Psychological Bulletin
2017, Vol. 143, No. 4, 347–383

0033-2909/17/$12.00 http://dx.doi.org/10.1037/bul0000091

Mental Health of Extremely Low Birth Weight Survivors: A Systematic Review and Meta-Analysis

Karen J. Mathewson, Cheryl H. T. Chow, Kathleen G. Dobson, Eliza I. Pope, Louis A. Schmidt, and Ryan J. Van Lieshout
McMaster University

Although individuals born at extremely low birth weight (ELBW; < 1,000 g) are the most vulnerable of all preterm survivors, their risk for mental health problems across the life span has not been systematically reviewed. The primary objective of this systematic review and meta-analysis was to ascertain whether the risk for mental health problems is greater for ELBW survivors than their normal birth weight (NBW) peers in childhood, adolescence, and adulthood. Forty-one studies assessing 2,712 ELBW children, adolescents, and adults and 11,127 NBW controls were reviewed. Group differences in mental health outcomes were assessed using random effects meta-analyses. The impacts of birthplace, birth era, and neurosensory impairment on mental health outcomes were assessed in subgroup analyses. Children born at ELBW were reported by parents and teachers to be at significantly greater risk than NBW controls for inattention and hyperactivity, internalizing, and externalizing symptoms. ELBW children were also at greater risk for conduct and oppositional disorders, autistic symptoms, and social difficulties. Risks for parent-reported inattention and hyperactivity, internalizing, and social problems were greater in adolescents born at ELBW. In contrast, ELBW teens self-reported lower inattention, hyperactivity, and oppositional behavior levels than their NBW peers. Depression, anxiety, and social difficulties were elevated in ELBW survivors in adulthood. Group differences were robust to region of birth, era of birth, and the presence of neurosensory impairments. The complex needs faced by children born at ELBW continue throughout development, with long-term consequences for psychological and social well-being.

Keywords: premature birth, mental disorders, child, adolescent, adult

Supplemental materials: http://dx.doi.org/10.1037/bul0000091.supp

Preterm births have increased by 30% over the last 2 decades and now make up about 8% of infants born in the United States and Canada (Blencowe et al., 2013; Joseph et al., 1998; Lau, Ambalavanan, Chakraborty, Wingate, & Carlo, 2013). Because of improvements in recent decades in neonatal intensive care, babies who are born at extremely low birth weight (ELBW; born ≤1,000 g) or extremely preterm (EP; ≤ 28 weeks gestational age) have a greater chance of surviving than ever before (e.g., survival rates rose from 25% in 1979 to 73% in 1997, Doyle, 2006; see also Fanaroff, Hack, & Walsh, 2003; Hack & Fanaroff, 1999). However, despite dramatically increased rates of survival, the risk for neurosensory impairment (NSI) has remained largely the same (Hack & Fanaroff, 1999) or has even increased (Hack & Fanaroff, 1999; Wilson-Costello, Friedman, Minich, Fanaroff, & Hack, 2005). As a result, the general focus of research in prematurity has shifted from survival to understanding morbidity and the subtler postnatal impairments associated with preterm birth, with the aim of improving the quality of life of those born extremely preterm (Aarnoudse-Moens, Weisglas-Kuperus, van Goudoever, & Oosterlaan, 2009). As the oldest studied individuals born at ELBW have reached their thirties, it is now possible to document their mental health status from childhood through adulthood. Knowledge of the long-term outcomes of preterm birth is necessary to inform survivors and their families about prognosis, to evaluate the effectiveness of intensive neonatal care (Farooqi, Hägglöf, Sedin, Gothefors, & Serenius, 2007; Szatmari, Saigal, Rosenbaum, & Campbell, 1993), and to plan intervention and prevention strategies.

Accumulating evidence suggests the presence of gradient effects for mental and physical health outcomes within preterm groups, with those born earlier in gestation experiencing higher rates of disorder and disability (Aarnoudse-Moens et al., 2009; Mulder, Pitchford, Hagger, & Marlow, 2009). For example, cognitive prob-

This article was published Online First February 13, 2017.

Karen J. Mathewson, Department of Psychology, Neuroscience and Behavior, McMaster University; Cheryl H. T. Chow, MiNDS Neuroscience Graduate Program, McMaster University; Kathleen G. Dobson, Department of Clinical Epidemiology and Biostatistics, McMaster University; Eliza I. Pope and Louis A. Schmidt, Department of Psychology, Neuroscience and Behavior, McMaster University; Ryan J. Van Lieshout, Department of Psychiatry and Behavioral Neurosciences, McMaster University.

This research was supported by a Team Grant from the Canadian Institutes of Health Research (CIHR, TMH-103145).

Correspondence concerning this article should be addressed to Karen J. Mathewson, Department of Psychology, Neuroscience and Behavior, McMaster University, Hamilton, ON L8S 4K1. E-mail: mathewkj@mcmaster.ca

lems (Hutchinson et al., 2013; Litt, Taylor, Klein, & Hack, 2005), attentional difficulties (e.g., Hanke et al., 2003; Horwood, Mogridge, & Darlow, 1998; Taylor, Klein, Minich, & Hack, 2000), hyperactivity (e.g., McCormick, Workman-Daniels, & Brooks-Gunn, 1996), internalizing (e.g., Greenley, Taylor, Drotar, & Minich, 2007; Horwood et al., 1998), and total psychological problems (e.g., Taylor et al., 2000) appear to increase with decreasing gestational age and/or birth weight. Born at the lowest weights of any preterm group, ELBW infants represent the segment of the preterm population at the greatest risk.

Despite these risks and the impact that psychological problems can have on ELBW survivors and their families, consensus concerning the rates and types of specific mental health difficulties faced by this group remains elusive. Other important questions also remain unanswered, including when these psychological risks first emerge, whether they vary in type and severity by developmental stage, and what factors may reduce or amplify risk, including the availability of high-quality health care after birth and education, changes in neonatal intensive care, and the presence of severe neurosensory impairments such as deafness, blindness, cerebral palsy, and intellectual disability.

Several factors may account for why we do not yet have a clearer picture of the emotional, behavioral, and social difficulties facing ELBW survivors across the life span. First, longitudinally followed samples of ELBW survivors remain small and studies are plagued by high rates of attrition (Saigal et al., 2006). Second, published findings of mental health outcomes in this population display considerable heterogeneity in terms of sample size, age, informants, when and where participants were born, and the tests and treatments to which they were exposed. Most of the extant literature has focused on childhood, with few reports available to help guide those who have survived to adolescence and adulthood. This is particularly unfortunate since more extremely preterm infants are surviving to adulthood than ever before. Finally, most studies have relied on subjective reports from geographically limited areas (e.g., single hospitals), with considerable variability in the measures used to assess mental health.

The literature on this topic has now reached a critical mass, and integration of existing findings is needed to understand the exposures and mechanisms underlying poor mental health in ELBW survivors, and to guide decision making for survivors, their families, and health care professionals. There is general consensus that one of the best and most reliable forms of evidence for addressing complex research questions is a systematic review, with meta-analysis where possible. Not only are quantitative meta-analyses capable of summarizing effects more precisely than descriptive accounts, but they afford examination of much larger samples than are typically possible in individual studies of persons born at ELBW. They also allow comparisons across regions much larger than those covered by single hospitals or geographic areas, thereby permitting questions related to regional disparities, which in turn, may address questions about the biological and social bases of mental health difficulties following preterm birth.

Mental health risks in childhood and adolescence for those born preterm have been examined in two previous systematic reviews, although outcomes were not examined separately for ELBW/EP groups. Both reviews reported an increased risk for cognitive problems in youth born preterm, as well as for attention-deficit/hyperactivity disorder (ADHD) and behavioral difficulties (Aarnoudse-Moens et al., 2009; Bhutta, Cleves, Casey, Cradock, & Anand, 2002). However, the latter review is now 15 years old and reported on cohorts of children born before 1990, before significant advances in neonatal care were widely adopted (e.g., antenatal corticosteroid and postnatal surfactant therapies). Accordingly, it included relatively few children born weighing <1,000 g, and focused mainly on children born at very low birth weight (VLBW; < 1,500 g). In addition, quantitative analyses were applied only to cognitive and attentional problems, in the latter case involving only seven studies. The comprehensive review by Aarnoudse-Moens and colleagues (2009) included studies published up to 2008 and again focused mainly on cognitive functioning in youth born at VLBW. Their quantitative estimates also relied heavily on studies of older cohorts born before 1990 (eight out of nine behavioral studies). Unfortunately, neither review examined child and adolescent outcomes separately, or reported specifically on youth at the highest risk—those born at ELBW—and neither reviewed studies of adults.

In addition to these reports, Johnson and colleagues have more recently provided narrative reviews of the cognitive and mental health problems of school-age children (Johnson, 2007), adolescents (Johnson & Wolke, 2013), and adults (Johnson & Marlow, 2014) born preterm. They concluded that former preterm school-age children and adolescents were more likely than their classmates to have behavioral, attentional, and communication problems, and emotional disorders appeared to remain elevated in emerging adulthood. However, no quantitative estimates of overall risk or severity were provided, VLBW and ELBW cohorts were not always distinguished, and the authors were not able to examine how important sociodemographic, developmental, or secular trends affected risk.

In the present review, we set out to examine factors related to developmental change, region of birth, availability of key neonatal intensive care interventions, the influence of neurosensory impairments on mental health in this population, and informant effects.

Developmental Stages and Change

Relatively few data exist on the developmental trajectory of mental health problems within ELBW individuals over time. However, some evidence suggests that preterm survivors manifest a distinct behavioral phenotype characterized by attention deficits, anxiety, and social problems (e.g., Johnson & Marlow, 2011). In the extremely preterm population, the type and course of emotional and behavioral problems has yet to be elucidated. The development of appropriate interventions and the optimal timing of their application require accurate information about the onset and course of these problems (Szatmari et al., 1993). Systematically gathering, quantifying, and summarizing this information are some of the goals of this review.

Accessibility and Changes in Health Care

Intensive neonatal care is a developing entity that may differ by region and has changed over time. In the developed world, some countries provide universal health care for their citizens and others require payment by insurance or other means. Free and timely access to health care in the neonatal period and beyond may have a major influence on the lifelong trajectory of mental health in

highly vulnerable infants, as can the prevailing social circumstances in the environments in which these individuals develop (Hack & Klein, 2006). Therefore, we wished to examine whether participants' region of birth had any influence on mental health outcomes.

Technological developments have also significantly altered neonatal intensive care practices over the last half century. Between the 1960s and the 1980s, intensive care of preterm newborns was limited to intravenous fluid administration, assisted ventilation (by endotracheal intubation), and oxygen monitoring (Hack & Costello, 2008). The 1990s saw widespread adoption of surfactant to treat or prevent severe respiratory distress syndrome in preterm neonates, and prenatal administration of synthetic glucocorticoids, to maximally develop lung function in fetuses at risk (e.g., Ferrara et al., 1991; Göpel et al., 2015; Hoekstra, Ferrara, Couser, Payne, & Connett, 2004; Pelkonen, Hakulinen, Turpeinen, & Hallman, 1998; Verder et al., 1994). Concomitantly, survival of preterm infants increased during the 1990s in the United States (Fanaroff et al., 1995; Hoekstra et al., 2004), Britain (Riley, Roth, Sellwood, & Wyatt, 2008), and Australia, (Doyle, 2006; Doyle, & the Victorian Infant Collaborative Study Group, 2004). This improved infant survival was largely attributed to the adoption of surfactant (Fanaroff et al., 1995; Horbar, Wright, & Onstad, & The Members of the National Institute of Child Health and Human Development Neonatal Research Network, 1993; Jobe, Mitchell, & Gunkel, 1993; Schwartz, Luby, Scanlon, & Kellogg, 1994), and steroid therapies (Victorian Infant Collective Study Group, 1997), as well as birth in large hospital centers (Hoekstra et al., 2004), and an increased willingness to treat the tiniest infants (Victorian Infant Collective Study Group, 1997).

Experimental trials of surfactant were undertaken in the later 1980s, with 35 randomized controlled trials reported between 1985 and 1992 (Jobe et al., 1993). In 1990, a meta-analysis of randomized controlled trials of antenatal corticosteroids was published (Crowley, Chalmers, & Keirse, 1990), demonstrating improved survival and reduced morbidity in infants born preterm. This was followed in 1994 by a National Institutes of Health recommendation that all women at risk of preterm delivery between 24 and 34 weeks be given antenatal steroids, only to be superseded six years later by a recommendation to restrict repeated administration of corticosteroids to women enrolled in randomized control trials, until it was determined that this now-common practice was both safe and effective (Bonanno & Wapner, 2012).

Although developments in neonatal intensive care continued in the 2000s, (e.g., positive airway pressure, widespread use of caffeine, decreased use of endotracheal intubation, protein enhancement of parenteral feeding, and treatments via the mother, such as antibiotic therapy and cesarean sections; Hack & Costello, 2008), there was no further dramatic increase in survival rates like those of the previous decades (Doyle, Roberts, & Anderson, & the Victorian Infant Collaborative Study Group, 2011; Field, Dorling, Manktelow, & Draper, 2008). As a result, we wished to compare mental health outcomes in cohorts born before and after surfactant and steroid therapies became widely available. To examine the question of whether these therapies may have resulted in improved mental health outcomes, we proposed to examine mental health in those born before and after their introduction, making a division at the year 1990.

Neurosensory Impairment (NSI)

A small but significant proportion of individuals born at ELBW have significant NSI. Although NSI may not strictly constitute an exclusion criterion for investigations of psychiatric functioning (Szatmari et al., 1993), higher rates of neurological disability (e.g., visual and hearing impairment, cerebral palsy, intellectual disability) may profoundly affect the risk for psychopathology in ELBW survivors, particularly depression and anxiety (Johnson & Wolke, 2013). On the other hand, when the goal of a study is to identify the consequences of low birth weight separately from any effects of major neurological damage (Breslau, 1995), then children born extremely preterm who have significant NSI may be excluded. We deemed it valuable to compare the outcomes from these two strategies, by contrasting rates and types of mental health problems in studies from which NSI participants were excluded versus those where they were included.

Informant Effects

Reports from different informants can provide converging evidence, but informants do not always agree, and this variance requires interpretation. First, discrepancies may be informative with respect to the different roles of informants (De Los Reyes, 2011). As the primary adults in their children's lives, parents may be particularly sensitive to their children's emotional problems or behavior, and they are likely to have had many more opportunities to observe them than have other significant adults such as teachers (Pharoah, Stevenson, Cooke, & Stevenson, 1994). Alternatively, discrepancies may accurately represent behavioral adaptations that are appropriate for different contexts, such as the home environment versus the classroom. While attentional and behavioral demands may be relaxed in the informal milieu of the home, attentional and behavioral control is required in situations of formal learning. Second, discrepant information may represent legitimate differences in informants' perspectives. Where a parent or teacher might see particular behaviors as problematic, the child or adolescent in the situation may not, or vice versa. Third, subjective views about mental health and social functioning are important primary sources of information for older children, adolescents, and adults, given that agentic potential over one's own well-being increases with age. To obtain a comprehensive picture of mental health outcomes in ELBW, it is therefore important to compare information from multiple informants, including parents and teachers, and where possible, children.

Developmental Origins of Health and Disease (DOHaD) Hypothesis and Developmental Psychopathology

There is much clinical and scientific interest in the etiology of mental health problems that appear to develop disproportionately in children born at ELBW. Following the evolutionary principle of biological fitness, the DOHaD hypothesis posits that fetal adaptation to environmental influences encountered during gestation may increase susceptibility to chronic health problems later in life. These later-arriving consequences may be viewed as the down side of a necessary trade-off, that is, as distal sequelae of the adaptations that ensured fetal survival during an earlier, more difficult

stage of development (Ellison, 2005; Gluckman, Hanson, & Buklijas, 2010; Nederhof & Schmidt, 2012). However, contrary to the view that prenatal conditions lead inexorably to chronic health problems in later life, the DOHaD framework suggests only that the risk for these outcomes may be increased, as a direct result of the early processes involved in adaptation. While being born at ELBW may increase the risk for chronic health problems, convergence of multiple risk factors over time may be necessary to precipitate actual manifestations of disease or disorder.

The degree of causality between developmental processes during gestation and chronic conditions in later life is of great interest to researchers of mental as well as physical health disorders. Because causality with respect to development cannot ethically be tested by experiment, researchers are left to extrapolate from observational data provided by intriguing "experiments of nature"—one of which is preterm birth—to aid in answering the question of causation. The DOHaD hypothesis might be best tested by prematurity in its most extreme form, in individuals born at the lower limits of viability, either extremely preterm or small for gestational age (SGA). If early, adverse, environmental conditions are likely to affect later psychological functioning, then infants born extremely preterm are more likely than their NBW peers to experience mental health difficulties during the course of development.

Objectives

To our knowledge, no meta-analytic review has attempted to examine mental health problems in children, adolescents, and adults born at ELBW/EP. The present review had six objectives: (1) to identify the mental health problems presenting the greatest risks for individuals born at ELBW/EP by qualitatively and quantitatively summarizing the existing literature on these individuals and age-matched controls; (2) to describe these risks by developmental stage (e.g., childhood, adolescence, adulthood); (3) to explore nonrandom variability across studies by determining whether mental health outcomes in this population varied by region of birth, birth era, or NSI; (4) to examine informant effects on these findings; (5) to use these findings to guide recommendations for future research in the field; and (6) to test the DOHaD hypothesis with respect to mental health outcomes in this unique population.

Method

This systematic review was registered with the PROSPERO International prospective register of systematic reviews CRD42014015491 (Dobson, Van Lieshout, Schmidt, Mathewson, Chow). The Preferred Reporting items for Systematic Reviews and Meta-Analyses statement served as a guideline for its preparation (Moher, Liberati, Tetzlaff, & Altman, & the PRISMA Group, 2009).

Study Selection Criteria

Observational designs such as case-control studies, and prospective longitudinal and cross-sectional studies (excluding reviews) were eligible for this systematic review. The primary exposure of interest was ELBW, defined as a group-wise mean birth weight of $\leq$1,000 g (2.2 lbs). Eligible studies contained participants with a mean age $\geq$5 years at the time of outcome assessment. Secondary exposures included extremely preterm birth (EP; i.e., with a group-wise mean gestational age $\leq 27^{+7}$ weeks). Three studies (Hille et al., 2001; Natalucci et al., 2013; Sternqvist & Svenningsen, 1999) included cohorts whose mean gestational ages were $\leq 28^{+7}$ weeks, but these were retained because they identified an extremely premature group on the basis of birth weight (<1,000 g; Hille et al., 2001; Natalucci et al., 2013), or because the mean gestational age (27.1 week) would have met criteria but for a wide standard deviation (1.03; Sternqvist & Svenningsen, 1999). Other eligibility criteria stipulated that individuals born at ELBW/EP be compared with a control group of participants born at or near normal birth weight (mean weight $\geq$2,500 g), or at, or near full term, but in any case, $\geq$37 weeks' gestational age. Finally, the control group could not constitute an entire population, as this would have biased our meta-analytic models.

Outcomes of primary interest were levels of symptoms of psychological problems (means, standard deviations), or the risk of developing a clinically significant psychological disorder in ELBW survivors (e.g., odds ratios, risk ratios). Where sufficient studies were available, we also examined group differences in specific behavioral and social problems, including conduct disorder, oppositional defiant disorder, and autism spectrum disorder. With respect to the possibility of Type I error, the level for significance was set at $p < .05$.

Information Sources and Search Strategy

A systematic search was performed of electronic databases (MEDLINE, EMBASE, PsycINFO, Web of Science, ERIC, and CINAHL), from their inceptions until May 2016. The list was complemented by hand-searches of the references from relevant reviews and the 41 selected studies. The electronic search strategy included medical subject heading terms (MeSH) in MEDLINE, where keywords and text words were combined.

The electronic search was based on five concepts: (1) term(s) related to ELBW: "extremely low birth weight"; (2) terms related to SGA: "small for gestational age"; (3) terms related to preterm birth: "premature birth," "very premature birth," "extremely premature birth," "preterm birth," "very preterm birth," "extremely preterm birth"; and (4) terms related to mental health: "mental disorders," "personality development," "human development"; and (5) those related to observational study designs: "cohort analysis," "longitudinal study," "epidemiology," "prospective studies," "retrospective studies," "case-control," "cross-sectional studies." Details of the search strategies used for each source are available in the online supplemental material.

The search strategy was developed by the reviewers in consultation with a university research librarian. A similar search strategy was used for each database with appropriate modification of the search terms. There were no search language restrictions. Case studies, abstracts, conference presentations, editorials, unpublished studies (e.g., theses and dissertations), and grey literature (e.g., government reports) were included in the search.

In addition, we conducted searches of several databases to help with identifying grey literature using (i.e., ERIC, Open Grey, The Health Care Management Information Consortium (HMIC) database, The National Technical Information Service, and PsycEXTRA). The progress of abstracts and papers or publications presented at con-

ferences was tracked, but no eligible unpublished studies were identified.

Hand searches of key journals within the field were also carried out, based on the reference lists from the published articles that were deemed eligible for inclusion in the study. These journals included *Journal of Pediatrics, JAMA, JAMA Pediatrics, JAMA Psychiatry, Acta Psychiatrica Scandinavica, Early Human Development, Child Development, Development and Psychopathology, Developmental Neuropsychology, European Journal of Child and Adolescent Psychiatry, European Journal of Pediatrics, Scandinavian Journal of Psychology, Acta Paediatrica, Journal of Developmental and Behavioral Pediatrics, Journal of the American Academy of Child and Adolescent Psychology, Archives of Clinical Neuropsychology, Neuropediatrics, Journal of Pediatric Psychology, Archives of Disease in Childhood, Psychological Medicine,* and the *Journal of the International Neuropsychological Society*.

Study Selection

The screening process was completed independently by three reviewers (CHTC, KGD, KJM) on the basis of the study inclusion criteria outlined above (see Figure 1). Initially, titles and abstracts of each study were screened, at which point duplicate and nonrelevant studies were removed. The full text of potentially relevant studies was further examined to determine whether inclusion criteria were met. In the case of disagreements, a fourth reviewer (RVL) was brought in to aid in resolution. All reviewers met and agreed on the final inclusion of studies ($n = 41$). All of the included studies were published between January 1990 and May, 2016.

Data Extraction

A data extraction form was specifically developed for this review. The form was pilot-tested using two randomly selected studies that met inclusion criteria, and refined if additional studies presented new information not covered by the pilot-tested version. The information extracted from each study included: (a) study methods, (b) sample demographics (e.g., year of birth, location, NSI, socioeconomic status), (c) outcome assessment details (i.e., types of mental disorders, scales used), (d) outcome data (e.g., means and standard deviations or odds ratios) and v) risk of bias assessments. Two reviewers (EIP, KGD) extracted data and information that were independently verified by one other reviewer (KJM).

Risk of Bias in Individual Studies

The risk of bias in the included studies was assessed using the Newcastle–Ottawa Scale (NOS) for cohort studies (Wells et al.,

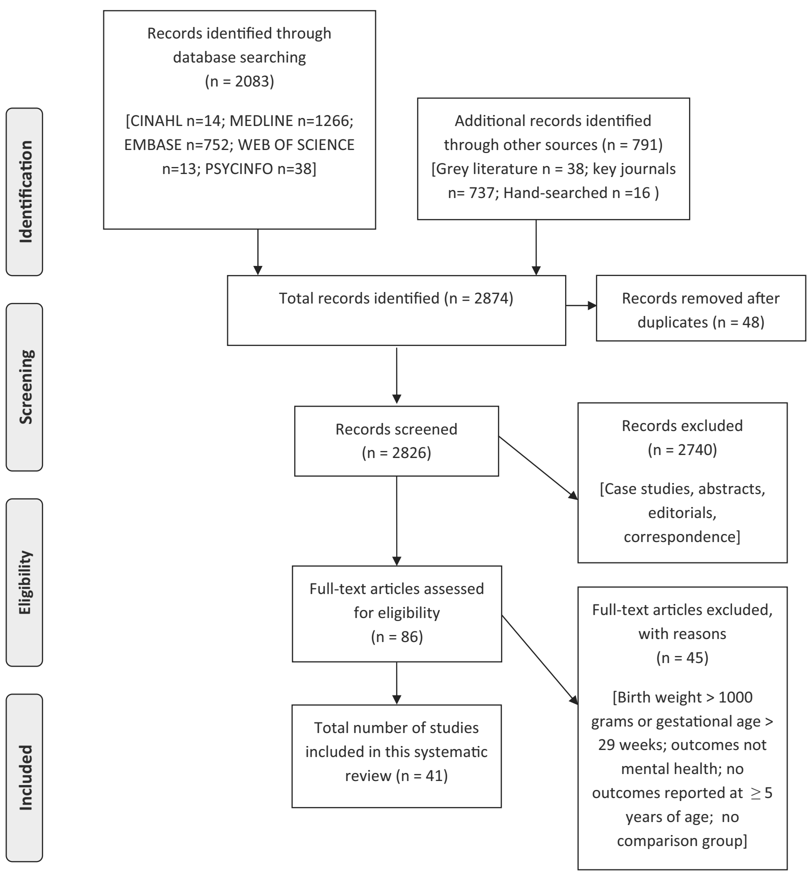

Figure 1. Flowchart of screening process for article selection. See the online article for the color version of this figure.

2000). Two independent reviewers (EIP, KJM) scored each of the 41 studies, and any discrepancies between reviewers were resolved by consensus reached after discussion, with the possibility of consulting a third reviewer if necessary. The median score given by each reviewer was 8, and agreement between the reviewers was calculated as 75.6%. Overall, the distribution of NOS scores ranged from 5 to 9 out of a possible maximum score of 9, with one study scoring 5, one scoring 6, six studies scoring 7, 28 studies scoring 8, and five studies scoring 9. The most common sources of bias in the studies related to the use of subjective reports to assess outcomes, followed distantly by drawing the control cohort from a different source, or inadequate follow-up of cohorts.

Data Synthesis

Meta-analyses. In some cases, multiple studies were published on the same cohort during the same developmental stage. To avoid inappropriate double-counting of participants which may have influenced study weighting, only one study per age category was entered in any meta-analysis. When a choice was required, those studies with the best profile, for example, the largest sample sizes and broadest concept coverage, were selected for inclusion in meta-analyses. Other criteria included the presentation of mental health outcomes using means (and standard deviations) or odds ratios for each group (ELBW, NBW), along with the sample sizes for each group, to enable the calculation of standardized mean differences.

