Hovedoppgave

NTNU
Norges teknisk-naturvitenskapelige universitet
Fakultet for samfunns- og utdanningsvitenskap
Institutt for psykologi

Ingrid Gausemel Remøy

Cognitive outcomes in children born preterm or with low birth weight

A review and meta-analysis of IQ outcomes

Hovedoppgave i Profesjonsstudiet i Psykologi
Veileder: Lars Morten Rimol
Medveileder: Siri Weider
Desember 2023

NTNU
Kunnskap for en bedre verden

Ingrid Gausemel Remøy

Cognitive outcomes in children born preterm or with low birth weight

A review and meta-analysis of IQ outcomes

Hovedoppgave i Profesjonsstudiet i Psykologi
Veileder: Lars Morten Rimol
Medveileder: Siri Weider
Desember 2023

Norges teknisk-naturvitenskapelige universitet
Fakultet for samfunns- og utdanningsvitenskap
Institutt for psykologi

NTNU
Kunnskap for en bedre verden

Forord

Det er mange mennesker jeg er dypt takknemlig for, og som har spilt en viktig rolle i fullføringen av denne oppgaven. Først og fremst vil jeg takke mine veiledere, Lars Morten Rimol og Siri Weider, for invitasjonen til å skrive om dette temaet, og for å gi meg muligheten til å jobbe med dem og lære av dem. Uten deres bidrag i søke- og screeningfasene, og Lars Mortens bidrag i utførelsen av de statistiske analysene, ville ikke dybden i oppgaven vært mulig å oppnå. Deres støtte og konstruktive tilbakemeldinger i skriveprosessen har også vært avgjørende. Jeg vil også takke Magnus Rom Jensen og Lisbeth Jahren for deres bidrag i utviklingen av søkestreng, og for å dele sin entusiasme for slikt møysommelig arbeid. En stor klem sendes til min medstudent Kristin Berg Johansson for hennes bidrag i screeningprosessen, og for all støtte i samtaler om hva vi hadde begitt oss ut på. Jeg må også sende en helt spesiell takk til Monica Martinussen for hjelp med tolkning av resultater i meta-analysen, og for å øse ut av sin kunnskap om statistiske analyser.

Bortsett fra de nevnte enkeltpersonene vil jeg også uttrykke min takknemlighet til alle mine hjelpsomme og svært dyktige venner som har stilt seg til disposisjon, det være for idémyldring eller en motiverende arbeidsøkt. Jeg håper jeg en dag klarer å betale dere alle tilbake. Jeg vil også si tusen takk til mine utrolig tålmodige foreldre og søsken som nok aldri trodde jeg skulle slutte å være student, men som likevel alltid heier på meg, jeg savner dere hver dag. Til sist vil jeg si tusen takk til min kjære Eirik for å holde ut med meg de siste månedene, for å bruke timer på millimeter i tabellene mine, og for å brygge helt enestående kaffe!

Takk for alle bidrag, veiledning og støtte.

Abstract

Background: There is a growing concern about the increasing prevalence of premature birth (GA <37 / <2500 g) and the cognitive outcomes in this population. The aim of this review and meta-analyses was to summarize available and updated empirical evidence on prematurity as a risk factor for cognitive development.

Methods: The conceptualization and methodology of this review followed the recommendations of Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA). The search was carried out in Medline, PsycInfo and Web Of Sciene databases. The search was limited to articles written in English or a Scandinavian language, and published between January 2000 and May 2023. The selection of research articles were restricted to those who provided an FSIQ estimate, as measured with WISC, in children born preterm or with low birthweight. The review included studies with cross-sectional or longitudinal cohorts, where FSIQ scores were compared to a control group of children born at term or to standardized scales.

Results: For the review, 42 articles met the inclusion criteria, which collectively included 5314 children from nine different countries born preterm or with low birthweight. Age of the participants ranged from 6-15 years. From the identified publications, 11 articles were included for meta-analyses, based on having controlgroups and sharing comparable inclusion criteria with at least two other studies eligible for the analysis. Lower cognitive scores for FSIQ was identified in children born with birth weight <1000 g (Hedges' g = -0.90, 95%CI [-22.9, -7.6]), <1500 g (Hedges' g = -0.86, 95%CI [-16.5, -8.9]) and with gestational age <32 weeks (Hedges' g = -0.81, 95%CI [-13.4, -9.1]). These meta-analyses and the review of updated empirical literature highlight the risks that children born prematurely or with low birth weight face in their subsequent cognitive development.

Conclusion: Being born preterm or with low birthweight is associated with a poorer FSIQ score in childhood compared with term-born peers. Recent research affirms the risk of prematurity on cognitive development through childhood despite medical advancements in care, and highlights the influence of biological and environmental factors.

Sammendrag

Bakgrunn: Det er økende bekymring knyttet til den stigende forekomsten av for tidlig fødsel (GA <37 / <2500 g) og hvilke kognitive senvirkninger denne populasjonen opplever. Målet med denne gjennomgangen og meta-analysene var å oppsummere tilgjengelig og oppdatert empirisk kunnskap om for tidlig fødsel som risikofaktor for kognitiv utvikling.

Metode: Arbeidet med denne gjennomgangen fulgte de anbefalte retningslinjene fra Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA). Et litteratursøk ble gjennomført i databasene Medline, PsycInfo and Web of Sciene. Søket var begrenset til artkler publisert på engelsk eller et skandinavisk språk, og som var publiserte mellom januar 2000 og mai 2023. Artikler som ga FSIQ-score målt med en versjon av WISC på barn født premature eller med lav fødselsvekt ble inkludert i gjennomgangen. Det endelige utvalget av studier inneholdt tverrsnittstudier og longitudinelle studier hvor FSIQ-score ble sammenlignet med en kontrollgruppe eller med standardisert normgrunnlag.

Resultater: 42 artikler møtte inklusjonskriteriene og ble inkludert i gjennomgangen. Disse omfattet 5314 barn fra ni forskjellige land født premature eller med lav fødselsvekt. Deltakernes alder varierte fra 6–15 år. Av de identifiserte studiene ble elleve artikler med kontrollgruppe og hvor minst to andre studier delte sammenlignbare inklusjonskriterier inkludert i videre analyser. Meta-analysene fant lavere scorer for FSIQ hos barn født med fødselsvekt <1000 g, (Hedges' g = -0.90, 95%CI [-22.9, -7.6), <1500 g (Hedges' g = -0.86, 95%CI [-16.5, -8.9]) og med gestinasjonsalder <32 uker (Hedges' g = -0.81, 95%CI [-13.4, -9.1]). Disse meta-analysene og gjennomgangen av gjeldende empirisk litteratur fremhever risikoen barn som er født for tidlig eller med lav fødselsvekt møter i sin kognitive utvikling.

Konklusjon: For tidlig fødsel eller å bli født med lav fødselsvekt er assosiert med lavere FSIQ-score i barndom sammenlignet med jevnaldrende kontrollgrupper født til termin. Nyere forskning bekrefter den vedvarende risikoen ved prematuritet for kognitiv utvikling gjennom barndommen til tross for medisinske fremskritt, og fremholder innflytelsen av biologiske og miljømessige faktorer.

List of Abbreviations

Abbreviation	Definition
AGA	Appropriate for gestational age
BPD	Bronchopulmonary dysplasia
BW	Birth weight
CMA	Comprehensive Meta-Analysis Software
CI	Confidence interval
ELBW	Extremely low birth weight
EPT	Extremely preterm
FSIQ	Full Scale Intelligence Quotient
FT	Full-term
GA	Gestational age
IQ	Intelligence Quotient
LBW	Low birth weight
LGA	Large for gestational age
LPT	Late preterm
MeSH	Medical Subject Headings
MPT	Moderate Preterm
MRI	Magnetic resonance imaging
NBW	Normal birth weight
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PRI	Perceptual Reasoning Index
PSI	Processing Speed Index
SD	Standard deviation
SGA	Small for gestational age
VCI	Verbal Comprehension Index
VLBW	Very low birth weight
VPT	Very preterm
WISC	Wechsler Intelligence Scale for Children
WMI	Working Memory Index

Contents

List of Figures

List of Tables

1 Introduction

The past two decades have witnessed a significant rise in the number of premature births where babies are surviving at increasingly earlier stages of pregnancy (World Health Organization [WHO], 2023b). While the mortality in this population has decreased, the number of children surviving premature birth is on the rise. This increase has led to a growing public health concern as new questions have emerged about the long-term developmental consequences for this population (Bhutta et al., 2002; Twilhaar et al., 2018). Researchers have found higher rates of cognitive difficulties in these children. Many children who have been born preterm suffer from wide-ranging cognitive difficulties; in the most severe cases, multiple cognitive domains might be affected, leading to a high degree of impairment (Kaul et al., 2021). Also, in seemingly less severe cases, cognitive difficulties may be related to some degree of impairment even in seemingly healthy children with average performance (Johnson, 2007). It is essential to consider also these milder impairments, as low severity is not the same as minimal effects on child development and life outcomes (Kaul et al., 2021).

Many children born preterm present with difficulties at the primary school level (Hutchinson et al., 2013). Detecting and treating cognitive difficulties before school entry can significantly affect a child's academic performance and overall well-being. There are many standard tools for assessing neurodevelopment in infants. While many of these are highly specific, they often make poor predictions of school-age outcome and beyond (Potharst et al., 2012), and children with milder impairments who are more challenging to diagnose may not be detected (Wong et al., 2016). Cognitive abilities develop rapidly during this period and are believed to stabilise in early schoolyears (Schneider et al., 2014). Even if a child appears unaffected in their early years, high expectancies at school may reveal any hidden deficits (De Kleine et al., 2003). Cognitive impairment have been found to occur more frequenly in the preterm population, surpassing other known and common impairments such as motor, visual or hearing impairment (Allen, 2008). Over the last few decades, research on various cognitive functions in children born preterm has received increasing attention. Assessment of Intelligence Quotient (IQ) provides a numerical representation of intellectual functioning, a representation including skills like problem-solving, processing speed, working memory and verbal comprehension (Nisbett et al., 2012; Piaget, 2005). In addition to often being measurable by well standardised tools and consequently easily quantified and compared across populations, IQ has also been linked with a range of life outcomes from academic success (Hack et al., 2002) to mental health (Koenen et al., 2009), making it a highly relevant subject for study in the preterm population.

Insights into the cognitive outcomes of these children and the factors contributing to poorer cognitive outcomes are crucial, as this knowledge is fundamental to improving their overall outcomes. Clear prognostic information on long-term outcomes is essential for discussing clinical decisions with parents, coordinating services with healthcare providers and assisting education professionals when developing and providing necessary support. This information should be based on the best available data. Moreover, gaining an understanding of risk factors and other elements affecting cognitive development among this population, including perinatal and demographic factors or factors contributing to alterations in brain development, can provide valuable insights for informed clinical decision-making and parental guidance in the neonatal period. Additionally, it can lead to a more accurate identification of high-risk infants who require close monitoring and timely intervention.

As the cognitive development of children born preterm has gained increasing attention in recent years, there are concerns about study design and small sample sizes (Allen, 2008; Bhutta et al., 2002). To address these concerns, multiple meta-analyses have been conducted since 2000, aiming to provide a more comprehensive understanding of existing research. In this same period, medical advancements have progressed, making it increasingly important to continuously update knowledge on these advancements. Current research indicates that despite the advancements made in neonatal medical care, the prevalence of cognitive deficiency in preterm-born children remains relatively stable. As time goes on, evidence grows, and more children born after medical advancement reach childhood, adolescence and adulthood, it becomes crucial to update our knowledge and deepen our understanding to ensure the best possible outcome for these children and those born today. As a contribution to this the current thesis has aimed to estimate the effect of preterm birth on intelligence in studies published after year 2000, and to review available and updated empirical evidence on prematurity as a risk factor for cognitive development in children, specifically focusing on IQ scores.

2 Background

2.1 Preterm birth: definitions and risk factors

Preterm birth involves a child being born before it is fully matured and biologically prepared for life outside the uterus. While children born full-term usually have developed essential capacities such as reflexes, regulating body temperature, digestion, sustained breathing, and other complex physiological functions, children born prematurely might not have reached certain developmental milestones and might need additional support to survive and/or thrive. In 2020, the rates of premature birth around the world ranged between 5% and 18%, which sums up to an estimated 13.4 million babies worldwide, or more than 1 in 10 babies being born preterm (WHO, 2023b).

When talking about those born prematurely, it is important to be aware of the heterogeneity of this group and the complexity of the concept. The term preterm, or preterm birth, generally includes all children born at fewer than 37 weeks of gestational age (GA), in contrast to full-term birth, where the child is born at GA approximately 40 weeks. The result of this is that the term "preterm" covers a broad group with significant variations, ranging from the smallest infants with a gestational age ranging down to around 22 weeks to infants who have developed considerably further before being born. For this reason, it is common to describe the preterm group in different ways and also to divide the group into subcategories. The following section will briefly go through the most commonly used terminology.

Gestational age

Gestational age is a term used to describe how far along a pregnancy is, measured in weeks from the first day of the woman's last menstrual cycle to the current date. At 37 weeks of gestational age, a pregnancy is considered full-term (FT). If born before this, a baby is considered preterm (PT). WHO offers a widely used definition of preterm birth comprising three subcategories: extremely preterm (<28 weeks of GA, EPT), very preterm (28 to 32 weeks of GA, VPT) and moderate (32 to 34 weeks of GA, MPT) to late preterm (34 to 37 weeks of GA, LPT) (WHO, 2023b). As straightforward as this sounds, determining gestational age is not always easy. The primary methods used to estimate GA are dating based on the last menstrual period, neonatal estimates based on standardised scoring, and ultrasound-based dating (Lynch et al., 2007). The problem with using last menstrual period is that many women are not aware of when they ovulate or simply do not recall. Neonatal estimates are based on assessment of the newborn, usually following a standardised method assessing the physical and neuromuscular maturity of the infant. Still, these methods also come with a high risk of miscalculations (Alexander et al., 1992;

Lynch et al., 2007). In high-income countries, the most accurate method is considered to be ultrasound in early pregnancy, and this is commonly done as part of routine check-ups. Ultrasound-based gestational dating relies on norms established using data from fetuses where reliable data on last menstrual period was available, and critics point out that this reliance could imply a common bias (Lynch et al., 2007).

Birthweight

Historically, birth weight has been a more tangible and reliable measurement than estimates of gestational age. In scientific literature, the terms low birthweight and premature were used largely interchangeably up until the 1960s when accumulating epidemiological data made the distinction clear between being born before term and being born with lower birthweight (Wilcox, 2001). A commonly used sub-categorisation of birthweight is extremely low birth weight (<1000 g, ELBW); very low birth weight (1000–1499 g, VLBW); low birth weight (1500–2499 g, LBW) (Butler et al., 2007). However, when going through the literature, one can counter different variations of this categorisation, with some using other limits, e.g., birth weight less than 2000 g, 1750 g, 1250 g, 800 g or 500 g. Some researchers use a lower limit and define a segment, e.g., 1000 g-1500 g, others will have included all children less than 1500 g etc.

Measuring prematurity by birthweight also has some shortcomings. Even though the measurement itself can be made with precision and the population largely will include those born preterm, children born with a birth weight of 2500 g or 1500 g will still include infants at different stages in development or born at very different weeks of gestational age (Butler et al., 2007).

Small for gestational age

Typically, during pregnancy, the fetus follows a distinct growth trajectory, and most children will develop within a weight range expected for their gestational age, commonly referred to as appropriate for gestational age (AGA). If you are born with a birth weight of 1500 g at 28 weeks of gestational age, this would put you right on the norm, while if you were to be born at week 32 with the same weight, this would put you at the 5th percentile. Infants falling below the 10th percentile for gestational age are labelled small for gestational age (SGA), while on the other side, infants exhibiting higher birthweights exceeding the 10th percentile are classified as large for gestational age (LGA) (Damhuis et al., 2021). It has been demonstrated that these distinctions are relevant for mortality and morbidity among newborns where weight and gestational age do not align. Being born preterm and LGA might result in normal birth weight but is still associated with more complications (Cartwright et al., 2020). Being born SGA has been linked with a worse outcome compared to infants born AGA regardless of gestational age (Gidi et al., 2020).

Implications of this heterogeneity

A recent review and meta-analysis by Sentenac et al. (2022) found geographical and temporal patterns in inclusion criteria with BW-criterions more commonly used in North America than in Europe and GA more used in studies on cohorts born after 1990.Their review found similar results within preterm or low birthweight sub-groups, reflecting the degree of preterm birth, regardless of whether GA or BW criteria was used. They recommend that future reviews use broad inclusion criteria based on both GA and BW to increase the number and representativeness of primary studies included, i.e., data from settings where GA data are unavailable or of poor quality (Sentenac et al., 2022).

Risk factors for premature birth

Preterm birth can be broadly classified into two subtypes which will have different underlying causes and interacting factors: *provider-initiated* or *spontaneous* preterm birth. Provider-initiated preterm birth occurs through induction of labour or pre-labour cesarean delivery, usually due to a maternal and/or fetal indication. Spontaneous preterm birth includes preterm labour, preterm spontaneous rupture of membranes, premature rupture of membranes before the 37th week, and cervical insufficiency (Goldenberg et al., 2012). This distinction is especially relevant in countries with high pregnancy monitoring and caesarean birth rates (Blencowe, Cousens et al., 2013). Indications for provider-initiated preterm birth are usually maternal or fetal clinical conditions or complications such as preeclampsia, cholestasis, or severe fetal growth restriction, but some late preterm deliveries are carried out in the absence of a strong medical indication (Gyamfi-Bannerman et al., 2011; WHO, 2023b) or are unintended due to errors in gestational age assessment (Blencowe, Cousens et al., 2013). Pre-existing maternal conditions can also lead to complications, and some have pointed towards growing rates of obesity and diabetes as important contributors to global preterm birth (Blencowe, Cousens et al., 2013).

Spontaneous preterm births are a complex process influenced by various factors, and in many cases, it is not possible to conclude on a causal factor. Risk factors also vary based on gestational age, social status, and environmental factors and are thought to be the result of an interaction between genetic, epigenetic and environmental factors (Blencowe, Cousens et al., 2013). Still, strong links have been made to both young and advanced maternal age (Saccone et al., 2022), multiple pregnancy, previous preterm delivery, uterine and cervical anomalies and maternal medical complications or infections (WHO, 2023a). In addition, some lifestyle factors such as both active and passive smoking have been associated with increased risk of preterm birth (Cui et al., 2016; Shah et al., 2000), as well as maternal psychological health (WHO, 2023a), stress, or lack of social support (Kim, 2022).

2.2 Medical advancements in perinatal care

While prematurity is still the leading cause of death in children under the age of 5 years, there has been a substantial reduction in mortality during the last two decades (Norman et al., 2019; Perin et al., 2022). Technical and medical advances in prenatal, perinatal and postnatal care to the pregnant woman and her offspring, have primarily contributed to this decline (Stoll et al., 2015). These advancements help support the not yet fully developed capacities of the fetus or infant crucial for their survival. Examples of such advances are the introduction of exogenous surfactants in the early 1990s, improvements in the clinical use of mechanical ventilation, and increased use of cortisone to accelerate lung maturation before birth (Bancalari et al., 2015; Crowther et al., 2015; Halliday, 2008).

During the prenatal stage, interventions often take a proactive form with the goal of supporting the development of the unborn child (Klingenberg, Kaaresen et al., 2021; Medley et al., 2018). A crucial part of preparing for life outside the uterus is the development of the capacity to breathe, and one of the most common complications for preterm born infants is respiratory distress syndrome (RDS) caused by surfactant deficiency (Oza et al., 2014). Sometimes the underdeveloped lungs of the infants require the use of a ventilator or oxygen therapy for support, heightening the risk of Bronchopulmonary dysplasia (BPD), a form of chronic lung disease (Wendel et al., 2023). Treatment with oxygen have also been linked to Retinopathy of Prematurity, another common condition in those born premature (Hartnett et al., 2013). Administration of corticosteroids to expecting mothers to accelerate the development of the fetal lungs has significantly reduced the risk of RDS, BPD, ROP and other respiratory and neurological complications (Crowther et al., 2015; Roberts et al., 2017; Zeng et al., 2022). Steroids for lung maturation is given in two doses 24h apart with maximum impact on lowering risk of RDS reached at around 48 hours after the initial dose, but the effect on mortality and morbidity is significant within hours after administration of the first dose. For this reason, it is recommended to administer steroids even if delivery is expected before the second dose can be given (Haas et al., 2006). Steroid treatment is often used in combination with tocolytic medications meant to delay the delivery as long as this is safe for both the mother and child. This gives more time for fetal development and for the mother to be transported to a place where she and the infant can receive the appropriate level of treatment (Michelsen et al., 2020; Roberts et al., 2017). Norwegian official guidelines for initial treatment of premature emphasizes a close collaboration between obstetrician and neonatologist when dealing with the risk of premature birth to ensure the best possible outcome for mother and child. Steroid treatment for lung maturation with additional tocolytic treatment is generally recommended from <23 weeks GA, with individual assessment between <22+5 and <23+0 weeks of GA (Michelsen et al., 2020).

Upon birth, most infants born preterm will need specialized care and observation in a specialized unit where they can be monitored and provided care according to their needs. They may receive mechanical ventilation or non-invasive respiratory support such as continuous positive airway pressure (Markestad et al., 2007). In the early 1990s, we saw the introduction of exogenous surfactants, which has been a significant advancement in enhancing lung function and reducing the incidence of RDS (Sweet et al., 2023). In Norway, surfactants will often be used prophylactically on children born <26 weeks GA (Klingenberg, Guthe et al., 2021).

