of studies included in the correlational analyses, significant results were obtained. This bolsters our findings and underlines the importance of BW and GA as predictors for later development. Such an inverse relationship was previously demonstrated for the incidence of major disabilities in very preterm and/or VLBW children[67] and is related to the risk for disruption in cortical development (corticogenesis) and brain connectivity, which increases when BW and GA decrease.[68] For the extremely preterm or extremely low BW infants, adverse concomitant sequelae (such as abnormal cerebral ultrasound findings, chronic lung disease, and postnatal steroid administration), may explain abnormal neurodevelopmental outcomes in addition to BW and GA.[69,70]

It has been questioned whether academic underachievement, behavioral problems, and neurocognitive dysfunction in very preterm and/or VLBW children improve or worsen over time.[6] Some studies have found evidence in support for the idea that the gap between very preterm and/or VLBW children and term-born peers becomes smaller with increasing age.[50,71] Others have compared outcomes at school age and in young adulthood and have suggested that very preterm and/or VLBW teenagers and young adults continue to lag behind term-born peers in terms of cognitive and academic achievement.[25,29] Our results show that the strength of the studies' effect sizes was not significantly related to age at assessment, which suggests that the disadvantage in academic achievement, behavioral sequelae, and neurocognitive function, at least for the age range studied (5.0–22.3 years), remains stable during development and persists into young adulthood. It should be noted that the number of studies we retrieved that assessed very preterm and/or VLBW young adults is scarce ($n = 4$), and studies in this age group are greatly needed. At the same time, it has been found that very preterm young adults are not less satisfied with their lives and do not have lower self-esteem than their peers.[4] Possibly family and environmental factors might alter the subjective experience of the impairments faced by very preterm and/or VLBW young adults.[72]

This meta-analysis has some limitations that need to be considered. It should be noted that some of the correlational analyses were conducted on a small of number of studies and, therefore, have limited power; results may have changed if more studies had been included. For the purpose of this meta-analytic study, we assumed that academic-achievement test scores derived from different measures of academic achievement were comparable because of identical normative scales (mean: 100; SD: 15). This assumption, however, overlooks the possible differences between tests in terms of content and may possibly explain part of the heterogeneity among the effect sizes obtained. In addition, our exclusive focus on internalizing and externalizing problems, as well as attention problems, might have disregarded other types of behavioral problems. Our inclusion criteria did not take the attrition rates of studies into account; however, correlational analyses showed that there was no significant relationship between studies' effect sizes and attrition rates (data not reported; details are available from Mrs Aarnoudse-Moens). Finally, we included children on the basis of BW and GA, which may have caused heterogeneity between studies. However, inclusion of studies on the basis of BW or GA exclusively would have resulted in a limitation of the number of studies available for this meta-analysis.

CONCLUSIONS

This meta-analysis quantitatively aggregated studies into the outcomes of very preterm and/or VLBW children in terms of multiple indices of academic achievement, behavioral functioning, and EF. We combined results from different countries. Despite the cross-cultural differences that exist in such a comparison, this meta-analysis provides evidence from a large number of participants that very preterm and/or VLBW children show severe deficits in mathematics, reading, and spelling and poor EF, and they face behavioral sequelae in terms of symptoms of inattention and internalizing behavioral problems. These adverse outcomes were demonstrated to persist into young adulthood and were inversely related to BW and GA. Our findings highlight the need for long-term follow-up for prematurity and VLBW survivors. In addition, having clearly established these children's areas of weakness, research needs to be performed to study underlying dysfunctions and focus on feasibility and efficacy of intervention strategies to minimize the long-term impact of prematurity and VLBW.

ACKNOWLEDGMENTS

We are most grateful to Drs Farooqi, Laucht, Taylor, Saigal, and Nadeau for providing us with additional data, allowing us to conduct this meta-analysis.