Because the eligible studies varied in many respects, differences in mental health outcomes between ELBW and NBW were assessed using random effects meta-analyses, an approach that assumes the studies included in the analyses are random samples from a larger population of studies, and likely to exhibit different effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009). We calculated mean effect sizes and their confidence intervals, along with prediction intervals derived from the samples to represent the dispersion of effects among the included studies (i.e., the range in which the effect size for any new study is likely to be found; Borenstein et al., 2009; Neyeloff, Fuchs, & Moreira, 2012).

As the included studies assessed psychiatric problems with a variety of different measurement scales, we elected to use the standardized mean difference (SMD) from every study as the effect estimate in our meta-analyses (see Table 1). These effect sizes represent pooled estimates of the differences between ELBW and NBW participants. For studies presenting dichotomous outcomes (e.g., frequencies, odds ratios), SMDs were calculated using the following formula, $\text{SMD} = \sqrt{3}/\pi^{*} \ln(\text{OR})$. The standard error was then calculated by multiplying the SMD by $\sqrt{3}/\pi$. Statistical formulae for converting dichotomous data into SMDs were suggested by Deeks, Higgins, and Altman (2011). These estimates were then converted to *d* for each symptom type, and adjusted to remove any bias accruing from the use of small samples, producing estimates of Hedges' *g* (Borenstein et al., 2009). Standardized mean differences were entered into Review Manager (RevMan 5.3) software as *g* values and their standard errors, for ELBW and NBW groups separately, to compare group differences across studies. Each of the meta-analyses, subgroup, and sensitivity analyses analyzed estimates of *g* from the mental health outcomes data eligible for analysis.

We examined symptoms of the following emotional and behavioral problems in separate analyses: attention problems (e.g., symptoms of the hyperactive/impulsive, inattentive, and combined subtypes of ADHD), internalizing, externalizing, conduct disorder, oppositional disorder, social problems and autism spectrum disorder. The three ADHD subtypes are distinct entities, with the combined subtype incorporating symptoms of both hyperactivity/impulsivity and inattention (Barkley, 1997). Internalizing problems were defined by the internalizing scales of the appropriate instruments (e.g., Child Behavior Checklist [CBCL]; the Strengths and Difficulties Questionnaire [SDQ]), and typically represent a composite of anxiety, withdrawal, and depressive symptoms (Achenbach, 1991; Goodman, 2001). Externalizing is characterized by measures of disruptive and aggressive behavior (e.g., Achenbach, 1991; Goodman, 2001). Hyperactivity that is functionally aligned with attention deficits is thought to be distinct from delinquency and aggression, and so was examined separately from externalizing problems. Given their identity as distinct and meaningful *Diagnostic and Statistical Manual* (*DSM*)-based clinical syndromes (Achenbach, Bernstein, & Dumenci, 2005), and because they were reported separately from externalizing problems in many of the studies in this review, conduct problems and oppositional defiant disorder difficulties were also examined in separate meta-analyses.

The reasons for performing separate analyses by problem-type were several-fold. First, there may have been heterotypic continuity in this population with respect to some disorders, and homotypic continuity for others. Second, each of the specific disorders is separated diagnostically in authoritative references such as the DSM and International Classification of Diseases because they are thought to represent different entities, both theoretically and clinically. Therefore, they were examined separately because we were concerned that combining them could obscure important effects and differences. Third, and perhaps most importantly, we wished this review to not only inform theory, but also include information on specific diagnoses that can guide clinical detection and treatment decision-making for clinicians, survivors, and families, and provide anticipatory guidance for these groups. Understanding the specific problems faced by individuals born preterm and their families can also serve as the basis for resource allocation by health care systems.

For each category of emotional functioning, data were analyzed from as many studies as possible (ranging from 2 to 13), with each unique cohort represented only once in any given analysis. Assessment measures were deemed eligible for meta-analysis by the comparability of their items and concepts. Comparability was based on careful examination of the individual items in each scale from each study, to determine whether measures from different studies were functionally equivalent. Analyses were stratified and examined by age group (child, adolescent, adult), to account for changes related to development. In addition, because some cohorts were represented by studies from more than one developmental period, some studies would have been eliminated from any of the analyses if we had we combined studies from all age groups. For the eligible studies, the period of childhood spanned an age range of age 5 to 13 years, adolescents included youth aged 14 to 18, and adulthood was defined as 19 years or older. Adult outcomes were not meta-analyzed because five of the six available adult studies represented the same cohort, and the data provided for the remain-

Table 1
Demographic Information by Age Group, Informant, and Study Measures

Study	Country of birth	Year(s) of inception of birth cohort	ELBW/EP NBW (*n* tested)	Male/Female (*n* or %)	Mean GA (wks)	Mean BWT (*g*)	Age at assessment (years)	Informant(s)	NSI and/or cognitive impairment	Assessment measures used	Mental health outcomes eligible for meta-analysis	Estimates of effect SMD (*g*) and *SE* (*g*)
							Childhood					
Anderson, Doyle, et al., 2003	Australia	1991–1992	257 ELBW	47%	26.7 (1.9)	884 (162)	8.7 (.3)	Parent	Included	BASC	Internalizing	.26 (.09)
			213 NBW	47%	Term	>2,499	8.9 (.4)				Externalizing	.10 (.09)
											Behavior symptoms	.22 (.09)
											Adaptive skills	−.26 (.09)
								Teacher			Internalizing	.24 (.09)
											Externalizing	.14 (.09)
											Behavior symptoms	.22 (.09)
											Adaptive skills	−.26 (.09)
Anderson, De Luca, et al., 2011	Australia	1997	189 ELBW	100/89	26.5 (2.0)	833 (164)	8.1 (.4)	Parent	Included	CADS-P	Any ADHD	.36 (.11)
			173 NBW	92/81	39.3 (1.1)	3,507 (453)	8.0 (.4)				Hyperactivity	.35 (.11)
											Inattention	.31 (.11)
Conrad et al., 2010	United States	1992–1997	31 ELBW	53%	27.2 (1.8)	810 (126)	12.0 (1.63)	Parent	Excluded	PBS-30	Hyperactivity/Inattention	1.07 (.24)
			55 NBW	27/28	38–42	3,640 (474)	10.9 (2.5)				Depression/anxiety	.89 (.24)
											Aggression/opposition	.68 (.23)
								Teacher			Hyperactivity/Inattention	.62 (.23)
											Depression/anxiety	.96 (.24)
											Aggression/opposition	.30 (.22)
Elgen et al., 2012	Norway	1999–2000	255 ELBW	167/88	26	851 (170)	5.8 (.3)	Parent	Included	SDQ	Attention problems	.67 (.07)
			1,089 NBW	NR	NR	NR	5.8 (.4)				Emotion problems	.72 (.07)
											Conduct problems	.34 (.07)
											Peer problems	.59 (.07)
											Prosocial behavior	−.14 (.07)
Farooqi et al., 2007	Sweden	1990–1992	83 ELBW	47%	24.6 (.7)	765 (111)	10.9 (.8)	Parent	Excluded	CBCL	Attention problems	.51 (.16)
			86 NBW	47%	39.2 (2.7)	3,520 (601)	11.6 (.9)				Internalizing	.65 (.16)
											Externalizing	.13 (.15)
											Anxiety	.62 (.16)
											Depression	.57 (.16)
											Conduct problems	.08 (.15)
											Oppositional	.00 (.13)
								Teacher		TRF	Attention problems	.34 (.15)
											Internalizing	.26 (.15)
											Externalizing	.16 (.15)
											Anxiety	.51 (.16)
											Depression	.39 (.16)
											Conduct problems	.16 (.15)
											Oppositional	.86 (.16)
								Child		DSRS	Depression	.86 (.16)
Farooqi et al., 2013	Sweden	1990–1992	83 ELBW	47%	24.6 (.7)	765 (110)	10.9 (.8)	Parent	Excluded	FTF	Attention problems	.58 (.33)
			86 NBW	47%	39.2 (2.7)	3,520 (601)	11.6 (.9)				Hyperactivity/Impulsivity	.52 (.09)
								Teacher			Attention problems	.34 (.15)
											Hyperactivity/Impulsivity	.53 (.36)
Fevang et al., 2016	Norway	1999–2000	216 ELBW	49%	27.0 (2.0)	868 (164)	11.0	Parent	Excluded	SDQ	Total problems	.68 (.18)
			1,767 NBW	48%	NR	NR	11.0			ASSQ (95%)	Autistic symptoms	.52 (.24)
										SNAP-IV (H)	Hyperactivity	.62 (.22)
										SNAP-IV (I)	Inattention	.90 (.21)
										SCARED	Anxiety	.56 (.22)
										SOCD	OCD symptoms	.60 (.22)
								Teacher		SDQ	Total problems	.67 (.18)
										ASSQ (95%)	Autistic symptoms	1.02 (.20)
										SNAP-IV (H)	Hyperactivity	.48 (.25)
										SNAP-IV (I)	Inattention	.86 (.21)

(*table continues*)

Table 1 (*continued*)

Study	Country of birth	Year(s) of inception of birth cohort	ELBW/EP NBW (*n* tested)	Male/Female (*n* or %)	Mean GA (wks)	Mean BWT (*g*)	Age at assessment (years)	Informant(s)	NSI and/or cognitive impairment	Assessment measures used	Mental health outcomes eligible for meta-analysis	Estimates of effect SMD (*g*) and *SE* (*g*)
Greenley et al., 2007[a]	United States	1982–1986	48 ELBW	16/32	25.8 (2.4)	660 (73)	11.2 (1.5)	Parent	Included	CBCL	Internalizing	.04 (.20)
			51 NBW	19/32	NR	NR	11.2 (1.23)				Externalizing	.07 (.20)
										CBQ	Conflict total score	.05 (.20)
								Teacher		TRF	Internalizing	.18 (.17)
											Externalizing	.02 (.17)
								Child		CDI	Depression	.33 (.20)
Hack et al., 2009	United States	1992–1995	219 ELBW	89/130	26.4 (2.0)	810 (124)	8.0	Parent	Included	CSI-4	ADHD combined	.57 (.10)
			176 NBW	65/111	≥37	3,300 (513)	8.0				ADHD hyperactive	.38 (.10)
											ADHD inattentive	.64 (.10)
											Anxiety	.34 (.10)
											Depression	−.06 (.10)
											Conduct problems	−.05 (.10)
											Oppositional	.16 (.10)
											Social phobia	.22 (.10)
											Autistic disorder	.39 (.10)
Hanke et al., 2003	Germany	1994–1995	21 ELBW	44%	27.3 (1.7)	804 (108)	6.2 (.7)	Parent	Included	CBCL	Attention problems	.80 (.26)
			68 NBW	30/38	Term	NBW	6.2 (.7)				Internalizing	.18 (.25)
											Externalizing	.26 (.25)
										HKS	Hyperactivity	.75 (.25)
Hille et al., 2001	United States	1984–1987	80 ELBW	32/48	27 (2.3)	853 (114)	8.0–19.0	Parent	Included	CBCL	Total problems	
(4 cohorts)	Canada	1977–1982	150	67/83	27 (2.3)	834 (126)						
	Germany	1985–1986	78	35/43	29 (2.0)	888 (101)					United States	NA[b]
	Netherlands	1983	100	42/58	29 (2.3)	882 (105)					Canada	NA[b]
											Germany	NA[b]
			1,200 NBW		Term	NR	4.0–11.0				Netherlands	NA[b]
			145		Term		NR					
			335		Term		NR					
			1172		Term		4.0–11.0					
Hoff et al., 2004	Denmark	1994–1995	194 ELBW	87/107	27.5 (1.8)	925 (169)	5.3 (.1)	Parent	Included	CASQ-P	Hyperactivity	.04 (.19)
			72 NBW	31/41	40.1 (1.1)	3,524 (511)	5.1 (.1)		Excluded		Outward reacting	.04 (.35)
											Anxious/withdrawn	.08 (.31)
											Poor social skills	.44 (.35)
Horwood et al., 1998	New Zealand	1986	77 ELBW	48%	27.1 (1.9)	823 (114)	7.6 (.4)	Parent/	Included	CPRS	Attention problems	.67 (.30)
			1092 NBW	51%	39.6 (1.6)	3,360 (534)	8.0	Teacher		CTRS	Conduct problems	.24 (.37)
								(combined)			Anxious/withdrawn	.82 (.29)
Hutchinson et al., 2013	Australia	1997	189 ELBW	100/89	26.5 (2.0)	833 (164)	8.5 (.4)	Parent	Included	SDQ	Attention problems	.02 (.21)
			173 NBW	92/81	39.3 (1.1)	3,506 (1,455)	8.5 (.4)				Emotion problems	.57 (.11)
											Conduct problems	.19 (.11)
											Peer problems	.31 (.11)
											Prosocial behavior	−.17 (.11)
Johnson et al., 2010a	United Kingdom, Ireland	1995	183 EPT	46%	<26	NR	10.9 (10.1–12.1)	Parent	Included	SCQ	Autism spectrum disorder	.76 (.12)
			137 Term	NR	Term	NR	10.9 (9.8–12.3)	interview				
Johnson et al., 2010b	United Kindgom, Ireland	1995	219 EPT	NR	<26	NR	10.9 (10.1–12.1)	Parent	Included	DAWBA	ADHD combined	.40 (.69)
			152 Term	NR	Term	NR	10.9 (9.8–12.3)	interview			ADHD inattention	1.29 (1.04)
											Emotional disorder	.00 (.63)
Johnson et al., 2016	United Kingdom, Ireland	1995	189 EPT	45%	<26	NR	11.0	Parent	Included	ADHD RS-IV	ADHD hyperactivity	.54 (.11)
			140 Term	42%	Term	NR	11.0				ADHD inattention	.88 (0.12)
								Teacher			ADHD hyperactivity	.32 (0.11)
											ADHD inattention	.73 (0.11)
Samara et al., 2008	United Kingdom, Ireland	1995	224 EPT	50%	<25	NR	6.3 (5.3–7.3)	Parent	Included	SDQ	Attention problems	.93 (.11)
			162 Term	NR	Term	NR	6				Emotion problems	.25 (.11)
											Conduct problems	.41 (.11)
											Peer problems	.62 (.11)
											Prosocial behavior	−.44 (.11)

Table 1 (*continued*)

Study	Country of birth	Year(s) of inception of birth cohort	ELBW/EP NBW (*n* tested)	Male/Female (*n* or %)	Mean GA (wks)	Mean BWT (*g*)	Age at assessment (years)	Informant(s)	NSI and/or cognitive impairment	Assessment measures used	Mental health outcomes eligible for meta-analysis	Estimates of effect SMD (*g*) and *SE* (*g*)
								Teacher		SDQ	Attention problems	1.06 (.11)
											Emotion problems	.37 (.11)
											Conduct problems	.22 (.11)
											Peer problems	.73 (.11)
											Prosocial behavior	−.49 (.11)
Scott et al., 2012	United States	2001–2003	148 ELBW	80/68	26 (2)	818 (174)	6.0 (.4)	Parent	Included	CBCL (DSM)	Affective problems	.20 (.57)
			111 NBW	51/60	≥37	3,382 (446)	6.0 (.3)				Anxiety	.49 (.46)
											ADHD	.61 (.48)
											Conduct problems	.23 (.46)
											Opposition disorder	.32 (.55)
								Teacher		TRF	Affective problems	.26 (.59)
											Anxiety	.39 (.56)
											ADHD	.63 (.46)
											Conduct problems	.45 (.56)
											Opposition disorder	.29 (.53)
										SSBS-2	Social competence	.99 (.56)
											Peer relations	1.08 (.56)
								Assessor		P-ChIPS	ADHD combined	.52 (.31)
											ADHD inattentive	.55 (.39)
											ADHD hyperactive	.53 (.34)
Shum et al., 2008	Australia	NR	45 ELBW	22/23	26.4 (1.9)	838 (152)	8.3 (1.0)	Parent	Excluded	ADHD RS-IV	ADHD combined	.61 (.21)
			49 NBW	25/24	39.9 (1.5)	3,578 (516)	8.2 (.9)				ADHD inattention	.54 (.21)
											ADHD hyperactivity	.38 (.21)
								Teacher			ADHD combined	.23 (.21)
											ADHD inattention	.43 (.21)
											ADHD hyperactivity	.38 (.20)
Stahlmann et al., 2009	Germany	1997–1999	75 ELBW	42/33	<27	790 (430–1,165)	8.0 (.6)	Parent	Included	SDQ	Attention problems	.90 (.13)
			4,000 NBW	NR	Term		7–10				Emotion problems	.60 (.24)
						NR					Conduct problems	.19 (.16)
											Peer problems	.27 (.25)
											Prosocial behavior	.33 (.33)
Stjernqvist & Svenningsen, 1999	Sweden	1985–1986	52 ELBW	41%	27.1 (1.0)	1,042 (242)	10.5 (.6)	Parent	Excluded	CBCL (empirical)	Internalizing	.73 (.19)
			61 NBW	26/35	40.1 (1.4)	3,648 (533)	10.6 (.6)				Externalizing	.37 (.19)
											Attention problems	.83 (.20)
											Social competence	1.07 (.21)
Szatmari et al., 1990	Canada	1980–1982	82 ELBW	37/45	27.4 (2.7)	835 (125)	5.0	Parent	Included	OCHS-R	Attention problems	.94 (.73)
			208 NBW	102/106	Term	NR	5.0				Emotion problems	.01 (.60)
											Conduct disorder	.87 (1.22)
								Teacher		OCHS-R	Attention problems	.41 (.52)
											Emotion problems	.06 (.60)
											Conduct disorder	−.39 (1.07)
Szatmari et al., 1993	Canada	1977–1981	129 ELBW	59/70	27 (2.0)	839 (124)	7.8 (.4)	Parent	Included	OCHS-R	Attention problems	1.50 (1.07)
			145 NBW	66/79	Term	3,369 (495)	8.1 (.5)				Emotion problems	.69 (.66)
											Conduct problems	.69 (1.19)
								Teacher		OCHS-R	Attention problems	.37 (.48)
											Emotion problems	−.36 (.61)
											Conduct problems	−.78 (.83)
								Child		PHSES	Self-esteem	−.19 (.13)
Taylor et al., 2000	United States	1982–1986	60 ELBW	19/41	25.7 (1.8)	666 (68)	6.7 (.9)	Parent	Included	CBCL	Attention problems	.76 (.20)
			49 NBW	16/33	Term	NR	7.0 (1.0)				Behavior competence	−.47 (.20)
										BHI	Hyperactivity	.58 (.20)
								Teacher		TRF	Attention problems	48 (.20)
											Social skills rating	−.14 (.20)
										BHI	Hyperactivity	.25 (.20)

(*table continues*)

Table 1 (*continued*)

Study	Country of birth	Year(s) of inception of birth cohort	ELBW/EP NBW (*n* tested)	Male/Female (*n* or %)	Mean GA (wks)	Mean BWT (*g*)	Age at assessment (years)	Informant(s)	NSI and/or cognitive impairment	Assessment measures used	Mental health outcomes eligible for meta-analysis	Estimates of effect SMD (*g*) and *SE* (*g*)
								Child		CDI	Depression	.31 (.20)
Whitfield et al., 1997	Canada	1974–1985	90 ELBW	36%	26.0 (23–38)	731 (520–800)	8.6	Psychologist	Included	SB-BRS	Distractibility	1.04 (.56)
			50 NBW	20/30	40 (38–42)	3,488 (2,614–4,706)	9.0 (6.5–12.1)				High activity level	1.94 (1.03)
											Anxious	1.70 (1.03)
											Gives up easily	1.36 (.75)
							Adolescence					
Burnett et al., 2014	Australia	1991–1992	205 ELBW	45%	26.6 (1.0)	889 (159)	17.9 (.9)	Assessor	Included	P-ChIPS	Any ADHD	.28 (.47)
			154 NBW	41%	39.2 (1.5)	3,408 (460)	18.1 (.8)				Any anxiety	−.11 (.45)
											Any mood disorder	−.13 (.38)
								Teen		APSD	Antisocial problems	−.21 (.11)
Gardner et al., 2004	United Kingdom, Scotland	1983	147 ELBW	77/68	27.1 (1.0)	1,048 (209)	15.6 (.5)	Parent	Excluded	SDQ	Attention problems	1.20 (1.05)
			108 NBW	59/41	NR	NR	16.0 (.5)				Emotion problems	.59 (.45)
											Conduct problems	.43 (.53)
											Peer problems	.81 (.50)
								Teacher		SDQ	Attention problems	.75 (.65)
											Emotion problems	1.27 (1.05)
											Conduct problems	−.23 (.54)
											Peer problems	1.53 (1.03)
								Teen		SDQ	Attention problems	.41 (.60)
											Emotion problems	.80 (.78)
											Conduct problems	.03 (.51)
											Peer problems	.44 (1.16)
Greenley et al., 2007[a]	United States	1982–1986	48 ELBW	16/32	25.8 (2.4)	660 (73)	16.8 (1.2)	Parent	Included	CBCL	Internalizing	.22 (.17)
			51 NBW	19/32	NR	NR	16.9 (1.3)				Externalizing	−.00 (.17)
								Teacher		TRF	Internalizing	.18 (.17)
											Externalizing	.02 (.17)
								Teen		CDI	Depression	.10 (.17)
Grunau et al., 2004	Canada	1981–1986	53 ELBW	32%	25.8 (23–29)	719 (520–800)	17.3 (16.3–19.7)	Parent	Excluded	CBCL	Internalizing	.66 (.23)
			31 NBW	55%	40 (39–42)	3,506 (3,068–4,196)	17.8 (16.5–19.0)				Externalizing	.58 (.23)
											Attention problems	.29 (.23)
											Anxious/depressed	.33 (.23)
											Social problems	.28 (.23)
											Withdrawal	.53 (.23)
											Thought problems	.15 (.22)
											Delinquency	.18 (.22)
											Aggression	.11 (.22)
Hallin et al., 2011a	Sweden	1985–1986	49 ELBW	20/32	27 (1.0)	1,002 (234)	18.4 (.2)	Teen	Excluded	YSR	Internalizing	.01 (.20)
			54 NBW	23/31	40 (1.5)	3,612 (525)	18.3 (.2)				Externalizing	−.56 (.20)
											Attention problems	.00 (.20)
											Anxious/depressed	.00 (.20)
											Withdrawn	.14 (.20)
											Social problems	.19 (.20)
											Delinquency	−.49 (.20)
											Aggression	−.51 (.20)
											Thought problems	−.49 (.20)
Hallin et al., 2011b	Sweden	1985–1986	51 ELBW	19/32	27 (1.0)	1,020 (232)	18.4 (.2)	Teen	Excluded	BYI	Anxiety	.15 (.19)
			54 NBW	23/31	40 (1.5)	3,610 (524)	18.3 (.2)				Depression	.12 (.19)
											Disruptive behavior	−.39 (.20)
											Self-concept	−.31 (.20)
											Anger	−.11 (.20)

Table 1 (*continued*)

Study	Country of birth	Year(s) of inception of birth cohort	ELBW/EP NBW (*n* tested)	Male/Female (*n* or %)	Mean GA (wks)	Mean BWT (*g*)	Age at assessment (years)	Informant(s)	NSI and/or cognitive impairment	Assessment measures used	Mental health outcomes eligible for meta-analysis	Estimates of effect SMD (*g*) and *SE* (*g*)
Methúsalemsdóttir et al., 2013	Iceland	1991–1995	29 ELBW	4/25	NR	500–999	17.0 (14–19)	Teen	Included	Kidscreen-52	Self-perception	−.66 (.55)
			30 NBW	7/23	Term	>2,500 g	17.0 (14–19)				Psychological well-being	−1.06 (.64)
											Mood/emotion problems	−.92 (.61)
Saigal et al., 2003	Canada	1977–1982	143 ELBW	58/83	27 (2.4)	838 (123)	14.1 (1.5)	Parent	Included	OCHS-R	Attention problems	.36 (.12)
			122 NBW	56/66	Term	3,391 (480)	14.4 (1.2)				Overanxious	.21 (.12)
											Depression	.26 (.12)
											Conduct problems	−.10 (.12)
											Oppositional	.03 (.12)
								Teen			Attention problems	.00 (.12)
											Over anxious	−.09 (.12)
											Depression	.02 (.12)
											Conduct problems	−.08 (.12)
											Oppositional	−.24 (.12)
Taylor et al., 2015	United States	1992–1995	169 ELBW	63/106	26.4 (2.0)	815 (122)	14.7 (.7)	Parent	Included	ASI-4	Attention problems	.34 (.12)
			115 NBW	42/73	≥37	3,260 (524)	14.8 (.8)				Hyperactivity	.20 (.12)
											Inattention	.38 (.12)
											Anxiety	.32 (.12)
											Depression	.11 (.12)
											Conduct problems	−.19 (.12)
											Oppositional	−.10 (.12)
											Social phobia	.33 (.12)
								Teen		YI-4	Attention problems	−.29 (.12)
											Hyperactivity	−.28 (.12)
											Inattention	−.23 (.12)
											Anxiety	−.06 (.12)
											Depression	−.20 (.12)
											Conduct problems	−.23 (.12)
											Oppositional	−.44 (.12)
											Social phobia	.14 (.12)
Wilson-Ching et al., 2013	Australia	1991–1992	193 ELBW	99/129	26.6 (2.0)	844 (161)	17.0 (1.5)	Parent	Included	CADS-P	Attention problems	.37 (.11)
			150 NBW	66/100	NR	NR	17.4 (1.6)				Hyperactivity	.32 (.11)
											Inattention	.42 (.11)
								Teen		CADS-S	Attention problems	−.24 (.11)
											Hyperactivity	−.20 (.11)
											Inattention	−.23 (.11)
							Adulthood					
Boyle et al., 2011	Canada	1977–1982	142 ELBW	62/80	23	835 (126)	23.2 (1.1)	Self	Included	YASR	Internalizing	.36 (0.12)
			133 NBW	60/73	Term	3,392 (480)	23.6 (1.0)				Externalizing	.01 (.12)
											Depression	.32 (.12)
											Anxiety	.30 (.12)
											Avoidant	.35 (.12)
											ADHD problems	.17 (.12)
											Anti-social	−.08 (.12)
										RSES	Self-esteem	−.24 (.12)
										STAI	Trait anxiety	.28 (.12)
										ADHD-RS	Hyperactivity	.11 (.12)
Lahat, Van Lieshout, Saigal, et al., 2014	Canada	1977–1982	121 ELBW	54/67	27.1	845 (124)	22–26	Self	Excluded	ADHD-RS	ADHD Total	.11 (.13)
			128 NBW	56/72	Term	3,380 (493)	22–26				Hyperactivity	−.01 (.13)
											Inattention	.22 (.13)
										YASR	Attention problems	.17 (.13)
Natalucci et al., 2013	Switzerland	1983–1985	55 ELBW	18/37	28.7	930	23.3	Self	Included	SF-36	Mental health	−.29 (.14)
			638 NBW	NR	NR	NR	(21.8–25.9)				Social functioning	−.47 (.14)