Not all advances are solely a result of new technology. One interesting example of this progress is the Kangaroo Mother Care method, defined as continuous skin-to-skin contact of the infant with the chest of the caregiver and feeding exclusively with breast milk (WHO, 2003). As recently as in 2022, WHO updated its guidelines to improve survival and health outcomes for babies born preterm (<37 weeks GA) or with low birth weight (<2500 g. The perhaps most significant change in the guidelines was going from recommending short, intermittent sessions of Kangaroo Mother Care after the infant had been stabilised to recommending skin-to-skin contact immediately after birth, without any initial period in an incubator (WHO, 2022). The new recommendations were a result of multiple studies showing clear benefits of Kangaroo Mother Care, such as lower mortality, fewer infections, and improved thermoregulation (Lode-Kolz et al., 2023; WHO, 2021). In WHOs study, the benefits of this method were so pronounced that the trial was stopped early on as it was considered unethical to continue separating mother and child at birth for an initial stabilisation of the child in an incubator. Originally, the Kangaroo Mother Care method was developed to deal with problems arising from a shortage of incubators. For a while, it was considered an option for societies or cases where modern equipment was not available. In the end, it outshined the capabilities of modernised neonatal care (Charpak et al., 2001; Ruiz-Peláez et al., 2004), and has since contributed to reducing morbidity rates in children born preterm (Sivanandan et al., 2023).

As a result of the remarkable progress in neonatal care, the number of infants surviving being born preterm has increased, and the chance of survival has moved to an increasingly earlier gestational age. From the 1960s to the beginning of the 2000s the gestational age at which at least half of the newborns survive has gone from 30–31 to 23–24 weeks. The decision whether to provide active care for the smallest and youngest preterm born infants in these weeks can be ethically challenging and demands careful and continuous considerations; these weeks are often referred to as "the grey zone of uncertainty". Insights from a systematic review looking at national guidelines for management of extremely premature deliveries in highly developed countries showed a general agreement for comfort care at 22 weeks GA and active care from 25 weeks GA. At 23–24 weeks of GA, practices differed substantially between centres, regions and countries (Guillén et al., 2015; Seri

et al., 2008).

2.3 Brain development and risk factors

Looking at what we know about development throughout gestation, specifically focusing on growth and the maturation of the brain, it becomes clear how disruptions during this critical period, such as premature birth, can lead to unfavourable outcomes. A fundamental process in the development of the brain and the central nervous system is neuronal proliferation and migration to the cerebral cortex, which is considered complete at around 20-24 weeks of GA. However, measured at 34 weeks, the brain still weighs only around 65% of what it will at term and has only about half of the cortical volume of the term brain (Volpe, 2009). Within this growth during the second half of gestation lies multiple complex and critical events in the brain development, such as white matter development and myelinisation, development of neurons and glia, axonal elongation, dendritic arborisation and synaptogenesis (Kinney, 2006). The brain development of infants born too early and during this vulnerable period has increased risk of disruption and subsequent neuropathology (Volpe, 2009).

Furthermore, infants born preterm face additional hurdles post-birth found to affect the brain development and outcomes of the newborn. Some of these are inherent to their underdevelopment such as bronchopulmonary dysplasia, intraventricular haemorrhage or periventricular leukomalacia (Kinney, 2006; Rees et al., 2023; Twilhaar et al., 2018). Others are a result of environmental factors including stress and strain resulting from standard neonatal care or as part of critical lifesaving procedures (Simons et al., 2003), such as postnatal steroids (Barrington, 2001; Smith et al., 2011), sedation (Ng et al., 2017) or other painful procedures (Brummelte et al., 2012)

In the last decades, modern research using MRI and subsequent analyses (e.g. Tract-Based Spatial Statistics) has been able to demonstrate disrupted brain development in children born preterm or with LBW, such as a decrease in brain volume and areas of thinner or thicker cortex and in white matter integrity (de Kieviet et al., 2012; Rimol et al., 2019; Sølsnes et al., 2015), and alterations in functional connectivity (Cho et al., 2022). Associations have also been made between maldevelopment in cortical areas related to cognitive functioning and negative effects on a range of neurocognitive outcomes, including IQ (de Kieviet et al., 2012; Sølsnes et al., 2015). Treatments meant to reduce postnatal mortality and morbidity such as postnatal steroids (Baud, 2004; Doyle et al., 2014), routine doses of hydrocortisone or dexamethasone (Tam et al., 2011) or painful procedures (Brummelte et al., 2012) have all been seen to affect brain development (Modi et al., 2001), possibly giving unfavourable outcomes long term.

Researchers in this field are continually working to determine the right balence and to give these children the best possible outcome (Daskalakis et al., 2023), and the need for active care should be under continuous individual consideration (Klingenberg, Kaaresen et al., 2021).

2.4 Long term consequences and intelligence outcomes

While medical advances in the field have affected the mortality rate and moved the threshold for likely survival, morbidity rates among those born preterm or with LBW are still high (Allen et al., 2011; Twilhaar et al., 2018). The shift caused by these advances has brought about new questions about what the outcomes of preterm births are today, as more and more of these children survive and reach adulthood (Saigal et al., 2008).

Being born premature has been, and still is, associated with a range of risks in biological, developmental and cognitive domains. Children who survive being born premature have an increased likelihood of experiencing challenges in motor outcomes, have poor linguistic development, behavioural problems, or other developmental issues (Allotey et al., 2018; Moreira et al., 2014; van Noort-van der Spek et al., 2012). These ris They are also at risk of developing problems with cognitive function, including intelligence, learning and memory, processing speed, and executive function (Aarnoudse-Moens et al., 2009; Arpi et al., 2019). Often, outcomes are adverse, such as cerebral palsy, blindness, deafness, severe psychiatric disorders as well as epilepsy (Marret et al., 2013; Sarda et al., 2021). However, an increasing body of research reveals the occurrence of more subtle deficits also in children without major impairment (Blencowe, Lawn et al., 2013; Bolk et al., 2018; Jansen et al., 2021), with cognitive impairments being one of the most common outcomes (Allen, 2008; Jansen et al., 2021; Johnson, 2007; Kerr-Wilson et al., 2012). Subtle deficits may only become evident when a child meets challenges that underscore their limitations (De Kleine et al., 2003), such as academic or behavioural expectations at school, and many studies on the prevalence of milder deficits have focused on outcomes in school age. Capabilities such as processing speed and working memory have been linked to academic achievement (Mulder et al., 2011), and compared to term-born, children who are born preterm and/or with low birth weight have been seen to perform significantly worse on achievement tests in school, such as math and measures of reading and spelling (Aarnoudse-Moens et al., 2009; Kovachy et al., 2015; McBryde et al., 2020; Samuelsson et al., 2006), and to have a higher need of educational support at school (Tommiska et al., 2020). The rate of EPT children experiencing cognitive, educational or behavioural impairment at early school age has been suggested to be as high as 70%, compared to around 40% in their NBW peers (Hutchinson et al., 2013).

2.4.1 Prematurity and intelligence

Cognitive impairment stands out as one of the most frequently occurring challenges among those born preterm (Allen, 2008). Intelligence quotient (IQ) is easy to quantify and compare across different populations, and is a widely used metric when evaluating developmental outcomes in this population. The impact of intelligence extends to several vital life outcomes in this group, such as mental health (Koenen et al., 2009), education, occupation and career success (Hack et al., 2002; Strenze, 2007). Multiple studies of children born PT or with LBW include IQ when assessing cognitive outcomes, either as their primary attention or as a predicting, mediating or cooccurring factor in studies assessing other factors such as learning difficulties (Carmo et al., 2022; Farooqi et al., 2016), motor impairment (Fjørtoft et al., 2013; Vermeulen et al., 2022), memory function (Aanes et al., 2019; Løhaugen et al., 2011), attention deficit (Campbell et al., 2015; de Kieviet et al., 2012; Johnson et al., 2015), executive function (Dai et al., 2020; Farooqi et al., 2016; Jin et al., 2020), or behavioural problems (Domellof et al., 2020; Jin et al., 2020).

Recent meta-analyses on cognitive outcomes in children have found to a correlation between being born preterm and deficits in intelligence (Allotey et al., 2018; Arpi et al., 2019; Bhutta et al., 2002; Brydges et al., 2018). Already at age three to five years, Arpi et al. (2019) found a three times higher risk of having low IQ in children born VPT compared with term-born controls. The PT born children scored 0.77 SD or 11.5 points lower than term-born controls. However, the authors found no association between GA and IQ. Brydges et al. (2018) performed a meta-analysis examining outcomes in VPT children with data from 44 studies. The findings revealed strong associations between being born VPT and intelligence outcomes, with children born VPT scoring 0.82 SDs (95%CI [-0.90, -0.74]) below their term-born peers, corresponding to 12.30 points (Brydges et al., 2018). The authors found a small but significant association between birth weight and IQ in those assessed at a younger age (4-10 years), and no association between IQ and GA. In another meta-analysis published by Allotey et al. (2018), children born VPT, MPT and LPT all showed significantly lower FSIQ scores than their term-born peers. Here, a significant correlation between gestational age at birth and test scores was observed, with children born VPT and MPT scoring lower than those born LPT. A fourth meta-analysis from the same period performed on LBW children <10 years in South Asia reported a negative association between birthweight and IQ score (Upadhyay et al., 2019). The study found a weighted mean difference of -4.56 (95%CI [-6.38, -2.74]) in favour of the NBW controls, and with a greater deficit in the part of the subgroup born with lowest birth weight.

Two decades ago, Bhutta et al. (2002) studied the negative effect of preterm birth on intelligence, analysing cohorts of children born before 1990. Here, they found a weighted mean difference of 10.9 (95%CI [9.2, 12.5]) in favour of the control and a significant

correlation between both birth weight and GA and the mean cognitive test scores. When comparing their results with the above-mentioned meta-analyses, all based on cohorts of children born in the era of antenatal corticosteroid and surfactant treatment, almost all find a similarly large effect, indicating that despite increased survival rates, in terms of long-term cognitive outcomes, there has not been a notable improvement since the 1990s (Twilhaar et al., 2018). Cheong et al. (2017) found similar results for a cohort born from 1991–2005, even reporting poorer academic performance at early school age in EPT children born in 2005 than in earlier eras.

When Brydges et al. (2018) split their study group into younger children (aged 4–10 years) and older children and adolescents (aged 11–17 years) the researchers found very similar effect sizes between the groups, implying cognitive impairment was not associated with age of assessment. Similarly, Allotey et al. (2018) observed that deficits persisted beyond primary school age. These findings support a hypothesis that children born preterm suffer from a lasting deficit in cognition, not a delay. In a life-long perspective, several studies of adolescents and adults with a history of prematurity claim to see a pattern of cognitive deficit or lower IQ among this group (Brydges et al., 2018; Linsell et al., 2018). A recent individual patient data meta-analysis concluded that adults born VPT or VLBW scored a mean of 12 points lower than their term-born peers (Eves et al., 2021), aligning with findings in childhood, adding to the hypothesis that this deficit may persist into adulthood.

Even mild intelligence deficit could have multiple consequences for the PT/LBW group. Low severity does not mean minimal impact (Arpi et al., 2019; De Kleine et al., 2003). In childhood it could affect school readiness and lead to educational delay (Pritchard et al., 2014; Roberts et al., 2011), underly behavioural problems (Fan et al., 2013), or affect self-esteem and quality of life (Gire et al., 2019; Tosello et al., 2021). In long-term, intelligence is a powerful predictor of educational, occupational and economic success (Hack et al., 2002; Strenze, 2007).

It is important to note that there are also studies adding nuances to this picture, concluding that preterm adolescents exhibit intellectual performance on levels with their term-born peers (Jensen et al., 2015) and that the majority of those born preterm reach adulthood without major comorbidities (Crump et al., 2019).

Intelligence is a complex trait, influenced by both genetic and environmental factors, e.g. parental factors (Eves et al., 2021). It is thought to be relatively stable from preschool age, although this stability increases with age (Doyle et al., 2012; Girault et al., 2018; Schneider et al., 2014). Some also point to the stabilisation of intelligence possibly being delayed by prematurity or associated risk factors (Nagy et al., 2022). In the assessment of preterm infants, it is common practice to correct for gestational age, and the difference to

chronological age are substantial in early childhood (Gould et al., 2021). At school age, the difference is less substantial, but still significant (Gould et al., 2021), In clinical contexts, correcting is recommended in infancy and up to two years of age (Helsedirektoratet, 2007). When assessing for research purposes, it is recommend using corrected age into adolescence (Doyle et al., 2016; Gould et al., 2021; van Veen et al., 2016; Wilson-Ching et al., 2014). Correcting for gestational age can give a more accurate comparison of development. However, as the child ages, it may introduce biases potentially leading to a higher classification of children with cognitive disabilities. Additionally, there is a risk of overshadowing individual strengths or the need for additional support (Gould et al., 2021; Roberts et al., 2013).

2.4.2 Wechsler Intelligence Scale for Children

When assessing cognitive function there is a need for objective measures, and in the field of intelligence assessment in children the Wechsler Intelligence Scale for Children (WISC) stands out as a widely employed, standardised tool. Since its first edition, WISC has been and continues to be one of the most significant and relevant cognitive tests in both intelligence research and clinical settings. As the primary emphasis of this thesis is on papers using measurements from WISC, it is appropriate to provide a brief introduction to the tool. For clarity moving forward, the term *FSIQ (Full-Scale Intelligence Quotient)* will in this thesis be used specifically for referencing results obtained from the WISC while *IQ* may also refer to scores or results from studies using other measurements of IQ.

The first edition of the WISC was published in 1949, developed by psychologist David Wechsler. It served as a downward extension of his Wechsler Bellevue Intelligence Scale, a battery primarily based on approches to mental testing and performance testing on e.g. military recruits. The goal of the WISC was to provide a comprehensive measure of a child's cognitive abilities and functioning across various domains. It was standardised over a five-year period on a sample of 1100 boys and 1100 girls aged from 5-15 years, 100 from each of the eleven age groups. Totalling at 2200 cases it set the bar for standardisation which all upcoming revisions followed (Seashore et al., 1950). Since its launch and to this date the WISC has had four revisions. The first (*WISC-R*) was done by Wechsler himself in 1974, with the aim of revising or removing subtests which came to be seen to be ambiguous, outdated or biased, as well as changing the range from 5 through 15 years to 6 through 16 years, and updating the previous normative sample from only white children to include what was then referred to as a more proportional nonwhite representation (Kaufman et al., 1986). The next revisions was done by the publisher in 1991 (*WISC-III*), 2003 (*WISC-IV*) and in 2014 (*WISC-V*) (Wechsler, 1991, 2003, 2014). In the WISC-III edition, the perhaps most significant update was on the normative sample where, in addition to the required

re-norming for the Flynn-effect, it opted for a more representative and nuanced sample by expanding the definition of ethnicity to include white, black, Hispanic and other instead of the WISC-R's white/nonwhite distinction. In addition, parents' level of education replaced parents' occupation as the socioeconomic variable, as the latter had shown to be difficult to categorise objectively (Kaufman, 1993). The content of the test was also revised, including changes like the removal or modification of items with more emotional or clinical context, putting a greater emphasis on speed of responding, as well as adding a new, four-factor system of interpretation as an alternative to the three-factor system used up to and including WISC-III. Despite mainly praising the revision, Kaufman in 1993 also addresses some weaknesses, primarily criticising what he saw as a too extensive emphasis on speed of responding, and the predominance of school and cultural-related subtests, especially in the Verbal scale, which he said penalises children with school problems and children from different cultural or linguistic backgrounds (Kaufman, 1993). This criticism is to a large extent met in the 2003 revision, WISC-IV. In this fourth edition, multiple subtests were eliminated or revised resulting in a reduced emphasis on response speed and school knowledge as a prerequisite for success. The new subtests largely emphasised fluid reasoning and/or working memory, which up until then had been underrepresented (Kaufman et al., 2006). It also provided scores on multiple indeces allowing for a more comprehensive assessment of a child's strenghts or weaknesses. The most recent version of WISC, WISC-V, was published in 2014. The goal for this edition was to take in the advances in structural models of intelligence, cognitive neuroscience, neurodevelopmental research and psychometrics. WISC-V presents a five-factor structure by adding the Fluid Intelligence Index, and in addition, it offers multiple ancillary and complimentary indices (Kaufman et al., 2015).

WISC structure and administration

WISC-III consisted of ten mandatory tests from which scores on Verbal IQ, Performance IQ and Full Scale IQ were derived. Verbal IQ aimed to measure the verbal and linguistic abilities of the child, while Performance IQ was designed to assess non-verbal cognitive abilities and problem-solving skills. With two supplemental tests, it offered scores on four factors: Verbal Comprehension, Perceptual Organization, Freedom from Distractibility, and Processing Speed. The first two factors consist of verbal subtests, and the latter consisted of performance subtests. The verbal mandatory subtests were information, similarities, arithmetic, vocabulary, and comprehension. The performance mandatory subtests were picture completion, picture arrangement, block design, object assembly, and coding (Wechsler, 1991).

WISC-IV consisted of 15 subtests where ten of these made up the core tests. In this revision, the distinction between Verbal and Performance IQ was entirely abandoned, and

the four-factor structure was made primary. The core subtests were Similarities, Vocabulary and Comprehension (Verbal Comprehension Index), Block Design, Picture Concepts and Matrix Reasoning (Perceptual Reasoning Index), Digit Span and Letter-Number Sequencing (Working Memory Index), and Coding and Symbol Search (Processing Speed Index). Verbal Comprehension Index (VCI) assessed childrens ability to understand and use verbal information while Perceptual Reasoning Index (PRI) assessed the childs ability to use visual perception and organisation, non-verbal reasoning abilities and visual-motor skills. Working Memory Index (WMI) evaluated capacity to hold and manipulate information in the mind, and Processing Speed Index (PSI) now contained the tasks where speed of processing visual information and completing simple motor tasks was measured (Wechsler, 2003).

WISC-V also consists of 15 subtests, and even though only seven of these are needed to calculate FSIQ, ten are considered the core tests and are required in order to derive scores on all five primary indices: Verbal Comprehension (comprised of the subtests Similarities and Vocabulary), Visual-Spatial (Block Design and Visual Puzzles), Fluid Reasoning (Matrix Reasoning and Figure Weights), Working Memory (Digit Span and Picture Span) and Processing Speed (Coding and Symbol Search). primary Processing Speed subtests. While VCI, WMI and PSI are concidered continuations from WISC-IV, PRI is now divided in Visual-Spatial Index, assessing ability to understand and analyze visual-spatial information, and Fluid Reasoning Index, evaluating non-verbal reasoning and problem-solving (Kaufman et al., 2015; Wechsler, 2014).

For all versions, the FSIQ is derived from a combination of the indices providing a global score intended to represent a childs overall intellectual abilities. Administration of all versions of WISC is individual and should be carried out by a trained professional such as a clinical psychologist, neuropsychologist or other professionals trained in its use. A full assessment takes approximately 60-85 minutes for most children. In clinical contexts, this amount of time is sometimes not available or appropriate because of considerations such as client fatigue, pressure on time or other restrictions in resources with the patient or the professional, and as a result, short versions of WISC are often constructed. Some of these short forms give only an estimated score on FSIQ (Sistiaga et al., 2021), others on all indices (Crawford et al., 2010). Sometimes as little as two subtests are used, but the accuracy of the short forms has been found to decrease when fewer subtests are included (Hrabok et al., 2014). Several short forms have shown high reliability and can be recommended for use in clinical settings where a full assessment is hard to achieve (Crawford et al., 2010; van Ool et al., 2018).

Considerations for research

In order to use WISC in linguistic and cultural contexts outside of the US, language adaptations and norming for other countries or cultures are needed. Translation and adaptation of test materials, instructions, and items while maintaining the intended measurement of cognitive abilities can be challenging and demanding, and the same can be said for the process of establishing representative scores for specific populations. Today, the WISC has been translated and adapted to over a dozen countries. Most of these adaptations have been thoroughly verified and show good reliability and validity. Still, challenges such as limited geographical and socioeconomical stratification in normative samples or a small normative sample size are common, sometimes due to limited funding.

As we have seen, between the WISC-III and WISC-IV, the FSIQ measurement went through substantial change. In the previous revisions, the ten subtests making out the FSIQ remained constant, but in the WISC-IV revision five of these subtests were replaced to make a measure of FSIQ, which represented the new indices and concepts in a better way (Kaufman et al., 2006). In WISC-V, FSIQ is calculated from seven subtests, six of which is continued from WISC-IV. This development is important to note as it makes generalization in IQ-based research across the different editions more challenging.

3 Method

The methodology of this meta-analysis is grounded in the standards of *Preferred Reporting Items for Systematic Reviews and Meta-analyses* (PRISMA). This tool is an evidence-based minimum set of guidelines meant to promote and improve transparency in and replicability of the work of preparing and reporting findings in systematic reviews and meta-analyses (Liberati et al., 2009). As will follow from this chapter, some limtations were applied in the search process to align the work with the scope of a thesis. For this reason the author has chosen not to refer to this review as a systematic review. Still, an extensive effort has been put into following the standards and methodology expected from more comprehensive systematic reviews and meta-analyses.

3.1 Search strategy and selection procedure

The first step in conducting a systematic review and meta-analysis is to search the literature for relevant studies. This was accomplished by initially reviewing a list of concepts of cognitive measures that could be pertinent to the search. The list of potential concepts included input from supervisors and from published meta-analyses and systematic reviews related to similar subjects. These concepts were then sorted and examined to identify which could be relevant search terms. Variations of terms, such as intelligence, IQ and general ability, were considered in this evaluation. The final concepts included very low birth weight, very preterm, cognitive development, intelligence, executive function and Wechsler Scales, and are documented in Appendix A.