REFERENCES

1. Hille ET, Weisglas-Kuperus N, van Goudoever JB, et al. Functional outcomes and participation in young adulthood for very preterm and very low birth weight infants: the Dutch Project on Preterm and Small for Gestational Age Infants at 19 years of age. *Pediatrics.* 2007;120(3). Available at: www.pediatrics.org/cgi/content/full/120/3/e587
2. Rodrigues MC, Mello RR, Fonseca SC. Learning difficulties in schoolchildren born with very low birth weight. *J Pediatr (Rio J).* 2006;82(1):6–14
3. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA.* 2002;288(6):728–737
4. Hack M, Youngstrom EA, Cartar L, et al. Behavioral outcomes and evidence of psychopathology among very low birth weight infants at age 20 years. *Pediatrics.* 2004;114(4):932–940
5. Saigal S, Pinelli J, Hoult L, Kim MM, Boyle M. Psychopathology and social competencies of adolescents who were extremely low birth weight. *Pediatrics.* 2003;111(5 pt 1):969–975
6. Aylward GP. Neurodevelopmental outcomes of infants born prematurely. *J Dev Behav Pediatr.* 2005;26(6):427–440
7. Saigal S, Doyle LW. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet.* 2008;371(9608):261–269
8. Pharoah PO, Stevenson CJ, Cooke RW, Stevenson RC. Prevalence of behaviour disorders in low birthweight infants. *Arch Dis Child.* 1994;70(4):271–274

9. Sommerfelt K, Ellertsen B, Markestad T. Personality and behaviour in eight-year-old, non-handicapped children with birth weight under 1500 g. *Acta Paediatr.* 1993;82(9):723–728
10. Assel MA, Landry SH, Swank P, Smit KE, Steelman LM. Precursors to mathematical skills: examining the roles of visual-spatial skills, executive processes, and parenting factors. *Appl Dev Sci.* 2003;7(1):27–38
11. Barkley RA. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol Bull.* 1997;121(1):65–94
12. Bull R, Scerif G. Executive functioning as a predictor of children's mathematics ability: inhibition, switching, and working memory. *Dev Neuropsychol.* 2001;19(3):273–293
13. Espy KA, McDiarmid MM, Cwik MF, Stalets MM, Hamby A, Senn TE. The contribution of executive functions to emergent mathematic skills in preschool children. *Dev Neuropsychol.* 2004;26(1): 465–486
14. Anderson V, Levin HS, Jacobs R. Executive functions after frontal lobe injury: a developmental perspective. In: Stuss DT, Knight RT, eds. *Principles of Frontal Lobe Function.* New York, NY: Oxford University Press; 2002:504–527
15. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 2000;283(15):2008–2012
16. Rosenthal R. Writing meta-analytic reviews. *Psychol Bull.* 1995;118:183–192
17. Woodcock RW, McGrew K, Mather N. *Woodcock-Johnson III Tests of Achievement.* Itasca, IL: Riverside; 2001
18. Wilkinson G. *The Wide Range Achievement Test-1993 Edition: Administration Manual.* Wilmington, DE: Wide Range; 1993
19. Wechsler D. *Wechsler Individual Achievement Test.* New York, NY: Psychological Corporation; 1992
20. Woodcock R. *Woodcock Reading Mastery Test-Revised: Examiner's Manual.* Circle Pines, MN: American Guidance Service Inc; 1987
21. Chaudhari S, Otiv M, Chitale A, Pandit A, Hoge M. Pune low birth weight study: cognitive abilities and educational performance at twelve years. *Indian Pediatr.* 2004;41(2):121–128
22. Anderson PJ, Doyle LW. Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. *JAMA.* 2003;289(24):3264–3272
23. Grunau RE, Whitfield MF, Davis C. Pattern of learning disabilities in children with extremely low birth weight and broadly average intelligence. *Arch Pediatr Adolesc Med.* 2002;156(6):615–620
24. Grunau RE, Whitfield MF, Fay TB. Psychosocial and academic characteristics of extremely low birth weight ($\leq$800 g) adolescents who are free of major impairment compared with term-born control subjects. *Pediatrics.* 2004;114(6). Available at: www.pediatrics.org/cgi/content/full/114/6/e725
25. Hack M, Flannery DJ, Schluchter M, Cartar L, Borawski E, Klein N. Outcomes in young adulthood for very-low-birth-weight infants. *N Engl J Med.* 2002;346(3):149–157
26. Litt J, Taylor HG, Klein N, Hack M. Learning disabilities in children with very low birthweight: prevalence, neuropsychological correlates, and educational interventions. *J Learn Disabil.* 2005; 38(2):130–141
27. Kilbride HW, Thorstad K, Daily DK. Preschool outcome of less than 801-gram preterm infants compared with full-term siblings. *Pediatrics.* 2004;113(4):742–747
28. Rickards AL, Kelly EA, Doyle LW, Callanan C. Cognition, academic progress, behavior and self-concept at 14 years of very low birth weight children. *J Dev Behav Pediatr.* 2001;22(1):11–18
29. Saigal S, Hoult L, Streiner DL, Stoskopf BL, Rosenbaum P. School difficulties at adolescence in a regional cohort of children who were extremely low birth weight. *Pediatrics.* 2000;105(2):325–331
30. Short EJ, Klein NK, Lewis BA, et al. Cognitive and academic consequences of bronchopulmonary dysplasia and very low birth weight: 8-year-old outcomes. *Pediatrics.* 2003;112(5). Available at: www.pediatrics.org/cgi/content/full/112/5/e359
31. Taylor HG, Klein N, Drotar D, Schluchter M, Hack M. Consequences and risks of $<$1000-g birth weight for neuropsychological skills, achievement, and adaptive functioning. *J Dev Behav Pediatr.* 2006;27(6):459–469
32. Taylor HG, Klein N, Minich NM, Hack M. Middle-school-age outcomes in children with very low birthweight. *Child Dev.* 2000;71(6):1495–1511
33. Downie AL, Frisk V, Jakobson LS. The impact of periventricular brain injury on reading and spelling abilities in the late elementary and adolescent years. *Child Neuropsychol.* 2005;11(6):479–495
34. Gross SJ, Mettelman BB, Dye TD, Slagle TA. Impact of family structure and stability on academic outcome in preterm children at 10 years of age. *J Pediatr.* 2001;138(2):169–175
35. Achenbach TM. *Manual for the Child Behavior Checklist/4-18 and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry; 1991