(*table continues*)

Table 1 (*continued*)

Study	Country of birth	Year(s) of inception of birth cohort	ELBW/EP NBW (*n* tested)	Male/Female (*n* or %)	Mean GA (wks)	Mean BWT (*g*)	Age at assessment (years)	Informant(s)	NSI and/or cognitive impairment	Assessment measures used	Mental health outcomes eligible for meta-analysis	Estimates of effect SMD (*g*) and *SE* (*g*)
							NR				Role-emotional	−.35 (.14)
										BSI	Interpersonal sensitivity	NA[b]
											Depression	NA[b]
											Anxiety	NA[b]
Schmidt et al., 2008	Canada	1977–1982	70 ELBW	30/40	27.6 (2.2)	874 (11)	23.3	Self	Excluded	SSS	Shyness	.38 (.16)
			83 NBW	35/48	Term	3,395 (470)	23.6				Sociability	−.34 (.16)
										CSES	Socialization	.55 (.16)
											Emotional well-being	−.33 (.16)
Schmidt et al., 2010	Canada	1977–1982	71 ELBW	30/41	27.6 (2.2)	874 (11)	23.3	Self	Excluded	YASR	Internalizing	.37 (.16)
			83 NBW	35/48	Term	3,395 (470)	23.6					
Van Lieshout et al., 2015a	Canada	1977–1982	84 ELBW	31/53	27.1 (2.3)	839 (132)	29–36	Assessor	Included	MINI	Depression	.67 (.83)
			90 NBW	36/54	40	3,411 (473)	29–36			(current risk)	Anxiety	.11 (.51)
											Social phobia	.44 (.72)
											Panic	1.06 (1.09)
											OCD	.20 (.76)
											Inattention	.84 (.80)
											Any alcohol or substance use disorder	−.51 (.41)
											Any nonsubstance-related psychiatric problems	.52 (.37)

Note. This table includes only those outcomes pertaining to mental health that were eligible for the review. Two articles that met criteria for inclusion in this review (Ganella et al., 2015; Taylor, Hack, & Klein, 1998) were omitted because the same mental health data from the same cohort at the same age were reported in other articles that were included (Burnett et al., 2014; Taylor et al., 2000). ELBW = extremely low birth weight; EP = extremely preterm; NBW = normal birth weight; GA = gestational age; NR = not reported; BWT = birth weight; NSI = neurosensory impairment; SMD = standardized mean difference; BASC = Behavior Assessment System for Children; CADS-P = Conners' ADHD/DSM-IV Scale for Parents; ADHD = attention-deficit/hyperactivity disorder; PBS-30 = Pediatric Behavior Scale-30; SDQ = Strengths & Difficulties Questionnaire; CBCL = Child Behavior Checklist; TRF = Teacher Report Form; DSRS = Depression Self-Rating Scale; FTF = Five to Fifteen; ASSQ = Autism Spectrum Screening Questionnaire; SNAP-IV = Swanson, Noland, and Pelham Questionnaire, Revision IV; CBQ = Conflict Behavior Questionnaire; SOCD = Symptoms of Obsessive-Compulsive Disorder; OCD = obsessive-compulsive disorder; SCARED = Screen for Child Anxiety Related Emotional Disorders; SCQ = Social Communication Questionnaire; CDI = Child Depression Inventory; CSI-4 = Child Symptom Inventory-4; HKS = Questionnaire of Hyperactivity Symptoms (German); CASQ-P = Conner's Abbreviated Symptom Questionnaire for Parents; CPRS = Conners' Parent Rating Scale; CTRS = Conners' Teacher Rating Scale; DAWBA = Development and Well-being Assessment; ADHD-RS = ADHD Rating Scale; ADHD RS-IV = ADHD Rating Scale-IV; DSM = Diagnostic and Statistical Manual; SSBS-2 = School Social Behavior Scale; second edition; P-ChIPS = Children's Interview for Psychiatric Syndromes; OCHS-R = Ontario Child Health Study-Revised; PHSES = Piers-Harris Self-Esteem Scale; BHI = Barkley Hyperactivity Index; SB-BRS = Stanford Binet Behavior Rating Scale; APSD = Antisocial Process Screening Device; YSR = Youth Self Report; BYI = Beck Youth Inventories; Kidscreen-52 = self-report for children and adolescents; ASI-4 = Adolescent Symptom Inventory-4; YI-4 = Youth Inventory-4; CADS-S = Conners' ADHD/DSM-IV Scale for Teens; YASR = Young Adult Self-Report; RSES = Rosenberg Self-Esteem Scale; STAI = State-Trait Anxiety Scale; SF-36 = Medical Outcomes Study Short Form-36-item questionnaire; BSI = Brief Symptoms Inventory; SSS = Shyness and Sociability Scale; CSES = Coopersmith Self-Esteem Scale; MINI = Mini-International Psychiatric Interview.

[a] Greenley et al. (2007) is cited twice because it was used twice. [b] Data provided did not permit the calculation of SMDs.

ing cohort did not permit calculation of standardized mean differences. Analyses were also stratified by informant (parent, teacher, or self-report), to examine possible informant discrepancies that have been reported in some instances (e.g., Szatmari et al., 1993). Where an analysis included 10 or more studies, funnel plots were created to assess the risk of small-study effects among the studies in that analysis (Sterne, Egger, & Moher, 2011). As the degree of funnel plot asymmetry (representing the extent to which the findings are biased) is difficult to assess visually, the funnel plots were evaluated using Egger's test for bias (Egger, Davey Smith, Schneider, & Minder, 1997).

Subgroup analyses. In addition to meta-analysis, tests of nonrandom variability across studies are critical to the interpretation of the results, because they may indicate the reliability of an effect. A major purpose of the review was to examine how important sociodemographic, secular treatment, and physical health-related factors affected risk for adverse mental health outcomes. Accordingly, outcomes were analyzed by geographical region (i.e., Australia, Europe, North America), by birth era (e.g., birth prior to 1990 vs. in 1990 or later), and by the inclusion vs. exclusion of ELBW individuals with significant NSI in each study. We also compared reports of mental health outcomes completed by different participating informants, to ascertain whether any discrepancies were meaningful and important. Finally, we performed a series of sensitivity analyses to assess whether studies at risk of greater or lesser bias generated different findings.

Results

Study Sample Characteristics

The 41 studies of mental health outcomes suitable for inclusion represented a total of 13,839 participants, 2,712 of whom were born at ELBW, and 11,127 who were born at NBW. (Each participant who contributed data was counted once across all studies). Sample sizes for the ELBW groups ranged from 21 to 408 individuals ($Mdn = 83$), and for NBW controls, from 30 to 4,000 individuals ($Mdn = 108$). Mean birth weight for ELBW participants was 840 (92) g, and for NBW controls, 3,343 (387) g. Mean gestational age was 26.6 (1.0) weeks for participants born at ELBW, versus 39.7 (0.7) weeks for NBW controls.

Controls were matched with cases on at least one demographic variable, usually age at the time of a childhood assessment. Controls were also matched on sex, race, geographic region of birth or birth hospital, and/or familial socioeconomic status, in many studies. Mean ELBW age by category was 8.4 (3.9) years across all child studies, 16.7 (3.7) years for adolescent studies, and 24.7 (8.9) years across all adult studies. ELBW sex distributions by category were 47 (0.9)% males across all child studies, 37 (0.9)% males for studies of adolescents, and 40 (0.4)% males across all adult studies. Demographic characteristics and information on study measures are presented in Table 1.

Forty-one studies reported on 24 unique samples born in Australia ($n = 3$), New Zealand ($n = 1$), North America (Canada, $n = 3$; United States, $n = 5$), and Europe ($n = 12$), (including Germany [$n = 3$], Switzerland [$n = 1$], the Netherlands [$n = 1$], the United Kingdom [$n = 2$] and the Nordic countries [$n = 5$]). Nineteen studies assessed cohorts born before 1990, and 22 assessed cohorts born after 1990. Sixteen of the 41 studies provided analyses that included individuals with significant NSI, 15 studies provided analyses from which individuals with significant NSI were excluded, and 10 studies reported both. Meta-analytic results are presented separately below by age group (children, adolescents, adults), informant (parent, teacher, self), and condition (e.g., attention, internalizing, externalizing, conduct problems, oppositional problems, social problems, and autistic symptoms). Twenty-five studies of children (aged 5 to 12 years), 10 studies of adolescents (aged 14 to 18 years), and six studies of adults (≥19 years) were included in this systematic review.

Meta-Analytic Results

Childhood. Results from random effects meta-analyses for children (including summary effect sizes and their 95% CIs, the variance among the study effect sizes (T^2), the ratio of observed variance to within-study error (Q), the proportion of the observed variance that is due to between-study differences in effect sizes (I^2), and the dispersion of the effect sizes (95% prediction intervals) are presented in Table 2. Effect sizes were interpreted as small (<0.30), moderate (0.30 to 0.60), or large (>0.60; Borenstein et al., 2009).

Childhood attention problems. Parents reported significantly higher scores for all three types of ADHD in ELBW children relative to NBW controls. Summary effect sizes across ADHD types were moderate-to-large, being greatest for combined ADHD ($g = 0.68$, 95% CI = 0.56 to 0.80), followed by inattentive ADHD ($g = 0.58$, 95% CI = 0.39 to 0.77), and hyperactive ADHD ($g = 0.46$, 95% CI = 0.37 to 0.55). These findings suggested that symptoms of ADHD are elevated in ELBW children relative to their NBW peers. The analyses indicated significant nonrandom variability among the effect sizes for combined ADHD ($T^2 = 0.02$, $Q = 22.15$, $p = .04$, $I^2 = 46\%$) and inattentive ADHD ($T^2 = 0.03$, $Q = 12.07$, $p = .03$, $I^2 = 59\%$), where about half of the variation in effect sizes was attributable to substantive or methodological differences among studies. For hyperactive ADHD, findings across studies were more consistent ($T^2 = 0.00$, $Q = 7.22$, $p > .50$, $I^2 = 0\%$). (See Figure 2A, and Figures 13A and 14A in the online supplemental materials). Standardized mean differences and standard errors from the 13 children's studies of combined ADHD were analyzed for publication bias by funnel plot. The results formed a symmetrical, inverted funnel-shaped pattern about a mean effect size of approximately 0.70. Although one outlier study showed more extreme scores, it was nonetheless represented within the funnel outlined by the other studies (see Figure 3A). Egger's test (Egger et al., 1997) revealed no significant asymmetry in the funnel plot, indicating a relatively low risk of bias (small-study effects) among these studies (intercept = 0.25, 95% CI = −1.65 to 2.16, $p > .75$).

Summary effects for teacher reports of ELBW children were small-to-moderate with respect to combined ADHD ($g = 0.54$, 95% CI = 0.29 to 0.79), hyperactive ADHD ($g = 0.35$, 95% CI = 0.19 to 0.50) and inattentive ADHD, ($g = 0.54$, 95% CI = 0.27 to 0.82). The analyses for combined ADHD ($T^2 = 0.07$, $Q = 18.78$, $p \leq .01$, $I^2 = 68\%$) and inattentive ADHD ($T^2 = 0.04$, $Q = 6.20$, $p = .10$, $I^2 = 52\%$) suggested substantial variability across studies, whereas effect sizes were more consistent for the hyperactive type of ADHD ($T^2 = 0.00$, $Q = 0.89$, $p > .30$, $I^2 = 0\%$). (See Figure 2B, and Figures 13B and 14B in the online supplemental materi-

Table 2
Results of Meta-Analyses for Children Born at Extremely Low Birth Weight (ELBW) Versus Normal Birth Weight (NBW) by Informant and Psychological Syndrome

	Parent report							Teacher report						
Psychological syndrome	Effect size (g)	95% CI	Number of studies	τ^2	Q	I^2	Prediction interval	Effect size (g)	95% CI	Number of studies	τ^2	Q	I^2	Prediction interval
Combined ADHD	.68[a]	.56 to .80	13	.02	22.15[a]	46%	.34 to 1.01	.54[a]	.29 to .79	7	.07	18.78[a]	68%	.39 to .69
Hyperactive ADHD	.46[a]	.37 to .55	9	.00	7.22	0%	−.03 to .95	.35[a]	.19 to .50	5	.00	.89	0%	−.44 to 1.13
Inattentive ADHD	.58[a]	.39 to .77	6	.03	12.07[a]	59%	.07 to 1.08	.54[a]	.27 to .82	4	.04	6.20	52%	−.23 to 1.32
Internalizing	.42[a]	.26 to .58	11	.05	36.62[a]	73%	−.09 to .93	.32[a]	.12 to .52	6	.03	12.50[a]	60%	.17 to .47
Externalizing	.15[a]	.02 to .28	5	.00	2.05	0%	−.57 to .86	.14[a]	.00 to .29	3	.00	.03	0%	−1.03 to 1.31
Conduct disorder	.23[a]	.09 to .37	9	.02	18.95[a]	58%	−.15 to .61	.19	−.01 to .38	5	.01	5.31	25%	−.23 to .60
ODD	.14	−.01 to .28	4	.00	.54	0%	−.73 to 1.01	.79[a]	.40 to 1.17	2	.02	1.11	10%	−2.23 to 3.81
Social Problems	.46[a]	.31 to .61	5	.02	9.01[†]	56%	.06 to .87	—	—	—	—	—	—	—
Autistic symptoms	.56[a]	.29 to .83	3	.04	5.83[a]	66%	.31 to .80	—	—	—	—	—	—	—

Note. Dashes indicate that there were insufficient data for meta-analysis. g = standard mean difference; τ^2 = between-studies variance; Q = ratio of variation to within-study error; I^2 = the proportion of total observed variation attributed to between-study effects; Prediction interval = the dispersion of true effect sizes; ADHD = attention-deficit/hyperactivity disorder; ODD = oppositional defiant disorder.
[a] $p \leq .05$, significant differences between the ELBW and NBW groups or significant Q values.
[†] $p \leq .06$.

als.) The patterns of results for parent-rated and teacher-rated symptoms across ADHD types were generally comparable, in that both informants located greater attention problems in ELBW groups.

Childhood internalizing and externalizing problems. Parents also reported significantly higher internalizing and externalizing scores in ELBW children than in NBW controls, with a moderate summary effect for internalizing (g = 0.42 95% CI = 0.26 to

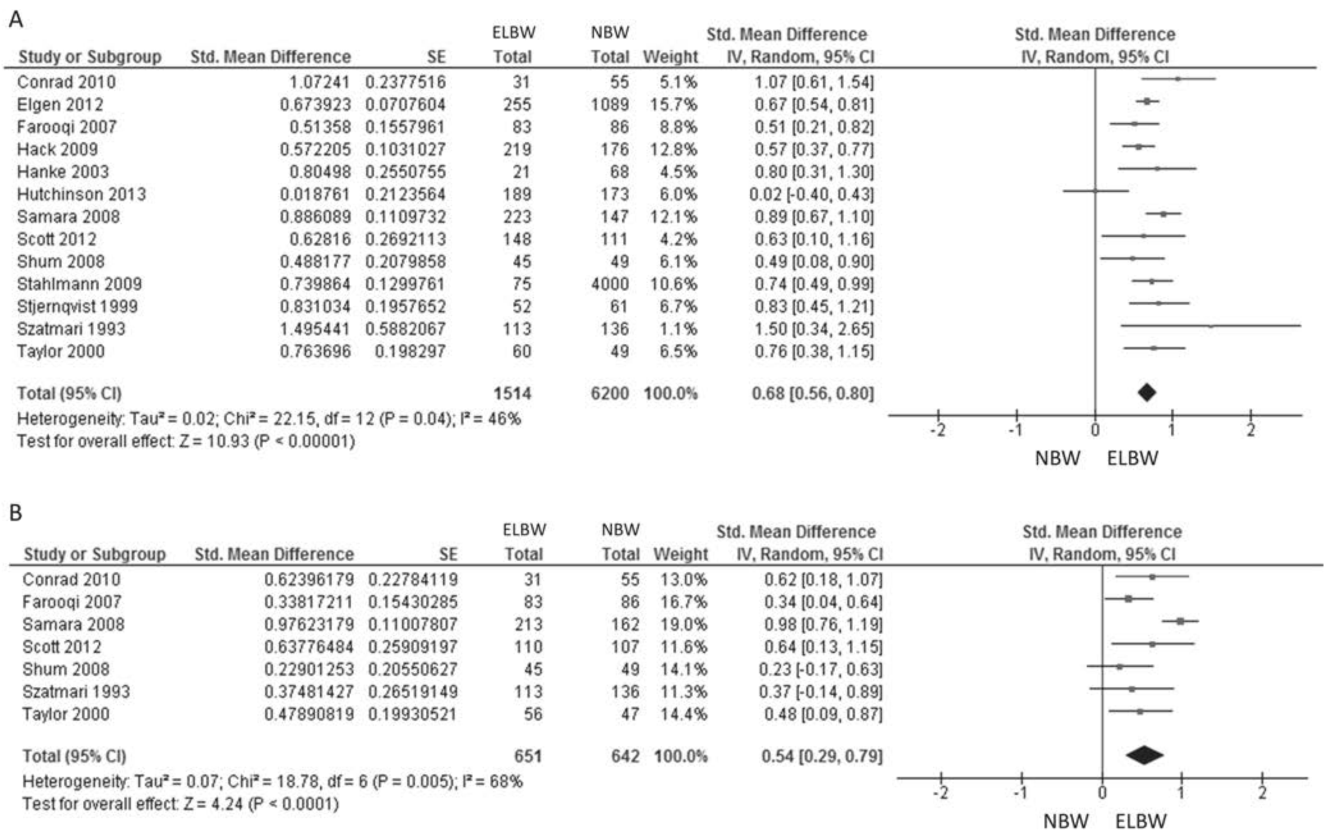

Figure 2. Forest plots of (A) parent- and (B) teacher-reported attention-deficit/hyperactivity disorder (ADHD; combined type) in children. ELBW = extremely low birth weight; NBW = normal birth weight. See the online article for the color version of this figure.

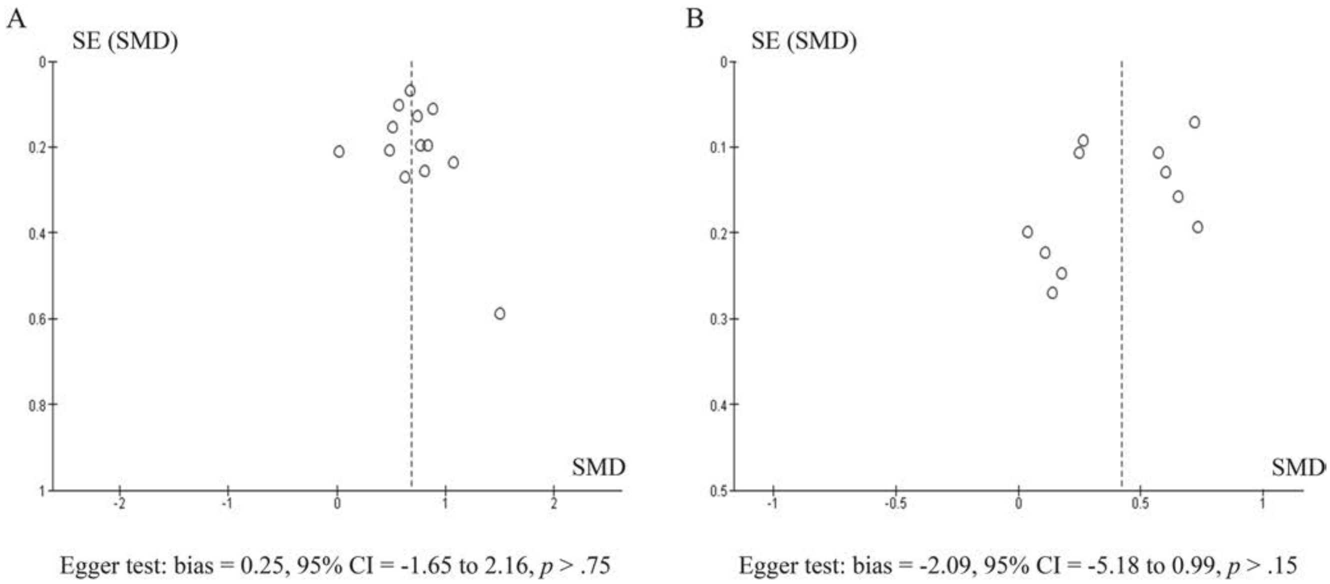

Figure 3. Funnel plots for parent ratings of (A) attention-deficit/hyperactivity disorder (ADHD; combined type) and (B) internalizing in children. See the online article for the color version of this figure.

0.58), and a small effect for externalizing ($g = 0.15$, 95% CI = 0.02 to 0.28). (See Figure 4A, and Figure 15A in the supplemental materials.) The analysis of parent-reported internalizing symptoms indicated significant nonrandom variability across the 11 studies ($T^2 = 0.05$, $Q = 36.62$, $p < .0001$, $I^2 = 73\%$), suggesting meaningful methodological or other differences among them. The standardized mean differences and standard errors for internalizing data from these studies were analyzed by funnel plot, where they formed a symmetrical pattern about a mean effect size of approximately 0.4 (see Figure 3B). Egger's test revealed no significant statistical asymmetry in this meta-analysis, consistent with a relatively low risk of bias (small-study effects) among these studies (intercept = −2.09, 95% CI = −5.18 to 0.99, $p > .15$). Effect sizes for parent reports of externalizing problems in ELBW children were consistent ($T^2 = 0.00$, $Q = 2.05$, $p > .70$, $I^2 = 0\%$).

Effect sizes for teachers were similar, with greater internalizing ($g = 0.32$, 95% CI = 0.12 to 0.52) and externalizing ($g = 0.14$, 95% CI = 0.00 to 0.29) difficulties in ELBW than NBW children. (See Figure 4B, and Figure 15B in the online supplemental materials.) The analysis suggested some variation in the effect sizes for internalizing ($T^2 = 0.03$, $Q = 12.50$, $p = .03$, $I^2 = 60\%$). Although effect sizes from individual studies of teacher-reported externalizing were consistent, ($T^2 = 0.00$, $Q = 0.03$, $p > .95$, $I^2 = 0\%$), only three studies contributed to the analysis, suggesting that teacher-rated externalizing in ELBW children reported should be interpreted with caution.

Childhood conduct and oppositional behaviors. Parents reported small but significant effects of birth weight status for conduct disorder ($g = 0.23$, 95% CI = 0.09 to 0.37), whereby children born at ELBW showed more symptoms of conduct disorder relative to NBW controls (see Figure 5A). The analysis indicated significant variation in the effect sizes ($T^2 = 0.02$, $Q = 18.95$, $p = .02$, $I^2 = 58\%$). In contrast, groups did not differ in parent ratings of oppositional defiant disorder ($g = 0.14$, 95% CI = −0.01 to 0.28). (See Figure 16A in the online supplemental materials.) The effect sizes for oppositional defiant disorder were consistent across four studies ($T^2 = 0.00$, $Q = 0.54$, $p > .90$, $I^2 = 0\%$). However, we note that the confidence interval for this summary effect included zero, suggesting inconsistency in the group findings across studies.

Teachers also reported small, marginally significant effects of birth weight status for conduct problems ($g = 0.19$, 95% CI = −0.01 to 0.38), although the confidence interval again included zero. The effect size for one study indicated more conduct problems in NBW controls ($g = -0.78$, 95% CI = −1.68 to 0.12; Szatmari et al., 1993), in contrast to four others that found conduct problems primarily in ELBW children (gs ≥ 0.16), ($T^2 = 0.01$, $Q = 5.31$, $p > .25$, $I^2 = 25\%$). (See Figure 5B.) Teachers also reported significant levels of oppositional defiant disorder ($g = 0.79$, 95% CI = 0.40 to 1.17) in ELBW children relative to their NBW peers ($T^2 = 0.02$, $Q = 1.11$, $p = .29$, $I^2 = 10\%$), but the variance in one study (95% CI = −0.86 to 1.35) was greater than in the other (95% CI = 0.55 to 1.17), even though both studies produced positive effects (gs ≥ 0.25). Interpretation of this analysis should be cautious, as only two studies contributed to the analysis. (See Figure 16B in the online supplemental materials.)

Childhood social problems and autistic symptoms. Parents reported significant, moderate effects of social problems ($g = 0.46$, 95% CI = 0.31 to 0.61) in ELBW children relative to their NBW peers. (See Figure 6A.) The analysis indicated some variation in the effect sizes ($T^2 = 0.02$, $Q = 9.01$, $p = .06$, $I^2 = 56\%$) related to methodological or substantive differences among the studies.