Subsequently, a search strategy was developed from these concepts. The search strategy comprised free keywords combined with Medical Subject Headings (MeSH) terms combined using "AND" or "OR". The final search string was developed in cooperation with and translated by experienced librarians for each database. Complete search log and strings for different databases is documented in Appendix B.

Eligibility criteria

We aimed to include studies fulfilling the following criteria: (1) participants in the premature cohort should be born premature (GA <37) and/or with a birth weight <2500 g; (2) cohorts free of intervention that may affect cognitive outcome; (3) reporting data on cognitive outcomes measured with standardized/validated tools/tests; (4) children 5 years old or older at point of being tested; (5) written in English or a Scandinavian language

After the initial search and review process the following criteria were added: (6) the study must include a measurement done with a full version of WISC, any version; (7) FSIQ

measurement must be provided in the article. These criteria were established due to time restrictions and a need for limitation when the preliminary results from the original search proved too extensive for a thesis. The additional criteria were thoughtfully created not to intervene with the original criteria or to encompass articles outside of the original search. By specifying WISC as a criterion this also automatically limited the age range to 6-16 years as WISC is designed and normed for this age group specifically.

Studies concerning parents' or siblings' health were excluded. Studies focusing on a maternal illness or conditions during pregnancy were excluded. Studies of treatment effect or early intervention were excluded, except if the study included a baseline measurement and otherwise met all inclusion criteria. Cohorts chosen explicitly because of suspicion of cognitive difficulties, or on a cohort of premature children with specific syndromes (e.g. Down syndrome) were excluded, as was studies on children who became ill or injured before, during, or after birth. Study designs such as register studies, commentaries, book reviews, protocol articles, and reviews were excluded.

Information sources

The search string was adapted by experienced librarians, and the final search was conducted in Medline (OvidSP), PsycInfo (OvidSP) and Web of Sciene databases in June 2023. After limiting the search to articles published in or after the year 2000, the final result from the search was 19396 articles.

3.2 Selection process

All articles were transferred to EndNote, a tool for reference management, and duplicates were removed using a modified version of Bramers method (Bramer et al., 2016). After duplicate removal, the remaining 12912 articles were uploaded to Rayyan for screening. Rayyan is an online tool developed for organizing systematic screening (Ouzzani et al., 2016). With Rayyan, multiple people can review the same articles independently. Due to time limits and the limitations on the scope the screening was performed as a single screening, meaning each article initially was screened by one person. In this case six persons were involved in the initial screening. Two were librarians who only excluded articles with obvious reasons for exclusion, such as animal experiments. The remaining four were two students of psychology and two professors of psychology (my supervisors). Title and abstract were assessed on all articles, and articles were included or excluded based on the criteria already mentioned. Any article where one person was unsure about whether it should be excluded was labeled and reviewed again in consultation with one of the professors. If an article could not be excluded based on the title or abstract it was

included to the next step. 668 articles were found to meet the initial criteria. At this point, the additional criteria, (6) and (7), were applied, and a second round of screening was performed in Rayyan by one person (IGR). All eligible articles were screened in full text and included if use of the relevant measurement, WISC, was reported. In addition, three articles were suggested by the author's supervisor. Two of these were duplicates and already excluded. One had not been identified by the search and met all criteria for inclusion. 217 articles were included in the next part of the reviewing process. The included studies were transferred from Rayyan to EndNote for thorough full text review to see if all earlier mentioned criteria were met. If multiple studies included the same cohort or if cohorts were overlapping the one with the most recent measurement was chosen, and if this was not relevant the study with largest cohort was included. Some studies had excluded children with neurological and/or cognitive impairment, mainly with the argument that the subjects were too impaired to complete the assessment. Articles who had excluded children with FSIQ <70 or less were included, articles excluding children with FSIQ higher than this were excluded. Finally, 42 articles were included in the review, and 11 articles were included in a meta-analysis. A summary of the complete search process can be seen in the PRISMA flow diagram for systematic search presented in Figure 1.

3.3 Data collection process

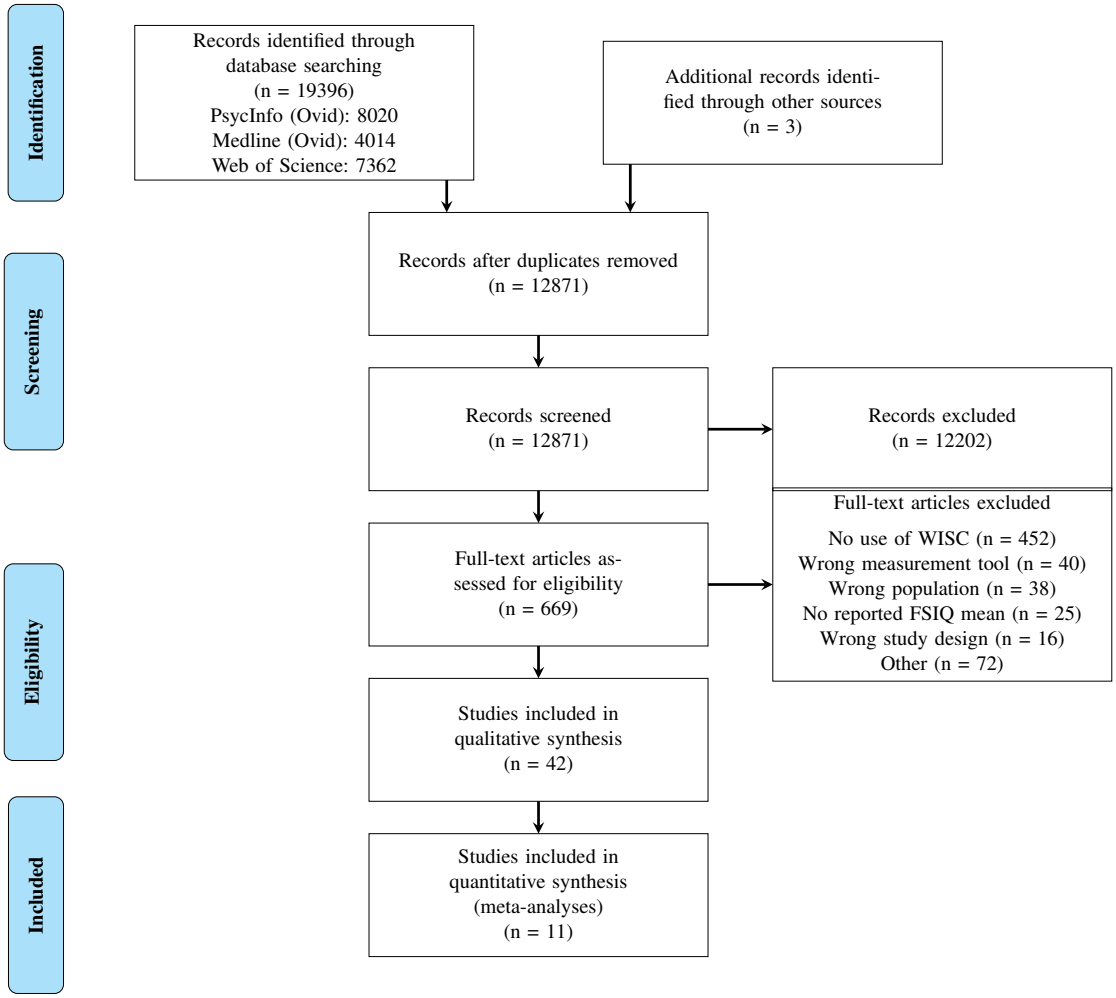
Data extraction and management were carried out using Microsoft Excel and carried out by one person (IGR). Data extracted included: (1) general information such as author, publication date, country; (2) data about study design, assessment tool and aim; (3) preterm cohort characteristics such as criterion used to define preterm or VLBW cohort, mean GA and/or BW and age of assessment and any exclusion criteria; (4) control group characteristics if applicable; (5) main and secondary results with respect to the aims of this thesis including mean FSIQ and SD.

3.4 Statistical analysis

Three seperate *random effect* meta-analyses were conducted in the *Comprehensive Meta-Analysis Software* (CMA) to combine the results of the included articles eligible for analysis. Effect sizes were calculated in three separate analyses. The random effects model was chosen to consider possible heterogeneity among the studies included. This model is considered a more conservative approach, acknowledging uncertainty associated with

Figure 1

Prisma Flow Diagram of the search process



between-study variability and is concidered more appropriate when dealing with a small number of studies, or when the sources of variation are not fully known. Heterogeneity was assessed and quantified calculating Q and I^2. The Q statistics is a measure of the variability in effect sizes beyond what would be expected by chance alone. The I^2 statistics represents the percentage of total variability in effect sizes that is due to heterogeneity rather than chance. Using both of these statistics allows for a more nuanced and comprehensive assessment of heterogeneity.

Publication bias was assessed using Eggers funnel plot (Sterne et al., 2001), a graphical tool for visualising the relationship between effect sizes and their precision. For this meta-analyses the funnel plot was generated by depicting Hedges'g on the x-axis and standard error on the y-axis.

Hedges' g was chosen as measure of effect size as it is preferred to alternatives such as

Cohen's d when analysing small sample sizes or when the sample sizes are significantly different. In addition, as the WISC is a standarised tool, the difference in means were of interest.

In one article, the mean FSIQ for the sample had been split into two groups. The overall sample mean was calculated by multiplying the mean and *n* for each group and dividing on total *n*. In another article with missing standard deviation (SD), a method from the *Cochrane Handbook for Systematic Reviews of Interventions* was employed to estimate the SD from a 95% confidence interval (CI) (Higgins et al., 2023). SD was calculated using the formula:

$$\text{SD} = \sqrt{n} \times \frac{\text{(Upper Limit} - \text{Lower Limit)}}{3.92}$$

where n represents the sample size in the group.

4 Results

The initial search identified 19399 publications, of which 6486 were excluded as they were duplicates. Three articles were suggested by the authors supervisor, of which two were excluded as duplicates. After screening title and abstract, another 12244 were excluded. An additional screening was done in full text of the remaining 669 articles, and another 452 were excluded. The remaining 217 articles were reviewed in full text, and finally, 42 articles met all criteria for the review.

From the included studies, three groups of articles were chosen for analysis after data extraction: (1) four articles on children born <1000 g; (2) four articles on children born <1500 g; (3) three articles on children born at <32 weeks of GA. These articles were chosen as they included comparable results from a control group that had been assessed with a full version of the WISC, and they were grouped according to the articles' inclusion criteria for PT/LBW. Some eligible articles had control groups assessed with a different cognitive assessment tool, or the control group consisted of children with proven ADHD, these were excluded from the analysis. Others had included children born at GA <33, GA <31 or <1750 g. These were excluded from the analysis due to insufficient articles using the same criteria to do a meaningful analysis. One article with inclusion criteria <1501 g was included in the group with articles <1500 g as it was considered within reasonable proximity for a meaningful comparison. Articles with inclusion criteria GA <33 or <31 weeks were not included in the GA <32 group as the one week timespan was considered too large for meaningful comparison.

4.1 Meta analysis results

Three meta-analyses were conducted for The results of all three meta-analyses are presented in Table 1. In all analyses a significantly mean effect size was detected in favour of the study groups compared to the control groups as indicated by the confidence intervals which did not include 0.

Birthweight <1000g

Based on the analysis of four studies (ELBW group, n = 370; control group, n = 302), children born ELBW performed significantly worse than their NBW peers. The effect size, measured as Hedges' g, was -0.90 (95%CI [-1.34, -0.46]). Effect sizes 0.80 or higher is considered a large effect (Cohen, 1988), indicating a substantial difference between these two groups (p <.001). As the WISC is a standardised tool, the difference in means is also an interesting measure, representing the gap in IQ points between the two samples.

Table 1

Results of meta-analyses

	Study	Diff. in means	CI (95%)	Cases (*n*)	Controls (*n*)	Hedges' *g*	*Q*	I^2
<1000g	Løhaugen (2011)	-22.0	[-30.3, -13.7]	16	19	-1.725		
	Grunewaldt (2013)	-7.0	[-17.7, 3.7]	23	33	-0.343		
	Doyle (2015)	-10.0	[-12.8, -7.2]	209	220	-0.668		
	Tommiska (2020)	-22.0	[-29.6, -14.4]	122	30	-1.153		
	Total	-15.2	[-22.9, -7.6]	370	302	-0.902	12.76 *	76.49
<1500g	Mu (2008)	-17.9	[-22.8, -13.0]	130	59	-1.12		
	Samuelson (2006)	-12.2	[-17.9, -6.6]	61	56	-0.78		
	McNicholas (2014)	-11.6	[-16.3, -6.8]	52	48	-0.95		
	Rickards (2001)	-8.8	[-14.1, -3.5]	17	18	-0.59		
	Total	-12.7	[-16.5, -8.9]	260	181	-0.86	5.11	41.25
GA <32	Cook (2003)	-11.1	[-13.7, -8.5]	268	198	-0.79		
	Cho (2022)	-15.9	[-23.0, -8.8]	36	26	-1.12		
	Linden (2011)	-9.9	[-14.4, -5.4]	100	50	-0.75		
	Total	-11.3	[-13.4, -9.1]	404	274	-0.81	1.44	0.00

Note. * p <.05

The difference in FSIQ between groups was highest in this sample, with ELBW infants scoring 15.2 points lower than their peers. However, it is important to note that the confidence interval for this difference is quite wide (95%CI [-22.87, -7.59]), reflecting some uncertainty around the exact magnitude of the effect.

Birthweight <1500g

Based on four studies included in the analysis (VLBW group, n = 363; control group, n = 204), children born VLBW performed significantly worse in the WISC assessment than their NBW peers. The effect size was found to be -0.86 (95%CI [-1.10, -0.62]), indicating a large effect also in this population. Difference in means in the <1500 g group compared to the controls was -12.70 (95%CI [-16.54, -8.86]).

GA <32

A large effect was also identified in the VPT group. Based on three studies included in the analysis (VPT group, n = 404; control group, n = 274) the group scored significantly worse than their control group (Hedges g = -0.81, 95%CI [-0.97, -0.65]). Children born VPT scored, on average, -11.27 points lower (95%CI [-13.4, -9.1]).

Heterogenity and publication bias

All three analyses were tested for heterogeneity in order to assess the presence of variability

among the included studies. For children born with a birth weight <1000 g, the I^2 statistic was 76.5, indicating substantial heterogeneity. The Q statistic was 12.76, and it was statistically significant ($p < .05$). Notably, this analysis was based on only three articles. While the observed heterogeneity is substantial, the interpretation should be cautious given the small number of studies included.

For the analysis of children born with a birth weight less than 1500 g, the I^2 statistic was 41.3, suggesting moderate heterogeneity. However, the Q statistic was not statistically significant ($Q(3) = 5.106, p = .164$). This analysis included four articles, providing a relatively small but slightly larger sample size compared to the <1000 g subgroup. The lack of significance in the Q statistic suggests that the observed variability may be within the range of chance variation, but the limited number of studies warrants cautious interpretation.

For children born with a gestational age less than 32 weeks, the Q statistics was 1.44 with 2 degrees of freedom. Since the Q-value is less than the degrees of freedom, the amount of between-study variance in the observed effects is less than would be expected based on sampling error alone. As a consequence, I^2 is automatically set to zero. The combination of a non-significant Q statistic and I^2 set to zero indicates relative homogeneity in effect sizes among the included studies, but the small number of studies poses a challenge to the robustness and generalizability of these findings.

Funnel plots were constructed for all three analyses to visually assess the potential presence of publication bias. The plots depicts the relationship between effect sizes (Hedges' g) and the corresponding standard error. Each point in the plot represents an individual study. The plots can be found in Appendix C. All three funnel plots showed symmetrical patterns. No clear evidence of publication bias or outliers was observed in the funnel plots. These findings suggest a relatively homogeneous distribution of effect sizes across the included studies. Given the limited number of studies, the assessment of symmetry and funnel shape should be interpreted with caution, as visual inspection may be challenging with a small number of studies.

4.2 Results of the review

Basic characteristics of all identified publications included in the review, such as inclusion criteria and study design, can be found in Table 2 on page 24. The 42 articles included in the review came from nine different countries. 55% of the studies were from countries in Europe, with Sweden being the most represented (five studies), followed by the UK (four studies) and the Netherlands (four studies). The rest of the articles were from Asia

Table 2

Basic study characteristics

Reference	Country	Birth year	Inclusion criteria	Study design, recruitment	Control group recruitment
Breslau (2001)	USA	1983–1985	<2500g	Case-control, Multi-center	Recruited from the same hospital
Carmo (2022)	Brazil	2003–2012	<36	Cross-sectional Single-center	n/a
Chaudhari (2004)	India	1987–1989	<2000g	Cohort study Single-center	Enrolled at same hospital
Cho (2022)	South-Korea	2010–2013	<32	Cross-sectional Single-center	Recruited from the local community
Cooke (2003)	UK	1991–1992	<32	Case-control Multi-center	Recruited from same class at school, class teacher choose the child of same sex and first language in the class whose birthday was closest to that of the index child
Dai (2020)	New Zealand	2005–2008	<1500g <30	Cohort study Single-center	n/a
Domellöf (2020)	Sweden	2000–2005	<35	Cohort study Single-center	Recruited from the same hospital, healthy at birth children born term, matched by sex and birth week
Doyle (2015)	Australia	1991–1992	<1000g <28	Cohort study Multi-center	Randomly selected from eligible births on the day an EP/ELBW survivor was due to be born, matched for gender, maternal country of origin
Emond (2006)	Brazil	1993–1994	1500–2499g $\leq$37	Cohort study Multi-center	Recruted from same hospitals/centers
Fan (2013)	Brazil	1999–2000	<2500g <37	Cross-sectional Single-center	n/a
Farooqi (2016)	Sweden	1992–1998	23-25	Case-control Multi-center	Recruited from national birth register, matched hospital and time to EPT children
Gire (2019)	France	2004–2007	<28	Cross-sectional Multi-center	n/a
Grunewaldt (2014)	Norway	1999–2001	<1000g	Cohort study Single-center	Healthy children recruited from four different local schools, age-matched

Continued on next page

Table 2 – Continued from previous page

Reference	Country	Birth year	Inclusion criteria	Study design, recruitment	Control group recruitment
Herber-Jonat (2014)	Germany	1999–2003	<25	Cohort study Single-center	n/a
Hutchinson (2013)	Australia	1997	<1000g <28	Cohort study Multi-center	Randomly selected from births on the expected date of birth for each EP/ELBW child and matched for gender and mothers country of birth
Iai (2022)	Japan	2008–2013	<1500g <32	Case-control Multi-center	n/a
Ionio (2022)	Italy	2012–2015	<37	Case-control Single-center	n/a
Jansen (2021)	Netherlands	2006–n.r.	<32	Cohort study Single-center	n/a
Jin (2020)	South-Korea	2006–2011	32-36	Cross-sectional Single-center	n/a
Kaul (2021)	Sweden	2004–2007	<27	Cohort study Multi-center	Matched control group born full term recruited from national birth register
Kim (2021)	South Korea	2008–2009	<1000g <30	Cohort study Single-center	Recruited via an in-hospital announcement
Kochukhova (2022)	Sweden	2004–2007	<32	Cohort study Multi-center (regional)	Recruited from different parallell study, age-matched and from same region
Kroll (2019)	UK	1979–1984	<33	Cohort study Single-center	n/a
Lahat (2014)	Canada	1977–1982	<1000g	Cohort study Multi-center (regional)	Selected from class lists provided by local school boards, group-matched with the ELBW cohort on age, sex, and socioeconomic status
Linden (2015)	Canada	2001–2004	$\leq$32	Cohort study Single-center	Recruited from same region, born in same period
Løhaugen (2011)	Norway	1992–1993	<1000g	Case-control, experimental Single-center	Recruited from local schools, and same hospital, matched by age (<80)
McNicholas (2014)	Ireland	1995–1997	<1500g	Cohort study Single-center	The next born NBW infant after each of the low birth weight participants and at same center, matched for gender

Continued on next page

Table 2 – Continued from previous page

Reference	Country	Birth year	Inclusion criteria	Study design, recruitment	Control group recruitment
Mu (2008)	Taiwan	1995–1997	<1500g	Cross-sectional Multi-center	Enrolled by random number sampling from same hospitals, born in same period
Mulder (2011)	UK	1997–1999	<31	Cohort study Multi-center	VPT children in the study were asked to invite a classmate to take part in the study as a control, matched on age and gender
Nagy (2022)	Hungary	NR	<1500g	Case-control Single-center and local recruitment	Recruited from mainstream schools
Pinto-Martin (2004)	USA	1984–1987	500–2000g	Cohort Multi-center	n/a
Rickards (2001)	Australia	1980–1982	<1501g	Cohort study Single-center	Randomly selected from children born in the same hospital at the same time
Rose (2011)	USA	1995–1997	<1750g <37	Cohort study Multi-center	Recruited from consecutive births from the same hospitals
Samuelsson (2006)	Sweden	1987–1988	<1500g	Cohort study Multi-center	Recruited from same hospital, born at the same time and matched by sex and parity
Squarza (2017)	Italy		<1000g <32	Cohort study Single-center	n/a
Takayanagi (2013)	Japan	1999–2006	1000-1499g <1000g	Case-control Single-center	n/a
Tinelli (2015)	n.r.	2001–2003	$\leq$1500g $\leq$32	Cross-sectional n.r.	Recruited from the local schools and matched to the preterm group in both gender and age
Tommiska (2020)	Finland	1996–1997	<1000g >22	Cohort study Multi-center (national)	Randomly selected from cohort participating in standardization of the NEPSY II, recruited from a local school and hospital personell with children, matched by age
van Veen (2020)	Netherlands	2008–2010	<1000g <30	Cohort study Single-center	n/a
van't Westende (2020)	Netherlands	2006–2007	<32	Cross-sectional Single-center	n/a
Vermeulen (2022)	Netherlands	2007–2011	<30	Cohort study Single-center	n/a

Continued on next page

Table 2 – Continued from previous page

Reference	Country	Birth year	Inclusion criteria	Study design, recruitment	Control group recruitment
Aanes (2019)	Norway	2003-2007	<1500g	Cohort study Single-center	Recruited from different study, matched for age and geographical (WASI)

Note. n/a = not applicable; n.r. = not reported

(17%), North America (12%), Australia (10%) and South America (7%), Africa was not represented. Combined, the samples include 5198 PT/LBW children assessed with a version of WISC, the largest sample being from Pinto-Martin et al. (2004) with 645 LBW participants. Despite only being represented in four studies (12% of the total), cohorts from North America make up 24% of the total children. The study with the smallest sample size had 23 participants (Grunewaldt et al., 2014).