36. Farooqi A, Hägglöf B, Sedin G, Gothefors L, Serenius F. Mental health and social competencies of 10- to 12-year-old children born at 23 to 25 weeks of gestation in the 1990s: a Swedish national prospective follow-up study. *Pediatrics*. 2007;120(1):118–133
37. Nadeau L, Boivin M, Tessier R, Lefebvre F, Robaey P. Mediators of behavioral problems in 7-year-old children born after 24 to 28 weeks of gestation. *J Dev Behav Pediatr*. 2001;22(1):1–10
38. Weindrich D, Jennen-Steinmetz C, Laucht M, Schmidt MH. Late sequelae of low birthweight: mediators of poor school performance at 11 years. *Dev Med Child Neurol*. 2003;45(7):463–469
39. Stjernqvist K, Svenningsen NW. Ten-year follow-up of children born before 29 gestational weeks: health, cognitive development, behaviour and school achievement. *Acta Paediatr*. 1999;88(5):557–562
40. Greenley RN, Taylor HG, Drotar D, Minich NM. Longitudinal relationships between early adolescent family functioning and youth adjustment: an examination of the moderating role of very low birth weight. *J Pediatr Psychol*. 2007;32(4):453–462
41. Benton ALH. *Multilingual Aphasia Examination*. Iowa City: University of Iowa; 1976
42. Lezak MD, Howieson DW, Loring DW. *Neuropsychological Assessment*. 4th ed. New York, NY: Oxford University Press; 2004
43. Spreen O, Strauss E. *A Compendium of Neuropsychological Tests*. New York, NY: Oxford University Press; 1991
44. Wechsler D. *Wechsler Intelligence Scale for Children*. 3rd ed. San Antonio, TX: Psychological Corporation; 1991
45. Wechsler D. *Wechsler Memory Scale*. 3rd ed. London, United Kingdom: Psychological Corporation; 1998
46. Reitan RM, Wolfson D. *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. Tucson, AZ: Neuropsychology Press; 1993
47. Sattler JM. *Assessment for Children, WISC-III and WPPSI-R Supplement*. San Diego, CA: Sattler Publisher, Inc; 1992
48. Kortte K, Horner M, Windham W. The Trail Making Test, part B: cognitive flexibility or ability to maintain set? *Appl Neuropsychol*. 2002;9(2):106–109
49. Allin M, Walshe M, Fern A, et al. Cognitive maturation in preterm and term born adolescents. *J Neurol Neurosurg Psychiatry*. 2008;79(4):381–386
50. Rushe TM, Rifkin L, Stewart AL, et al. Neuropsychological outcome at adolescence of very preterm birth and its relation to brain structure. *Dev Med Child Neurol*. 2001;43(4):226–233
51. Taylor HG, Minich NM, Klein N, Hack M. Longitudinal outcomes of very low birth weight: neuropsychological findings. *J Int Neuropsychol Soc*. 2004;10(2):149–163
52. Shum D, Neulinger K, O'Callaghan M, Mohay H. Attentional problems in children born very preterm or with extremely low birth weight at 7–9 years. *Arch Clin Neuropsychol*. 2008;23(1):103–112
53. Narberhaus A, Segarra D, Caldú X, et al. Corpus callosum and prefrontal functions in adolescents with history of very preterm birth. *Neuropsychologia*. 2008;46(1):111–116
54. Nosarti C, Giouroukou E, Micali N, Rifkin L, Morris RG, Murray RM. Impaired executive functioning in young adults born very preterm. *J Int Neuropsychol Soc*. 2007;13(4):571–581
55. Caldú X, Narberhaus A, Junqué C, et al. Corpus callosum size and neuropsychologic impairment in adolescents who were born preterm. *J Child Neurol*. 2006;21(5):406–410
56. Giménez M, Junqué C, Narberhaus A, Botet F, Bargalló N, Mercader JM. Correlations of thalamic reductions with verbal fluency impairment in those born prematurely. *Neuroreport*. 2006;17(5):463–466
57. Kulseng S, Jennekens-Schinkel A, Naess P, et al. Very-low-birthweight and term small-for-gestational-age adolescents: attention revisited. *Acta Paediatr*. 2006;95(2):224–230
58. Anderson PJ, Doyle LW; Victorian Infant Collaborative Study Group. Executive functioning in school-aged children who were born very preterm or with extremely low birth weight in the 1990s. *Pediatrics*. 2004;114(1):50–57
59. Foulder-Hughes LA, Cooke RW. Motor, cognitive, and behavioural disorders in children born very preterm. *Dev Med Child Neurol*. 2003;45(2):97–103
60. Olsén P, Vainionpää L, Pääkkö E, Korkman M, Pyhtinen J, Järvelin MR. Psychological findings in preterm children related to neurologic status and magnetic resonance imaging. *Pediatrics*. 1998;102(2 pt 1):329–336
61. Borenstein M, Rothstein H. *Comprehensive Meta-analysis, A Computer Program for Research Synthesis*. 2nd ed. Englewood, NJ: Biostat Inc; 1999
62. Cohen J. *Statistical Power Analyses for the Behavioral Sciences*. 2nd ed. Hillsdale, NY: Erlbaum; 1988