Parents also reported significantly more symptoms of autism spectrum disorders (ASD) in ELBW than NBW children ($g = 0.56$, 95% CI = 0.29 to 0.83). (See Figure 6B.) There was significant variation in the effect sizes among the three studies that contributed to the analysis ($T^2 = 0.04$, $Q = 5.83$, $p = .05$, $I^2 =$

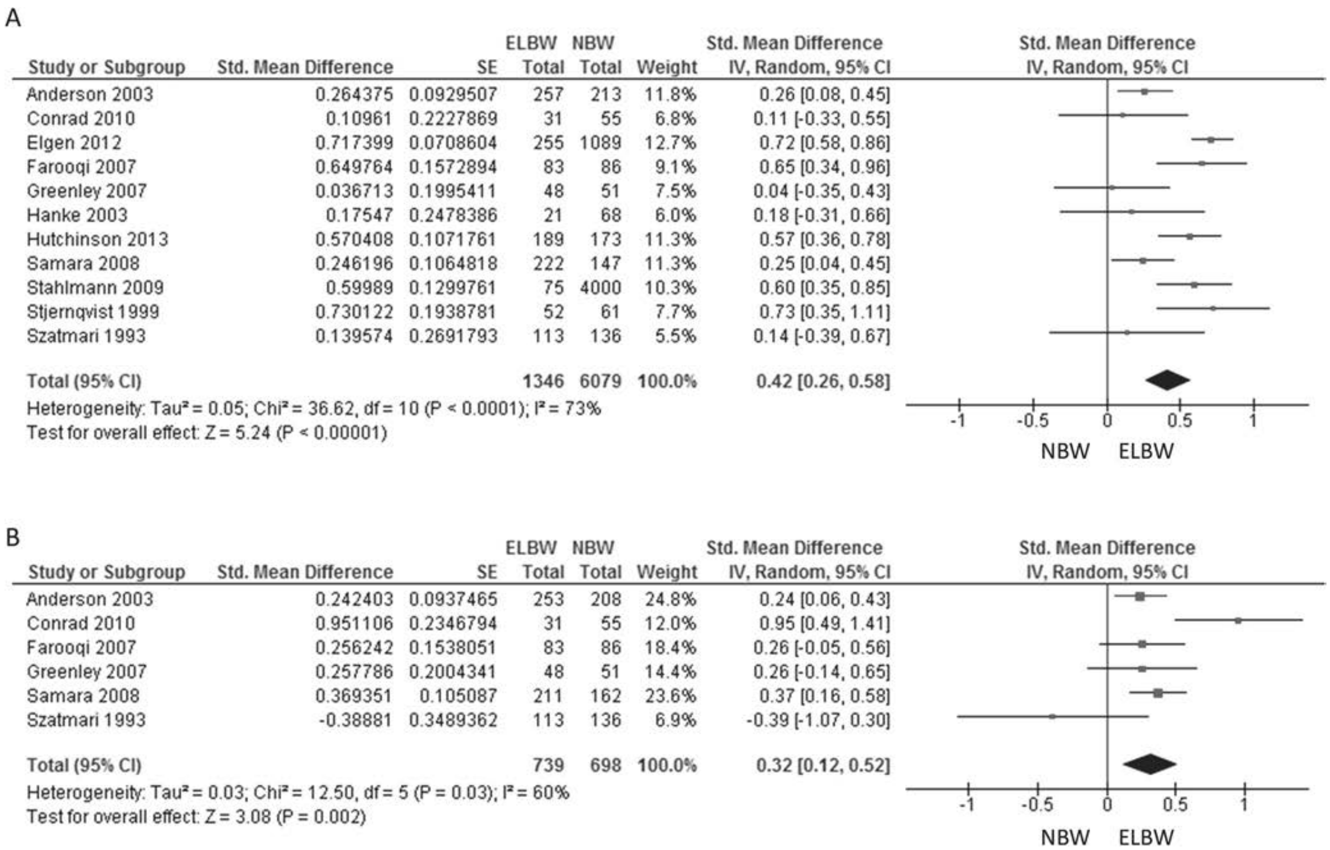

Figure 4. Forest plots of (A) parent- and (B) teacher-reported internalizing in children. ELBW = extremely low birth weight; NBW = normal birth weight. See the online article for the color version of this figure.

66%). Too few teacher reports of social problems and symptoms of ASD were available to meta-analyze studies of these problems.

Childhood prediction intervals. We calculated 95% prediction intervals to indicate the range in which the effect sizes for each childhood disorder were found, with narrower intervals indicating greater consistency among findings. (See Table 2.) In parent ratings of children, the smallest prediction interval—for symptoms of ASD—was still relatively wide (0.31 to 0.80), suggesting considerable variation across studies in parent reports of these symptoms. The widest prediction intervals for parent ratings of children were for externalizing (−0.57 to 0.86) and oppositional defiant disorder (−0.73 to 1.01). More importantly, five of the prediction intervals from parent ratings crossed the zero mark (hyperactive ADHD, internalizing, externalizing, conduct disorder, and oppositional defiant disorder), indicating that some of the studies in each of these meta-analyses identified negative effects (where NBW children have more problems), whereas others identified positive effects (where ELBW children have more problems). In teacher ratings, analyses of combined ADHD (0.39 to 0.69) and internalizing disorders (0.17 to 0.47) generated the most consistent findings, with the smallest prediction intervals. The widest prediction interval for teacher ratings was for childhood externalizing problems (−1.03 to 1.31).

Like parent ratings, some of the findings from teacher reports appeared to be unstable, as five of the prediction intervals from teacher ratings crossed the zero mark (hyperactive and inattentive ADHD, externalizing, conduct disorder, and oppositional defiant disorder), suggesting teachers in some studies identified more of these problems in NBW children, whereas others found more problems in ELBW children.

Adolescence. Results from random effects meta-analyses of mental health outcomes in adolescents are presented in Table 3. As teacher reports were provided in only two studies (20%) of this age group, we elected to analyze mental health ratings provided only by parents and teens.

Adolescent attention problems. Meta-analyses revealed significant, small to moderate effects of birth weight status for parental assessments of combined ADHD ($g = 0.52$, 95% CI = 0.19 to 0.85), hyperactive ADHD ($g = 0.26$, 95% CI = 0.10 to 0.43), and inattentive ADHD ($g = 0.40$, 95% CI = 0.24 to 0.56). (See Figures 17A to 19A in the supplemental materials.) Although these results suggested that ELBW teens were more likely than NBW controls to exhibit symptoms specific to the combined type of ADHD, there was considerable variation among the study effect sizes ($T^2 = 0.12$; $Q = 30.23$, $p < .00001$, $I^2 = 87\%$). There was little variation among studies of parent-reported hyperactive ADHD ($T^2 = 0.00$, $Q = 0.50$, $p > .45$, $I^2 = 0\%$) and inattentive ADHD ($T^2 = 0.00$, $Q = 0.05$, $p > .80$, $I^2 = 0\%$), but each of these analyses was based on two studies. As such, these findings should be interpreted with caution.

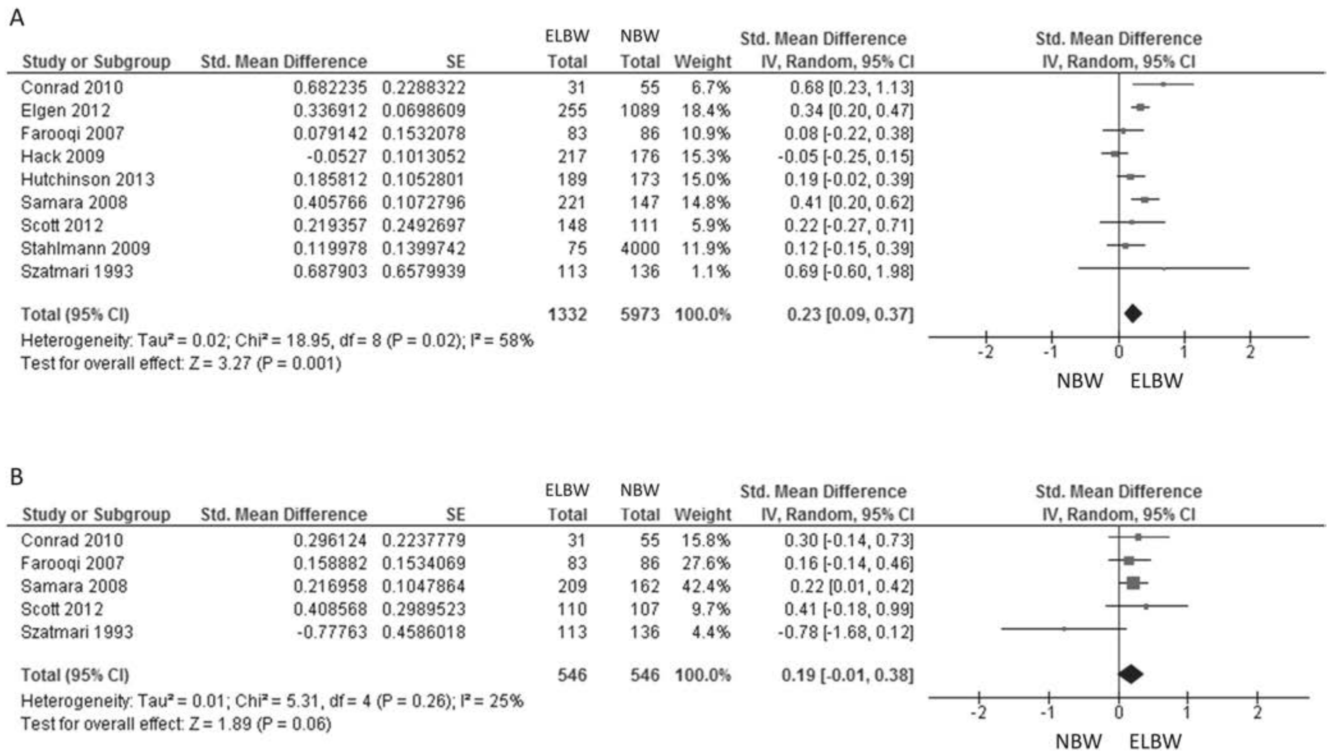

Figure 5. Forest plots of (A) parent- and (B) teacher-reported conduct disorder in children. ELBW = extremely low birth weight; NBW = normal birth weight. See the online article for the color version of this figure.

In contrast, adolescent self-reports revealed no consistent group differences in problems related to combined ADHD ($g = -0.03$, 95% CI = -0.28 to 0.23). Notably, the confidence interval for this summary effect crossed the zero mark, indicating that some studies suggested that NBW teens had greater problems related to combined ADHD, whereas others suggested that ELBW teens had more difficulties. (See Figures 17B to 19B in the online supplemental materials.) This analysis revealed considerable methodological or other variability among the individual effects related to adolescent reports ($T^2 = 0.07$, $Q = 19.95$, $p < .0005$, $I^2 = 80\%$). In addition, adolescent self-reports revealed significant effects of birth weight status for both hyperactive ($g = -0.23$, 95% CI = -0.39 to -0.07), and inattentive ADHD problems (g = -0.23, 95% CI = -0.38 to -0.07), that pointed to greater attentional difficulties in *NBW* teens. Within these analyses, the effect sizes from the individual studies were very consistent (hyperactive ADHD; $T^2 = 0.00$, $Q = 0.25$, $p > .60$, $I^2 = 0\%$; inattentive ADHD; $T^2 = 0.00$, $Q = 0.00$, $p > .90$, $I^2 = 0\%$), although in each case, summary effects were based on only two studies, suggesting these findings should be interpreted with caution.

Adolescent internalizing and externalizing problems. Parents reported that ELBW teens had significantly more problems with internalizing than did NBW teens ($g = 0.51$, 95% CI = 0.26 to 0.76). However, they indicated no group difference for externalizing, generating only a small, nonsignificant effect of birth weight status ($g = 0.29$, 95% CI = -0.26 to 0.84). (See Figures 20A and 21 in the online supplemental materials.) There was little variation among the individual effects of the three studies of the internalizing ($T^2 = 0.00$, $Q = 1.24$, $p > .50$, $I^2 = 0\%$), whereas the individual effects of the two studies of externalizing tended to diverge ($T^2 = 0.11$, $Q = 3.43$, $p = .06$, $I^2 = 71\%$). The nonsignificant effect for parent-reported externalizing in ELBW adolescents should be interpreted with caution due to the small number of studies contributing to the analysis, and because the confidence interval for this effect crossed the zero line, indicating that the two studies reported greater externalizing in different groups.

Adolescents themselves reported no significant group differences in internalizing behavior ($g = 0.31$, 95% CI = -0.44 to 1.06). The confidence interval for this summary effect also crossed the zero mark, reflecting inconsistency in the effect sizes of the two studies, indicated by a similar trend in the meta-analysis ($T^2 = 0.20$, $Q = 2.77$, $p = .10$, $I^2 = 64\%$). (See Figure 20B in the supplemental materials.) Only one study provided teens' self-reported estimates of externalizing, precluding meta-analysis. More data are needed before firm conclusions can be drawn about internalizing and externalizing tendencies in ELBW adolescents.

Adolescent conduct and oppositional behaviors. Meta-analyses indicated no significant effects of birth status for either parent-reported conduct disorder ($g = -0.30$, 95% CI = -1.58 to 0.98) or oppositional defiant disorder ($g = -0.03$, 95% CI = -0.21 to -0.14) among adolescents. (See Figures 22A and 23A in the online supplemental materials). Similar to parent-reported externalizing, the confidence interval for the summary effect for conduct problems crossed the zero mark, reflecting disparity among the studies (particularly between the Gardner ($g = 0.43$, 95% CI = -0.14 to 1.00) and Taylor studies ($g = -1.69$, 95% CI = -1.96 to -1.42), and wide within-study variation in the

Saigal study (g = −0.02, 95% CI = −1.41 to 1.37). The wide disparity among effect sizes was confirmed by analysis (T^2 = 1.56, Q = 77.58, p < .0001, I^2 = 96%). In contrast, the analysis of parent-reported oppositional defiant disorder suggested no significant difference between studies, (T^2 = 0.00, Q = 0.52, p > .45, I^2 = 0%), although this null finding rests on only two studies, and should be interpreted with caution.

Adolescents themselves reported no significant group difference between ELBW and NBW teens with respect to conduct disorder (g = −0.17, 95% CI = −0.38 to 0.05), but the confidence interval for the conduct disorder effect crossed the zero mark, where three studies reported greater risk of conduct disorder in NBW teens (gs ≤ −0.05), while one study reported no mean difference but showed wide within-study variation (Gardner: g = 0.03, 85% CI = −0.61 to 0.67). (See Figure 22B in the supplemental materials.) There was some variation among the study effect sizes (T^2 = 0.02, Q = 6.58, p = 0.09, I^2 = 54%). With respect to oppositional defiant disorder, adolescents reported higher levels among *NBW* teens (g = −0.34, 95% CI = −0.54 to −0.13). (See Figure 23B in the online supplemental materials.) Effect sizes for oppositional defiant disorder were consistent (g = −0.44 vs. −0.23; T^2 = 0.01, Q = 1.45, p > .20, I^2 = 31%), although again, this finding was based on two studies and merits cautious interpretation.

Adolescent social problems. Parents reported significantly more social difficulties in ELBW than NBW teens (g = 0.52, 95% CI = 0.00 to 1.03), with a marginal difference between the effects from these two studies (T^2 = 0.08, Q = 2.18, p > .10, I^2 = 54%). (See Figure 24A in the online supplemental materials.) Adolescents themselves indicated no significant differences between ELBW and NBW teens with respect to social problems (g = 0.21, 95% CI = −0.16 to 0.57). (See Figure 24B in the online supplemental materials.) Effect sizes for adolescent-reported social problems were small and consistent (T^2 = 0.00, Q = 0.15, p > .65, I^2 = 0%), but we note wide variation within the Gardner study (g = 0.44, 95% CI = −0.81 to 1.69).

Adolescent prediction intervals. All results from the parent and self-reports of adolescents should be interpreted with caution, however, as they are based on small numbers of studies. Moreover, the 95% prediction intervals for the summary effects are very wide, suggesting considerable variation among the effect sizes from different studies. Importantly, for every meta-analyses for this age group, the prediction intervals for parent-ratings of adolescents and adolescent self-ratings crossed the zero mark. (See Table 3.) The smallest prediction interval—for parent ratings of social problems (−0.21 to 1.24)—was still substantial, and the largest prediction intervals—for parent-rated conduct disorder (−4.79 to 4.19) and externalizing (−5.26 to 5.85)—were very wide. In adolescent self-ratings, the smallest prediction interval—for conduct disorder (−0.77 to 0.44)—was substantial, yet contrasted with much larger prediction intervals for social problems (−5.30 to 5.71) and internalizing (−7.12 to 7.75). Thus, associations between birth weight status and either parent- or self-rated mental health problems in adolescence appeared to be unstable and contradictory.

Narrative findings. The majority (77%) of eligible studies contributed data to the meta-analyses. Only five studies of children (Hille et al., 2001; Horwood et al., 1998; Johnson et al., 2010b; Szatmari, Saigal, Rosenbaum, Campbell, & King, 1990; Whitfield, Grunau, & Holsti, 1997) and three studies of adolescents, (Burnett et al., 2014; Hallin & Stjernqvist, 2011b; Methúsalemsdóttir, Egilson, Guðmundsdóttir, Valdimarsdóttir, & Georgsdóttir, 2013) were not analyzed. Reasons for exclusion included data that were redundant with those from another related study which were analyzed (Burnett et al., 2014; Hallin et al., 2011b; Johnson et al., 2010b; Szatmari et al., 1990), insufficient information to calculate effect sizes (Hille et al., 2001), averaged parent/teacher responses (Horwood et al., 1998), assessment by a clinician only (Whitfield et al., 1997), and in one case, measurement instruments that were significantly different from those selected for meta-analysis (Methúsalemsdóttir et al., 2013).

Effect sizes for these omitted studies are available in Table 1 for comparison with effect sizes of those that were meta-analyzed. In line with the findings from the meta-analyses, Horwood et al. (1998) reported large effect sizes for attention problems (g = 0.67, SE = 0.30) and anxious withdrawal (g = 0.82, SE = 0.29), and a

Table 3
Results of Meta-Analyses for Adolescents Born at Extremely Low Birth Weight (ELBW) Versus Normal Birth Weight (NBW) by Informant and Psychological Syndrome

	Parent report							Self-report						
Psychological syndrome	Effect size (g)	95% CI	Number of studies	τ^2	Q	I^2	Prediction interval	Effect size (g)	95% CI	Number of studies	τ^2	Q	I^2	Prediction interval
Combined ADHD	.52[a]	.19 to .85	5	.12	30.23[a]	87%	−.55 to 1.59	−.03	−.28 to .23	5	.07	19.95[a]	80%	−.83 to .78
Hyperactive ADHD	.26[a]	.10 to .43	2	.00	.50	0%	−3.38 to 3.91	−.23[*]	−.39 to −.07	2	.00	.25	0%	−3.85 to 3.38
Inattentive ADHD	.40[a]	.24 to .56	2	.00	.05	0%	−3.25 to 4.05	−.23[*]	−.38 to −.07	2	.00	.00	0%	−3.84 to 3.39
Internalizing	.51[a]	.26 to .76	3	.00	1.24	0%	−1.04 to 2.06	.31	−.44 to 1.06	2	.20	2.77	64%	−7.12 to 7.75
Externalizing	.29	–.26 to .84	2	.11	3.43[†]	71%	−5.27 to 5.85	—	—	—	—	—	—	—
Conduct disorder	−.30	−1.58 to .98	4	1.56	77.58[a]	96%	−4.79 to 4.19	−.17	−.38 to .05	4	.02	6.58[‡]	54%	−.77 to .44
ODD	−.03	−.21 to .14	2	.00	.52	0%	−3.79 to 3.72	−.34[a]	−.54 to −.13	2	.01	1.45	31%	−2.03 to 1.35
Social problems	.52[*]	.00 to 1.03	2	.08	2.18	54%	−.21 to 1.24	.21	−.16 to .57	2	.00	.15	0%	−5.30 to 5.71

Note. Dashes indicate that there were insufficient data for meta-analysis. g = standard mean difference; τ^2= between-studies variance; Q = ratio of variation to within-study error; I^2 = the proportion of total observed variation attributed to between-study effects; Prediction interval = the dispersion of true effect sizes; ADHD = attention-deficit/hyperactivity disorder; ODD = oppositional defiant disorder.
[a] $p \leq .05$, significant differences between the ELBW and NBW groups or significant Q values.
[†] $p \leq .06$. [‡] $p \leq .09$.

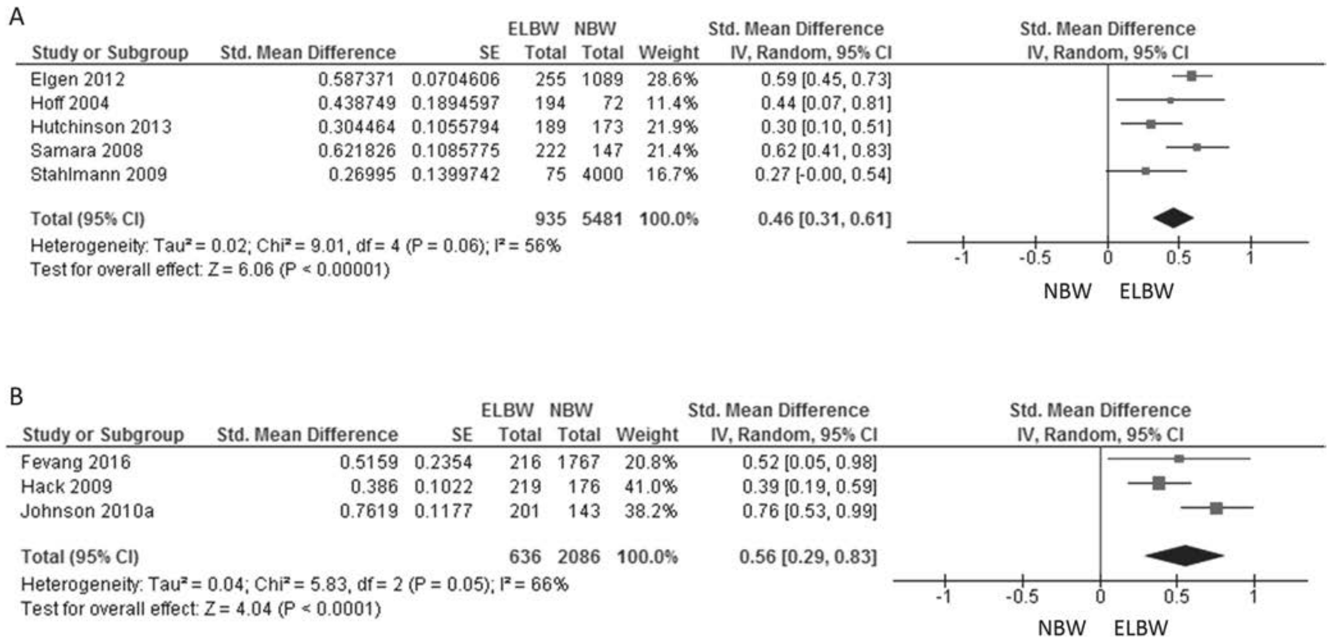

Figure 6. Forest plots of parent-reported (A) social problems and (B) autism spectrum disorder (ASD) symptoms in children. ELBW = extremely low birth weight; NBW = normal birth weight. See the online article for the color version of this figure.

small effect size for conduct problems (g = 0.24, SE = 0.37) in 8-year-old ELBW children born in New Zealand. Johnson et al. (2010b) reported a large effect size (g = 1.29, SE = 1.04) for inattention, but no effect of emotional disorder (g = 0.00, SE = 0.63), in 11-year-old ELBW children born in the United Kingdom. Szatmari et al. (1990) reported large effect sizes for attention (g = 0.94, SE = 0.73) and conduct problems (g = 0.87, SE = 1.22) in Canadian 5-year-olds. Whitfield et al. (1997) reported very large effects of birth weight status on distractibility (g = 1.04, SE = 0.56), hyperactivity (g = 1.94, SE = 1.03), and anxiety (g = 1.70, SE = 1.03) in Canadian 8-year-olds born weighing <800 g. In contrast, Burnett et al. (2014) reported small positive effects of birth weight status for ADHD (g = 0.28, SE = 0.47), mood disorders (g = −0.13, SE = 0.38), and antisocial problems (g = −0.21, SE = 0.11), but not anxiety (g = 0.01, SE = 0.10), in 18-year-old Australians. In a follow-up to their earlier study, Hallin et al. (2011b) reported small positive effects for anxiety (g = 0.15, SE = 0.19), and depression (g = 0.12, SE = 0.19), along with a moderate negative effect of birth weight status on disruptive behavior (g = −0.39, SE = 0.20) in Swedish 18-year-olds. In contrast, Methúsalemsdóttir et al. (2013) reported large group differences in quality of life between in 17-year-old Icelandic teens born at ELBW and their NBW peers, which favored the normal group: psychological well-being (g = −1.06, SE = 0.64) and overall mood (g = −0.92, SE = 0.61) were notably lower in ELBW teens.

Adulthood. Six studies of adults (aged 22 to 36 years) met inclusion criteria for the review. Data were collected from self-reports or assessments by trained interviewers (see Table 1). These were not meta-analyzed because five of the studies collectively represented the same unique cohort, and the data provided for the remaining cohort did not permit calculation of standardized mean differences. A narrative description of these findings follows to summarize the findings for the reader.

Adult attention problems. No group differences in the risks for ADHD problems reached significance in studies of adult ELBW survivors (e.g., Boyle et al., 2011; d = 0.17, p > .15; Lahat, Van Lieshout, Saigal, Boyle, & Schmidt, 2014, ps > .09; Van Lieshout, Boyle, Saigal, Morrison, & Schmidt, 2015a: OR = 4.58, 95% CI = 0.94 to 22.22, p > .05). However, for adults born at ELBW and exposed to antenatal corticosteroids (Van Lieshout et al., 2015a: OR = 11.45, 95% CI = 2.06 to 63.50, p < .05) or who had poor fluid intelligence (Lahat, Van Lieshout, Saigal, et al., 2014: p < .05), the risk of inattentive ADHD was elevated compared with their NBW peers.

Adult internalizing, externalizing, and social problems. The risk for self-reported internalizing problems was significantly greater in young adults born at ELBW than in their NBW peers (d = 0.42, p < .01; Boyle et al., 2011). Similarly, shyness persisted in the ELBW group in adulthood (d = 0.38, p < .03; Schmidt, Miskovic, Boyle, & Saigal, 2008), and adult survivors were at higher risk for self-reported anxiety problems (d = 0.34, p < .02; Boyle et al., 2011). If they had been exposed to corticosteroids prenatally, they were also at higher risk for clinical levels of anxiety, for example, generalized anxiety disorder (OR = 3.42, 95% CI = 1.06 to 11.06) and social anxiety disorder (OR = 5.80, 95% CI = 1.20 to 27.99, ps < .05), as assessed by structured psychiatric interview (Van Lieshout et al., 2015a).

In a European cohort (Natalucci et al., 2013), self-reported mental health (d = −0. 43, p < .01) and social functioning (d = −0.40, p < .05) were poorer in young adults born at ELBW, in comparison to community norms. In addition, distress related to interpersonal sensitivity (d = 0. 21, p < .05) was higher. However,

adults from this cohort reported no group differences in anxiety problems ($d = 0.24$, $p > .10$; Natalucci et al., 2013).

As in adolescence, externalizing problems at age 22 to 26 did not differ between groups (Boyle et al., 2011: $d = 0.01$, $p > .90$). The adult risk for substance use disorders was significantly lower in ELBW survivors than in NBW controls (Van Lieshout et al., 2015a: OR = 0.38, 95% CI = 0.17 to 0.86, $p < .05$).

In general, mental health problems related to internalizing, anxiety, and/or interpersonal sensitivity appeared to persist in adults born at ELBW, particularly those exposed to antenatal corticosteroids, whereas the risk for substance-related disorders was lower than in NBW controls. Indices of mental well-being were also lower in adult survivors.

Subgroup Analyses

To explore the possibility that nonrandom variability may account for some of the disparities among study effects, regional variations, differences in birth era (pre-1990 vs. post-1990), and the effects of including or excluding individuals with NSI were assessed in subgroup analyses. Parental reports were numerous enough for analysis of most of the childhood problems reported. Parental and self-reports for combined ADHD in adolescents also permitted meaningful analysis.