All except one study were observational and had either a Cohort, Cross-sectional- or Case Control design. The only study not having an observational design was Løhaugen et al. (2011), where a baseline measurement from a trial for computerised cognitive training program was used. Twenty five of the studies used a BW inclusion criterion, 12 of these used an additional GA criterion, while 16 used only a GA inclusion criterion. None of the studies from North America or Australia used only a GA-criteria compared to over half of the European studies, aligning with the findings of Sentenac et al. (2022).

If we apply the categorisation recommended by WHO (2023b), five articles cover a cohort which would fall into the EPT subcategory, 14 cover VPT and eight cover M-LPT. Despite being represented with the fewest articles, the total EPT sample size is larger than the VPT or M-LPT, and some of the articles covering VPT or M-LPT may also have included children born EPT in their sample if the study did not set a lower-end cut-off for gestational age. Nine of the publications cover cohorts corresponding to the ELBW category, 11 cover VLBW and seven cover LBW children. Here, the LBW subgroups holds the largest total sample size compared to VLBW and ELBW groups, but some of these samples will have included children born VLBW or ELBW.

A total of 24 articles have included a control group, the majority defining full-term as >36 weeks of GA and/or with birth weight >2499 g. Two of these control cohorts were assessed with tools other than WISC, and one was a control group with confirmed ADHD. These control groups have not been included for comparison in this review. Eighteen articles report on how the control group was matched with the study cohort, with 17 matching for age, nine for gender, six reporting some level of geographical matching and only two reporting matching for socioeconomic factors. The total comparable full-term sample size assessed with a version of WISC was 2309. All articles report assessing control group

subjects at the same age or in the same age period as the study subjects.

The studies identified in the review all provided FSIQ measurements from a WISC assessment of children born either PT or with LBW, but their primary research aims were diverse. Some had goals similar to that of this thesis and explored the effect of LBW or premature birth on cognitive outcomes. Others investigated different factors such as parental stress, or computerised working memory training. Table 3 on page 29 provides a short summary of the specific aims for each study. Partly due to their diverse primary objectives, the studies also applied very different criteria for exclusion. Some explored the prevalence or effect of other biological, social or environmental variables, others found such variables could confound their results and opted for exclusion. Two of the studies reported using no exclusion criteria, others did not specify any specific criteria or report only excluding those too impaired to complete the assessment. Of the 27 studies specifying their exclusion criteria in any degree, 13 list degrees of congenital anomaly, six have excluded children with cerebral palsy specifically, and seven excluded one or more sensory impairments such as blindness or deafness. One article excluded children with FSIQ <65 (Gire et al., 2019), and another <70 (Linden et al., 2015). All reported exclusion criteria are listed in Table 3 .

An overview of sample characteristics from all studies can be found in Table 4 on page 34. In cases where the sample size (n) for the underlying information such as BW or GA differs from the number of participants who completed the WISC, this is specified. Reported GA for the preterm samples range from 23 to 36 weeks of GA, while total mean BW where this was reported range from 718 g to 2346 g. Among those reporting BW range, the lowest reported BW was 396 g. Jin et al. (2020) reports the highest measured individual BW at 3110 g and serves as an example of how being born preterm and being born with low birth weight does not always go hand in hand.

The lowest reported mean age at assessment was 6.3 years (Takayanagi et al., 2013) and the highest was 15.39 (Kroll et al., 2019), but many studies did not report a mean or did not specify beyond stating in general in what year or years of age the assessment was done. Carmo et al. (2022) reports the most extensive range, including participants from 6–14 years. Looking at who the authors have focused on, 26 publications have assessed the children somewhere in the range 6–10 years old, twice as many as have focused on the range 11–16 years old. Observing the means in each of these groups did not reveal any obvious trend. In articles reporting gender distribution, male participants represented between 61% (Chaudhari et al., 2004) and 35% (Kroll et al., 2019) of the population.

Individual studies reported using corrected age when assessing children at seven (Dai et al., 2020), eight (van Veen et al., 2020) and even up to 18 years of age (Doyle et al.,

Table 3

Comprehensive study characteristics

Reference	WISC-version	Exclusion criteria	Aims
Breslau (2001)	WISC-R	Severe neurologic impairment	To examine the contributions of familial factors and disadvantaged community to IQ change from the beginning of schooling to 5 years later, and to estimate and control for the potential effects of low birth weight on IQ change in the general population
Carmo (2022)	WISC-IV	Patients who presented with clear ID, autism or another neurodevelopmental disorder, patients enrolled in special schools	To evaluate the cognitive and academic profile of PT children at school age, and to identify sociodemographic and premature factors that influence these outcomes
Chaudhari (2004)	WISC-R (Indian)	Cerebral palsy or mental retardation too severe to complete assessment	To assess the intelligence, visuo-motor perception, motor competence and school performance of children with BW <2000g, at the age of 12 years
Cho (2022)	WISC-IV	Patients with major brain injuries (other than isolated grade I intraventricular hemorrhage on cranial ultrasound) or major disabilities, such as cerebral palsy, mental retardation, deafness, blindness, or congenital abnormalities	To identify cognitive function differences in VPT versus term-born children, and investigate alterations in white matter microstructure and functional connectivity (using TBSS and resting-state functional MRI)
Cooke (2003)	WISC-III	Mothers were not resident within a Liverpool postal district at the time of birth	To assess a cohort of preterm born infants at the age of 7 years for growth, motor, and cognitive measures, and investigate the effects of growth impairment on school performance
Dai (2020)	WISC-IV	Significant congenital anomaly, change in nutritional protocol during first 7 days, lack of relevant blood glucose concentration measurements	To examine the associations between intelligence, executive function and academic achievement in children born VPT
Domellöf (2020)	WISC-IV (Swedish)	n.r.	To investigate cognitive and behavioral outcomes in relation to GA in school-aged children born PT
Doyle (2015)	WISC-III	n.r.	To determine the relative contributions of biological and social exposures to outcomes into adolescence in extremely preterm survivors

Continued on next page

Table 3 – Continued from previous page

Reference	WISC-version	Exclusion criteria	Aims
Emond (2006)	WISC-III	Multiple births, infants with congenital anomalies, signs of neurological abnormalities in the first 24h of life	To investigate the development and behaviour of LBW term infants compared with matched term infants of appropriate birthweight
Fan (2013)	WISC-III (Brazilian)	Previous history of neonatal disorders such as the use of mechanical ventilation for more than two weeks, an intracranial hemorrhage grade III or IV, a congenital defect, neonatal seizures, bacterial meningitis, clinical evidence of perinatal asphyxia or the presence of periventricular leukomalacia	To assess the cognitive and behavioral development of PT and LBW newborns from disadvantaged social and economic environments
Farooqi (2016)	WISC-III (Swedish)	n.r.	To assess the cognitive and behavioral aspects of executive functioning (EF) and learning skills at 10 to 15 years in EPT children compared with term control children
Gire (2019)	WISC-IV	Blindness or amblyopia, deafness, severe cerebral palsy, FSIQ <65	To determine the quality of life (QoL) of school-aged children who were born <28+0 weeks of GA and who have no resultant major disabilities
Grunewaldt (2014)	WISC-III	Diagnosed congenital syndromes	To investigate functional outcome and cerebral MRI morphometry at 10 years in ELBW children without CP compared to healthy controls, and to examine any relationship with the quality of infant-motor-repertoire included in the GMA
Herber-Jonat (2014)	WISC-IV	n.r.	To determine the long-term neurodevelopmental outcome in EPT infants born 22-23 weeks of GA as compared to infants of 24 weeks GA with immediate postnatal life support
Hutchinson (2013)	WISC-IV	n.r.	To investigate cognitive, academic, and behavioral outcomes at age 8 years for a cohort of children born EP or ELBW
Iai (2022)	WISC-IV (Japanese)	n.r.	To compare WISC-IV profiles in schoolchildren born VPT/VLBW with peers having ADHD, and to identify specific neurocognitive traits in VLBW/VPT children
Ionio (2022)	WISC-IV (Italian)	Presence of congenital anomalies, severe sensory impairments, severe brain injuries and other neurological complications, parents lack of Italian language skills	To examine the effects of preterm birth on childrens cognitive, behavioral and socioemotional development

Continued on next page

Table 3 – Continued from previous page

Reference	WISC-version	Exclusion criteria	Aims
Jansen (2021)	WISC-III	Congenital anomalies of the central nervous system, severe other congenital anomalies, chromosomal disorders, metabolic disorders, neonatal meningitis	To investigate the rate and stability of impairments in children born preterm by assessing early and school-age outcome and individual changes in outcome
Jin (2020)	WISC-IV (Korean)	Chromosomal abnormalities, congenital anomalies, and severe neurodevelopmental problems that would make it difficult to perform tests	To explore the cognitive function, cognitive visual function, executive function, and behavioral problems at school age in moderate to late preterm infants
Kaul (2021)	WISC-IV (Swedish)	n.r.	To study possible cognitive profiles in EPT children by exploring strengths/weaknesses beyond Full-Scale IQ, identifying overlaps in deficits, and determining the proportion of EPT children with multiple impairments
Kim (2021)	WISC-IV (Korean)	No exclusions	To evaluate the cognitive and behavioral outcomes of children born EPT and to analyze biological or socioeconomic risk factors for poor cognitive outcomes
Kochukhova (2022)	WISC-V (Swedish)	No exclusions	To examine neurodevelopmental outcomes in VPT children considering perinatal, neonatal, and socioeconomic factors and to assess the persistence of antenatal steroid effects on cognition at 12 years
Kroll (2019)	WISC-R	n.r.	To construct a comprehensive model of predictors of IQ and its developmental trajectories in survivors of very preterm birth from childhood to adult life
Lahat (2014)	WISC-R	Neurosensory impairments (cerebral palsy, blindness, deafness, mental retardation, and microcephaly), child could not complete the assessments	To examine whether fluid intelligence moderates the link between birth weight and later ADHD symptoms
Linden (2011)	WISC-IV	No major congenital anomalies, autism, major sensory, motor or cognitive (<70) impairment	To examine factors which predict parenting stress in a longitudinal cohort of children born very preterm
Løhaugen (2011)	WISC-III (Scandinavian)	Not completing required training sessions	To evaluate effect of a computerized working memory program on different aspects of memory functions in children born ELBW
McNicholas (2014)	WISC-IV	Severe chromosomal anomalies, children whose twin died at birth	To examine the medical, cognitive and academic outcomes of VLBW children

Continued on next page

Table 3 – Continued from previous page

Reference	WISC-version	Exclusion criteria	Aims
Mu (2008)	WISC-III (Taiwan)	Intrauterine growth restriction	To identify delays across major areas of development
Mulder (2011)	WISC-IV	Not attending mainstream school, congenital abnormalities, severe disabilities causing and unable to perform the behavioral tests	To investigate the development of executive function and attention in VPT children compared to term controls in middle childhood, and to study whether processing speed mediated the effect of VPT birth on executive function and attention test performance
Nagy (2022)	WISC-IV (Hungary)	Perinatal complications, lack of typical developmental course	To assess the school-age out- comes of Hungarian VLBW/ELBW preterm children in intelligence and executive function as compared to typically developing, full-term children
Pinto-Martin (2004)	WISC-III	n.r.	To examine the prevalence of special educational placement and its relationship to grade retention, verbal and performance scores on tests of general intelligence, reading and maths achievement scores and classroom hyperactivity among LBW children
Rickards (2001)	WISC-III	Cerebral palsy	To compare cognition, academic progress, behavior, and self-concept children of VLBW with randomly selected NBW controls
Rose (2011)	WISC-III	Obvious congenital, physical, or neurological abnormalities	To ascertain whether the deficits found in attention, processing speed, memory, and representational competence found in toddler years persist into adolescence, to determine the role of these abilities in mediating PT/FT differences in IQ, and to determine whether the cascade model that fit in infancy would apply at 11 years
Samuelsson (2006)	WISC-III	n.r.	To examine the development of reading skills among VLBW children and to what extent reading difficulties at 9 years of age persist or have changed at 15 years of age
Squarza (2017)	WISC-III (Italian)	Multiple birth, presence of genetic abnormalities, severe neurofunctional impairment, neurosensory disabilities (blindness, deafness)	To investigate the association between early neurodevelopmental assessment and the risk of adverse cognitive outcome in extremely low birth weight children
Takayanagi (2013)	WISC-III	n.r.	To determine the characteristics of the cognitive function in VLBW at 6 years of age and investigate significant factors during neonatal intensive care unit admission that affect cognitive outcomes

Continued on next page

Table 3 – Continued from previous page

Reference	WISC-version	Exclusion criteria	Aims
Tinelli (2015)	WISC-III	Any major cerebral damage, intraventicular hemorrhage grade I, motor impairment or other specific disorders at neurological examination, congenital malformations, major ocular anomalies such as cataracts, optic atrophy and retinopathy of prematurity, auditory impairment	To evaluate the impact of prematurity (in the absence of severe brain lesions) on parietal functions in VLBW school-aged children
Tommiska (2020)	WISC-III (Finnish)	No exclusions (18 children (9% of the survivors) with severe cognitive impairment could not be assessed and were excluded from the analysis)	To investigate cognitive and motor outcomes, ADHD-behaviour, school performance, and overall outcomes in ELBW children at preadolescence, and minor neuromotor impairments in a subpopulation
van Veen (2020)	WISC-III (Dutch)	Those too disabled to be assessed	To examine verbal IQ and performance IQ scores and associations with sociodemographic factors, neonatal risk factors, early cognitive outcomes and academic achievement scores in VPT/EPT children at 8 years
van't Westende (2020)	WISC-III (Dutch)	Congenital anomalies of the central nervous system, severe other congenital anomalies, chromosomal disorders, metabolic disorders, neonatal meningitis	To compare quantitative measures derived from electroencephalography (EEG) between EPT and VPT born children at 9-10 years of age
Vermeulen (2022)	WISC-III (Dutch)	Children from parents living outside the adherence area of MMC, referrals from other NICUs, children with congenital malformations	To evaluate whether in preterm born children motor performance at two years was associated with PIQ at 8 years
Aanes (2019)	WISC-IV	Severe cerebral palsy, severe sensory impairments (deafness or blindness), any contraindications for MRI	To explore VLBW/NBW differences in volumes of hippocampal subfields and memory function, to to examine if hippocampal subfield volumes were associated with neonatal risk factors, and to investigate any significant structure-function relationships between hippocampal subfields and memory test scores in the VLBW group

Note. n.r. = not reported

Table 4

Sample characteristics

Reference	Sample size		Study group			
	PT/LBW (n)	Controls (n)	Weeks of GA Mean (SD, range)	Birth weight, g Mean (SD, range)	Age at follow-up, y Mean (SD, range)	FSIQ score Mean (SD)
Breslau (2001)	411 [1]	306 [1]	n.r.	n.r.	11	96.7 (n.r.) [1]
Carmo (2022)	54	n/a	30.0 (3.5) [a]	1354.0 (623.5) [a]	6–14	96.0 (14.9)
Chaudhari (2004)	180	90	NR (32–FT)	1549 (242.3, 860–1999)	12	89.5 (16.9)
Cho (2022)	36	26	27.5 (2.4)	1086.9 (350.5)	6.3	83.94 (15.89)
Cooke (2003)	268	198	29.8 (23–32)	1467 (512–2860)	6.8–8.4	89.4 (14.2)
Dai (2020)	76	n/a	26 [2] (25–28)	919 (206)	7.21 (0.2)	90 (16)
Domellöf (2020)	51	57	31.1 (3.5, 23–35)	1637 (690, 404–2962)	7.8 (0.6, 7.0–8.7)	94.4 (11.1)
Doyle (2015)	258	262	26.7 (1.9) [b]	n.r.	8	95.5 (16.0)
Emond (2006)	202	81	38.8 (1.4)	2346 (148, 1840–2490)	8.2 [2] (8.0–9.1)	75.2 (13.3)
Fan (2013)	97	n/a	33.6 (2.0)	1890 (4.9)	6–7	98.7 (15.8)
Farooqi (2016)	132	103	24.4 (0.7, 23–25)	718 (129)	11.96 (1.7, 10.1–15.6)	80.3 (18.7)
Grunewaldt (2014)	23	33	25.8	778	10.16	98 (19.57)
Gire (2019) [3]	114	n/a	26.25 (0.87)	879.03 (181.73)	7–10	91.54 (15.35)
Herber-Jonat (2014)	60	n/a	24.3 (0.4, 22–24) [c]	645 (118) [c]	8.7 [2] (7.0–10.8)	87.37 (n.r.)
Hutchinson (2013)	189	173	26.5 (2.0)	833 (164)	8.45 (0.41)	93.1 (16.1)
Iai (2022)	50	n/a	27.8 (2.20, 23–31)	986 (270, 396–1492)	7.74 (0.563)	99.6 (10.9)

Continued on next page

Table 4 – Continued from previous page

Reference	Sample size		Study group			
	PT/LBW (n)	Controls (n)	Weeks of GA Mean (SD, range)	Birth weight, g Mean (SD, range)	Age at follow-up, y Mean (SD, range)	FSIQ score Mean (SD)
Ionio (2022)	185	n/a	28.46 (2.08)	1045 (267.37)	7	104.68 (14.55)
Jansen (2021)	71	n/a	29.2 (2.0)	1234 (365)	10	95.3 (17)
Jin (2020)	37	n/a	34.6 (7.5, 32–36)	2229.2 (472.8, 1100–3110)	9.1 (1.2, 7.2–10.9)	92.9 (11.9)
Kaul (2021)	359	367	25.0 (1.0, 22–26)	782 (168)	6.6 (0.2)	83.9 (14.6)
Kim (2021)	71	40	27.5 (2.2)	885 (238)	7.4 (0.5)	89.1 (18.3)
Kochukhova (2022)	78	50	29 (2)	1204 (343)	12	94 (17)
Kroll (2019)	161	n/a	28.95 (2.31)	1286.6 (277.06)	15.39 (0.47)	97.02 (17.2)
Lahat (2014)	121	128	n.r.	841.91 (123.91)	8	92.41 (14.66)
Linden (2011) [4]	100	50	29.6 (2.4)	1322 (434)	7	100.7 (13.7)
Løhaugen (2011)	16	19	25.8 (1.8)	778 (118)	14–15	78 (13)
McNicholas (2014)	52	48	29.9 (2.8) [d]	1172 (219, 650–1500) [d]	11.6 (1.0)	89.71 (12.51)
Mu (2008)	130	59	29.54 (2.72)	1165.04 (238.90)	8	93.14 (16.33)
Mulder (2011)	56	22	27.6 (1.8, 25.0–30.9)	n.r.	9.75	90.8 (12.6)
Nagy (2022)	72	33	ELBW: 27.59 (n.r.) VLBW: 29.8(n.r.)	ELBW: 840.9 (n.r.) VLBW: 1262(n.r.)	9–10	ELBW: 102.7 (n.r.) VLBW: 109 (n.r.)
Pinto-Martin (2004)	645	n/a	n.r.	n.r.	9.7 (0.55)	97.1 (17.5) ELBW: 88.6 (20.6) VLBW: 94.0 (16.9) 1501–2000g: 99.3 (16.3)
Rickards (2001)	120	41	29.3 (2.0)	1167 (215)	14	92.2 (15.5)
Rose (2011)	44	86	29.7 (2.8)	1165.2 (268.4)	11	85.68 (14.26)

Continued on next page

Table 4 – Continued from previous page

Reference	Sample size		Study group			
	PT/LBW (n)	Controls (n)	Weeks of GA Mean (SD, range)	Birth weight, g Mean (SD, range)	Age at follow-up, y Mean (SD, range)	FSIQ score Mean (SD)
Samuelsson (2006)	61	56	31.2 (n.r.) [e]	1234 (n.r.) [e]	15	84.9 (17.5)
Squarza (2017)	99	n/a	27.7 (2.3)	769.7 (165.5)	7	103.0 (15.7)
Takayanagi (2013)	ELBW: 93	n/a	ELBW: 26.8 (2.3)	ELBW: 759 (146)	6	ELBW: 85.3 (13.4)
	VLBW: 96		VLBW: 30.0 (1.9)	VLBW: 1281 (148)		VLBW 91.8 (9.7)
Tinelli (2015)	29	n/a	28.3 (24–32)	1107 (490–1480)	8.6 (6.5–11.1)	93.7 (16.6)
Tommiska (2020)	122	n/a	27.3	802	11	90 (20)
van Veen (2020)	120	n/a	28 (1.6)	1031 (252)	8	94.5 (16.9)
van't Westende (2020)	53	n/a	29 [2] (25–32)	1271 [2] (650–1960)	n.r.	97.36 (15.75)
			VPT: 30 (28–31)	EPT: 1024 (650–1150)		EPT: 93.82 (15.04)
			EPT: 26 (25–27)	VPT: 1423 (720–1960)		VPT: 99.03 (16.01)
Vermeulen (2022)	88	n/a	28 [2] (27–29)	1078 (275)	8	98.4 (13.5)
Aanes (2019)	34	n/a	29.2 (2.5)	1031 (305)	8.7 (1.7)	98 (12)

Note. n/a = not appliccable; n.r. = not reported; FT = full term (>37 weeks)

[1] Calculated from two groups

[2] Median

[3] Excluded participants with FSIQ <65

[4] Excluded participants with FSIQ <70

[a] n = 83

[b] n = 298

[c] n = 79

[d] n = 64

[e] n = 56

2015). Cooke et al. (2003) chose not to correct at seven years of age in their group, siting the controversies of correcting beyond school age, and concluding that the differences observed in their study were greater than what ten weeks of correction could account for.