63. Cochran WG. The combination of estimates from different experiments. *Biometrics.* 1954;10:101–129
64. Cohen J. A power primer. *Psychol Bull.* 1992;112:155–159
65. Levine T, Asada K. Sample sizes and effect sizes are negatively correlated in meta-analyses: evidence and implications of a publication bias against nonsignificant findings. Presented at: annual meeting of the International Communication Association, May 23, 2007; San Francisco, CA
66. Pennington BF, Ozonoff S. Executive functions and developmental psychopathology. *J Child Psychol Psychiatry.* 1996;37(1):51–87
67. Bennett FC, Scott DT. Long-term perspective on premature infant outcome and contemporary intervention issues. *Semin Perinatol.* 1997;21(3):190–201
68. Huttenlocher PR, Dabholkar AS. Regional differences in synaptogenesis in human cerebral cortex. *J Comp Neurol.* 1997;387(2):167–178
69. Piecuch RE, Leonard CH, Cooper BA, Sehring SA. Outcome of extremely low birth weight infants (500 to 999 grams) over a 12-year period. *Pediatrics.* 1997;100(4):633–639
70. Vohr BR, Wright LL, Dusick AM, et al. Neurodevelopmental and functional outcomes of extremely low birth weight infants in the National Institute of Child Health and Human Development Neonatal Research Network, 1993–1994. *Pediatrics.* 2000;105(6):1216–1226
71. Curtis WJ, Lindeke LL, Georgieff MK, Nelson CA. Neurobehavioural functioning in neonatal intensive care unit graduates in late childhood and early adolescence. *Brain.* 2002;125(pt 7):1646–1659
72. Saigal S, Rosenbaum P. What matters in the long-term: reflections on the context of adult outcomes versus detailed measures in childhood. *Semin Fetal Neonatal Med.* 2007;12(5):415–422