Influence of regional differences. Tests for significant regional differences in mental health outcomes were conducted in RevMan 5.3. If the overall regional test for a mental health outcome was significant, we then conducted separate pairwise Z tests of the effect sizes from Australia versus Europe, Europe versus North America, and North America versus Australia for that outcome. Results are presented in Table 4, Figures 7 to 9, and Figures 25 to 32 in the online supplemental materials. There were significant overall regional differences only for parent-rated childhood inattentive ADHD, ($Q = 6.32$, $p = .04$, $I^2 = 68\%$; see Figure 26 in the online supplemental materials), internalizing, ($Q = 7.58$, $p = .02$, $I^2 = 74\%$; see Figure 8), and ASD symptoms ($Q = 4.96$, $p = .03$, $I^2 = 80\%$; see Figure 31 in the online supplemental materials), and also for parent-rated combined ADHD in adolescence ($Q = 30.19$, $p < .00001$, $I^2 = 93\%$; Figure 9).

None of the pairwise Z tests was significant for parent-rated inattentive ADHD (all ps $> .24$). In contrast, the effect sizes for both parent-rated childhood internalizing ($Z = 2.69$, $p < .01$) and ASD symptoms ($Z = 2.19$, $p < .03$), were significantly larger in Europe than in North America. Effect sizes for parent-rated adolescent combined ADHD were also significantly larger in Europe than North America ($Z = 5.29$, $p < .001$), or Australia ($Z = -4.62$, $p < .001$).

No other regional differences reached significance. Results from the adolescent analyses should be interpreted cautiously due to the small numbers of studies available for each region (especially Australia, which is represented by one or two studies in these subgroup analyses).

Influence of birth era. We also conducted pairwise Z tests to ascertain whether the effect sizes for birth cohorts born before 1990 differed from those born in 1990 or later. Results of these subgroup analyses are presented in Table 5, Figure 10, and Figures 33 to 38 in the online supplemental materials. Only one significant difference emerged ($Q = 5.91$, $p = .02$, $I^2 = 83\%$; see Figure 10).

Table 4

Results of Subgroup Analyses (Pair-Wise Analyses) for Regional Differences by Age Group, Informant and Psychological Syndrome

Psychological Syndrome	Number of Studies	Australian Effect Estimate (g)	*SE*	European Effect Estimate (g)	*SE*	North American Effect Estimate (g)	*SE*	Australian-European Difference in Means	*SE* Difference	*Z*-score	European-North American Difference in Means	*SE* Difference	*Z*-score	North American-Australian Difference in Means	*SE* Difference	*Z*-score
Children																
Parent report																
Combined ADHD	A 2 E 6 NA 5	.26	.24	.72	.05	.75	.11	−.46	.24	−1.88†	−.03	.13	−.24	−.49	−.27	−1.84
Hyperactive ADHD	A 2 E 5 NA 2	.36	.10	.49	.10	.35	.09	−.13	.14	−.95	.14	.13	1.05	.01	.13	.08
Inattentive ADHD	A 2 E 3 NA 1	.36	.10	.71	.13	.64	.10	−.35	.30	−1.16	.07	.17	.42	−.28	.14	−.97
Internalizing	A 2 E 6 NA 3	.41	.15	.54	.20	.09	.13	−.13	.18	−.71	.45	.17	2.69**	.32	.20	1.58
Externalizing	A 1 E 3 NA 1	.10	.09	.23	.11	.07	.20	−.13	.14	−.92	.16	.23	.69	.03	.22	.13
Conduct Disorder	A 1 E 4 NA 4	.19	.11	.27	.07	.29	.21	−.08	.13	−.62	−.02	.23	−.09	−.10	.24	−.41
ODD	E 2 NA 2	—	—	.09	.12	.17	.10	—	—	—	−.08	.15	−.53	—	—	—
Social Problems	A 1 E 4	.31	.11	.52	.08	—	—	−.21	.13	−1.59	—	—	—	—	—	—
ASD	E 2 NA 1	—	—	.71	.10	.39	.10	—	—	—	.32	.14	2.19*	—	—	—
Adolescents																
Combined ADHD (parent report)	A 1 E 1 NA 3	.37	.11	1.21	.14	.34	.08	−.84	.18	−4.62***	.87	.16	5.29***	.03	.14	.22
Combined ADHD (self-report)	A 1 E 2 NA 2	−.24	.11	.23	.20	−.14	.14	−.47	.23	−2.04*	.37	.25	1.47	−.10	.18	−.55

Note. — denotes insufficient data for meta-analysis. A = Australia; E = Europe; NA = North America.
† $p < .06$. * $p < .05$. ** $p < .01$. *** $p < .001$, outcomes differ between regions.

Study or Subgroup	Std. Mean Difference	SE	ELBW Total	NBW Total	Weight	Std. Mean Difference IV, Random, 95% CI
North America						
Conrad 2010	1.0821	0.2399	31	55	5.0%	1.08 [0.61, 1.55]
Hack 2009	0.5733	0.1033	219	176	12.8%	0.57 [0.37, 0.78]
Scott 2012	0.63	0.27	148	111	4.2%	0.63 [0.10, 1.16]
Szatmari 1993	1.5	0.59	113	136	1.1%	1.50 [0.34, 2.66]
Taylor 2000	0.7691	0.1997	60	49	6.5%	0.77 [0.38, 1.16]
Subtotal (95% CI)			**571**	**527**	**29.6%**	**0.75 [0.52, 0.98]**
Heterogeneity: Tau² = 0.02; Chi² = 6.04, df = 4 (P = 0.20); I² = 34%						
Test for overall effect: Z = 6.31 (P < 0.00001)						
Europe						
Elgen 2012	0.6743	0.0708	255	1089	15.7%	0.67 [0.54, 0.81]
Farooqi 2007	0.5159	0.1565	83	86	8.8%	0.52 [0.21, 0.82]
Hanke 2003	0.812	0.2573	21	68	4.5%	0.81 [0.31, 1.32]
Samara 2008	0.8879	0.1112	223	147	12.1%	0.89 [0.67, 1.11]
Stahlmann 2009	0.74	0.13	75	4000	10.6%	0.74 [0.49, 0.99]
Stjernqvist 1999	0.8367	0.1971	52	61	6.6%	0.84 [0.45, 1.22]
Subtotal (95% CI)			**709**	**5451**	**58.3%**	**0.72 [0.63, 0.82]**
Heterogeneity: Tau² = 0.00; Chi² = 4.89, df = 5 (P = 0.43); I² = 0%						
Test for overall effect: Z = 14.86 (P < 0.00001)						
Australia						
Hutchinson 2013	0.0188	0.2128	189	173	6.0%	0.02 [-0.40, 0.44]
Shum 2008	0.4922	0.2097	45	49	6.1%	0.49 [0.08, 0.90]
Subtotal (95% CI)			**234**	**222**	**12.1%**	**0.26 [-0.21, 0.72]**
Heterogeneity: Tau² = 0.07; Chi² = 2.51, df = 1 (P = 0.11); I² = 60%						
Test for overall effect: Z = 1.09 (P = 0.28)						
Total (95% CI)			**1514**	**6200**	**100.0%**	**0.68 [0.56, 0.80]**
Heterogeneity: Tau² = 0.02; Chi² = 22.21, df = 12 (P = 0.04); I² = 46%						
Test for overall effect: Z = 10.91 (P < 0.00001)						
Test for subgroup differences: Chi² = 3.86, df = 2 (P = 0.15), I² = 48.2%						

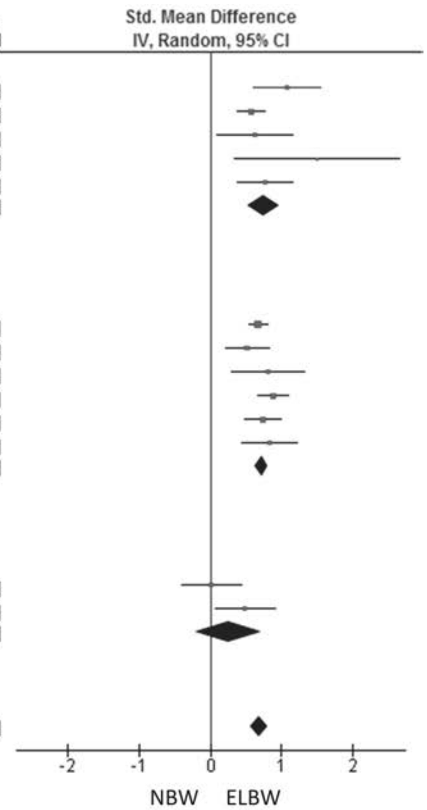

Figure 7. Regional effects on parent ratings of childhood attention-deficit/hyperactivity disorder (ADHD; combined type). NBW = normal birth weight; ELBW = extremely low birth weight. See the online article for the color version of this figure.

In adolescent self-reports, teens born prior to 1990 were more likely to identify high levels of ADHD (combined) in ELBW groups, whereas teens born in 1990 or later were more likely to identify high levels of ADHD (combined) in NBW groups, a contrast that was statistically significant ($Z = -2.43$, $p < .02$). However, the overall pattern of tests related to birth era suggested that most mental health outcomes did not differ between children or adolescents who were born before versus after 1990.

Influence of NSI. Pairwise Z tests were conducted to test whether mental health outcomes differed depending on the inclusion of individuals with NSI. Results from these subgroup analyses are presented in Table 6, Figure 11, and Figures 39 to 48 in the online supplemental materials. Only one significant difference emerged ($Q = 14.62$, $p = .0001$, $I^2 = 93\%$; see Figure 11). In adolescent self-reports, the exclusion of teens with NSI from the ELBW group was likely to result in higher rates of self-reported ADHD (combined) in the ELBW group, whereas the inclusion of teens with NSI in the ELBW group was likely to result in slightly lower rates of combined ADHD, a contrast that was statistically significant ($Z = 3.90$, $p < .001$). However, only one study contributed data that included teens with NSI, suggesting cautious interpretation, and the overall pattern of findings suggested that the risks for childhood and adolescent mental health outcomes were not increased by the inclusion of ELBW survivors with significant NSI for any of the outcomes examined.

Informant comparisons. Ratings of mental health outcomes were compared between informants in subgroup analyses, using the same Z test strategy. Results are presented in Table 7, Figure 12 and Figures 49 to 51 in the online supplemental materials. Parents' and teachers' ratings of mental health outcomes in ELBW and NBW children did not differ for any of the outcomes for which data were available: ADHD (all types), internalizing, externalizing, and conduct disorder (all ps $> .20$). In contrast, ratings by parents and adolescents differed for several adolescent mental health outcomes. Parents were more likely to report ADHD problems in ELBW teens, (combined ADHD: $Q = 6.49$, $p = .01$, $I^2 = 85\%$; see Figure 12; hyperactive ADHD: $Q = 18.40$, $p = .0001$, $I^2 = 95\%$, see Figure 49 the online supplemental materials; inattentive ADHD: $Q = 29.24$, $p = .0001$, $I^2 = 97\%$; see Figure 50 in the online supplemental materials). In contrast, adolescents were more likely to identify attention problems in NBW controls (combined ADHD: $Z = -2.60$, $p < .01$; hyperactive ADHD: $Z = -4.24$, $p < .001$; inattentive ADHD, $Z = -.5.63$, $p < .001$). In addition, although parents reported no group difference in oppositional defiant disorder between ELBW and NBW adolescents, ($Q = 5.01$, $p = .03$, $I^2 = 80\%$), adolescents themselves were more likely to identify oppositional defiant disorder among NBW control teens, ($Z = -2.26$, $p < .03$). (See Figure 51 in the online supplemental materials.)

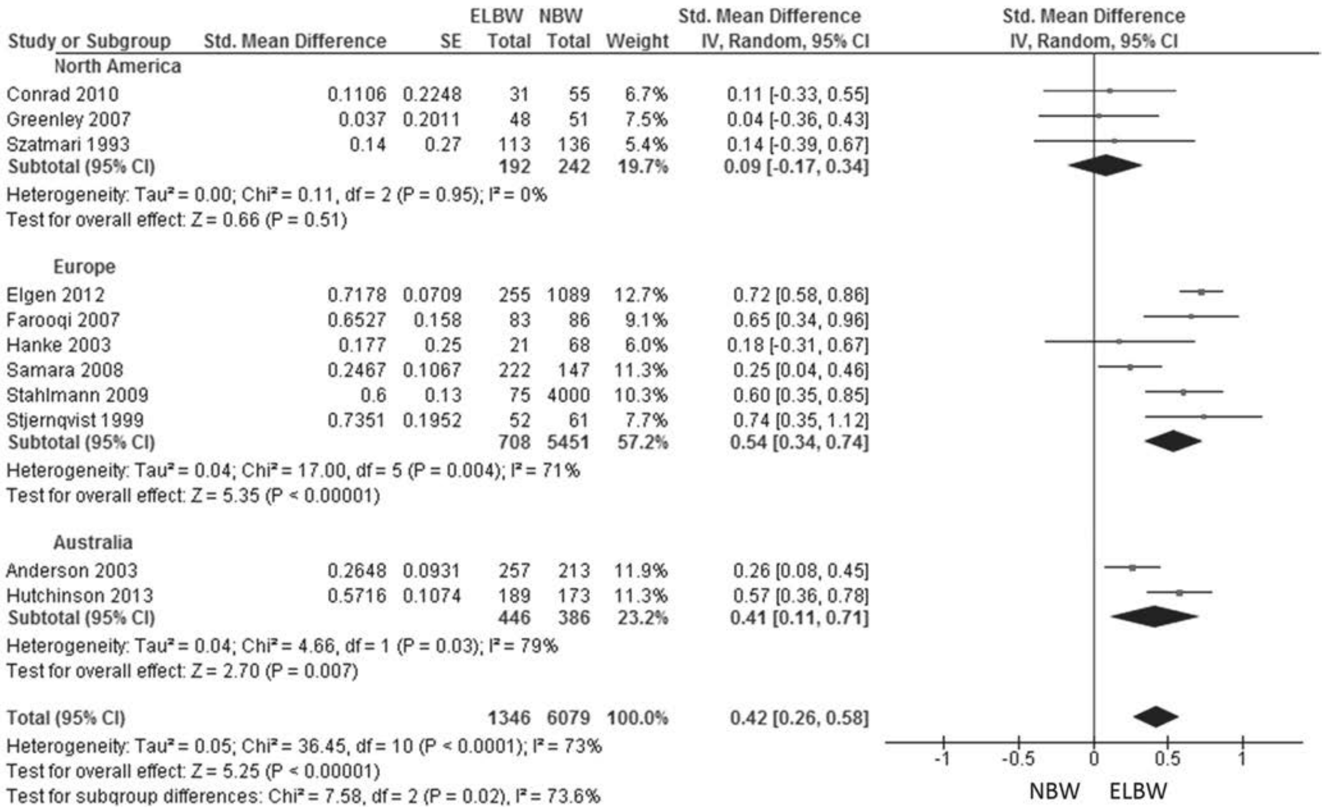

Figure 8. Regional effects on parent ratings of childhood internalizing. NBW = normal birth weight; ELBW = extremely low birth weight. See the online article for the color version of this figure.

Sensitivity Analyses

A series of sensitivity analyses was carried out to assess whether findings differed between studies that were deemed to be at greater versus lesser risk of bias, according to their NOS ratings. Results are presented in Table 8 and in Figures 52 to 59 in the online supplemental materials. None of these analyses revealed significant differences in the findings between studies at risk of greater or lesser bias (meta-analyses: all *p*s > .13; *Z* scores: all *p*s > .06). The differences in the effect sizes were small (<0.24) except for internalizing (0.29) and externalizing (0.56). However, estimates of mental health from studies at high risk of bias had larger standard errors and appeared to be more unstable than those from studies at low risk of bias. These included parent-rated childhood hyperactive ADHD (*Z* score = 0.37; *M* = 0.38, 95% CI = −0.03 to 0.78), internalizing (*Z* score = 0.20; *M* = 0.34, 95% CI = −0.21 to 0.89), externalizing (*Z* score = 0.42; *M* = 0.07, 95% CI = −0.33 to 0.46), and conduct disorder (*Z* score = 0.81; *M* = 0.12, 95% CI = −0.15 to 0.39). Similar analyses of parent-rated adolescent internalizing (*Z* score = 1.11; *M* = 0.34, 95% CI = −0.05 to 0.74) and externalizing (*Z* score = 1.84, *M* = 0.02, 95% CI = −0.37 to 0.41) suggested unstable estimates from studies at higher risk of bias. In these analyses, the 95% CIs for the summary effects from all of the groups of studies at higher risk of bias happened to cross the zero mark. In contrast, confidence intervals for the summary effects from groups of studies at lower risk of bias did not include zero.

Discussion

This systematic review summarizes all available literature on the mental health of individuals born at ELBW or EP up until mid-2016. Evidence from a large number of participants suggested that these individuals are at increased risk for particular mental health problems, beginning in childhood and extending into the fourth decade of life.

Analyses of ADHD studies revealed that children born at ELBW were more likely to have attention difficulties than their NBW peers, with moderate to large effects in parent ratings and small to moderate effects in teacher ratings. Parents reported significantly more ADHD problems in these children in almost every study included in this review. Children born extremely preterm were also significantly more prone to internalizing, externalizing, conduct problems, social problems, and symptoms of ASD, relative to age-matched controls. Effect sizes for parent-rated internalizing, social problems, and ASD symptoms were moderate, and small but significant for externalizing and conduct disorder. Teachers, as well as parents, identified small but significant group differences in internalizing and externalizing problems. Prediction intervals of the effect sizes for some parent and teacher ratings crossed the zero line, (e.g., externalizing, conduct, oppositional defiance, and ADHD), indicating that findings were sometimes contradictory across studies. While the majority of studies identified greater difficulties in

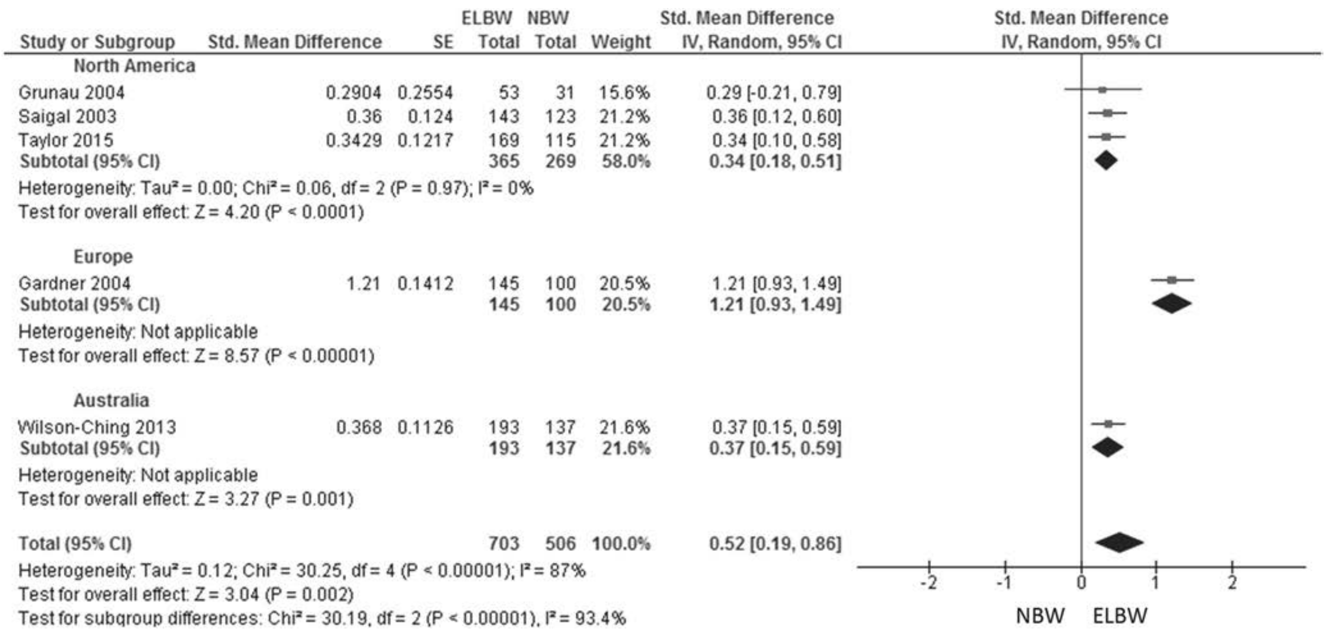

Figure 9. Regional effects on parent ratings of adolescent attention-deficit/hyperactivity disorder (ADHD; combined type). NBW = normal birth weight; ELBW = extremely low birth weight. See the online article for the color version of this figure.

ELBW children, a few found more problems in NBW children. (See Table 2.)

Meta-analyses of parent reports of adolescents also suggested significantly greater risk for ADHD (all three types), internalizing (especially anxiety), and social difficulties in the ELBW group, with moderate effect sizes, excepting a small effect for hyperactive ADHD. Parent-reported group differences in externalizing, conduct, and oppositional problems did not reach statistical significance at this life stage. In contrast to parental reports, adolescents reported significantly lower risks of combined, hyperactive, and inattentive ADHD and oppositional behavior in ELBW teens than in NBW teens, although the effects were small. Levels of internalizing, externalizing, and social problems were similar across groups. The sparse numbers of adolescent studies in our analyses and the wide dispersion of their effects, (seen in the large prediction intervals for both types of informant) require a cautious approach to interpretation,

Table 5
Results of Subgroup Analyses for Individuals Born at Extremely Low Birth Weight Versus Normal Birth Weight by Age Group, Informant and Psychological Syndrome: Birth Era

Psychological syndrome	Number of studies	Effect estimate if born before 1990	*SE*	Effect estimate if born 1990 and later	*SE*	Difference in means	*SE* difference	*Z* score
Parent report				Children				
Combined ADHD	3 10	.83	.13	.65	.07	−.18	.15	−1.19
Hyperactive ADHD	1 8	.25	.19	.47	.05	.22	.20	1.10
Internalizing	3 8	.31	.23	.45	.09	.14	.25	.56
Externalizing	2 3	.22	.15	.12	.08	−.10	.17	−.60
Conduct disorder	1 8	.69	.66	.23	.08	−.46	.66	−.69
				Adolescents				
Combined ADHD (parent report)	3 2	.63	.32	.36	.09	−.27	.33	−.82
Combined ADHD (self-report)	3 2	.15	.15	−.26	.08	−.41	.17	−2.43*

Note. ADHD = attention-deficit/hyperactivity disorder.
* $p < .05$. (Outcome differs by birth era.)

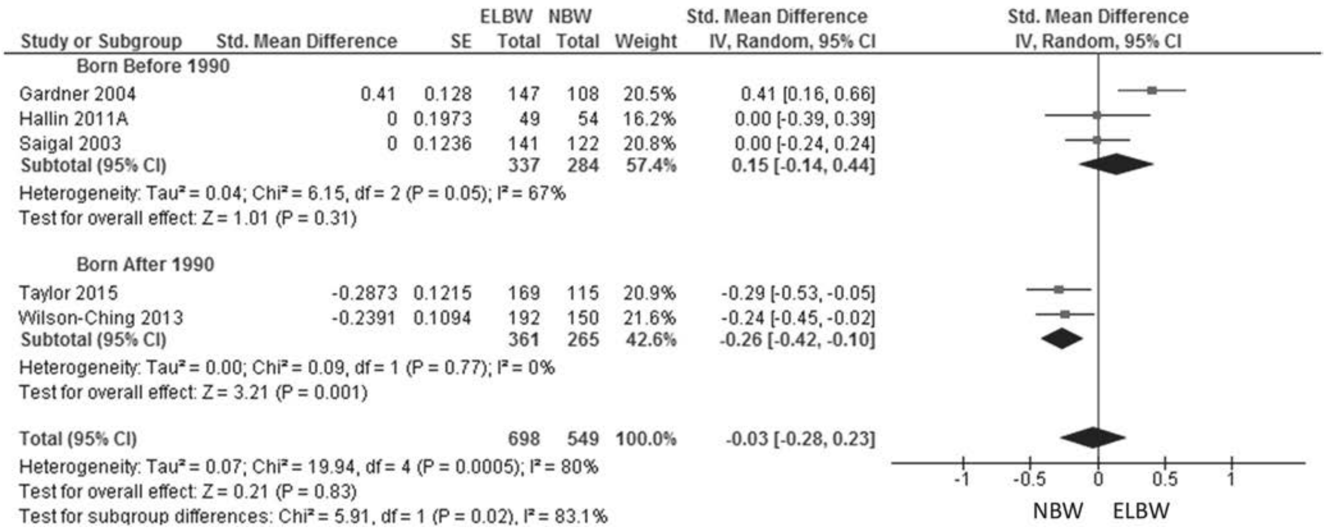

Figure 10. Effects of birth era on self-ratings of adolescent attention-deficit/hyperactivity disorder (ADHD; combined type). NBW = normal birth weight; ELBW = extremely low birth weight. See the online article for the color version of this figure.

but they provide a general outline of the current status of mental health research in this age group. (See Table 3.)