4.2.1 Intelligence outcome

As seen in Table 3, the most represented versions of the WISC were WISC-III and WISC-IV, with eighteen and seventeen articles, respectively. Four studies used WISC-R, and only one article had measurements done with the newest version of the WISC, WISC-V.

A complete list of obtained mean FSIQ score is found in Table 4, including scores for subgroups of cohorts where this was provided. Lowest reported mean FSIQ in a cohort born PT/LBW was 75.2 (Emond et al., 2006). It is worth noting that the cohorts in the study by Emond et al. (2006) were from a poor area in northeast Brazil, and while the study design and inclusion criteria used in the publication were acceptable for this thesis, the population this cohort was drawn from was growing up in a notably poorer area than other studies represented in this review. This study also obtained the lowest difference in means, with only a 4.2 difference in FSIQ between the LBW group and the controls, and the controls scoring a mean of 79.4. The second lowest reported mean was 80.3 (Kaul et al., 2021). For PT/LBW the highest obtained mean FSIQ was 104.68 (Ionio et al., 2022), and leaving out Emond et al. (2006) mentioned above, the studies with control group reported a mean FSIQ for these between 95.44 and 116.6.

Most of the included articles identified significantly lower scores in their study group compared to controls. This was true both when compared to a control group and when compared to normative values. Significant differences were reported in almost all studies with comparable control groups, and in studies ranging from ELBW (Doyle et al., 2015; Hutchinson et al., 2013; Kim et al., 2021; Lahat et al., 2014; Nagy et al., 2022; Tommiska et al., 2020), EPT (Farooqi et al., 2016; Hutchinson et al., 2013; Kaul et al., 2021), VLBW (Dai et al., 2020; McNicholas et al., 2014; Mu et al., 2008; Rickards et al., 2001; Samuelsson et al., 2006), VPT (Cho et al., 2022; Cooke et al., 2003; Kochukhova et al., 2022; Linden et al., 2015; Mulder et al., 2011), to M-LPT/LBW samples (Breslau et al., 2001; Chaudhari et al., 2004; Domellof et al., 2020; Emond et al., 2006; Rose et al., 2011). However, in some studies, even though they found lower FSIQ in the study sample, the difference failed to reach significance (Fan et al., 2013; Grunewaldt et al., 2014; Jin et al., 2020; Nagy et al., 2022). A complete overview of studies with comparison between FSIQ results of study groups and control groups can be found in Table 6 on page 40. In cases where articles reported measures for subgroups in their samples these were included in this table instead of the combined results. As the publications provided different measures

of variance and difference in group means, the reported measure and significance have been included in the table.

Many of the included articles have examined how the samples score compared to the normative range. When using standardised tools for intelligence testing such as the WISC, the normative range is considered to be an FSIQ score between 85 and 115, that is, not exceeding 1SD above or below the normative mean of 100. As seen in Table 6 only five of the control groups obtained a mean below the normative mean. In contrast, only two out of 25 of the PT/LBW groups obtained a mean above 100, one being a study where all participants with FSIQ scores below <70 have been excluded. Also counting the studies without control group, the number would be four out of 42. It is worth noting that, as many of the authors point out, the mean FSIQ scores for PT/LBW samples at group level mainly fall into the normal range. This was also found in samples containing exclusively EPT/ELBW children (Doyle et al., 2015; Grunewaldt et al., 2014; Hutchinson et al., 2013; Kim et al., 2021; Lahat et al., 2014; Nagy et al., 2022; Tommiska et al., 2020). When examining studies reporting mean FSIQ below -1SD, the sample categories represented also range from EPT/ELBW to PT/LBW (Cho et al., 2022; Emond et al., 2006; Farooqi et al., 2016; Kaul et al., 2021; Samuelsson et al., 2006)

When investigating occurrences of cognitive deficit in their sample, the studies used different criteria to define impairment. An overview of the classification used and the occurrence of impairment found in the study groups is listed in Table 5 on page 39. Most studies utilised the -1SD criteria, setting the cut-off for normal range at 85 and above, with FSIQ scores <70 indicating moderate, major or severe impairment, and with scores between 70–84 indicating mild impairment. Some studies used the lowest 10th percentile, setting cut-off at 80. The prevalence of low FSIQ or some degree of impairment ranged from 9.3% to 40%. Severe/major impairment with FSIQ <70 was observed in 5% to 19% cases. Some studies specified that the percentage of children with lower scores in FSIQ found in their study was significantly higher in premature children than in full-term children, both compared to control groups (Chaudhari et al., 2004; Fan et al., 2013; Hutchinson et al., 2013; Kim et al., 2021; Mulder et al., 2011; Tommiska et al., 2020) and compared to norm (Dai et al., 2020). In addition to the studies presented in Table 5, McNicholas et al. (2014) also identified a high prevalence of cognitive impairment. In their study, 20% of VLBW children scored more than 2SD below their NBW peers, and the overall group averaged 1SD below their NBW counterparts. In contrast, Squarza et al. (2017) reported a low or not higher than expected percentage of impairment in their study samples.

Table 5

Occurrence of cognitive deficit in preterm or low birth weight group

Reference	Classification	Score	% of participants
Carmo (2022)	Borderline	<80	10.9%
Chaudhari (2004)*	Borderline	70–84	24.4%
	Mental retardation	<70	15.4%
Dai (2020)*	Low	<85	36%
	–	<70	5%
Fan (2013) *	Borderline	70–80	9.3%
Gire (2019)	Mild	77–89	28.4%
	Moderate	65–77	19.60%
Herber-Jonat (2014)	Mild	70–84	n.r.
	Severe	<70	8%
Hutchinson (2013)*	Mild	<85	35.8%
	Major	<70	17.6%
Jansen (2021)	Mild	70–84	21%
	Moderate-severe	<70	7%
Jin (2020)	Borderline	70–84	24.3%
Kim (2021)*	Mild	70–84	24%
	Severe	<70	13%
Kochukhova (2022)	Mild	70–84	40%
	Severe	<70	19%
Mulder (2011)*	Neuropsych. deficits	<85	25%
Rickards (2001)	n.r.	<85	18%
	–	70–84	12.5%
	–	<70	5%
Tommiska (2020)*	Mild	70–85	20%
	Moderate	55–69	14%
	Severe	<55	3%

Note. * Reports significant difference
[1] Only significant results in<70
[2] Excluded FSIQ<65

Table 6

Comparison between FSIQ study group and control group from studies with a control group as reported in included publications

	Reference	PT Mean IQ (SD)	FT Mean IQ (SD)	Difference in Mean	*p*
GA (weeks)					
<35	Domellöf (2020)	94.4 (11.1)	102.6 (10.3)	F = 14.39	<.001
≤32	Linden (2011)[1]	100.7 (13.7)	110.6 (12.0)	n.r.	<.001
<32	Cho (2022)	83.94 (15.89)	98.65 (10.85)	14.71	<.001
<32	Cooke (2003)	89.4 (14.2)	100.5 (13.7)	11.1	<.001
28–32	Kochukhova (2022)	96 (16)	102 (11)	6	<.05
<31	Mulder (2011)	90.8 (12.6)	104.6 (9.4)	13.8	<.001
<28	Kochukhova (2022)	90 (13)	102 (11)	12	<.01
<27	Kaul (2021)	83.9 (14.6)	100.3 (11.7)	16.4	<.001
23-25	Farooqi (2016)	80.3 (18.7)	104.6 (15.7)	24.3	<.001
BW (g)					
<2500	Breslau (2001)[2]	96.73 (n.r.)	104.06 (n.r.)	7.33	<.001
<2000	Chaudhari (2004)	89.5 (16.9)	97.2 (14.1)	7.7	<.05
<1501	Rickards (2001)	92.2 (15.5)	105.0 (13.3)	8.9	<.005
<1500	McNicholas (2014)	89.71 (12.51)	101.27 (11.73)	11.56	<.001
<1500	Mu (2008)	93.14 (16.33)	111.05 (14.81)	17.91	<.001
<1500	Nagy (2022)	109 (n.r.)	116.6 (12.03)	7.6	.028
<1500	Samuelsson (2006)	84.9 (17.5)	97.1 (13.2)	t = -4.26	<.01
<1500	Aanes (2019)	98 (12)	WASI	n/a	n/a
<1000	Grunewaldt (2014)	98 (19.57)	105 (20.52)	7	.208
<1000	Lahat (2014)	92.41 (14.66)	103.78 (12.44)	t = -6.58	<.005
<1000	Nagy (2022)	102.7 (n.r.)	116.6 (12.03)	13.9	<.0001
BW (g) & GA (weeks)					
<1750 and <37	Rose (2011)	85.68 (14.26)	95.44 (11.62)	F = 17.55	<.001
1500–2500 — and ≤37	Emond (2006)	75.2 (13.3)	79.4 (14.2)	4.2	<.05
≤1500 and ≤32	Tinelli (2015)	93.7 (16.6)	Ravens	n/a	n/a
<1000 and <30	Kim (2021)	89.1 (18.3)	107.1 (12.7)	18	<.001
<1000 and <28	Doyle (2015)	95.5 (16.0)	104.9 (14.1)	9.4	n.r.
<1000 and <28	Hutchinson (2013)	93.1 (16.1)	105.6 (12.4)	12.5	<.001
<1000 and >22	Tommiska (2020)	90 (20)	112 (14)	t = 6.6	<.001

Note. F = result of ANOVA; t = result of t-test; n/a = not appliccable; n.r = not reported; WASI = Wechsler Abbreviated Scale of Intelligence; Raven = Raven's Progressive Matrices

[1] Excluded children with FSIQ <70

[2] PT and FT mean calculated from two groups

4.2.2 WISC-indices

In addition to FSIQ scores, many studies have analyzed and reported scores on other indices provided by the WISC. Results from the studies reporting scores on the Verbal Comprehension Index, Perceptual Reasoning Index, Working Memory Index and Processing Speed Index from WISC-IV are listed in Table 7 on page 42. Most of the studies found significant differences between the study group and control group in all indices (Cho et al., 2022; Hutchinson et al., 2013; Kaul et al., 2021; Kim et al., 2021; McNicholas et al., 2014; Mulder et al., 2011; Nagy et al., 2022). Some could only find significant associations in the ELBW subsection of their cohort (Nagy et al., 2022). Others only found a significant outcome in one index (Grunewaldt et al., 2014; Nagy et al., 2022). Most of the mean scores for the PT/LBW groups are within the normal range compared to the norm (85–115), though in the EPT group studied by Kaul et al. (2021), the mean Processing Speed Index and Working Memory Index scores were 85 and 78.3, respectively. In their findings, Carmo et al. (2022) observed scores below 80 in 5.5% of their PT cohort for Verbal Comprehension Index, 10.9% for Perceptual Reasoning Index, 16.4% for Processing Speed Index and 25.5% for Working Memory Index.

Looking closer at each index, some nuances emerge. In 11 of the 15 publications reporting scores on these indices, Verbal Comprehension Index was the index with the highest reported mean for the PT/LBW group, ranging from 89.6 to 112.6. The results for the FT/NBW groups are more scattered, with six out of ten articles reporting the highest mean score in the Perceptual Reasoning Index. For the PT/LBW cohorts, the reported mean for this index ranged from 87.47 to 106.1, while in the FT/NBW cohorts, scores range from 99.7 to 113.5. Jin et al. (2020) found that the PT childrens score on the Perceptual Reasoning Index was significantly higher than scores on all other indices, reporting it as a relative strength. In contrast, Kaul et al. (2021) identified this index as the one where most of their PT-born participants had a deficit. In their study, Nagy et al. (2022) found that participants born ELBW scored lower than the control group in all indices, and lower than the VLBW group in the Processing Speed Index. Significant differences for VLBW participants were observed only in the Perceptual Reasoning Index. Kaul et al. (2021) found Working Memory Index to be the weakest index in their EPT cohort, supported by the results of Gire et al. (2019). The Working Memory Index also the broadest range of mean scores within the PT/LBW group. Interestingly, it is also the index with the lowest upper mean score obtained for a PT/LBW group. For Processing Speed Index, PT/LBW means ranged from 85 to 107.1, while the FT/NBW cohorts' mean scores range from 96.9 to 109.2, putting the lowest reported mean of a PT/LBW group a full SD below the lowest of a FT/NBW group.

Table 7

Means and comparisons between preterm or low birthweight groups (PT) and full-term groups (FT) by WISC-indices as reported in included publications

		Reference	VCI			PRI		
			PT	FT	Mean Diff. (*p*)	PT	FT	Mean Diff. (*p*)
GA	<37	Ionio (2022)	99.66 (20.96)			99.22 (21.93)		
	<36	Carmo (2022)	100.0 (13.9)			99.0 (15.6)		
	<35	Domellöf (2020)	96.0 (10.1)	102.8 (10.3)	10.21 (.002)	101.1 (14.2)	109.5 (11.3)	11.43 (<.001)
	32–36	Jin (2020)	92.70 (10.29)			101.5 (11.10)		
	<32	Cho (2022)	89.64 (15.94)	99.83 (12.90)	n.r. (.013)	87.47 (16.36)	101.70 (12.54)	n.r. (.001)
	<31	Mulder (2011)	94.0 (11.1)	102.6 (10.3)	8.7 (.002)	90.8 (14.6)	104.7 (10.3)	13.9 (<.001)
	<30	Kim (2021)	94.2 (16.5)	108.5 (11.5)	n.r. (<.001)	91.7 (19.7)	108.5 (15.1)	n.r. (<.001)
	<28	Gire (2019)	98.24 (16.19)			91.30 (15.31)		
	<27	Kaul (2021)	92.2 (14.4)	104.0 (11.5)	10.0 (<.001)	89.7 (14.2)	104.8 (12.7)	13.8 (<.001)
BW	<1500g	McNicholas (2014)	92.04 (15.03)	99.65 (13.50)	7.61 (<.05)	92.08 (12.14)	99.69 (10.57)	7.61 (<.05)
	<1500g	Nagy (2022)	112.6 (9.39)	117 (11.68)	n.r. (.265)	106.1 (9.93)	113.5 (11.5)	n.r. (.024)
	<1000g	Nagy (2022)	108 (11.49)	117 (11.68)	n.r. (.004)	100.7 (13.36)	113.5 (11.5)	n.r. (<.001)
	<1000g	Grunewaldt (2014)	102 (n.r.)	106 (n.r.)	n.r. (.374)	98 (n.r.)	104 (n.r.)	n.r. (.227)
GA & BW	<32 & <1500g	Iai (2022)	102 (15.7)			99.0 (11.8)		
	<28 & <1000g	Hutchinson (2013)	93.1 (14.3)	103.2 (12.6)	7.8 (<.001)	93.1 (14.3)	108.2 (12.8)	10.1 (<.001)

		Reference	WMI			PSI		
			PT	FT	Mean Diff. (*p*)	PT	FT	Mean Diff. (*p*)
GA	<37	Ionio (2022)	93.94 (23.05)			88.38 (23.80)		
	<36	Carmo (2022)	90.0 (13.3)			93.0 (14.4)		
	<35	Domellöf (2020)	87.3 (12.0)	92.6 (11.2)	5.63 (.019)	95.3 (14.0)	97.4 (12.2)	0.46 (.501)
	32–36	Jin (2020)	92.86 (14.61)			92.43 (13.12)		
	<32	Cho (2022)	86.53 (17.19)	97.13 (11.00)	n.r. (.005)	86.44 (18.34)	98.30 (15.29)	n.r. (.013)
	<31	Mulder (2011)	92.5 (11.0)	102.4 (8.5)	9.9 (<.001)	93.8 (13.2)	102.2 (13.3)	8.4 (.013)
	<30	Kim (2021)	91.2 (16.4)	103.0 (12.6)	n.r. (<.001)	86.3 (18.1)	99.0 (15.1)	n.r. (<.001)
	<28	Gire (2019)	90.95 (14.40)			92.30 (14.48)		
	<27	Kaul (2021)	78.3 (13.1)	90.7 (11.0)	11.2 (<.001)	85.0 (14.4)	96.9 (12.7)	11.5 (<.001)
BW	<1500g	McNicholas (2014)	93.70 (15.55)	104.40 (14.16)	10.69 (<.05)	88.24 (12.35)	100.19 (10.73)	11.96 (<.05)
	<1500g	Nagy (2022)	105.9 (11.36)	110 (13.6)	n.r. (.524)	107.1 (14.42)	109.2 (11.9)	n.r. (1.000)
	<1000g	Nagy (2022)	99.3 (12.94)	110 (13.6)	n.r. (.003)	97.3 (14.13)	109.2 (11.9)	n.r. (.002)
	<1000g	Grunewaldt (2014)	91 (n.r.)	101 (n.r.)	n.r. (.038)	97 (n.r.)	103 (n.r.)	n.r. (.301)
GA & BW	<32 & <1500g	Iai (2022)	95.0 (15.3)			99.7 (13.5)		
	<28 & <1000g	Hutchinson (2013)	94.0 (16.3)	102.4 (12.9)	7.1 (<.001)	94.7 (15.9)	101.1 (11.9)	5.7 (<.001)

Note. n.r. = not reported

Ionio et al. (2022) found this to be the weakest index in their PT group, while for Domellof et al. (2020), this was the only index where results were insignificant. For those who analyzed subdivisions within their sample, both Nagy et al. (2022) and Domellof et al. (2020) found the effect particularly strong between children born EPT and FT.

A few authors had interest in specific subtests. Rickards et al. (2001) identified significant differences between VLBW and NBW in specific subtests reflecting educational achievements such as verbal learning and problem-solving skills, namely Information and Arithmetic. Compared to children with ADHD, Iai et al. (2022) found a significant difference in only one subtest, Cancellation, in favour of the ADHD control group. The profiles of VLBW/VP children and ADHD groups were similar. Kaul et al. (2021) found no clear profile within their EPT group. They noted a considerable heterogeneity in their study group and pointed to the relative strengths of individuals. In the subjects with the most severe overall cognitive impairment the proportion of individuals with no or mild deficit ranged from 28% to 55% for the different indices. While Hutchinson et al. (2013) proposes that finding poorer outcomes across all domains points to prematurity causing global cognitive deficit rather than impairments in selective domains, Kaul et al. (2021) advocates for these individual variations and relative strengths and weaknesses hidden in the FSIQ.

4.2.3 Relations between intelligence and BW or GA

Many studies have explored a potential relationship between BW or GA and intelligence. In many cases GA predicted cognitive functioning in terms of short gestation being related to poorer cognitive outcomes in children born PT/LBW (Domellof et al., 2020; Doyle et al., 2015; Hutchinson et al., 2013; Kim et al., 2021; Kroll et al., 2019; Takayanagi et al., 2013; van Veen et al., 2020). Tommiska et al. (2020) found that none of the children in their cohort born at 22 to 23 weeks of GA fell within the normative range. From week 24, the number of ELBW children presenting normal cognitive development increased gradually to a third between week 24 to 26, and to over half of the sample at week 27. In line with this, Kroll et al. (2019) found a linear association between IQ and GA, with every completed gestational week after week 24 being associated with a 1.4-point increase in Full-scale IQ. Hutchinson et al. (2013) found that the children born at 26 weeks achieved lower scores than those born at weeks 26 to 27, though only the difference in the Perceptual Reasoning Index reached significance. While not statistically significant, Jin et al. (2020) also noted that the mean FSIQ in their MPT group was 4.9 points lower than their LPT peers. Some of the studies did not identify a significant relationship between FSIQ and gestational age variables (Fan et al., 2013; Nagy et al., 2022).

Birth weight was also a widely explored variable. Many identified a significant relationship between birth weight and FSIQ score (Hutchinson et al., 2013; Kim et al., 2021; Kroll et al., 2019; Pinto-Martin et al., 2004; Takayanagi et al., 2013), Others did not find this difference significant (Fan et al., 2013; Nagy et al., 2022; Rickards et al., 2001), or found that weeks of GA was a better predictor (Kroll et al., 2019). In the study of Pinto-Martin et al. (2004), the authors found that IQ-score dropped by 5 points between the LBW group and the VLBW group, with an additional drop of 5 points to the ELBW group. Kim et al. (2021) also reported a link between discharge weight z-score and low FSIQ in their EPT group, with a 1-point increase in the z-score decreasing the risk of a low FSIQ score by approximately half. According to some studies, those being born both PT and SGA had the poorest outcomes (Kim et al., 2021; van Veen et al., 2020).

4.2.4 Perinatal factors

The articles in this review presented varied outcomes regarding perinatal factors. As mentioned earlier, several studies have deliberately excluded participants with certain conditions such as intraventricular haemorrhage or other perinatal complications, and other conditions common in children born premature. Still, many of the publications found several perinatal factors having adverse associations with cognitive outcomes. In their study on EPT/ELBW survivors born after surfactant was introduced, Doyle et al. (2015) found strong and persistent associations between intraventricular haemorrhage and lower cognitive scores in EPT/ELPW children. In addition they identified negative associations between postnatal corticosteroids and cognitive outcomes, and also concluded that these findings did not diminish over time. In line with this, Takayanagi et al. (2013) found a significant relationship between the total dose of steroids (dexamethasone) and unfavourable outcomes in FSIQ in general and Verbal IQ especially. Kochukhova et al. (2022) investigated the effect of administered antenatal steroids on neurodevelopment. In their follow-up of children born PT at 12 years of age, this treatment was associated with a higher FSIQ, but only in the EPT part of their cohort. The higher FSIQ score was predicted predominantly through better scores on Verbal Comprehension Index and Perceptual Reasoning Index. Kim et al. (2021) found a higher occurrence of longer duration of antibiotic among those with FSIQ <85. The rate of laser treatment for the eye disease retinopathy of prematurity was also significantly higher in this group; EPT children who received laser treatment for retinopathy of prematurity had an 8.8 times higher risk of having a low FSIQ score. Other authors could not conclude with a significant relationship between FSIQ and any active prenatal or postnatal care. Carmo et al. (2022) found no association between prolonged invasive mechanical ventilation and changes in cognitive development in their cohort, and Mu et al. (2008) did not identify any correlations between

cognitive outcome and perinatal outcomes, including Apgar score, intermittent positive pressure ventilation, days of oxygen use or length of hospital stay.

4.2.5 Growth and brain development

Five studies investigated links between growth and intelligence. Three of these found significant associations between head circumference and IQ (Cooke et al., 2003; Emond et al., 2006; Kim et al., 2021; Takayanagi et al., 2013). According to Emond et al. (2006), head circumference at six months and at eight years old had a significant independent effect on FSIQ at eight years. No relation was found between the WISC measurements and birth length or corpulence index. However, Kim et al. (2021) in their EPT cohort observed significantly lower z-scores for weight, height and head circumference in the EPT group with FSIQ <85 compared to the group with normal FSIQ (85–115). In their study, Cooke et al. (2003) found intrauterine growth restriction to be an unlikely explanation for poor cognitive performance, but noted that changes in the relative size of head circumference indicating postnatal growth failure was linked with poorer outcomes. They suggested poorer performances by preterm children may be a consequence of postnatally restricted cerebral and mainly cortical growth, following haemorrhage, periventricular leucomalacia, neonatal steroid treatment or poor nutrition (Cooke et al., 2003). Takayanagi et al. (2013) found an association between restricted intrauterine brain growth and an unfavourable outcome, mainly affecting Performance IQ score. Other than Emond et al. (2006), only Domellof et al. (2020) reports on growth measurements done at the age of WISC assessment. Their findings revealed significant positive correlations between height and FSIQ, Perceptual Reasoning Index, Working Memory Index and Processing Speed Index, and between weight and FSIQ and Working Memory Index. Their study did not observe significant differences based on whether the children were born SGA or AGA. Kochukhova et al. (2022) did also not find any significant associations between FSIQ and SGA in preterm born individuals. These results are contrasted by the findings of Chaudhari et al. (2004), who found that children born both preterm and SGA were also those with the most unfavourable FSIQ outcomes, obtaining much lower scores than the preterm AGA or full-term SGA groups.

Only a few of the studies included have data on brain development. A reason for this may be that this review excluded articles studying cohorts with specific injuries or suspected anomalies where this kind of data is more common. Still, four of the included articles have in some way investigated associations between cognitive outcomes in PT/LBW cohorts and brain development. Grunewaldt et al. (2014) found reduced brain volumes and cortical surface area in ELBW children at ten years of age. In their cohort, larger volumes of putamen and globus pallidus correlated positively with FSIQ. Cho et al.

(2022) investigated alteration in white matter microstructure using tract based spatial statistics and functional connectivity using resting-state functional MRI at six years. The spatial statistics revealed no significant differences in white matter between the VPT group and the controls, implying a possible catch-up. However, they found significant changes in functional connectivity between specific regions of higher-order networks in the VPT cohort, which they suggest could reflect underlying deficits associated with cognitive function. Specifically, they found an increase in functional connectivity between frontoparietal and language networks and a decrease between nodes in the right salience network. Using Quantitative Electroencephalography to analyse electrical activity of the brain, van't Westende et al. (2020) identified a reduced relative power and functional connectivity in the upper alpha frequency in children born EPT compared to VPT. While they found a relationship with some of the variables they studied such as attention and time perception, contrary to their expectations, they did not to find a significant relationship between the quantitative measures and IQ. Kroll et al. (2019) found no evidence for a negative impact of neonatal brain injury in their sample at 12 years of age, suggesting that the VPT brain is able to reorganise and regain function after such injury.

4.2.6 Sociodemographic factors

Over half of the studies have examined possible relationships between cognitive outcome and some sociodemographic factor, such as parental education, occupation or income, social class, type of school or mothers' country of birth. Some studies found that adjusting for such variables could slightly reduce the mean difference in FSIQ between the study group and the control group, but that the difference remained statistically significant (Dai et al., 2020; Doyle et al., 2015; Hutchinson et al., 2013), even when the significance of the sociodemographic variable itself stood for a substantial and significant degree of the variance (McNicholas et al., 2014). Parental education was the most frequently reported sociodemographic variable among the included articles. Association were found between parental education and global cognitive functioning in children born VLBW or PT and FT participants (Domellof et al., 2020; Mu et al., 2008), with maternal education showing the strongest association (Domellof et al., 2020). Many of the studies used only maternal education and not paternal. Strong associations with FSIQ score were identified in children born both VLBW, ELBW, VPT and EPT (Emond et al., 2006; Fan et al., 2013; Kochukhova et al., 2022; Nagy et al., 2022; Vermeulen et al., 2022). Nagy et al. (2022) also reported significant effects of maternal education on each index of the WISC-IV, while Vermeulen et al. (2022) identified a 10.6 point difference in Performance IQ in children of mothers with high compared to low education in their study. Parental occupation at birth was associated with FSIQ score (Kroll et al., 2019; McNicholas et al., 2014), in one

study explaining up to 24% of the variance (McNicholas et al., 2014). Rickards et al. (2001) found that all three composite scales of the WISC-III were significantly higher in children from a higher social class, based on the occupation of the primary income earner in the family. Living in a multilingual household was identified as an earlier predictor of cognitive functioning, but this association disappeared or decreased to a negligible size in adolescence (Doyle et al., 2015; van Veen et al., 2020). Looking at in-group differences, Kim et al. (2021) found no significant difference in parental education level or family income in EPT children with an FSIQ score of 85 or above or those with lower scores. Two of the studies controlling for parental education level did not find that this independently predicted cognitive functioning (Rickards et al., 2001; van Veen et al., 2020). In their study of children in northeast Brazil, Emond et al. (2006) found the negative effects on development caused by an unfavourable social background and health problems had a more significant impact on those born LBW than on NBW peers.

4.2.7 Comorbidities

Many authors have looked at comorbidities concerning intelligence or other outcomes in their study groups. A frequently occurring variable was school-related abilities such as reading, writing or mathematical skills, with PT/LBW groups performing significantly lower than their peers in studies including a control group (Chaudhari et al., 2004; Hutchinson et al., 2013; McNicholas et al., 2014; Samuelsson et al., 2006). Some also confirm a strong link between academic performance and IQ (Chaudhari et al., 2004; Dai et al., 2020). For instance, Dai et al. (2020) found it 9 to 12 times more likely for VPT/VLBW children with low IQ to have poorer teacher-reported academic achievement at age seven than those with normal IQ. A low FSIQ was related to below-expected performance in reading, writing and mathematics even after adjusting for socioeconomic status. Samuelsson et al. (2006) found indications of a catch-up in reading skills in VLBW adolescents, but concluded that a presence of persistent global cognitive impairment would still have a moderating effect on their reading abilities. McNicholas et al. (2014) identified the most robust results in mathematical achievement. Adding low birth weight as a variable in their model explained 11% of the variance in mathematical attainment scores. In general, more VLBW and ELBW children than NBW children have academic special educational needs (McNicholas et al., 2014; Samuelsson et al., 2006; Tommiska et al., 2020), and the association between IQ, achievement scores and birth weight appears to be present also in children not receiving this support (Pinto-Martin et al., 2004). A possible additional challenge is highlighted by McNicholas et al. (2014), who noted that VLBW children missed more days at school because of illness or other medical reasons than their NBW peers.

Chaudhari et al. (2004) highlighted visuo-motor perception as an important factor in many school-related skills and found that the VPT and VLBW children in their cohort had indeed both poor visuo-motor perception and low performances in math and writing. Children with abnormal motor repertoire in infancy presented with lower scores in Working Memory Index and Processing Speed Index, though not in FSIQ at 10 years of age (Grunewaldt et al., 2014). Likewise, Vermeulen et al. (2022) identified a link between fine motor performance measured at two years and Performance IQ score at eight years. The study suggested early assessment with Bayley Scales of Infant Development-III could be valuable for identifying children at risk for lower Performance IQ scores. Similarly, Squarza et al. (2017) found associations between the Griffiths Mental Development Scale-III sub-quotients measuring fine- and gross motor skills and school-age scores on Verbal IQ or Performance IQ, respectively. Both studies point to impairments in motor performance that may hinder active exploration and obstruct subsequent development.

Some of the publications have investigated the relationship between intelligence and other cognitive abilities. Two studies found higher levels of attention difficulties in children born VPT/VLBW (Fan et al., 2013; Iai et al., 2022), even if they were not diagnosed with ADHD (Iai et al., 2022). In children born PT, Jin et al. (2020) revealed a significant discrepancy between the General Ability Index and the Cognitive Proficiency Index in children born preterm. The General Ability Index combines the scores of the Verbal Comprehension and Perceptual Reasoning indices, while the Cognitive Proficiency Index combines Working Memory Index and Processing Speed Index scores. A General Ability Index score that is higher than Cognitive Proficiency Index usually indicates that the child has difficulty in processing information. These observation aligns with the findings of Rose et al. (2011), indicating that lower FSIQ in preterm born could largely be attributed to difficulties with basic cognitive processes such as information processing. In their work, Rose et al. (2011) suggests a model where being born preterm leads to difficulties in elementary cognitive processes, subsequently affecting complex cognitive processes, ultimately contributing to poorer FSIQ outcomes.

In their study, Domellof et al. (2020) identified GA as a significant predictor for behavioural problems, specifically affective and ADHD problems, as measured by the Child Behavior Checklist (6-18). They did not find that FSIQ independently predicted any of the behavioural outcomes. Their finding is supported by Ionio et al. (2022), who found a negative correlation between IQ and social problems in children born PT. In contrast, using the same tool Fan et al. (2013) reports finding associations between FSIQ and social competence domain. Emond et al. (2006) found that teachers and mothers of LBW children reported significantly fewer problems with peer relations, and they speculate that the LBW group grows up to be smaller, less confident and less aggressive than their peer group. In their study, Ionio et al. (2022) also identified a positive correlation between a high and low

IQ and Theory of Mind, which they suggest could underly social difficulties. Finally, Gire et al. (2019) revealed a significantly lower Quality of Life in school-aged EPT children as compared to the reference population, despite an absence of major disability. This serves as a reminder about the hidden challenges urging a more comprehensive approach to support these children.

5 Discussion

The present thesis contains a review and a meta-analyses aimed to estimate the impact of preterm birth on cognitive outcome, specifically in regards to intelligence and FSIQ estimates, as measured with WISC. The meta-analysis assessed the magnitude of the cognitive impairment children born PT/LBW experience compared to their FT/NBW counterparts. Through a systematic search and a thorough review, this thesis brings together data from 42 articles published since 2000 involving 7642 children, including data from the general population and comparison groups. The process of and results from this work underline the complexity and the challenges the field of research on those born premature faces. While the results should be interpreted with caution because of limitations in sample and methodology, this thesis gives a small contribution to support the growing evidence showing an association between premature birth and intelligence outcomes. Despite advancements in medical care, the studies identified in this review underscores a persistent link between premature birth and cognitive challenges.

Intelligence outcomes

The results of the meta-analyses showed that children who are born VPT, with VLBW or with ELBW tend to score a lower FSIQ as measured by WISC compared to their FT or NBW peers. In the meta-analysis the differences observed in total FSIQ score showed that children born VPT scored 11.3 points lower than controls, while children born VLBW scored 12.7 points lower. The score deficit was most substantial in those born with ELBW, scoring 15.2 points lower than their NBW peers. The findings in the analyses align with previous research done since 2000, including Twilhaar et al. (2018), which found a 13-point difference in FSIQ score between children born EPT/VPT and FT in favour of control children. Others have reported a gap in cognitive outcomes between children born premature and FT peers also in early childhood and through adolescence (Arpi et al., 2019; Brydges et al., 2018; Kerr-Wilson et al., 2012; Upadhyay et al., 2019), and this gap seems to maintain into adulthood (Eves et al., 2021). Additionally, some researchers have examined the relevance of age at assessment and concluded that the persisting differences in cognitive function at all ages are likely due to a deficit rather than a delay (Brydges et al., 2018). In the publications included in this review, authors have focused largely on early school age (6–10) with twice as many articles assessing children at this age than in the 11–16 range. It was interesting to find a seemingly even distribution in mean FSIQ scores in these groups with both higher and lower means occurring on both sides. While no conclusions may be drawn from observing the FSIQ means alone, this could be interpreted as in alignment with the hypothesis of a persisting deficit. A difference in FSIQ results is still a highly relevant subject despite improvements in neonatal care

practices (Cheong et al., 2017; Kerr-Wilson et al., 2012; Twilhaar et al., 2018). While survival rates after premature birth have significantly increased due to scientific and technological progress, the cognitive outcomes for these vulnerable children still have a considerable potential for improvement. A cognitive disadvantage may have a negative impact on further development and areas like behaviour problems, school performances, occupational and economic success and quality of life (Fan et al., 2013; Roberts et al., 2017; Strenze, 2007; Tosello et al., 2021).

Despite a significant gap in favour of children born FT/NBW, it is essential to note that the general cognitive performance of the study cohorts falls mainly into the average normative range (FSIQ 85–114). The results of the reviewed studies show that the mean FSIQ score of PT/LBW groups rarely falls more than 1SD below the normative mean of 100. However, looking at the results from the present review and meta-analysis, it is reasonable to conclude that the cognitive deficit the PT/LBW group suffers from has a high prevalence and a severity of clinical significance. Their representation rate in the medium/borderline impairment range is higher than one would expect at their respective age, as is the percentage of children born PT/LBW falling into the lowest part of the curve representing low cognitive functioning or more severe impairment (Chaudhari et al., 2004; Dai et al., 2020; Fan et al., 2013; Hutchinson et al., 2013; Kim et al., 2021; Mulder et al., 2011; Tommiska et al., 2020). In addition, it has been demonstrated that the rate of impairment may be significantly underestimated when determined according to a normative mean as opposed to when compared to local controls (Hutchinson et al., 2013).

While intelligence tests such as the WISC play a crucial role in assessing cognitive functioning, assessment may sometimes fall short in revealing specific cognitive challenges. FSIQ score might hide specific deficits, which the progress in intelligence assessment and introducing indices might help shed light on. This review has attempted to take this into consideration, and while these scores are far less frequently reported, results from the subdomains provided by the WISC-IV were extracted from publications where this was available. Overall, the studies in this review mainly had consistent findings on these indices, identifying significantly lower scores in all domains. Interestingly, almost all studies reported the highest mean score in the Verbal Comprehension Index. Concluding without comprehensive analysis is impossible, but this could suggest that having a relative verbal strength is not uncommon in this population. While some reported more robust findings in specific indices or certain subgroups, these results were scarce and not consistent. One reason for this could be that an LBW or PT profile does not exist, another could lie in the heterogeneity of the groups studied in this review, both in age at assessment, region, exclusion criteria for the study, and others. An important topic for discussion is the impact of mild deficits in multiple areas, as opposed to more severe impairments in just a few, adding to the point brought up and investigated by Kaul et al. (2021). Based on the

findings in this review, the latter is highly relevant for this population. The mean scores for all indices were within the normal range but significantly lower in the study groups compared to their term-born peers in most of the studies. Having shortcomings across various areas could be more likely to have a pronounced impact on daily functioning due to the limited possibility for compensatory strategies. Children with severe or multiple moderate impairments might be more visible and easier to identify and give appropriate support. In contrast, those with multiple but mild impairments might go unnoticed and not receive the support they need from teachers, parents or healthcare providers. Another noteworthy aspect of this review is the diversity in study aims among the included studies. While the heterogeneity of these studies makes it challenging to navigate when synthesising current knowledge, it also reflects the multifaceted applications of intelligence measures in the existing literature and highlights the broad applicability of the WISC as a tool for examining various aspects of child development and cognitive function.

One of the main challenges when navigating research on children born preterm is the wide range of definitions. This variety was also evident in the current thesis, where, despite a pool of 42 articles, only a maximum of four could be combined in the meta-analyses. The consequences of this heterogeneity in inclusion criteria still need to be clarified. Using broad inclusion criteria when doing meta-analyses and systematic reviews have been recommended to increase the number of studies included and to improve the generalisability by avoiding potential biases, e.g. from geographical preference in choice of criteria (Sentenac et al., 2022). This approach was used in the present thesis during the search process and for the review, while the approach to the meta-analysis was far more conservative.

Many of the studies in this review found associations between weeks of GA and IQ score (Domellof et al., 2020; Doyle et al., 2015; Hutchinson et al., 2013; Kim et al., 2021; Takayanagi et al., 2013; van Veen et al., 2020), mainly in the most vulnerable subgroups of their cohorts being born at around week 22-23 (Tommiska et al., 2020) and up to week 33 (Kroll et al., 2019). While this is in line with other recent findings (Joseph et al., 2022), others again have not been able to replicate this association (Brydges et al., 2018; Linsell et al., 2018), suggesting this link might be most relevant for the earliest born. Being born at an earlier gestational age will imply being born more immature and with a higher risk of more severe medical complications, such as BPD (Jensen et al., 2014; Twilhaar et al., 2018). Part of understanding why this effect varies so much could be the interaction between individual biological and social or environmental factors.

Perinatal factors & brain development

Many of the articles included made an effort to analyse perinatal conditions or treatments

and their effect on cognitive outcome (Carmo et al., 2022; Doyle et al., 2015; Kochukhova et al., 2022; Mu et al., 2008; Takayanagi et al., 2013). Others excluded all participants with any history of injuries or conditions such as BPD or periventricular leukomalacia (Fan et al., 2013; Nagy et al., 2022; Tinelli et al., 2015). The variation in how they approached these factors was considerable, and the diversity made it hard to identify any trend in the publications. Nevertheless, many risk factors identified by other researchers also come up in many of the included publications, such as perinatal conditions, the risks and benefits of steroid treatment, growth and brain development. Disturbance in perinatal development and lack of maturity due to being born too soon give those born PT/LBW a poorer outcome than their peers. This disturbance can result from prematurity itself, associated risk factors, or interventions meant to improve the child's outcome. In the second half of gestation, complex and critical events in brain development, such as neuronal proliferation and migration to the cerebral cortex take place. Being born during these fundamental events exposes these children to a notable risk of injuries such as cerebral white matter injury and disturbance in functional connectivity. This type of injury is common in this population and has been associated with cognitive deficits (Cho et al., 2022; Grunewaldt et al., 2014; Tinelli et al., 2015). Moreover, being born prematurely is associated with risk for additional diseases such as BPD, which in turn also has been strongly linked to deficits in intelligence and has been identified as one of the most crucial factors for cognitive outcome in children born EPT and VPT (Cheong et al., 2018; Twilhaar et al., 2018). Looking further, some treatments for BPD, such as postnatal corticosteroids, have also been strongly linked to unfavourable neurologic outcomes such as cerebral palsy. For this reason, the treatment is now used with more consideration for the risks. However, the associations should not be interpreted as evidence to stop using postnatal corticosteroids in infants where this is needed (Daskalakis et al., 2023). The evaluation of risks, benefits, morbidity and mortality continues to be more appropriately examined in randomised trials studying the treatment as opposed to observational studies. Nevertheless, in this discussion, BPD and postnatal corticosteroids serve as a clear example of the complexity of risk factors and how risks and the outcome of interventions meant to save and better the lives of these infants cannot be seen in isolation.

Another biological factor analysed in some of the articles in this review was perinatal and childhood growth. Where the links between growth and intelligence were analysed in any way, there was a large variety in how the subject was approached. Intrauterine growth, postnatal growth failure and smaller head circumference were found to have significant associations with FSIQ. Interventions aimed at better growth before and after term is often an available intervention and has been seen to positively impact these infants (Belfort et al., 2011).

Environmental factors

In addition to biological factors, sociodemographic or environmental factors may have independent effects on cognition. While over half of the included studies assessed this in some way, the variation in how this was assessed made it hard to find clear trends. On the other hand, one could also interpret this as a result of the complex interplay between biological and sociodemographic factors. Among the included publications, parental education, especially maternal education, appeared most frequently as a factor positively correlated with children's cognitive functioning. Although appearing in very few of the publications, parental occupation and family income were also found to be associated with FSIQ. It has been suggested that the impact and interplay of biological and environmental variables change with time (Bendersky et al., 1994; Nisbett et al., 2012). Examples of such changes have also been observed in publications included in this review. For instance, Doyle et al. (2015) found that lower maternal education had a more significant impact on outcomes in later childhood and adolescence than in early childhood. On the other hand, living in a multilingual household was associated with lower performance until eight years of age, after which the association disappeared by late adolescence (Doyle et al., 2015; van Veen et al., 2020). This interplay effect may also have contributed to the varying study results.