Higher levels of anxiety, depression, and shyness, and lower self-rated social functioning were self-identified by adults born at ELBW. In adulthood, any risks for combined or inattentive ADHD were higher in ELBW groups only under special circumstances, namely, exposure to antenatal corticosteroids, or having lower levels of fluid intelligence. Prenatal exposure to corticosteroids among ELBW survivors appeared to amplify group differences in social anxiety disorder and generalized anxiety disorder (Van Lieshout et al., 2015a). Although levels of antisocial behavior were similar in both groups, adults born

Table 6
Results of Subgroup Analyses for Individuals Born at Extremely Low Birth Weight Versus Normal Birth Weight by Age Group, Informant and Psychological Syndrome: Neurosensory Impairment (NSI)

Psychological syndrome	Number of studies	Effect estimate for NSI included (*g*)	*SE*	Effect estimate for NSI excluded (*g*)	*SE*	Difference in means	*SE* difference	*Z* score
Parent report				Children				
Combined ADHD	8 5	.63	.08	.75	.10	.12	.13	.94
Hyperactive ADHD	6 3	.44	.07	.51	.08	.07	.10	.69
Inattentive ADHD	3 3	.56	.13	.61	.17	.05	.22	.23
Internalizing	7 4	.41	.10	.43	.14	.02	.18	.11
Externalizing	3 2	.11	.08	.23	.12	.12	.15	.82
Conduct Disorder	6 3	.17	.08	.36	.15	.19	.17	1.12
ODD	3 1	.14	.09	.12	.15	−.02	.18	−.11
Social Problems	4 1	.42	.09	.62	.11	.20	.14	1.42
ASD	2 1	.57	.19	.52	.24	−.05	.31	−.16
				Adolescent				
Combined ADHD (parent report)	3 2	.36	.07	.77	.45	.41	.46	.89
Combined ADHD (self-report)	4 1	−.16	.07	.41	.13	.57	.15	3.90***

*** $p < .001$. (Outcome differs when participants with NSI are included versus excluded from analysis.)

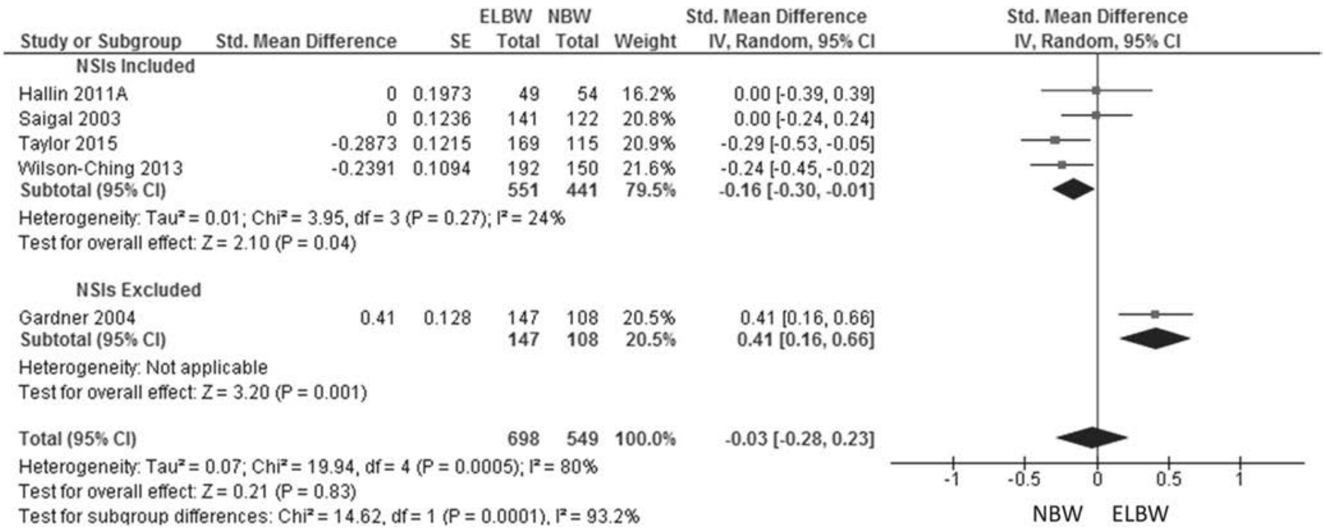

Figure 11. Effects of neurosensory impairment (NSI) on self-ratings of adolescent attention-deficit/hyperactivity disorder (ADHD; combined type). NBW = normal birth weight; ELBW = extremely low birth weight. See the online article for the color version of this figure.

at ELBW actually manifested a lower risk of alcohol and substance use than those born at NBW.

These findings concurred with extant reviews of children born at VLBW (Aarnoudse-Moens et al., 2009; Bhutta et al., 2002; Breslau, 1995). On the basis of a relatively small set of studies, Bhutta and colleagues (2002) reported a higher prevalence of ADHD, internalizing, and externalizing in children born preterm (<2,500 g). Here, we report moderate to large effect sizes for risks of ADHD, internalizing, social problems, and ASD symptoms in children born at ELBW.

Table 7
Comparisons of Informants' Ratings of Individuals Born at Extremely Low Birth Weight Versus Normal Birth Weight by Age Group and Psychological Syndrome

Psychological syndrome	Number of studies	Effect estimate for parent ratings (*g*)	*SEM*	Effect estimate for teacher or self ratings (*g*)	*SEM*	Difference in means	SE difference	Z score
				Children				
Combined ADHD	13 7	.68	.06	.54	.13	−.14	.14	−.99
Hyperactive ADHD	9 5	.46	.05	.35	.08	−.11	.09	−1.77
Inattentive ADHD	6 4	.58	.10	.54	.14	−.04	.17	−.24
Internalizing	11 6	.42	.08	.32	.10	−.10	.13	−.77
Externalizing	5 3	.15	.07	.14	.07	−.01	.10	−.10
Conduct disorder	9 5	.23	.07	.19	.10	−.04	.12	−.32
				Adolescents				
Combined ADHD	5 5	.52	.17	−.03	.13	−.55	.21	−2.60**
Hyperactive ADHD	2 2	.26	.08	−.23	.08	−.49	.12	−4.24***
Inattentive ADHD	2 2	.40	.08	−.23	.08	−.63	.11	−5.63***
Internalizing	2 2	.51	.13	.31	.38	−.20	.40	−.50
Conduct disorder	4 4	−.30	.65	−.17	.11	.13	.66	.20
ODD	2 2	−.03	.09	−.34	.10	−.31	.14	−2.26*
Social problems	2 2	.52	.27	.21	.19	−.31	.33	−.95

* $p < .05$. ** $p < .01$. *** $p < .001$. (Outcomes differ by informant.)

Our results were also broadly consistent with recent narrative reviews of children (Johnson, 2007), adolescents (Johnson & Wolke, 2013) adults (Johnson & Marlow, 2014). These authors noted a significantly higher prevalence of attention and social problems in children and adolescents born preterm, but found little consensus in the literature with regard to internalizing and externalizing difficulties. For individuals born extremely preterm, the higher risk of psychiatric symptoms was likely to endure into adulthood, based on higher prescription usage and hospital admissions for psychiatric and addiction problems in this population (Johnson & Marlow, 2014). Findings from our systematic review of prospective studies confirmed this general trend in nonhospitalized adults born at ELBW.

Some researchers have worried that disparities between children born at ELBW versus NBW would persist over time or even widen with age (e.g., Taylor et al., 2000; see also Whitfield et al., 1997), while others argued that postnatal experiences may be able to compensate for behavioral and developmental effects related to prematurity (Aylward, 2005; Murphy, Barkley, & Bush, 2002; Wilson-Ching et al., 2013). Between childhood and adulthood, increases in attention span and declining attention problems are reported both for adults born at full-term (Murphy et al., 2002) and those born very preterm (Breeman, Kaekel, Baumann, Bartmann, & Wolke, 2015). Despite putative maturational changes, however, vulnerability to attention problems may not entirely disappear. Clinically significant attention problems may be outgrown by adulthood in term-born individuals, but ADHD diagnoses have shown greater stability between childhood and adulthood in adults born very preterm (Breeman et al., 2015; Halmøy, Klungsøyr, Skjærven, & Haavik, 2012).

In contrast to attention problems, the risk of internalizing did not appear to diminish in adolescence or adulthood, as ELBW participants continued to note difficulties with depression, anxiety, and avoidant personality problems at these maturing life stages. Parents reported higher levels of internalizing and social difficulties in ELBW adolescents than in their NBW peers, and unremitted as well as remitted anxiety symptoms between early and late adolescence (Taylor, Margevicius, Schluchter, Andreias, & Hack, 2015). Adolescents themselves endorsed a preoccupied attachment style (Hallin et al., 2011b); higher levels of emotional problems (Gardner et al., 2004); and reduced athletic, vocational, and romantic confidence, in comparison to NBW controls (Grunau, Whitfield, & Fay, 2004). Research on adults born at ELBW suggests that internalizing problems may worsen in the transitions between childhood and adolescence, or adolescence and adulthood, (Schmidt, Miskovic, Boyle, & Saigal, 2010). Internalizing (Boyle et al., 2011) and non-substance-related psychiatric problems may remain higher over time in ELBW cohorts (Van Lieshout et al., 2015a), and social functioning is impaired (Natalucci et al., 2013). Although these kinds of problems do not appear to resolve with age, given the paucity of adolescent and adult assessments of mental health in individuals born extremely preterm, our present understanding of the developmental trajectories of mental health issues in this population remains necessarily tentative.

The consequences of mental health and social problems that do not diminish with increased age may be far-reaching. Poor social skills may lead to reduced social acceptance in childhood and adolescence (e.g., bullying; Methúsalemsdóttir et al., 2013; Yau et al., 2013). Adequate social skills continue to underpin academic and vocational success in adulthood, even in the general population. Not only psychological well-being, but educational attainment and even income level may be adversely affected by social difficulties related to extremely preterm birth (Allin et al., 2006; Mathiasen, Hansen, Nybo Anderson, & Greisen, 2009; Moster, Lie, & Markestad, 2008; Saigal et al., 2016), suggesting that the costs of enduring mental health problems and social deficits may well increase with age, as young adults assume greater personal and financial responsibilities than were required in their youth.

Subgroup Analyses

In addition to issues of developmental change, we attempted to evaluate the effects of regional and cohort differences on mental health outcomes in ELBW across studies and to ascertain whether group differences in the risk for psychiatric problems related to the presence of NSI in ELBW samples. Surprisingly, group differences for most syndromes were relatively robust to all three factors: birth region, birth era (i.e., significant developments in neonatal intensive care), and the presence of NSI, with a few exceptions.

Regional distributions. Regional disparities were relatively rare. In children, a pattern of widely distributed, moderate effect sizes were found for most disorders, other than inattention, internalizing, and ASD symptoms. The general consistency of the findings across regions suggests that attentional, behavioral and social outcomes in individuals born at ELBW may have a developmentally programmed, biological basis. However, regional disparities in corticosteroid use may account, at least in part, for regional differences in attention difficulties. Repeated antenatal exposure to corticosteroids may result in more distractible and hyperkinetic behavior during childhood (French, Hagan, Evans, Mullan, & Newnham, 2004; Rieger et al., 2004; for a review see Van den Bergh, Mulder, Mennes, & Glover, 2005). Unfortunately, few of the studies provided information on maternal treatment with corticosteroids, except for two Norwegian studies of the same cohort (Elgen et al., 2012; Fevang, Hysing, Markestad, & Sommerfelt, 2016), where 70% to 71% of ELBW neonates had received prenatal steroids, and one Australian study (Hutchinson et al., 2013), where 88% of ELBW had received prenatal steroids.

Birth era. One drawback to examining studies of infants born extremely preterm over a 29-year period (1974 [Whitfield et al., 1997] to 2003 [Scott et al., 2012]) was that prenatal and neonatal care did not remain uniform during that time (Bhutta et al., 2002). We attempted to account for changes in neonatal care by comparing pre- and post-1990 cohorts because surfactant and steroid therapies were widely adopted in the late 1980s and early 1990s. The majority of mental health outcomes appeared to be independent of birth era, with the exception of combined ADHD in adolescent self-reports, where ADHD was likely to be slightly higher in the ELBW group among teens born before 1990 and slightly higher in the NBW group among those born after 1990. (See Table 5 and Figure 10.) Given that risks for other kinds of impairments (e.g., NSI) have remained strong in the postsurfactant era (Anderson, & Doyle, & the

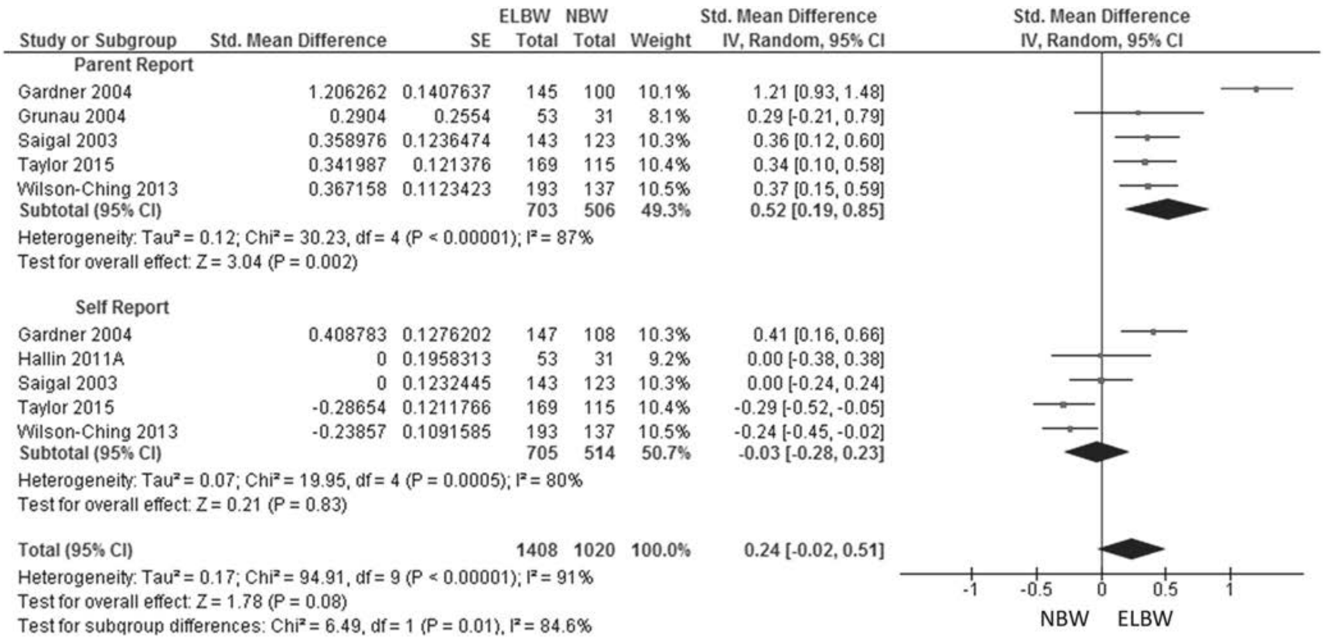

Figure 12. Parent–adolescent informant discrepancies for attention-deficit/hyperactivity disorder (ADHD; combined type). NBW = normal birth weight; ELBW = extremely low birth weight. See the online article for the color version of this figure.

Victorian Infant Collaborative Study Group, 2003; Hack et al., 2009; Hanke et al., 2003), further research is needed to determine whether changes to modern neonatal care have led to mental health benefits, apart from any secular improvements.

Neurosensory impairment. Recently, it has been suggested that the higher risk for mental health problems in individuals born at ELBW might be driven by a subgroup of the population with severe handicaps (e.g., Johnson & Wolke, 2013). Indeed,

Table 8
Sensitivity Analyses for Parent Ratings of Individuals Born at Extremely Low Birth Weight Versus Normal Birth Weight by Age Group and Psychological Syndrome: High Versus Low Risk of Bias

Psychological syndrome	Number of studies	Effect size for studies at high risk of bias (*g*)	*SEM*	Effect size for studies at low risk of bias (*g*)	*SEM*	Difference in means	*SE* difference	*Z* score
				Children				
Combined ADHD	2 11	.67	.12	.69	.08	.02	.14	.15
Hyperactive ADHD	1 8	.38	.21	.46	.05	.08	.21	.37
Inattentive ADHD	1 5	.54	.21	.58	.11	.04	.24	.17
Internalizing	2 9	.34	.45	.43	.09	.09	.11	.20
Externalizing	1 4	.07	.20	.16	.07	.09	.22	.42
Conduct disorder	1 8	.12	.14	.25	.08	.13	.16	.81
Social problems	1 4	.27	.14	.50	.08	.23	.16	1.46
				Adolescents				
Internalizing	1 2	.34	.20	.63	.17	.29	.26	1.11
Externalizing	1 1	.02	.20	.58	.23	.56	.30	1.84

Note. All *Z* scores were nonsignificant.

most early studies excluded children with NSI, in a bid to identify the relatively subtle consequences of low birth weight in survivors apart from any effects of major neurological damage (Breslau, 1995). Here, the inclusion of children with NSI did not widen the group differences in psychiatric outcomes, suggesting that any detrimental effects of birth weight status on childhood mental health were not related to the presence of significant handicaps in ELBW cohorts, but to prematurity itself. (See Table 6.) The only exception was a significantly higher risk of self-reported combined ADHD in studies that excluded teens with NSI. (See Figure 11). This finding requires corroboration, as only one study in which NSI was excluded was available for analysis. Like the higher rates of NSI that accompanied improved survival (Fanaroff et al., 2003), higher rates of mental health difficulties in ELBW samples may simply reflect the fact that more extremely preterm babies with high biological risk now survive.

Informant Discrepancies

One of the advantages of collecting reports from multiple informants is that congruent evidence from independent sources may support the stability of theorized effects. On the other hand, informant discrepancies may allude to real differences in informants' perspectives of particular behaviors (Achenbach, 2011). Informant discrepancies may also reflect legitimate contextual differences, for example, between the attention and behavioral control required for the classroom setting versus expectations at home. Accordingly, researchers have begun to analyze such discrepancies for latent information that is not contained in the individual ratings of either group (De Los Reyes, 2011).

In the childhood studies of mental health problems in ELBW survivors reported here, group differences between children born at ELBW and NBW were similar in parent and teacher-ratings of ADHD (all types), internalizing, externalizing, and conduct disorder (all Z scores <1.77, ns; cf. Kohen, Brooks-Gunn, McCormick, & Graber, 1997; Zeiner, 1997). In contrast to the similarity between parent and teacher ratings, adolescent self-ratings indicated lower levels of hyperactive ADHD, inattentive ADHD, and oppositional defiant problems in ELBW teens than their NBW peers, whereas parents reported the opposite. Parent-teen reporting discrepancies were epitomized in a study where parents rated ELBW teens as having higher attention problem scores, while adolescents gave ELBW teens lower scores (Wilson-Ching et al., 2013). However, effects of birth weight status in parent and youth reports were similar for internalizing, conduct disorder, and social problems.

How are such parent-adolescent discrepancies to be interpreted? Assuming they contain real information, possible explanations for adult–teen discordance in the reporting of mental health problems include teenage optimism (Gardner et al., 2004; Hallin & Stjernqvist, 2011a), lack of insight (Burnett et al., 2014), downward revision of expectations (Saigal, Pinelli, Hoult, Kim, & Boyle, 2003), and the fact that teens are likely to have access to information about themselves that parents do not have (Gardner et al., 2004). Parents may have significant concerns about behavior that their adolescent offspring do not regard as problematic (De Los Reyes, 2011). Alternatively, underreporting of problems may be related to social acceptability in ELBW adolescents, who tend to be more cautious (Hack et al., 2002), shy (Schmidt et al., 2008; Waxman, Van Lieshout, Saigal, Boyle, & Schmidt, 2013), and conforming (Allin et al., 2006; Pesonen et al., 2008) than their NBW counterparts. This may be particularly true for sensitive parents (Bilgin & Wolke, 2015), who see their ELBW children as good, fragile, or in need of assistance. The issue of parent-adolescent discrepancy is not trivial, as unresolved differences in perspective may lead to negative outcomes: for instance, differing interpretations of conflict by teens and parents within the family may predict later adjustment difficulties for the teens (Greenley et al., 2007). Sorting out the relevance and underlying causes of parent–adolescent discrepancies in mental health reports merits further study.

Risk of Bias

The strength of a meta-analytic review depends heavily on the studies available for assessment. Therefore, it is critical to assess potential reporting biases in the published findings. One potential problem is the tendency of smaller studies to produce effects that differ from (and are often larger than) those of larger studies—known as *small-study effects* (Sterne, Gavaghan, & Egger, 2000). Indeed, significant findings from meta-analyses that are based primarily on small studies have occasionally been contradicted by larger, better-controlled studies at a later date (Egger et al., 1997). Accordingly, where appropriate, we created funnel plots to determine whether the meta-analytic results relied on small-study effects. Because funnel plots are recommended when meta-analyses include at least 10 studies, we performed these analyses for only two syndromes, namely, childhood combined ADHD and internalizing. Funnel plots and Egger tests of the findings from studies of parent-rated combined ADHD and internalizing yielded no evidence of small-study effects, suggesting that parents of children born at ELBW are more likely than parents of children born at NBW to report significant attentional and internalizing problems in their offspring.

Second, sensitivity analyses were performed to determine whether studies produced discrepant findings, depending on whether they carried greater or lesser risk of bias according to the NOS ratings. None of these analyses revealed significant differences in the findings. Notably, the summary effects from high-bias studies were numerically smaller than those from studies at lower risk of bias, suggesting that high bias studies tended to provide more conservative estimates of mental health problems for youth born at ELBW than low-risk studies. However, the findings from studies at higher risk of bias tended to be inconsistent, as indicated by the 95% CIs for summary effects that included zero. The inconsistency may be partly attributable to the very small numbers of high-bias studies and wide within-study variability in some studies.

Implications for Theories of Development

Developmental origins of health and disease. In addition to describing the current state of knowledge regarding mental health in ELBW survivors, the present findings also have implications for theories of development. The DOHaD hypothesis posits that fetal adaptation to environmental influences encountered during gestation may increase susceptibility to chronic health problems later in

life. Findings that individuals born extremely preterm are more likely than their NBW peers to experience mental health problems appear to be consistent with this general framework.

Within the DOHaD hypothesis, different explanations have emerged, including the mismatch hypothesis and the cumulative stress hypothesis. The mismatch hypothesis is illustrated by fetal adaptation to aversive conditions that confer a relatively immediate advantage in utero, but do not fit the demands of the environment encountered after birth, with negative long-term consequences for the individual (Gluckman & Hanson, 2004). The cumulative stress model is similar to models of allostatic load. This hypothesis suggests that individuals who encounter aversive conditions during gestation may be sensitized by the exposure and become more vulnerable to aversive challenges that arise later in life (e.g., Davis, Waffarn, & Sandman, 2011).

Nederhof and Schmidt (2012) proposed an integration of these models, whereby both interpretations relate to individual differences in fetal sensitivity to programming effects—that is, fetal propensity to phenotypic plasticity. In this integrated model, fetuses whose endophenotypes are relatively more sensitive to programming will manifest higher levels of prenatal adaptation. If the environment to which they adapted becomes substantially different from the one anticipated, these fetuses are likely to experience an environmental mismatch. Those low-reactive fetuses whose endophenotypes are relatively less sensitive to programming are likely to show additive effects of stress when adversity is encountered again later in life. Developmental programming effects may thus represent specific manifestations of a more general biological mechanism, namely, a continuum of developmental plasticity (e.g., Barker, 2004; Gluckman et al., 2010).

If a fetus were to adapt extensively to adverse prenatal conditions, that could well result in an individual who thrives in environments where attention-switching and fast reactions are adaptive and ensure survival (match), but fares poorly in environments where sustained concentration is required for success, and impulsive behavior is detrimental (mismatch). In this case, the “adapted” fetus would be at increased risk of disease only if the environment failed to match the adaptation. On the other hand, if the fetus were less sensitive to programming, it might show limited adaptation, but still “take a hit” from prenatal adversity. In this scenario, weak adaptation may increase the risk of disease only slightly, depending on the number of additional “hits” experienced over time. As demonstrated in animal models, the effects of early adversity and exposure to stressors later in life may be additive (e.g., Choy, de Visser, & van den Buuse, 2009). For individuals with greater physiological or cognitive vulnerabilities than their peers, a modicum of cautiousness may be generally adaptive, but when additional life stressors arrive (such as the challenges of adolescence and young adulthood), dispositional cautiousness might intensify to the point where it becomes pathological, resulting in differentially higher levels of internalizing like those seen in individuals born at ELBW.

One of the predictions of the Nederhof and Schmidt (2012) model is high variability in the outcomes when studies fail to account for individual differences in sensitivity to programming. In the present study, the prediction intervals for both childhood and adolescent mental health problems varied considerably by syndrome. We suggest that the wide prediction intervals among ELBW mental health outcomes are more consistent with differential sensitivity to programming among individuals and do not preferentially support accounts of either cumulative stress or mismatch.

Mechanisms for programming sensitivity: Epigenetic modifications. In current thinking, one of the most plausible mechanisms for plasticity is epigenetic modification (Gluckman et al., 2010). Although genotypes are critical determinants of physiology, phenotypic variations may be brought about by environmental factors that can influence gene expression without altering the genotype itself. Processes such as DNA methylation (the addition of methyl groups to genetic material) often dampen gene expression, making characteristics that are related to “silenced” genes less likely to develop. On the other hand, changes in histone structure that facilitate transcription factor access to DNA tend to promote gene expression, increasing the likelihood of characteristics that are controlled by those genes (Moore, 2015). Importantly, nutritional and hormonal conditions experienced by the fetus during gestation may initiate epigenetic changes that are central to developmental plasticity (Gluckman, Hanson, & Beedle, 2007; Gluckman et al., 2010).

Epigenetic changes related to hormonal regulation may constitute a proximal physiological mechanism by which the fetus adapts to maternal stress (Meaney, 2010). In a recent human study, Monk et al. (2016) demonstrated an association between mothers’ perceived stress and an important index of neurobehavioral development in the fetus, namely, coupling between fetal heart rate and movement (DiPietro et al., 2010). The association was mediated by epigenetic effects: increased methylation of placental genes involved in the glucocorticoid pathway predicted reduced coupling in fetuses of stressed mothers. One of the placental genes examined by these researchers (HSD11β2) is associated with an enzyme that metabolizes maternal cortisol to produce inactive cortisone. Methylation of that gene may impede synthesis of the protective enzyme, thereby exposing the fetus to high levels of maternal cortisol.

As described by Ellison (2005), cortisol is heavily involved in mechanisms related to energy mobilization and allocation, and in particular, energy allocation trade-offs. Therefore, cortisol levels are likely to change in response to environmental conditions that affect energy allocation. Powerful gestational effects of cortisol are demonstrated in the final trimester of pregnancy when fetal cortisol levels rise (Donaldson, Nicolini, Symes, Rodeck, & Tannirandorn, 1991) and in the subsequent metabolic crisis experienced by the fetus near the end of gestation (Smith et al., 2005). When the energy demands of the developing fetus begin to exceed the energetic resources supplied through the placenta, hypothalamic-pituitary-adrenocortical (HPA) axis activity in the fetus increases to enhance the energetic supply. However, the ensuing rise in corticosteroids from the fetal side of the placenta may lead to a cascade of events that initiates labor. Importantly, the timing of this metabolic shift involving enhanced HPA activity may be influenced by conditions such as maternal diabetes, undernutrition, multiple pregnancies, or other energetic stressors, such that birth onset may be early or delayed (Ellison, 2005). Thus, the same mechanism—change in cortisol levels—appears to be able to affect the timing of delivery and developmental outcomes (e.g., Glover, O’Connor, & O’Donnell, 2010; Quesada, Tristão, Pratesi, & Wolf, 2014; Wüst, Entringer, Federenko, Schlotz, & Hellhammer, 2005). Exposure of the developing brain to conditions that

maintain high levels of cortisol during gestation may lead to poor outcomes, including mental health problems (Huang, 2011; Schlotz & Phillips, 2009).

Of course, prenatal programming effects do not account for all of the mental health risks associated with being born at ELBW. Equally important exposures occur when the preterm neonate is exposed to the world ex utero while still in a vulnerable state of development (Quesada et al., 2014). Premature exposure to physical sensations (e.g., light, sound, and handling), mechanical monitors, and painful, sometimes tissue-damaging treatments are stressful for preterm infants (Johnston, Barrington, Taddio, Carbajal, & Filion, 2011). However, in infants born too early—before the rise in fetal cortisol levels in late gestation (Donaldson et al., 1991)—the HPA axis is likely to be still in a suppressed state, resulting in adrenal insufficiency (Fernandez & Watterberg, 2009; Watterberg & Scott, 1995), whereby the amount of cortisol produced in response to a stressor is too little to maintain homeostasis. The ability to mount an effective cortisol response to stressors when it is needed is essential for survival, and adrenal insufficiency exposes preterm neonates to heightened risk of both short-term and long-term consequences of stress (Mörelius, He, & Shorey, 2016). Moreover, endocrine set points may be altered in children born extremely preterm (Grunau et al., 2007; Quesada et al., 2014). In both EP and full-term infants, basal cortisol levels decline continuously until 18 months of age (Grunau et al., 2007). Because the decline is noticeably greater in full-term infants, EP infants showing depressed basal levels of cortisol at 3 months show higher levels of baseline cortisol than do term-born children by 8 months. These relatively higher levels are maintained up to at least 18 months in EP infants, suggesting an enduring shift in cortisol regulation. We submit that durable effects of extremely preterm birth on endocrine and other systems, in combination with additional life stressors, may be sufficient to increase the risk for psychiatric problems in individuals born at ELBW.

Strengths and Limitations

This review provides a qualitative and quantitative overview of the current state of knowledge concerning mental health outcomes in individuals born extremely preterm. The review followed a formal protocol, examining 41 studies from 24 unique cohorts distributed across the world over an extensive period and a wide range of ages (5 to 36 years). Because we limited this review to individuals born at ELBW, individual differences in participants' baseline weight and gestational age did not vary much across the selected studies, ensuring a relatively homogenous study population. Pooling information from multiple cohorts generated much larger samples in case and control groups within each age category than could feasibly be collected by individual hospitals or laboratories, thereby reducing the impact of low statistical power. In studies of children and adolescents, information was often available from multiple informants.

Our findings also suggest some testable hypotheses for examining the pathogenesis of psychopathology following extremely preterm birth, including possible biological bases for particular mental health problems, and the question of whether prenatal exposure to corticosteroids constitutes a mechanism that compromises mental health. With improved understanding of mechanisms that plausibly affect mental health outcomes, more focused therapeutic interventions may be developed to lessen the impact of psychiatric problems in individuals born at ELBW.

It is also important to note several limitations when interpreting these findings. First, while methodological quality was relatively high, there was a substantial range of heterogeneity across studies. As well, mental health outcomes were reported from a wide variety of scales and range of measurement time points, even within developmental categories. Among the 41 reviewed studies, well over 40 different instruments were used to assess mental health between ages 5 and 36 years. Most, but not all of the assessment measures were validated. Further, we assumed that assessment instruments from different studies had comparable sensitivity and specificity. This overlooks existing subtle differences between instruments, as well as variation in the administration of those tests. Although we tried to minimize this concern by analyzing odds ratios from directly comparable questionnaires, standardization in the instruments used to assess mental health problems and, importantly, in the reporting of the findings was generally lacking (Milfont & Fischer, 2010; Van Lieshout, Boyle, Schmidt, Saigal, & Ferro, 2015b).

Second, issues of measurement invariance with respect to the assessment instruments commonly administered to this population should be acknowledged. Although it is generally assumed in psychological testing that a given scale measures the same psychological construct in all of the groups to which it is administered, if this is not true, then comparisons across ELBW and control groups will not be valid. In the only ELBW study to address this issue to date, measurement invariance was established more clearly for attention-deficit hyperactivity problems than for emotional and conduct problems in ELBW children (e.g., Van Lieshout, et al., 2015b).

Third, the number of extant studies of mental health in individuals born at ELBW declined sharply beyond childhood. While reports of mental health sequelae of extremely preterm birth were available in 25 studies of children, eligible adolescent studies numbered fewer than half that number (i.e., 10 studies). Most of the cohorts will have now reached adulthood, yet studies of adults were rare, with only two unique cohorts represented among the six adult studies eligible for this review. Part of the problem stems from attrition over long-term follow-up. Attrition itself is unlikely to overestimate any effects of ELBW because those lost to follow-up tend to have poorer outcomes (Breeman et al., 2015; Breslau, 1995). However, sample size reductions in cohorts that were already limited when first established make it more difficult to conduct studies in adolescence and adulthood. Therefore, our conclusions for adolescents and adults were based on less evidence than was available for children, and findings in these groups may be more variable or less stable. Given the limited availability of mental health outcomes data at older ages, we were unable to adequately address questions of whether particular mental health problems declined with the attainment of maturity or the nature of their trajectories.

Fourth, the number of studies entered in meta-analyses was typically a subset of all the studies that assessed similar outcomes. Some meta-analyses included only two studies, especially in analyses of adolescent data, allowing only cautious interpretations of the findings. However, the narrative findings from reports that

were not analyzed were similar to those derived from the meta-analyses, as described earlier.

In this review, we were also unable to parse any effects of being born SGA from those associated with being born at ELBW. Although researchers sometimes reported the percentage of their sample that was born SGA or experienced intrauterine growth restriction (15 of 41 studies), only two of these studies stratified their analyses by size for gestational age (SGA vs. appropriate for gestational age (AGA; Boyle et al., 2011; Van Lieshout et al., 2015a). Proportions of ELBW cohorts born SGA were relatively low, ranging from 7% (n = 6; Stahlmann, Rapp, Herting, & Thyen, 2009) to 47% (n = 26; Natalucci et al., 2013), with most studies falling between 12% and 25%. The etiology of mental health problems in individuals born at ELBW and SGA warrants further investigation (e.g., Abel et al., 2010; Lahti et al., 2015), particularly in light of differences within the SGA population between infants that experienced asymmetric prenatal growth (in which the body is disproportionately small at birth while the head is relatively large) and those that experienced symmetric but limited growth, in which constraints affected body and head more proportionately. Of interest, the Boyle et al. (2011) and Van Lieshout et al. (2015a) studies (of the same cohort at different adult ages) reported dose effects of preterm status on mental health risk: Those born at ELBW and SGA were more likely than adults born at ELBW and AGA to have psychiatric problems, who, in turn, were at greater risk than were term-born controls born at NBW.

Recommendations for Further Research

Limitations in the literature to date are considerable barriers to a scientific understanding of the pathogenesis and treatment of psychological problems in ELBW and EP survivors. They also impede clinical efforts to prevent or treat mental disorders in this population. We therefore make the following recommendations in an attempt to advance the field. First, it is important that individuals born at ELBW be followed into adulthood if possible. We know that survival rates have increased without a comparable decrease in morbidity, and that difficulties with certain mental health problems appear to persist in young adulthood. If the effects of extremely preterm birth on mental health are enduring, it will be important to collect data on ELBW survivors as they reach young adulthood, mature, and age. Without follow-up data from older ages, it is impossible to know how and when to treat mental health concerns that change with age, or which concerns remain troublesome in adulthood.

Second, we recommend that individual studies attempt to use similar, valid measures across developmental epochs, if possible. Longitudinal studies not only constrain idiosyncratic variability, allowing a clearer view of symptoms and their trajectories over time, but also allow issues like homotypic or heterotypic continuity to be addressed. These kinds of models may be particularly useful in understanding the nature of mental health symptoms in individuals born at ELBW, especially if they are rooted in early occurring biological processes.

Third, despite the fact that the issue of wide variation in measurement instruments across different cohorts was noted in a review of low birth weight studies from 2 decades ago (Breslau, 1995), this problem continues to make direct comparisons across studies problematic, even for recent cohorts. To address it, the measures used to assess mental health should be widely adopted, and the use of gold standard instruments should be more widespread than the use of questionnaires better suited for screening for mental disorders. Given the clinical and resource implications of psychiatric diagnoses, as well as the stigma attached to mental health difficulties, assessments of mental health should be based on more than a few questions. More comprehensive assessments would allow greater confidence in estimated rates of risk, and more detailed descriptions from participants would help establish the phenomenology of mental health problems in ELBW.

Fourth, given differences in the brains of ELBW survivors and the unique behavioral phenotypes hypothesized to be present in preterm survivors, it is not necessarily true that applied measures tap the same constructs in this population and NBW controls. Therefore, it behooves researchers to establish measurement invariance for measures of mental health, and then to use the invariant measures to make accurate estimations of the mental health risks in ELBW.

Fifth, studies of individuals born at ELBW are frequently underpowered, due to the still limited numbers of survivors of extremely preterm birth. If common data collection protocols were established to allow for appropriate combining of data, data from multiple centers could be pooled to produce larger samples of this relatively rare population, permitting meta-analyses of individual participants' data. The use of population-based, linked registry data will also provide more accurate estimates of risk, although such estimates may be limited by diagnostic coverage or lack data on mediators and moderators of observed links.

Sixth, the time has come for studies in the field to progress from descriptive to explanatory accounts, another issue raised previously (e.g., McCormick, 1997). If some subgroups of the ELBW population drive the risk of adverse mental health outcomes more than others, then careful examination of potential social and biological moderators and mediators of these outcomes is indicated to facilitate discovery of factors that promote mental health (or not) following extremely preterm birth. More attention needs to be paid to identifying factors precipitating or mitigating risks of mental health problems in ELBW cohorts, including underlying pregnancy conditions and heterogeneity of preterm birth, parenting style (e.g., Huhtala et al., 2014; Hoff, Hansen, Munck, & Mortensen, 2004; Pyhälä et al., 2011), socioeconomic status (SES; e.g., Hack et al., 2009; Taylor et al., 2015), cognitive functioning (e.g., Hoff et al., 2004; Nadeau, Boivin, Tessier, Lefebvre, & Robaey, 2001), neonatal interventions (e.g., antenatal corticosteroid exposure, early nutrition), intermediate biological phenotypes (e.g., in EEG or neuroimaging, e.g., Schmidt et al., 2010), and genetic (e.g., Lahat, Van Lieshout, Mathewson, et al., 2016) and epigenetic (e.g., Cruickshank et al., 2013) markers. All of these factors are potential sources of heterogeneity in samples of individuals born preterm. A more complete understanding of these risk and resilience factors would not only allow us to better understand the pathophysiology of mental disorders in this population, but to predict those at highest risk. They would also aid in the development and targeting of preventive and treatment efforts.

Finally, more comprehensive findings should be translated into optimal care protocols and interventions for particular mental health problems, with the optimal time-points for administering them. Although the Neonatal Intensive Care Unit stay is an obvious time for some interventions, others may need to be implemented following the appearance of symptoms later during development. Our findings suggest that certain mental health problems in ELBW survivors may persist into adulthood, leading to an accumulation of morbidity over time. Therefore, it is important that families and clinicians be aware of these early problems and that appropriate treatment be made available to optimize development and quality of life. Follow-up of children born at ELBW should expand from physical and medical considerations to include psychological and family services programs (Farooqi et al., 2007), to inform planning with respect to mental health care during development.

Summary and Conclusions

Findings from this review and meta-analysis provide substantial evidence that individuals born at ELBW are at greater overall risk for psychological difficulties than are their NBW peers. This does not mean that, in general, infants born extremely preterm will ultimately develop mental health problems—only that the risk of developing such problems is higher in this group than in those born at full term. The difficulties most frequently involve attentional control, anxiety-related problems, and social interaction problems, including social withdrawal. Subgroup analyses suggested that these findings are relatively robust to region of birth, secular changes in care, and the presence of NSI. Regional differences, effects of birth era, and NSI effects on mental health problems principally involved attention problems, internalizing, and ASD. Parent and teacher ratings yielded comparable group differences in mental health symptoms between ELBW and NBW children. Although parents were also likely to report significant attentional and social difficulties in teens born at ELBW, teens themselves identified greater attention problems and oppositional behavior in typically developing adolescents born at NBW. The wide prediction intervals for most disorders remain to be explained, suggesting additional moderator variables be examined in future studies.

The complex needs faced by children born extremely preterm continue throughout development, and include mental health problems that are likely to have long-term consequences for personal relationships, vocational success, and psychological well-being. Innovation will be needed to develop interventions that are neuroprotective (Hutchinson et al., 2013; Stahlmann et al., 2009), and socially supportive (Greenley et al., 2007).

References

*References marked with an asterisk indicate studies included in the meta-analysis.

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics, 124,* 717–728. http://dx.doi.org/10.1542/peds.2008-2816

Abel, K. M., Wicks, S., Susser, E. S., Dalman, C., Pedersen, M. G., Mortensen, P. B., & Webb, R. T. (2010). Birth weight, schizophrenia, and adult mental disorder: Is risk confined to the smallest babies? *Archives of General Psychiatry, 67,* 923–930. http://dx.doi.org/10.1001/archgenpsychiatry.2010.100

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 profile.* Burlington, VT: Department of Psychiatry, University of Vermont.

Achenbach, T. M. (2011). Commentary: Definitely more than measurement error: But how should we understand and deal with informant discrepancies? *Journal of Clinical Child and Adolescent Psychology, 40,* 80–86. http://dx.doi.org/10.1080/15374416.2011.533416

Achenbach, T. M., Bernstein, A., & Dumenci, L. (2005). DSM-oriented scales and statistically-based syndromes for ages 18 to 59: Linking taxonomic paradigms to facilitate multitaxonomic to approaches. *Journal of Personality, 84,* 49–63. http://dx.doi.org/10.1207/s15327752jpa8401_10

Allin, M., Rooney, M., Cuddy, M., Wyatt, J., Walshe, M., Rifkin, L., & Murray, R. (2006). Personality in young adults who are born preterm. *Pediatrics, 117,* 309–316. http://dx.doi.org/10.1542/peds.2005-0539

*Anderson, P. J., De Luca, C. R., Hutchinson, E., Spencer-Smith, M. M., Roberts, G., & Doyle, L. W., & the Victorian Infant Collaborative Study Group. (2011). Attention problems in a representative sample of extremely preterm/extremely low birth weight children. *Developmental Neuropsychology, 36,* 57–73. http://dx.doi.org/10.1080/87565641.2011.540538

*Anderson, P., & Doyle, L. W., & the Victorian Infant Collaborative Study Group. (2003). Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. *Journal of the American Medical Association, 289,* 3264–3272. http://dx.doi.org/10.1001/jama.289.24.3264

Aylward, G. P. (2005). Neurodevelopmental outcomes of infants born prematurely. *Journal of Developmental and Behavioral Pediatrics, 26,* 427–440. http://dx.doi.org/10.1097/00004703-200512000-00008

Barker, D. J. P. (2004). The developmental origins of well-being. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 359,* 1359–1366. http://dx.doi.org/10.1098/rstb.2004.1518

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94. http://dx.doi.org/10.1037/0033-2909.121.1.65

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, 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. http://dx.doi.org/10.1001/jama.288.6.728

Bilgin, A., & Wolke, D. (2015). Maternal sensitivity in parenting preterm children: A Meta-analysis. *Pediatrics, 136,* e177–e193. http://dx.doi.org/10.1542/peds.2014-3570

Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A. B., . . . the Born Too Soon Preterm Birth Action Group. (2013). Born too soon: The global epidemiology of 15 million preterm births. *Reproductive Health, 10*(Suppl. 1), S2. Retrieved from http://www.reproductive-health-journal.com/content/10/S1/S2. http://dx.doi.org/10.1186/1742-4755-10-S1-S2

Bonanno, C., & Wapner, R. J. (2012). Antenatal corticosteroids in the management of preterm birth: Are we back where we started? *Obstetrics and Gynecology Clinics of North America, 39,* 47–63. http://dx.doi.org/10.1016/j.ogc.2011.12.006

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis.* Chichester, UK: Wiley. http://dx.doi.org/10.1002/9780470743386

*Boyle, M. H., Miskovic, V., Van Lieshout, R., Duncan, L., Schmidt, L. A., Hoult, L., . . . Saigal, S. (2011). Psychopathology in young adults born at extremely low birth weight. *Psychological Medicine, 41,* 1763–1774. http://dx.doi.org/10.1017/S0033291710002357

Breeman, L. D., Kaekel, J., Baumann, N., Bartmann, P., & Wolke, D. (2015). Attention problems in very preterm children from childhood to adulthood: The Bavarian longitudinal study. *Journal of Child Psychology and Psychiatry, 5792,* 132–140. http://dx.doi.org/10.1111/jcpp.12456

Breslau, N. (1995). Psychiatric sequelae of low birth weight. *Epidemiologic Reviews, 17,* 96–106. http://dx.doi.org/10.1007/BF01441637

*Burnett, A., Davey, C. G., Wood, S. J., Wilson-Ching, M., Molloy, C., Cheong, J. L. Y., . . . Anderson, P. J. (2014). Extremely preterm birth and adolescent mental health in a geographical cohort born in the 1990s. *Psychological Medicine, 44,* 1533–1544. http://dx.doi.org/10.1017/S0033291713002158

Choy, K. H., de Visser, Y. P., & van den Buuse, M. (2009). The effect of 'two hit' neonatal and young-adult stress on dopaminergic modulation of prepulse inhibition and dopamine receptor density. *British Journal of Pharmacology, 156,* 388–396. http://dx.doi.org/10.1111/j.1476-5381.2008.00008.x

*Conrad, A. L., Richman, L., Lindgren, S., & Nopoulos, P. (2010). Biological and environmental predictors of behavioral sequelae in children born preterm. *Pediatrics, 125,* e83–e89. http://dx.doi.org/10.1542/peds.2009-0634

Crowley, P., Chalmers, I., & Keirse, M. J. (1990). The effects of corticosteroid administration before preterm delivery: An overview of the evidence from controlled trials. *British Journal of Obstetrics and Gynaecology, 97,* 11–25. http://dx.doi.org/10.1111/j.1471-0528.1990.tb01711.x

Cruickshank, M. N., Oshlack, A., Theda, C., Davis, P. G., Martino, D., Sheehan, P., . . . Craig, J. M. (2013). Analysis of epigenetic changes in survivors of preterm birth reveals the effect of gestational age and evidence for a long term legacy. *Genome Medicine, 5,* 96. http://dx.doi.org/10.1186/gm500

Davis, E. P., Waffarn, F., & Sandman, C. A. (2011). Prenatal treatment with glucocorticoids sensitizes the hpa axis response to stress among full-term infants. *Developmental Psychobiology, 53,* 175–183. http://dx.doi.org/10.1002/dev.20510

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2011). Combining dichotomous and continuous outcomes. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* [Version 5.1.0] (updated March 2011). The Cochrane Collaboration. Retrieved January 3, 2017, from http://handbook.cochrane.org/

De Los Reyes, A. (2011). Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 40,* 1–9. http://dx.doi.org/10.1080/15374416.2011.533405

DiPietro, J. A., Kivlighan, K. T., Costigan, K. A., Rubin, S. E., Shiffler, D. E., Henderson, J. L., & Pillion, J. P. (2010). Prenatal antecedents of newborn neurological maturation. *Child Development, 81,* 115–130. http://dx.doi.org/10.1111/j.1467-8624.2009.01384.x

Donaldson, A., Nicolini, U., Symes, E. K., Rodeck, C. H., & Tannirandorn, Y. (1991). Changes in concentrations of cortisol, dehydroepiandrosterone sulphate and progesterone in fetal and maternal serum during pregnancy. *Horumon To Rinsho, 35,* 447–451. http://dx.doi.org/10.1111/j.1365-2265.1991.tb03564.x

Doyle, L. W. (2006). Evaluation of neonatal intensive care for extremely-low-birth-weight infants. *Seminars in Fetal & Neonatal Medicine, 11,* 139–145. http://dx.doi.org/10.1016/j.siny.2005.11.009

Doyle, L. W., Roberts, G., & Anderson, P. J., & the Victorian Infant Collaborative Study Group. (2011). Changing long-term outcomes for infants 500–999 g birth weight in Victoria, 1979–2005. *Archives of Disease in Childhood-Fetal and Neonatal Edition, 96*(6), F443–F447. http://dx.doi.org/10.1136/adc.2010.200576

Doyle, L. W., & the Victorian Infant Collaborative Study Group. (2004). Neonatal intensive care at borderline viability—Is it worth it? *Early Human Development, 80,* 103–113. http://dx.doi.org/10.1016/j.earlhumdev.2004.05.009

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629

*Elgen, S. K., Leversen, K. T., Grundt, J. H., Hurum, J., Sundby, A. B., Elgen, I. B., & Markestad, T. (2012). Mental health at 5 years among children born extremely preterm: A national population-based study. *European Child & Adolescent Psychiatry, 21,* 583–589. http://dx.doi.org/10.1007/s00787-012-0298-1

Ellison, P. T. (2005). Evolutionary perspectives on the fetal origins hypothesis. *American Journal of Human Biology, 17,* 113–118. http://dx.doi.org/10.1002/ajhb.20097

Fanaroff, A. A., Hack, M., & Walsh, M. C. (2003). The NICHD neonatal research network: Changes in practice and outcomes during the first 15 years. *Seminars in Perinatology, 27,* 281–287. http://dx.doi.org/10.1016/S0146-0005(03)00055-7

Fanaroff, A. A., Wright, L. L., Stevenson, D. K., Shankaran, S., Donovan, E. F., Ehrenkranz, R. A., . . . Verter, J. (1995). Very-low-birth-weight outcomes of the National Institute of Child Health and Human Development Neonatal Research Network, May 1991 through December 1992. *American Journal of Obstetrics and Gynecology, 173,* 1423–1431. http://dx.doi.org/10.1016/0002-9378(95)90628-2

*Farooqi, A., Hägglöf, B., Sedin, G., Gothefors, L., & Serenius, F. (2007). 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, 120,* 118–133. http://dx.doi.org/10.1542/peds.2006-2988

*Farooqi, A., Hägglöf, B., & Serenius, F. (2013). Behaviours related to executive functions and learning skills at 11 years of age after extremely preterm birth: A Swedish national prospective follow-up study. *Acta Paediatrica, 102,* 625–634. http://dx.doi.org/10.1111/apa.12219

Fernandez, E. F., & Watterberg, K. L. (2009). Relative adrenal insufficiency in the preterm and term infant. *Journal of Perinatology, 29*(Suppl. 2), S44–S49. http://dx.doi.org/10.1038/jp.2009.24

Ferrara, T. B., Hoekstra, R. E., Couser, R. J., Jackson, J. C., Anderson, C. L., Myers, T. F., & Raye, J. R. (1991). Effects of surfactant therapy on outcome of infants with birth weights of 600 to 750 grams. *The Journal of Pediatrics, 119,* 455–457. http://dx.doi.org/10.1016/S0022-3476(05)82062-1

*Fevang, S. K. E., Hysing, M., Markestad, T., & Sommerfelt, K. (2016). Mental health in children born extremely preterm without severe neurodevelopmental difficulties. *Pediatrics, 137,* 1–11. http://dx.doi.org/10.1542/peds.2015-3002

Field, D. J., Dorling, J. S., Manktelow, B. N., & Draper, E. S. (2008). Survival of extremely premature babies in a geographically defined population: Prospective cohort study of 1994–9 compared with 2000–5. *British Medical Journal, 336,* 1221–1223. http://dx.doi.org/10.1136/bmj.39555.670718.BE

French, N. P., Hagan, R., Evans, S. F., Mullan, A., & Newnham, J. P. (2004). Repeated antenatal corticosteroids: Effects on cerebral palsy and childhood behavior. *American Journal of Obstetrics and Gynecology, 190,* 588–595. http://dx.doi.org/10.1016/j.ajog.2003.12.016

Ganella, E. P., Burnett, A., Cheong, J., Thompson, D., Roberts, G., Wood, S., . . . Doyle, L. W. (2015). Abnormalities in orbitofrontal cortex gyrification and mental health outcomes in adolescents born extremely preterm and/or at an extremely low birth weight. *Human Brain Mapping, 36,* 1138–1150. http://dx.doi.org/10.1002/hbm.22692

*Gardner, F., Johnson, A., Yudkin, P., Bowler, U., Hockley, C., Mutch, L., . . . the Extremely Low Gestational Age Steering Group. (2004). Behavioral and emotional adjustment of teenagers in mainstream school who were born before 29 weeks' gestation. *Pediatrics, 114,* 676–682. http://dx.doi.org/10.1542/peds.2003-0763-L

Glover, V., O'Connor, T. G., & O'Donnell, K. (2010). Prenatal stress and the programming of the HPA axis. *Neuroscience and Biobehavioral Reviews, 35,* 17–22. http://dx.doi.org/10.1016/j.neubiorev.2009.11.008

Gluckman, P. D., & Hanson, M. A. (2004). Developmental origins of disease paradigm: A mechanistic and evolutionary perspective. *Pediatric Research, 56,* 311–317. http://dx.doi.org/10.1203/01.PDR.0000135998.08025.FB

Gluckman, P. D., Hanson, M. A., & Beedle, A. S. (2007). Early life events and their consequences for later disease: A life history and evolutionary perspective. *American Journal of Human Biology, 19,* 1–19. http://dx.doi.org/10.1002/ajhb.20590

Gluckman, P. D., Hanson, M. A., & Buklijas, T. (2010). A conceptual framework for the developmental origins of health and disease. *Journal of Developmental Origins of Health and Disease, 1,* 6–18. http://dx.doi.org/10.1017/S2040174409990171

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1337–1345. http://dx.doi.org/10.1097/00004583-200111000-00015

Göpel, W., Kribs, A., Härtel, C., Avenarius, S., Teig, N., Groneck, P., . . . the German Neonatal Network (GNN). (2015). Less invasive surfactant administration is associated with improved pulmonary outcomes in spontaneously breathing preterm infants. *Acta Paediatrica, 104,* 241–246. http://dx.doi.org/10.1111/apa.12883

*Greenley, R. N., Taylor, H. G., Drotar, D., & Minich, N. M. (2007). Longitudinal relationships between early adolescent family functioning and youth adjustment: An examination of the moderating role of very low birth weight. *Journal of Pediatric Psychology, 32,* 453–462. http://dx.doi.org/10.1093/jpepsy/jsl027

Grunau, R. E., Haley, D. W., Whitfield, M. F., Weinberg, J., Yu, W., & Thiessen, P. (2007). Altered basal cortisol levels at 3, 6, 8 and 18 months in infants born at extremely low gestational age. *Journal of Pediatrics, 150,* 151–156. http://dx.doi.org/10.1016/j.jpeds.2006.10.053

*Grunau, R. E., Whitfield, M. F., & Fay, T. B. (2004). Psychosocial and academic characteristics of extremely low birth weight (< or =800 g) adolescents who are free of major impairment compared with term-born control subjects. *Pediatrics, 114*(6), e725–e732. http://dx.doi.org/10.1542/peds.2004-0932

Hack, M., & Costello, D. W. (2008). Trends in the rates of cerebral palsy associated with neonatal intensive care of preterm children. *Clinical Obstetrics and Gynecology, 51,* 763–774. http://dx.doi.org/10.1097/GRF.0b013e3181870922

Hack, M., & Fanaroff, A. A. (1999). Outcomes of children of extremely low birthweight and gestational age in the 1990s. *Early Human Development, 53,* 193–218. http://dx.doi.org/10.1016/S0378-3782(98)00052-8

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. http://dx.doi.org/10.1056/NEJMoa010856

Hack, M., & Klein, N. (2006). Young adult attainments of preterm infants. *Journal of the American Medical Association, 295,* 695–696. http://dx.doi.org/10.1001/jama.295.6.695

*Hack, M., Taylor, H. G., Schluchter, M., Andreias, L., Drotar, D., & Klein, N. (2009). Behavioral outcomes of extremely low birth weight children at age 8 years. *Journal of Developmental and Behavioral Pediatrics, 30,* 122. http://dx.doi.org/10.1097/DBP.0b013e31819e6a16

*Hallin, A. L., & Stjernqvist, K. (2011a). Adolescents born extremely preterm: Behavioral outcomes and quality of life. *Scandinavian Journal of Psychology, 52,* 251–256. http://dx.doi.org/10.1111/j.1467-9450.2010.00850.x

*Hallin, A. L., & Stjernqvist, K. (2011b). Follow-up of adolescents born extremely preterm: Self-perceived mental health, social and relational outcomes. *Acta Paediatrica, 100,* 279–283. http://dx.doi.org/10.1111/j.1651-2227.2010.01993.x

Halmøy, A., Klungsøyr, K., Skjærven, R., & Haavik, J. (2012). Pre- and perinatal risk factors in adults with attention-deficit/hyperactivity disorder. *Biological Psychiatry, 71,* 474–481. http://dx.doi.org/10.1016/j.biopsych.2011.11.013

*Hanke, C., Lohaus, A., Gawrilow, C., Hartke, I., Köhler, B., & Leonhardt, A. (2003). Preschool development of very low birth weight children born 1994–1995. *European Journal of Pediatrics, 162,* 159–164.