The biological-environmental-interplay also becomes relevant in other areas. Intelligence involves understanding complex ideas, adjustment to surroundings, learning from experience, and solving problems through logical thinking (Piaget, 2005), capacities that are crucial for classroom learning and educational achievement. Children with lower IQ will have a greater risk for poorer school outcomes and possible subsequent challenges, as may children born PT/LBW. A link between academic performance and FSIQ in PT/LBW children has been confirmed (Schneider et al., 2014), and is supported by studies in this review (Chaudhari et al., 2004; Dai et al., 2020). Children born with PT/LBW are more likely to have academic special education needs, they are significantly more likely to be rated with poor achademic achievement by their teachers, and they perform lower than their peers in a range of school-related skills such as reading, writing and math. Petrill et al. (2000) conducted a study investigating why some children perform differently in intelligence tests and educational achievement. They concluded that both genetics and environmental factors influence intelligence and educational success. Genetics being more influential in the correlation between intelligence and academic performance, while environmental factors explain more of the discrepancies between intelligence test scores and academic attainment. With this in mind, an emphasis on environmental factors, providing a supportive and enriching environment could positively impact the long-term outcomes of children born PT/LBW.

The same goes for many other specific school-related abilities, many related to cognitive performance such as visuo-motor perception, motor performance, working memory and attention, processing speed and information processing. Some promising interventions aimed at supporting the development of specific weakened abilities do exist, such as computerised training interventions to promote working memory and visuo-motor performance (Grunewaldt et al., 2014), but the effectiveness of this kind of training has been subject to debate (Anderson et al., 2018). As indicated by Rose et al. (2011), it is plausible that early development deficits in elementary cognitive abilities could lead to worse outcomes in other functions or capabilities that rely on them. Rose et al. (2011) found that elementary infant attention and information processing abilities influenced more complex infant abilities such as representational competence, which, in turn, influenced IQ. As another example, visuospatial and visuo-motor performance difficulties may cause deficits in fine motor performance, resulting in lower scores in abilities like processing speed or other skills that rely on fine motor performance (van Veen et al., 2020). This implies that early discovery of such deficits and subsequent interventions directed at relevant abilities may be useful in supporting the development of more complex cognitive functions and processes.

5.1 Limitations of the evidence included in the review

There are multiple risks of selection bias in the included publications. Some studies have included national (Kaul et al., 2021) or regional (Doyle et al., 2015; Hutchinson et al., 2013; Kochukhova et al., 2022; Pinto-Martin et al., 2004) cohorts of children born preterm or with low birth weight. On the other side, Cho et al. (2022) only included children from one centre and only those agreeing to an MRI. Of 42 included publications, 32 were studies on cohorts from a single centre. These may be affected by variations in routine medical care, such as treatment strategy. Recruiting a control group from the same hospital could reduce some of this effect, but pooling data from a more extensive and diverse population will usually be more appropriate. More than half of the articles included in the review were from Europe, and combined, Europe, North America and Australia made up 76% of the articles and 81% of the PT/LBW cohorts. One of the articles was from what could be considered a developing country. This imbalance limits the generalisability of the findings both in regards to a global population, or to a more specific, high-income population. Individuals who choose to decline participation, not to enrol, or who actively opt out may also introduce a risk of over- or underestimating the observed effect in a study, as some, such as those with greater social disadvantage, may be more prone to be lost during follow-up (Wolke et al., 1995). Some of the publications in this review analysed differences between participants and non-participants from the

population if this was possible (Emond et al., 2006; Kim et al., 2021), or compared their study cohort to the general population to identify deviations and possible selection bias (Nagy et al., 2022). Many of the articles lacked relevant information in text, such as exclusion criteria, making it hard to evaluate whether the authors had made no exclusions or excluded any participants without reporting. Different exclusion criteria also make it challenging to compare groups and might obscure the true effects of different exposures. For example, some studies excluded participants with low cognitive scores, while others only excluded those who could not complete the assessment. The risks of selection bias also applies to the process of recruiting a control group. To avoid bias in the current review and meta-analysis, studies where participants were chosen because of their injury or an intervention were excluded. Excluding all studies with any exclusion criteria could introduce different kinds of bias, leaving few publications and limiting the generalisability to the real-world scenario.

Most articles report on using blinded assessment, implying that the assessor of the WISC was blinded to whether the child they assessed were in the study group or control group. A few studies report not being able to use blinded assessment (van Veen et al., 2020; Vermeulen et al., 2022), and at least in theory an observer-expectancy effect could have resulted in confirmation bias in test results in these studies.

The age of the cohorts is another possible source of bias. This is important due to the significant advancements in perinatal care and medical interventions in the last decades. Studies on cohorts born before these advancements will have questionable relevance today. By limiting the selection to studies published after 2000, the included studies mostly consisted of subjects born after the 1990s. However, five studies included subjects born around 1980 and could possibly be a source of contamination. A possible bias to be mindful of in the population discussed in this study is the use of corrected or chronological age. The studies in this review had applied this to a varying extent. Individual studies reported correcting for ages up to 18 years, but most of the included publications did not correct for age. This could confound the results when combining corrected and uncorrected results.

Given that intelligence tends to stabilise at an early school age, and may stabilise even later in preterm individuals, the age at assessment also poses a potential source of bias. Assessing children at a very young age or combining results across a wide age range could confound the results. In this review, the age range in cohorts was partly a result of the WISC age range of 6–16 years, which is a notable range in this case. The number of articles included in the current meta-analyses was not high enough to give estimates on the effect of an age variable.

As seen in this review, various biological and environmental factors are known to possibly affect IQ. Such factors are hard to control for when studying intelligence outcomes in children born preterm, and both may confound the results, as well as make it hard to compare results across cohorts.

5.2 Limitations in methodology

This work has several limitations. In the search- and data-extraction process, dual independent screening is considered the gold standard. Individual reviewers make mistakes, and these mistakes might affect the result and subsequently any conclusions drawn. A recent study revealed that relying solely on single abstract screening can result in overlooking up to 13% of relevant studies (Gartlehner et al., 2020). In this work, a single screening was performed because of time constraints. To mitigate this risk, articles where there was doubt about exclusion were screened by a second person. Only around 5% of the articles were screened by two persons. In addition, the author performed data extraction and synthesising of included articles for the review alone, which introduces another possible source for mistakes and bias (Drucker et al., 2016).

Only studies in a Scandinavian language or English were included. This might have introduced a bias, and potentially valuable studies in other languages might have been missed. Using WISC assessment as an inclusion criterion also has its limitations. The WISC is widely used, sets the standard for standardisation of intelligence tests, and allows for easier comparison between different studies and populations, but other standardised tests or short versions of WISC which often are more efficient, often also have high degree of validity and reliability. By limiting the review to one test and only full version, we might have excluded publications with relevant data. Data availability was also an issue, leading to the criteria of the article having to report FSIQ measurement. These criteria are stricter than ideal, but were necessary for this work to progress within the time limits.

A limitation of the meta-analysis is the limited number of studies included, which is reflected in the rather wide confidence intervals. I addition, there is a risk that positive findings are over-represented, as studies reporting negative or non-significant results are less likely to be published, also known as publication bias (Dwan et al., 2008). This was attempted controlled for by estimating statistical measures of heterogeneity, i.e. Q and I^2. For the analysis of studies including children born with a birth weight <1000 g, the Q statistic was significant, but measures of heterogeneity are less reliable when the number of studies is small and the result should be interpreted with caution. For the analyses of children born with birth weight <1500 g and GA <32 it did not reach statistical significance. However, a non- significant Q should not be interpreted as an absence of

heterogeneity, as a significant Q is hard to obtain with such a small number of included articles as was the case here. Including more studies could reduce the effect of such bias, as well as reduce the confidence interval and give a more exact estimate of true effect, assuming the additional studies do not introduce bias.

As the present thesis has shown, there are multiple challenges when researching the relationship between being born prematurely and cognitive outcomes. One of these challenges is the high variability in how different studies have defined low birthweight or premature birth and how they use these definitions when setting inclusion criteria for their study cohort. While not applied to the current meta-analysis, the recommendations of Sentenac et al. (2022) about including both BW- and GA-criteria were used in this work both during the search process and for the review. For the meta-analysis, strict criteria for comparability were set initially to make the work suitable for a thesis. In hindsight, adopting a more inclusive approach could have been beneficial for the thesis, with regards to Sentenac et al. (2022). A higher number of publications could also have been achieved by including articles using other tests for intelligence, such as Raven's Progressive Matrices, Wechsler Abbreviated Scale of Intelligence, or short versions of WISC. This review has included articles representing children ranging from EPT to LPT and from ELBW to LBW. This provided an opportunity to find results from the whole spectre. However, the many variations of how to define the study group lead to groups possibly overlapping, e.g. EPT subjects being either part of or excluded from a study investigating a VPT. Studies where this was or could be a relevant factor was not included in the analyses.

As per the PRISMA guidelines, it is common practice to use a tool for quality assessment of the included articles. The intended use of a tool like the Newcastle Ottawa Scale, was limited partly due to its applicability being restricted to only certain study designs. This made it difficult to maintain a consistent and standardized approach to quality evaluation. These limitations should be considered when interpreting the findings and conclusions of this study. While other studies have made modified versions of the Newcastle Ottawa Scale when encountering this challenge, this was not prioritised in this review because of limitations in time. In this regard it may be worth noting that many of the items of consideration used in this tool, such as those assessing representatives of the study cohort, selection and comparability of controls and certainty of exposure and outcome were parts of the inclusion and exclusion criteria in the selection process of the present review. Weaknesses which would give the article a lower score would most likely be in areas reported in this review, such as recruitment source, exclusion criteria and whether assessors were blinded for preterm or full-term status. Although this might make it more likely that many articles would have at least a moderately high score in such an assessment, this can not be said with certainty without a structured assessment of risk and quality.

Ethical concideration and conflicts of interest

A total of 32 articles explicitly reported collecting parental informed consent. Seven articles refer to agreement and list number of refusals. Of the remaining articles, one report on lack of interest (Samuelsson et al., 2006), and two does not give any information regarding consent (Iai et al., 2022; Rose et al., 2011). The ten articles not reporting about informed consent involved interviews and assessments suggests a level of active participation from the parents side, but the lack of specific details or documentation on this is unfortunate.

The author and other persons involved in this thesis had no conflicts of interest.

5.3 Conclusions and Implications for further research

Summing up the results from the current work, the findings support the hypothesis that being born PT/LBW is associated with a lower FSIQ score. A wide range of biological and environmental risk factors and variables contribute significantly to the complex landscape of cognitive outcomes in children born PT/LBW. These factors, both independently and combined, could significantly impact life outcomes for this vulnerable population. The findings underscore the importance of approaching the challenges these children face with a holistic understanding in order to provide specific, well-informed support and interventions that address their needs and foster positive developmental trajectories. Assessing cognitive functions in these children during childhood accommodates not only diagnosing a potential issue, it also enables parents, teachers or other helpers around the child to tailor support or specific interventions at an early point in time, which may positively influence the subsequent life outcomes of these high-risk children. Finding the child's strengths and weaknesses is essential, as different challenges require different solutions. Poor cognitive outcomes can be improved through interventions at many levels, from preventing or mitigating adverse outcomes to enhancing specific weaknesses. While the first category may imply interventions aimed at avoiding premature birth, supporting healthy growth and finding the balance between other perinatal risk factors or conditions and treatment, the latter could include specific training programs for these children or providing support and information to parents and teachers on how to encourage the development of vulnerable skills. Understanding the influence of biological, social, and environmental factors on intelligence outcomes, particularly in children born with PT/LBW, will provide insights and may help promote a better personal, educational and social adjustment for this population. Controlling for the numerous potential effects on IQ is challenging when studying intelligence outcomes in these children. Rather than looking for main effects, exploring the interactions among these variables might be a more effective approach to the

subject. Further research should also be done to identify the role of potentially protective environmental factors in shaping intellectual abilities, and to reveal critical periods during which different interventions may be most beneficial. A better understanding also aids in the identification of high-risk infants for close monitoring and early intervention. A more profound knowledge of the factors at play in cognitive development in these infants will hold significance for teachers and education specialists, providing the foundation for the development of targeted support or specialised teaching plans. Moreover, future findings may enhance clinical decision-making and parental counselling during the neonatal period. To achieve such an understanding, the cognitive performance and the factors related to cognitive functioning in this vulnerable population are important subjects for further research.

5.4 Support

The invitation to perform a meta-analysis on cognitive outcomes in children born premature as part of a thesis came from Professor Lars Morten Rimol and Associate Professor Siri Weider from the Institute of Psychology at Norwegian University of Science and Technology (NTNU). Senior Research Librarian Magnus Rom Jensen and Research Librarian Lisbeth Jahren, both also based at NTNU, played an important part in constructing and translating the search string, executing the search, managing search hits in Rayyan and EndNote, and they contributed in the initial screening process by excluding articles with obvious reason for exclusion. In the initial screening phase, Jensen and Jahren excluded 1506 articles, Rimol and Weider reviewed 1148 articles, while fellow student of psychology Kristin Berg Johansson reviewed 2945 articles. PhD Candidate Martin Brattmyr, also associated with NTNU, provided consultation in performing and interpreting the meta-analysis in CMA. Professor Monica Martinussen from the University of Tromsø significantly contributed by offering insights in interpretation of results, guidance in examination of heterogeneity and with constructive feedback on the result section concerning the meta-analyses.

References

Aanes, S., Bjuland, K. J., Sripada, K., Solsnes, A. E., Grunewaldt, K. H., Haberg, A., Lohaugen, G. C. & Skranes, J. (2019). Reduced hippocampal subfield volumes and memory function in school-aged children born preterm with very low birthweight (vlbw). *Neuroimage Clin*, *23*, 101857. https://doi.org/10.1016/j.nicl.2019.101857

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*(2), 717–728. https://doi.org/10.1542/peds.2008-2816

Alexander, G. R., de Caunes, F., Hulsey, T. C., Tompkins, M. E. & Allen, M. (1992). Validity of postnatal assessments of gestational age: A comparison of the method of ballard et al. and early ultrasonography. *American journal of obstetrics and gynecology*, *166*(3), 891–895. https://doi.org/10.1016/0002-9378(92)91357-G

Allen, M. C. (2008). Neurodevelopmental outcomes of preterm infants. *Current opinion in neurology*, *21*(2), 123–128. https://doi.org/10.21037/tp.2019.09.10

Allen, M. C., Cristofalo, E. A. & Kim, C. (2011). Outcomes of preterm infants: Morbidity replaces mortality. *Clinics in perinatology*, *38*(3), 441–454. https://doi.org/10.1016/j.clp.2011.06.011

Allotey, J., Zamora, J., Cheong-See, F., Kalidindi, M., Arroyo-Manzano, D., Asztalos, E., Van Der Post, J., Mol, B., Moore, D., Birtles, D. et al. (2018). Cognitive, motor, behavioural and academic performances of children born preterm: A meta-analysis and systematic review involving 64 061 children. *BJOG: An International Journal of Obstetrics & Gynaecology*, *125*(1), 16–25. https://doi.org/10.1111/1471-0528.14832

Anderson, P. J., Lee, K. J., Roberts, G., Spencer-Smith, M. M., Thompson, D. K., Seal, M. L., Nosarti, C., Grehan, A., Josev, E. K., Gathercole, S. et al. (2018). Long-term academic functioning following cogmed working memory training for children born extremely preterm: A randomized controlled trial. *The Journal of Pediatrics*, *202*, 92–97. https://doi.org/10.1016/j.jpeds.2018.07.003

Arpi, E., D'Amico, R., Lucaccioni, L., Bedetti, L., Berardi, A. & Ferrari, F. (2019). Worse global intellectual and worse neuropsychological functioning in preterm-born children at preschool age: A meta-analysis. *Acta Paediatrica*, *108*(9), 1567–1579. https://doi.org/10.1111/apa.14836

Bancalari, E. & Claure, N. (2015). Advances in respiratory support for high risk newborn infants. *Maternal health, neonatology and perinatology*, *1*(1), 1–10. https://doi.org/10.1186/s40748-015-0014-5

Barrington, K. J. (2001). The adverse neuro-developmental effects of postnatal steroids in the preterm infant: A systematic review of RCTs. *BMC pediatrics*, *1*(1), 1–9. https://doi.org/10.1186/1471-2431-1-1

Baud, O. (2004). Postnatal steroid treatment and brain development. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, *89*(2), F96–F100. https://doi.org/10.1136/adc.2003.028696

Belfort, M. B., Rifas-Shiman, S. L., Sullivan, T., Collins, C. T., McPhee, A. J., Ryan, P., Kleinman, K. P., Gillman, M. W., Gibson, R. A. & Makrides, M. (2011). Infant growth before and after term: Effects on neurodevelopment in preterm infants. *Pediatrics*, *128*(4), e899–e906. https://doi.org/10.1542/peds.2011-0282

Bendersky, M. & Lewis, M. (1994). Environmental risk, biological risk, and developmental outcome. *Developmental psychology*, *30*(4), 484. https://doi.org/10.1037/0012-1649.30.4.484

Bhutta, A. T., Cleves, M. A., Casey, P. H., Cradock, M. M. & Anand, K. J. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *Jama*, *288*(6), 728–737. https://doi.org/10.1001/jama.288.6.728

Blencowe, H., Cousens, S., Chou, D., Oestergaard, M., Say, L., Moller, A.-B., Kinney, M. & Lawn, J. (2013). Born too soon: The global epidemiology of 15 million preterm births. *Reproductive Health*, *10*(S1). https://doi.org/10.1186/1742-4755-10-s1-s2

Blencowe, H., Lawn, J. E., Vazquez, T., Fielder, A. & Gilbert, C. (2013). Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatric research*, *74*(1), 35–49. https://doi.org/10.1038/pr.2013.205

Bolk, J., Farooqi, A., Hafström, M., Åden, U. & Serenius, F. (2018). Developmental coordination disorder and its association with developmental comorbidities at 6.5 years in apparently healthy children born extremely preterm. *JAMA pediatrics*, *172*(8), 765–774. https://doi.org/10.1001/jamapediatrics.2018.1394

Bramer, W. M., Giustini, D., de Jonge, G. B., Holland, L. & Bekhuis, T. (2016). De-duplication of database search results for systematic reviews in endnote. *Journal of the Medical Library Association: JMLA*, *104*(3), 240. https://doi.org/doi.org/10.5195/jmla.2016.24

Breslau, N., Chilcoat, H. D., Susser, E. S., Matte, T., Liang, K. Y. & Peterson, E. L. (2001). Stability and change in children's intelligence quotient scores: A comparison of two socioeconomically disparate communities. *American Journal of Epidemiology*, *154*(8), 711–717. https://doi.org/doi:10.1093/aje/154.8.711

Brummelte, S., Grunau, R. E., Chau, V., Poskitt, K. J., Brant, R., Vinall, J., Gover, A., Synnes, A. R. & Miller, S. P. (2012). Procedural pain and brain development in premature newborns. *Annals of neurology*, *71*(3), 385–396. https://doi.org/10.1002/ana.22267

Brydges, C. R., Landes, J. K., Reid, C. L., Campbell, C., French, N. & Anderson, M. (2018). Cognitive outcomes in children and adolescents born very preterm: A meta-analysis. *Developmental Medicine & Child Neurology*, *60*(5), 452–468. https://doi.org/10.1111/dmcn.13685

Butler, A. S., Behrman, R. E. et al. (2007). *Preterm birth: Causes, consequences, and prevention*. National academies press. https://doi.org/10.17226/11622

Campbell, C., Horlin, C., Reid, C., McMichael, J., Forrest, L., Brydges, C., French, N. & Anderson, M. (2015). How do you think she feels? vulnerability in empathy and the role of attention in school-aged children born extremely preterm. *British Journal of Developmental Psychology*, *33*(3), 312–323. https://doi.org/10.1111/bjdp.12091

Carmo, A. L. S. d., Fredo, F. W., Bruck, I., Lima, J. d. R. M. d., Janke, R. N. R. G. H., Fogaça, T. d. G. M., Glaser, J. A., Riechi, T. I. J. d. S. & Antoniuk, S. A. (2022). Neurological, cognitive and learning evaluation of students who were born preterm. *Revista Paulista De Pediatria*, *40*, 11. https://doi.org/doi:10.1590/1984-0462/2022/40/2020252

Cartwright, R. D., Anderson, N. H., Sadler, L. C., Harding, J. E., McCowan, L. M. & McKinlay, C. J. (2020). Neonatal morbidity and small and large size for gestation: A comparison of birthweight centiles. *Journal of Perinatology*, *40*(5), 732–742. https://doi.org/10.1038/s41372-020-0631-3

Charpak, N., Ruiz-Peláez, J. G., Figueroa de C, Z. & Charpak, Y. (2001). A randomized, controlled trial of kangaroo mother care: Results of follow-up at 1 year of corrected age. *Pediatrics*, *108*(5), 1072–1079. https://doi.org/10.1542/peds.108.5.1072

Chaudhari, S., Otiv, M., Chitale, A., Pandit, A. & Hoge, M. (2004). Pune low birth weight study–cognitive abilities and educational performance at twelve years. *1*(2), 121–8.

Cheong, J. L., Anderson, P. J., Burnett, A. C., Roberts, G., Davis, N., Hickey, L., Carse, E., Doyle, L. W., Group, V. I. C. S. et al. (2017). Changing neurodevelopment at 8 years in children born extremely preterm since the 1990s. *Pediatrics*, *139*(6). https://doi.org/10.1542/peds.2016-4086

Cheong, J. L. & Doyle, L. W. (2018). An update on pulmonary and neurodevelopmental outcomes of bronchopulmonary dysplasia. *Seminars in perinatology*, *42*(7), 478–484. https://doi.org/10.1053/j.semperi.2018.09.013

Cho, H. J., Jeong, H., Park, C. A., Son, D. W. & Shim, S. Y. (2022). Altered functional connectivity in children born very preterm at school age. *Scientific Reports*, *12*(1), 8. https://doi.org/doi:10.1038/s41598-022-11184-x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Academic press.