Higgins, J. P. T., Altman, D. G., & Sterne, J. A. C. (2011). Assessing risk of bias in included studies. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* [Version 5.1.0] (updated March 2011). The Cochrane Collaboration. Retrieved January 3, 2017, from http://handbook.cochrane.org/

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

*Hille, E. T., den Ouden, A. L., Saigal, S., Wolke, D., Lambert, M., Whitaker, A., . . . Paneth, N. (2001). Behavioural problems in children who weigh 1000 g or less at birth in four countries. *Lancet, 357,* 1641–1643. http://dx.doi.org/10.1016/S0140-6736(00)04818-2

Hoekstra, R. E., Ferrara, T. B., Couser, R. J., Payne, N. R., & Connett, J. E. (2004). Survival and long-term neurodevelopmental outcome of extremely preterm infants born at 23–26 weeks' gestational age at a tertiary center. *Pediatrics, 133*(1), e1–e6. http://dx.doi.org/10.1542/peds.113.1.e1

*Hoff, B., Hansen, B. M., Munck, H., & Mortensen, E. L. (2004). Behavioral and social development of children born extremely premature: 5-year follow-up. *Scandinavian Journal of Psychology, 45,* 285–292. http://dx.doi.org/10.1111/j.1467-9450.2004.00407.x

Horbar, J. D., Wright, E. C., & Onstad, L., & The Members of the National Institute of Child Health and Human Development Neonatal Research Network. (1993). Decreasing mortality associated with the introduction of surfactant therapy: An observational study of neonates weighing 601 to 1300 grams at birth. *Pediatrics, 92,* 191–196.

*Horwood, L. J., Mogridge, N., & Darlow, B. A. (1998). Cognitive, educational, and behavioural outcomes at 7 to 8 years in a national very low birthweight cohort. *Archives of Disease in Childhood Fetal and Neonatal Edition, 79*(1), F12–F20. http://dx.doi.org/10.1136/fn.79.1.F12

Huang, L. T. (2011). The link between perinatal glucocorticoids exposure and psychiatric disorders. *Pediatric Research, 69,* 19R–25R. http://dx.doi.org/10.1203/PDR.0b013e318212c29b

Huhtala, M., Korja, R., Lehtonen, L., Haataja, L., Lapinleimu, H., & Rautava, P., & the PIPARI Study Group. (2014). Associations between parental psychological well-being and socio-emotional development in 5-year-old preterm children. *Early Human Development, 90,* 119–124. http://dx.doi.org/10.1016/j.earlhumdev.2013.12.009

*Hutchinson, E. A., De Luca, C. R., Doyle, L. W., Roberts, G., Anderson, P. J., & the Victorian Infant Collaborative Study Group. (2013). School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics, 131*(4), e1053–e1061. http://dx.doi.org/10.1542/peds.2012-2311

Jobe, A. H., Mitchell, B. R., & Gunkel, J. H. (1993). Beneficial effects of the combined use of prenatal corticosteroids and postnatal surfactant on preterm infants. *American Journal of Obstetrics and Gynecology, 168,* 508–513. http://dx.doi.org/10.1016/0002-9378(93)90483-Y

Johnson, S. (2007). Cognitive and behavioural outcomes following very preterm birth. *Seminars in Fetal & Neonatal Medicine, 12,* 363–373. http://dx.doi.org/10.1016/j.siny.2007.05.004

*Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., & Marlow, N. (2010a). Autism spectrum disorders in extremely preterm children. *Journal of Pediatrics, 156,* 525–531. http://dx.doi.org/10.1016/j.jpeds.2009.10.041

*Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., & Marlow, N. (2010b). Psychiatric disorders in extremely preterm children: Longitudinal finding at age 11 years in the EPICure study. *Journal of the American Academy of Child and Adolescent Psychiatry, 49,* 453–463.

*Johnson, S., Kochhar, P., Hennessy, E., Marlow, N., Wolke, D., & Hollis, C. (2016). Antecedents of attention-deficit/hyperactivity disorder symptoms in children born extremely preterm. *Journal of Developmental and Behavioral Pediatrics, 37,* 285–297. http://dx.doi.org/10.1097/DBP.0000000000000298

Johnson, S., & Marlow, N. (2011). Preterm birth and childhood psychiatric disorders. *Pediatric Research, 69*(5 Part 2), 11R–18R. http://dx.doi.org/10.1203/PDR.0b013e318212faa0

Johnson, S., & Marlow, N. (2014). Growing up after extremely preterm birth: Lifespan mental health outcomes. *Seminars in Fetal & Neonatal Medicine, 19,* 97–104. http://dx.doi.org/10.1016/j.siny.2013.11.004

Johnson, S., & Wolke, D. (2013). Behavioural outcomes and psychopathology during adolescence. *Early Human Development, 89,* 199–207. http://dx.doi.org/10.1016/j.earlhumdev.2013.01.014

Johnston, C., Barrington, K. J., Taddio, A., Carbajal, R., & Filion, F. (2011). Pain in Canadian NICUs: Have we improved over the past 12 years? *Clinical Journal of Pain, 27,* 225–232. http://dx.doi.org/10.1097/AJP.0b013e3181fe14cf

Joseph, K. S., Kramer, M. S., Marcoux, S., Ohlsson, A., Wen, S. W., Allen, A., & Platt, R. (1998). Determinants of preterm birth rates in Canada from 1981 through 1983 and from 1992 through 1994. *The New England Journal of Medicine, 339,* 1434–1439. http://dx.doi.org/10.1056/NEJM199811123392004

Kohen, D. E., Brooks-Gunn, J., McCormick, M., & Graber, J. A. (1997). Concordance of maternal and teacher ratings of school and behavior problems in children of varying birth weights. *Journal of Developmental and Behavioral Pediatrics, 18,* 295–303. http://dx.doi.org/10.1097/00004703-199710000-00003

Lahat, A., Van Lieshout, R. J., Mathewson, K. J., Mackillop, J., Saigal, S., Morrison, K. M., . . . Schmidt, L. A. (2016). Extremely low birth weight babies grown up: Gene–environment interaction predicts internalizing problems in the third and fourth decades of life. *Development and Psychopathology,* 1–7. http://dx.doi.org/10.1017/s0954579416000511

*Lahat, A., Van Lieshout, R. J., Saigal, S., Boyle, M. H., & Schmidt, L. A. (2014). ADHD among young adults born at extremely low birth weight: The role of fluid intelligence in childhood. *Frontiers in Psychology, 5,* 446. http://dx.doi.org/10.3389/fpsyg.2014.00446

Lahti, M., Eriksson, J. G., Heinonen, K., Kajantie, E., Lahti, J., Wahlbeck, K., . . . Räikkönen, K. (2015). Late preterm birth, post-term birth, and abnormal fetal growth as risk factors for severe mental disorders from early to late adulthood. *Psychological Medicine, 45,* 985–999. http://dx.doi.org/10.1017/S0033291714001998

Lau, C., Ambalavanan, N., Chakraborty, H., Wingate, M. S., & Carlo, W. A. (2013). Extremely low birth weight and infant mortality rates in the United States. *Pediatrics, 131,* 855–860. http://dx.doi.org/10.1542/peds.2012-2471

Litt, J., Taylor, H. G., Klein, N., & Hack, M. (2005). Learning disabilities in children with very low birthweight: Prevalence, neuropsychological correlates, and educational interventions. *Journal of Learning Disabilities, 38,* 130–141. http://dx.doi.org/10.1177/00222194050380020301

Mathiasen, R., Hansen, B. M., Nybo Anderson, A. M., & Greisen, G. (2009). Socio-economic achievements of individuals born very preterm at the age of 27 to 29 years: A nationwide cohort study. *Developmental Medicine and Child Neurology, 51,* 901–908. http://dx.doi.org/10.1111/j.1469-8749.2009.03331.x

McCormick, M. C. (1997). The outcomes of very low birth weight infants: Are we asking the right questions? *Pediatrics, 99,* 869–876. http://dx.doi.org/10.1542/peds.99.6.869

McCormick, M. C., Workman-Daniels, K., & Brooks-Gunn, J. (1996). The behavioral and emotional well-being of school-age children with different birth weights. *Pediatrics, 97,* 18–25.

Meaney, M. J. (2010). Epigenetics and the biological definition of gene x environment interactions. *Child Development, 81,* 41–79. http://dx.doi.org/10.1111/j.1467-8624.2009.01381.x

*Methúsalemsdóttir, H. F., Egilson, S. Þ., Guðmundsdóttir, R., Valdimarsdóttir, U. A., & Georgsdóttir, I. (2013). Quality of life of adolescents born with extremely low birth weight. *Acta Paediatrica, 102,* 597–601. http://dx.doi.org/10.1111/apa.12192

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3,* 111–121.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151,* 264–269. http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135

Monk, C., Feng, T., Lee, S., Krupska, I., Champagne, F. A., & Tycko, B. (2016). Distress during pregnancy: Epigenetic regulation of placenta glucocorticoid-related genes and fetal neurobehavior. *American Journal of Psychiatry, 173,* 705–713. http://dx.doi.org/10.1176/appi.ajp.2015.15091171

Moore, D. S. (2015). *The developing genome: An introduction to behavioral epigenetics* (pp. 65–74). New York, NY: Oxford University Press.

Mörelius, E., He, H. G., & Shorey, S. (2016). Salivary cortisol reactivity in preterm infants in neonatal intensive care: An integrative review. *International Journal of Environmental Research and Public Health, 13,* 337. http://dx.doi.org/10.3390/ijerph13030337

Moster, D., Lie, R. T., & Markestad, T. (2008). Long-term medical and social consequences of preterm birth. *The New England Journal of Medicine, 359,* 262–273. http://dx.doi.org/10.1056/NEJMoa0706475

Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Development of executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology, 34,* 393–421. http://dx.doi.org/10.1080/87565640902964524

Murphy, K. R., Barkley, R. A., & Bush, T. (2002). Young adults with attention deficit hyperactivity disorder: Subtype differences in comorbidity, educational, and clinical history. *Journal of Nervous and Mental Disease, 190,* 147–157. http://dx.doi.org/10.1097/00005053-200203000-00003

Nadeau, L., Boivin, M., Tessier, R., Lefebvre, F., & Robaey, P. (2001). Mediators of behavioral problems in 7-year-old children born after 24 to 28 weeks of gestation. *Journal of Developmental and Behavioral Pediatrics, 22,* 1–10. http://dx.doi.org/10.1097/00004703-200102000-00001

*Natalucci, G., Becker, J., Becher, K., Bickle, G. M., Landolt, M. A., & Bucher, H. U. (2013). Self-perceived health status and mental health outcomes in young adults born with less than 1000 g. *Acta Paediatrica, 102,* 294–299. http://dx.doi.org/10.1111/apa.12102

Nederhof, E., & Schmidt, M. V. (2012). Mismatch or cumulative stress: Toward an integrated hypothesis of programming effects. *Physiology & Behavior, 106,* 691–700. http://dx.doi.org/10.1016/j.physbeh.2011.12.008

Neyeloff, J. L., Fuchs, S. C., & Moreira, L. B. (2012). Meta-analyses and Forest plots using a microsoft excel spreadsheet: Step-by-step guide focusing on descriptive data analysis. *BMC Research Notes, 5,* 52. http://dx.doi.org/10.1186/1756-0500-5-52

Pelkonen, A. S., Hakulinen, A. L., Turpeinen, M., & Hallman, M. (1998). Effect of neonatal surfactant therapy on lung function at school age in children born very preterm. *Pediatric Pulmonology, 25,* 182–190. http://dx.doi.org/10.1002/(SICI)1099-0496(199803)25:3<182::AID-PPUL8>3.0.CO;2-O

Pesonen, A. K., Räikkönen, K., Heinonen, K., Andersson, S., Hovi, P., Järvenpää, A. L., . . . Kajantie, E. (2008). Personality of young adults

born prematurely: The Helsinki study of very low birth weight adults. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 49,* 609–617. http://dx.doi.org/10.1111/j.1469-7610.2007.01874.x

Pharoah, P. O., Stevenson, C. J., Cooke, R. W., & Stevenson, R. C. (1994). Prevalence of behaviour disorders in low birthweight infants. *Archives of Disease in Childhood, 70,* 271–274. http://dx.doi.org/10.1136/adc.70.4.271

Pyhälä, R., Räikkönen, K., Pesonen, A. K., Heinonen, K., Lahti, J., Hovi, P., . . . Kajantie, E. (2011). Parental bonding after preterm birth: Child and parent perspectives in the Helsinki study of very low birth weight adults. *Journal of Pediatrics, 158,* 251–256. http://dx.doi.org/10.1016/j.jpeds.2010.07.059

Quesada, A. A., Tristão, R. M., Pratesi, R., & Wolf, O. T. (2014). Hyper-responsiveness to acute stress, emotional problems and poorer memory in former preterm children. *Stress, 17,* 389–399. http://dx.doi.org/10.3109/10253890.2014.949667

Rieger, M., Pirke, K. M., Buske-Kirschbaum, A., Wurmser, H., Papoušek, M., & Hellhammer, D. H. (2004). Influence of stress during pregnancy on HPA activity and neonatal behavior. *Annals of the New York Academy of Sciences, 1032,* 228–230. http://dx.doi.org/10.1196/annals.1314.026

Riley, K., Roth, S., Sellwood, M., & Wyatt, J. S. (2008). Survival and neurodevelopmental morbidity at 1 year of age following extremely preterm delivery over a 20-year period: A single centre cohort study. *Acta Paediatrica, 97,* 159–165. http://dx.doi.org/10.1111/j.1651-2227.2007.00637.x

Saigal, S., Day, K. L., Van Lieshout, R. J., Schmidt, L. A., Morrison, K. M., & Boyle, M. H. (2016). Health, wealth, social integration, and sexuality of extremely low-birth-weight prematurely born adults in the fourth decade of life. *JAMA Pediatrics, 170,* 678–686. http://dx.doi.org/10.1001/jamapediatrics.2016.0289

*Saigal, S., Pinelli, J., Hoult, L., Kim, M. M., & Boyle, M. (2003). Psychopathology and social competencies of adolescents who were extremely low birth weight. *Pediatrics, 111*(5 Pt 1), 969–975. http://dx.doi.org/10.1542/peds.111.5.969

Saigal, S., Stoskopf, B., Streiner, D., Boyle, M., Pinelli, J., Paneth, N., & Goddeeris, J. (2006). Transition of extremely low-birth-weight infants from adolescence to young adulthood: Comparison with normal birth-weight controls. *Journal of the American Medical Association, 295,* 667–675. http://dx.doi.org/10.1001/jama.295.6.667

*Samara, M., Marlow, N., & Wolke, D., & the EPICure Study Group. (2008). Pervasive behavior problems at 6 years of age in a total-population sample of children born at </= 25 weeks of gestation. *Pediatrics, 122,* 562–573. http://dx.doi.org/10.1542/peds.2007-3231

Schlotz, W., & Phillips, D. I. (2009). Fetal origins of mental health: Evidence and mechanisms. *Brain, Behavior, and Immunity, 23,* 905–916. http://dx.doi.org/10.1016/j.bbi.2009.02.001

*Schmidt, L. A., Miskovic, V., Boyle, M. H., & Saigal, S. (2008). Shyness and timidity in young adults who were born at extremely low birth weight. *Pediatrics, 122*(1), e181–e187. http://dx.doi.org/10.1542/peds.2007-3747

*Schmidt, L. A., Miskovic, V., Boyle, M., & Saigal, S. (2010). Frontal electroencephalogram asymmetry, salivary cortisol, and internalizing behavior problems in young adults who were born at extremely low birth weight. *Child Development, 81,* 183–199. http://dx.doi.org/10.1111/j.1467-8624.2009.01388.x

Schwartz, R. M., Luby, A. M., Scanlon, J. W., & Kellogg, R. J. (1994). Effect of surfactant on morbidity, mortality, and resource use in newborn infants weighing 500 to 1500 g. *The New England Journal of Medicine, 330,* 1476–1480. http://dx.doi.org/10.1056/NEJM199405263302102

*Scott, M. N., Taylor, H. G., Fristad, M. A., Klein, N., Espy, K. A., Minich, N., & Hack, M. (2012). Behavior disorders in extremely preterm/extremely low birth weight children in kindergarten. *Journal of Developmental and Behavioral Pediatrics, 33,* 202–213. http://dx.doi.org/10.1097/DBP.0b013e3182475287

*Shum, D., Neulinger, K., O'Callaghan, M., & Mohay, H. (2008). Attentional problems in children born very preterm or with extremely low birth weight at 7–9 years. *Archives of Clinical Neuropsychology, 23,* 103–112. http://dx.doi.org/10.1016/j.acn.2007.08.006

Smith, R., Mesiano, S., Nicholson, R., Clifton, V., Zakar, T., Chan, E.-C., . . . Giles, W. (2005). The regulation of human parturition. In M. L. Power & J. Schulkin (Eds.), *Birth, distress, & disease: Placenta-brain interactions* (pp. 74–87). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511545658.004

*Stahlmann, N., Rapp, M., Herting, E., & Thyen, U. (2009). Outcome of extremely premature infants at early school age: Health-related quality of life and neurosensory, cognitive, and behavioral outcomes in a population-based sample in northern Germany. *Neuropediatrics, 40,* 112–119. http://dx.doi.org/10.1055/s-0029-1243166

Sterne, J. A. C., Egger, M., & Moher, D. (2011). Addressing reporting biases. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* [Version 5.1.0] (updated March 2011). The Cochrane Collaboration. Retrieved January 3, 2017, from http://handbook.cochrane.org/

Sterne, J. A. C., Gavaghan, D., & Egger, M. (2000). Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology, 53,* 1119–1129. http://dx.doi.org/10.1016/S0895-4356(00)00242-0

*Stjernqvist, K., & Svenningsen, N. W. (1999). Ten-year follow-up of children born before 29 gestational weeks: Health, cognitive development, behaviour and school achievement. *Acta Paediatrica, 88,* 557–562. http://dx.doi.org/10.1111/j.1651-2227.1999.tb00175.x

*Szatmari, P., Saigal, S., Rosenbaum, P., & Campbell, D. (1993). Psychopathology and adaptive functioning among extremely low birthweight children at eight years of age. *Development and Psychopathology, 5,* 345–357. http://dx.doi.org/10.1017/S0954579400004454

*Szatmari, P., Saigal, S., Rosenbaum, P., Campbell, D., & King, S. (1990). Psychiatric disorders at five years among children with birthweights less than 1,000 g. A regional perspective. *Developmental Medicine and Child Neurology, 32,* 954–962. http://dx.doi.org/10.1111/j.1469-8749.1990.tb08117.x

Taylor, G. H., Hack, M., & Klein, N. K. (1998). Attention deficits in children with < 750 gm birth weight. *Child Neuropsychology, 4,* 21–34. http://dx.doi.org/10.1076/chin.4.1.21.3188

*Taylor, H. G., Klein, N., Minich, N. M., & Hack, M. (2000). Middle-school-age outcomes in children with very low birthweight. *Child Development, 71,* 1495–1511. http://dx.doi.org/10.1111/1467-8624.00242

*Taylor, H. G., Margevicius, S., Schluchter, M., Andreias, L., & Hack, M. (2015). Persisting behavior problems in extremely low birth weight adolescents. *Journal of Developmental and Behavioral Pediatrics, 36,* 178–187. http://dx.doi.org/10.1097/DBP.0000000000000139

Van den Bergh, B. R., Mulder, E. J., Mennes, M., & Glover, V. (2005). Antenatal maternal anxiety and stress and the neurobehavioural development of the fetus and child: Links and possible mechanisms. A review. *Neuroscience and Biobehavioral Reviews, 29,* 237–258. http://dx.doi.org/10.1016/j.neubiorev.2004.10.007

*Van Lieshout, R. J., Boyle, M. H., Saigal, S., Morrison, K., & Schmidt, L. A. (2015a). Mental health of extremely low birth weight survivors in their 30s. *Pediatrics, 135,* 452–459. http://dx.doi.org/10.1542/peds.2014-3143

Van Lieshout, R. J., Boyle, M. H., Schmidt, L. A., Saigal, S., & Ferro, M. A. (2015b). Measurement invariance across parent and self-ratings of extremely low birth weight survivors and normal birth weight controls in childhood and adolescence on the Child Behavior Checklist and Youth Self-Report. *Journal of Pediatric Psychology, 40,* 825–833. http://dx.doi.org/10.1093/jpepsy/jsv043

Verder, H., Robertson, B., Greisen, G., Ebbesen, F., Albertsen, P., Lundstrøm, K., . . . the Danish-Swedish Multicenter Study Group. (1994). Surfactant therapy and nasal continuous positive airway pressure for newborns with respiratory distress syndrome. *The New England Journal of Medicine, 331,* 1051–1055. http://dx.doi.org/10.1056/NEJM199410203311603

Victorian Infant Collaborative Study Group. (1997). Improved outcome into the 1990s for infants weighing 500–999 g at birth. *Archives of Disease in Childhood, 77*(2), F91–F94. http://dx.doi.org/10.1136/fn.77.2.F91

Watterberg, K. L., & Scott, S. M. (1995). Evidence of early adrenal insufficiency in babies who develop bronchopulmonary dysplasia. *Pediatrics, 95,* 120–125.

Waxman, J., Van Lieshout, R. J., Saigal, S., Boyle, M. H., & Schmidt, L. A. (2013). Still cautious: Personality characteristics of extremely low birth weight adults in their early 30s. *Personality and Individual Differences, 55,* 967–971. http://dx.doi.org/10.1016/j.paid.2013.08.003

Wells, G. A., Shea, B., O'Connell, D., Peterson, J. E. A., Welch, V., Losos, M., & Tugwell, P. (2000). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses*. Montréal, Québec: McGill University. Retrieved January 2, 2017, from http://www.medicine.mcgill.ca/rtamblyn/Readings/The Newcastle-Scale for assessing the quality of nonrandomised studies in meta-analyses.pdf

*Whitfield, M. F., Grunau, R. V. E., & Holsti, L. (1997). Extremely premature (< or = 800 g) schoolchildren: Multiple areas of hidden disability. *Archives of Disease in Childhood Fetal and Neonatal Edition, 77*(2), F85–F90. http://dx.doi.org/10.1136/fn.77.2.F85

*Wilson-Ching, M., Molloy, C. S., Anderson, V. A., Burnett, A., Roberts, G., Cheong, J. L. Y., . . . Anderson, P. J. (2013). Attention difficulties in a contemporary geographic cohort of adolescents born extremely preterm/extremely low birth weight. *Journal of the International Neuropsychological Society, 19,* 1097–1108. http://dx.doi.org/10.1017/S1355617713001057

Wilson-Costello, D., Friedman, H., Minich, N., Fanaroff, A. A., & Hack, M. (2005). Improved survival rates with increased neurodevelopmental disability for extremely low birth weight infants in the 1990s. *Pediatrics, 115,* 997–1003. http://dx.doi.org/10.1542/peds.2004-0221

Wüst, S., Entringer, S., Federenko, I. S., Schlotz, W., & Hellhammer, D. H. (2005). Birth weight is associated with salivary cortisol responses to psychosocial stress in adult life. *Psychoneuroendocrinology, 30,* 591–598. http://dx.doi.org/10.1016/j.psyneuen.2005.01.008

Yau, G., Schluchter, M., Taylor, H. G., Margevicius, S., Forrest, C. B., Andreias, L., . . . Hack, M. (2013). Bullying of extremely low birth weight children: Associated risk factors during adolescence. *Early Human Development, 89,* 333–338. http://dx.doi.org/10.1016/j.earlhumdev.2012.11.004

Zeiner, P. (1997). Parent-reported symptoms of hyperactivity and attention deficits predict teacher-reported symptoms. *Acta Paediatrica, 86,* 178–182. http://dx.doi.org/10.1111/j.1651-2227.1997.tb08863.x

Received January 10, 2016
Revision received November 1, 2016
Accepted November 14, 2016 ■