Cooke, R. W. I. & Foulder-Hughes, L. (2003). Growth impairment in the very preterm and cognitive and motor performance at 7 years. *Archives of Disease in Childhood*, *88*(6), 482–487. https://doi.org/doi:10.1136/adc.88.6.482

Crawford, J. R., Anderson, V., Rankin, P. M. & MacDonald, J. (2010). An index-based short-form of the WISC-IV with accompanying analysis of the reliability and abnormality of differences. *British Journal of Clinical Psychology*, *49*(2), 235–258. https://doi.org/10.1348/014466509X455470

Crowther, C. A., McKinlay, C. J., Middleton, P. & Harding, J. E. (2015). Repeat doses of prenatal corticosteroids for women at risk of preterm birth for improving neonatal health outcomes. *Cochrane Database of Systematic Reviews*, (7). https://doi.org/10.1002/14651858.CD003935.pub4

Crump, C., Winkleby, M. A., Sundquist, J. & Sundquist, K. (2019). Prevalence of survival without major comorbidities among adults born prematurely. *Jama*, *322*(16), 1580–1588. https://doi.org/10.1001/jama.2019.15040

Cui, H., Gong, T.-T., Liu, C.-X. & Wu, Q.-J. (2016). Associations between passive maternal smoking during pregnancy and preterm birth: Evidence from a meta-analysis of observational studies. *PLoS One*, *11*(1). https://doi.org/10.1371/journal.pone.0147848

Dai, D. W. T., Wouldes, T. A., Brown, G. T. L., Tottman, A. C., Alsweiler, J. M., Gamble, G. D., Harding, J. E. & Piano Study, G. (2020). Relationships between intelligence, executive function and academic achievement in children born very preterm. *Early Hum Dev*, *148*, 105–122. https://doi.org/10.1016/j.earlhumdev.2020.105122

Damhuis, S. E., Ganzevoort, W. & Gordijn, S. J. (2021). Abnormal fetal growth: Small for gestational age, fetal growth restriction, large for gestational age: Definitions and epidemiology. *Obstetrics and Gynecology Clinics*, *48*(2), 267–279. https://doi.org/10.1016/j.ogc.2021.02.002

Daskalakis, G., Pergialiotis, V., Domellöf, M., Ehrhardt, H., Di Renzo, G. C., Koç, E., Malamitsi-Puchner, A., Kacerovsky, M., Modi, N., Shennan, A. et al. (2023). European guidelines on perinatal care: Corticosteroids for women at risk of preterm birth. *The Journal of Maternal-Fetal & Neonatal Medicine*, *36*(1).

De Kleine, M., Den Ouden, A., Kollée, L., Nijhuis-van Der Sanden, M., Sondaar, M., van Kessel-Feddema, B., Knuijt, S., Van Baar, A., Ilsen, A., Breur-Pieterse, R. et al. (2003). Development and evaluation of a follow up assessment of preterm infants at 5 years of age. *Archives of Disease in Childhood*, *88*(10), 870–875. https://doi.org/10.1136/adc.88.10.870

de Kieviet, J. F., Zoetebier, L., Van Elburg, R. M., Vermeulen, R. J. & Oosterlaan, J. (2012). Brain development of very preterm and very low-birthweight children in childhood and adolescence: A meta-analysis. *Developmental Medicine & Child Neurology*, *54*(4), 313–323.

Domellof, E., Johansson, A. M., Farooqi, A., Domellof, M. & Ronnqvist, L. (2020). Risk for behavioral problems independent of cognitive functioning in children born at

low gestational ages. *Front Pediatr*, *8*, 311. https://doi.org/10.3389/fped.2020.00311

Doyle, L. W., Cheong, J. L. Y., Burnett, A., Roberts, G., Lee, K. J. & Anderson, P. J. (2015). Biological and social influences on outcomes of extreme-preterm/low-birth weight adolescents. *Pediatrics*, *136*(6). https://doi.org/doi:10.1542/peds.2015-2006

Doyle, L. W. & Anderson, P. J. (2016). Do we need to correct age for prematurity when assessing children? *The Journal of pediatrics*, *173*, 11–12. https://doi.org/10.1016/j.jpeds.2016.03.038

Doyle, L. W., Davis, P. G., Schmidt, B. & Anderson, P. J. (2012). Cognitive outcome at 24 months is more predictive than at 18 months for IQ at 8–9 years in extremely low birth weight children. *Early human development*, *88*(2), 95–98. https://doi.org/10.1016/j.earlhumdev.2011.07.013

Doyle, L. W., Ehrenkranz, R. A. & Halliday, H. L. (2014). Early (<8 days) postnatal corticosteroids for preventing chronic lung disease in preterm infants. *Cochrane Database of Systematic Reviews*, (5). https://doi.org/10.1002/14651858.CD001146.pub4

Drucker, A. M., Fleming, P. & Chan, A.-W. (2016). Research techniques made simple: Assessing risk of bias in systematic reviews. *Journal of Investigative Dermatology*, *136*(11), 109–114.

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E., Easterbrook, P. J., Von Elm, E., Gamble, C. et al. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS one*, *3*(8). https://doi.org/10.1371/journal.pone.0003081

Emond, A. M., Lira, P. I., Lima, M. C., Grantham-McGregor, S. M. & Ashworth, A. (2006). Development and behaviour of low-birthweight term infants at 8 years in northeast Brazil: A longitudinal study. *Acta Paediatr*, *95*(10), 1249–57. https://doi.org/10.1080/08035250600615127

Eves, R., Mendonça, M., Baumann, N., Ni, Y., Darlow, B. A., Horwood, J., Woodward, L. J., Doyle, L. W., Cheong, J., Anderson, P. J. et al. (2021). Association of very preterm birth or very low birth weight with intelligence in adulthood: An individual participant data meta-analysis. *Jama Pediatrics*, *175*(8), e211058–e211058. https://doi.org/10.1001/jamapediatrics.2021.1058

Fan, R. G., Portuguez, M. W. & Nunes, M. L. (2013). Cognition, behavior and social competence of preterm low birth weight children at school age. *Clinics (Sao Paulo)*, *68*(7), 915–21. https://doi.org/10.6061/clinics/2013(07)05

Farooqi, A., Adamsson, M., Serenius, F. & Hagglof, B. (2016). Executive functioning and learning skills of adolescent children born at fewer than 26 weeks of gestation. *Plos One*, *11*(3), 20. https://doi.org/doi:10.1371/journal.pone.0151819

Fjørtoft, T., Grunewaldt, K. H., Løhaugen, G. C. C., Mørkved, S., Skranes, J. & Evensen, K. A. I. (2013). Assessment of motor behaviour in high-risk-infants at 3 months predicts motor and cognitive outcomes in 10 years old children. *Early human development*, *89*(10), 787–793. https://doi.org/10.1016/j.earlhumdev.2013.06.007

Gartlehner, G., Affengruber, L., Titscher, V., Noel-Storr, A., Dooley, G., Ballarini, N. & König, F. (2020). Single-reviewer abstract screening missed 13 percent of relevant studies: A crowd-based, randomized controlled trial. *Journal of Clinical Epidemiology*, *121*, 20–28. https://doi.org/10.1016/j.jclinepi.2020.01.005

Gidi, N. W., Goldenberg, R. L., Nigussie, A. K., McClure, E., Mekasha, A., Worku, B., Siebeck, M., Genzel-Boroviczeny, O. & Muhe, L. M. (2020). Comparison of neonatal outcomes of small for gestational age and appropriate for gestational age preterm infants born at 28–36 weeks of gestation: A multicentre study in ethiopia. *BMJ Paediatrics Open*, *4*(1). https://doi.org/10.1136/bmjpo-2020-000740

Girault, J. B., Langworthy, B. W., Goldman, B. D., Stephens, R. L., Cornea, E., Reznick, J. S., Fine, J. & Gilmore, J. H. (2018). The predictive value of developmental assessments at 1 and 2 for intelligence quotients at 6. *Intelligence*, *68*, 58–65. https://doi.org/10.1016/j.intell.2018.03.003

Gire, C., Resseguier, N., Brevaut-Malaty, V., Marret, S., Cambonie, G., Souksi-Medioni, I., Muller, J. B., Garcia, P., Berbis, J., Tosello, B., Auquier, P., Lemarchand, M. C., Mestre, N., Rebattel, M., Roze, J. C., Coudronnieres, C., Menard, G., Pache, M., Morando, C. & Einaudi, M. A. (2019). Quality of life of extremely preterm school-age children without major handicap: A cross-sectional observational study. *Archives of Disease in Childhood*, *104*(4), 333–+. https://doi.org/doi:10.1136/archdischild-2018-315046

Goldenberg, R. L., Gravett, M. G., Iams, J., Papageorghiou, A. T., Waller, S. A., Kramer, M., Culhane, J., Barros, F., Conde-Agudelo, A., Bhutta, Z. A. et al. (2012). The preterm birth syndrome: Issues to consider in creating a classification system. *American journal of obstetrics and gynecology*, *206*(2), 113–118. https://doi.org/10.1016/j.ajog.2011.10.865

Gould, J. F., Fuss, B. G., Roberts, R. M., Collins, C. T. & Makrides, M. (2021). Consequences of using chronological age versus corrected age when testing cognitive and motor development in infancy and intelligence quotient at school age for children born preterm. *PloS one*, *16*(9). https://doi.org/10.1371/journal.pone.0256824

Grunewaldt, K. H., Fjørtoft, T., Bjuland, K. J., Brubakk, A.-M., Eikenes, L., Håberg, A. K., Løhaugen, G. C. & Skranes, J. (2014). Follow-up at age 10 years in elbw children—functional outcome, brain morphology and results from motor assessments in infancy. *Early human development*, *90*(10), 571–578. https://doi.org/10.1016/j.earlhumdev.2014.07.005

Guillén, Ú., Weiss, E. M., Munson, D., Maton, P., Jefferies, A., Norman, M., Naulaers, G., Mendes, J., Justo da Silva, L., Zoban, P. et al. (2015). Guidelines for the management of extremely premature deliveries: A systematic review. *Pediatrics*, *136*(2), 343–350. https://doi.org/10.1542/peds.2015-0542

Gyamfi-Bannerman, C., Fuchs, K. M., Young, O. M. & Hoffman, M. K. (2011). Nonspontaneous late preterm birth: Etiology and outcomes. *American journal of obstetrics and gynecology*, *205*(5), 456–e1. https://doi.org/10.1016/j.ajog.2011.08.007

Haas, D. M., Haas, D. M., McCullough, W., McNamara, M. F. & Olsen, C. (2006). The first 48 hours: Comparing 12-hour and 24-hour betamethasone dosing when preterm deliveries occur rapidly. *The Journal of Maternal-Fetal & Neonatal Medicine*, *19*(6), 365–369. https://doi.org/10.1080/14767050600715873

Hack, M., Flannery, D. J., Schluchter, M., Cartar, L., Borawski, E. & Klein, N. (2002). Outcomes in young adulthood for very-low-birth-weight infants. *New England Journal of Medicine*, *346*(3), 149–157.

Halliday, H. (2008). Surfactants: Past, present and future. *Journal of perinatology*, *28*(1), S47–S56.

Hartnett, M. E. & Lane, R. H. (2013). Effects of oxygen on the development and severity of retinopathy of prematurity. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, *17*(3), 229–234.

Helsedirektoratet. (2007). Faglige retningslinjer for oppfølging av for tidlig fødte barn.

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. & Welch, V. (Eds.). (2023). *Cochrane handbook for systematic reviews of interventions* (version 6.4 (updated August 2023)). Cochrane.

Hrabok, M., Brooks, B. L., Fay-McClymont, T. B. & Sherman, E. M. (2014). Wechsler intelligence scale for children-(WISC-IV) short-form validity: A comparison study in pediatric epilepsy. *Child Neuropsychology*, *20*(1), 49–59.

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

Iai, Y., Shimakawa, S., Fukui, M., Okumura, T., Tsuda-Kitahara, H. & Ashida, A. (2022). A comparative analysis of children born with low birthweight and attention deficit hyperactivity disorder. *Pediatr Int*, *64*(1), e15298. https://doi.org/10.1111/ped.15298

Ionio, C., Lista, G., Veggiotti, P., Colombo, C., Ciuffo, G., Daniele, I., Landoni, M., Scelsa, B., Alfei, E. & Bova, S. (2022). Cognitive, behavioral and socioemotional development in a cohort of preterm infants at school age: A cross-sectional study. *Pediatric Reports*, *14*(1), 115–126. https://doi.org/doi:10.3390/pediatric14010017

Jansen, L., Peeters-Scholte, C., van den Berg-Huysmans, A. A., van Klink, J. M. M., Rijken, M., van Egmond-van Dam, J. C., Vermeiren, R. & Steggerda, S. J. (2021). Longitudinal follow-up of children born preterm: Neurodevelopment from 2 to 10 years of age. *Front Pediatr*, *9*, 674221. https://doi.org/10.3389/fped.2021.674221

Jensen, E. A. & Schmidt, B. (2014). Epidemiology of bronchopulmonary dysplasia. *Birth Defects Research Part A: Clinical and Molecular Teratology*, *100*(3), 145–157.

Jensen, R. B., Juul, A., Larsen, T., Mortensen, E. L. & Greisen, G. (2015). Cognitive ability in adolescents born small for gestational age: Associations with fetal growth velocity, head circumference and postnatal growth. *Early human development*, *91*(12), 755–760.

Jin, J. H., Yoon, S. W., Song, J., Kim, S. W. & Chung, H. J. (2020). Long-term cognitive, executive, and behavioral outcomes of moderate and late preterm at school age. *Clinical and Experimental Pediatrics*, *63*(6), 219–225. https://doi.org/10.3345/kjp.2019.00647

Johnson, K., Healy, E., Dooley, B., Kelly, S. & McNicholas, F. (2015). Children born with very low birth weight have difficulties with sustained attention, but not response inhibition. *European Child & Adolescent Psychiatry*, *24*, S188–S188.

Johnson, S. (2007). Cognitive and behavioural outcomes following very preterm birth. *Seminars in Fetal and Neonatal Medicine*, *12*(5), 363–373.

Joseph, R. M., Hooper, S. R., Heeren, T., Santos Jr, H. P., Frazier, J. A., Venuti, L., Foley, A., Rollins, C. K., Kuban, K. C., Fry, R. C. et al. (2022). Maternal social risk, gestational age at delivery, and cognitive outcomes among adolescents born extremely preterm. *Paediatric and Perinatal Epidemiology*, *36*(5), 654–664.

Kaufman, A. S. (1993). King WISC the third assumes the throne. *Journal of School Psychology*, *31*(2), 345–354.

Kaufman, A. S., Flanagan, D. P., Alfonso, V. C. & Mascolo, J. T. (2006). Test review: Wechsler intelligence scale for children, (WISC-IV). *Journal of psychoeducational assessment*, *24*(3), 278–295.

Kaufman, A. S., Long, S. W. & O'Neal, M. R. (1986). Topical review of the WISC-R for pediatric neuroclinicians. *Journal of child neurology*, *1*(2), 89–98.

Kaufman, A. S., Raiford, S. E. & Coalson, D. L. (2015). *Intelligent testing with the WISC-V*. John Wiley & Sons.

Kaul, F. Y., Johansson, M., Mansson, J., Stjernqvist, K., Farooqi, A., Serenius, F. & L, B. T. (2021). Cognitive profiles of extremely preterm children: Full-scale IQ hides strengths and weaknesses. *Acta Paediatr*, *110*(6), 1817–1826. https://doi.org/10.1111/apa.15776

Kerr-Wilson, C., Mackay, D., Smith, G. & Pell, J. (2012). Meta-analysis of the association between preterm delivery and intelligence. *Journal of public health*, *34*(2), 209–216.

Kim, E. S., Kim, E. K., Kim, S. Y., Song, I. G., Jung, Y. H., Shin, S. H., Kim, H. S., Kim, J. I., Kim, B. N. & Shin, M. S. (2021). Cognitive and behavioral outcomes of school-aged children born extremely preterm: A korean single-center study with long-term follow-up. *Journal of Korean Medical Science*, *36*(39), 14. https://doi.org/doi:10.3346/jkms.2021.36.e260

Kim, S. (2022). Different maternal age patterns of preterm birth: Interplay of race/ethnicity, chronic stress, and marital status. *Research in Nursing & Health*, *45*(2), 151–162.

Kinney, H. C. (2006). The near-term (late preterm) human brain and risk for periventricular leukomalacia: A review. *Seminars in perinatology*, *30*(2), 81–88.

Klingenberg, C., Guthe, H. J., Andresen, J. H. & Nissen, I. B. (2021). 5.12 surfaktantbehandling. https://www.helsebiblioteket.no/innhold/retningslinjer/pediatri/nyfodtmedisin-veiledende-prosedyrer-fra-norsk-barnelegeforening/5-lunge-og-respirasjon/5.12-surfaktantbehandling#-helsebiblioteket-innhold-retningslinjer-pediatri-nyfodtmedisin-veiledende-prosedyrer-fra-norsk-barnelegeforening-5-lunge-og-respirasjon-512-surfaktantbehandling

Klingenberg, C., Kaaresen, P. I., Songstad, N. T., Kaspersen, K.-H., Nordhov, S. M., Pettersen, Å. T., Hvingel, B., Mørkved, M. & Nervik, M. (2021). Initialbehandling av premature. https://www.helsebiblioteket.no/innhold/retningslinjer/pediatri/nyfodtmedisin-veiledende-prosedyrer-fra-norsk-barnelegeforening/2-initialbehandling-av-premature

Kochukhova, O., Fredriksson Kaul, Y., Johansson, M., Montgomery, C., Holmstrom, G., Strand Brodd, K. & Hellstrom-Westas, L. (2022). Antenatal steroids and neurodevelopment in 12-year-old children born extremely preterm. *Acta Paediatr*, *111*(2), 314–322. https://doi.org/10.1111/apa.16140

Koenen, K. C., Moffitt, T. E., Roberts, A. L., Martin, L. T., Kubzansky, L., Harrington, H., Poulton, R. & Caspi, A. (2009). Childhood IQ and adult mental disorders: A test of the cognitive reserve hypothesis. *American Journal of Psychiatry*, *166*(1), 50–57.

Kovachy, V. N., Adams, J. N., Tamaresis, J. S. & Feldman, H. M. (2015). Reading abilities in school-aged preterm children: A review and meta-analysis. *Developmental Medicine & Child Neurology*, *57*(5), 410–419.

Kroll, J., Karolis, V., Brittain, P. J., Tseng, C. E. J., Froudist-Walsh, S., Murray, R. M. & Nosarti, C. (2019). Systematic assessment of perinatal and socio-demographic factors associated with IQ from childhood to adult life following very preterm birth. *Intelligence*, *77*, 7. https://doi.org/ARTN10140110.1016/j.intell.2019.101401

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. *Front Psychol*, *5*, 446. https://doi.org/10.3389/fpsyg.2014.00446

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J. & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of internal medicine*, *151*(4), W–65.

Linden, M. A., Cepeda, I. L., Synnes, A. & Grunau, R. E. (2015). Stress in parents of children born very preterm is predicted by child externalising behaviour and parent coping at age 7 years. *Archives of Disease in Childhood*, *100*(6), 554–558. https://doi.org/doi:10.1136/archdischild-2014-307390

Linsell, L., Johnson, S., Wolke, D., O'Reilly, H., Morris, J. K., Kurinczuk, J. J. & Marlow, N. (2018). Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: A prospective, population-based cohort study. *Archives of disease in childhood*, *103*(4), 363–370.

Lode-Kolz, K., Hermansson, C., Linnér, A., Klemming, S., Hetland, H. B., Bergman, N., Lilliesköld, S., Pike, H. M., Westrup, B., Jonas, W. et al. (2023). Immediate skin-to-skin contact after birth ensures stable thermoregulation in very preterm infants in high-resource settings. *Acta Paediatrica*, *112*(5), 934–941.

Løhaugen, G. C., Antonsen, I., Haberg, A., Gramstad, A., Vik, T., Brubakk, A. M. & Skranes, J. (2011). Computerized working memory training improves function in adolescents born at extremely low birth weight. *J Pediatr*, *158*(4), 555–561. https://doi.org/10.1016/j.jpeds.2010.09.060

Lynch, C. D. & Zhang, J. (2007). The research implications of the selection of a gestational age estimation method. *Paediatric and perinatal epidemiology*, *21*, 86–96.

Markestad, T. & Halvorsen, B. *Faglige retningslinjer for oppfølging av for tidlig fødte barn*. 2007.

Marret, S., Marchand-Martin, L., Picaud, J.-C., Hascoët, J.-M., Arnaud, C., Rozé, J.-C., Truffert, P., Larroque, B., Kaminski, M., Ancel, P.-Y. et al. (2013). Brain injury in very preterm children and neurosensory and cognitive disabilities during childhood: The EPIPAGE cohort study. *PloS one*, *8*(5), e62683.

McBryde, M., Fitzallen, G. C., Liley, H. G., Taylor, H. G. & Bora, S. (2020). Academic outcomes of school-aged children born preterm: A systematic review and meta-analysis. *JAMA network open*, *3*(4), e202027–e202027.

McNicholas, F., Healy, E., White, M., Sherdian-Pereira, M., O'Connor, N., Coakley, S. & Dooley, B. (2014). Medical, cognitive and academic outcomes of very low birth weight infants at age 10-14 years in Ireland. *Irish Journal of Medical Science*, *183*(4), 525–532. https://doi.org/doi:10.1007/s11845-013-1040-9

Medley, N., Poljak, B., Mammarella, S. & Alfirevic, Z. (2018). Clinical guidelines for prevention and management of preterm birth: A systematic review. *BJOG: An International Journal of Obstetrics & Gynaecology*, *125*(11), 1361–1369.

Michelsen, T. M., Bergøy, Ø., Ellingsen, L., Klingenberg, C., Lang, A., Morken, N.-H., Salvesen, K. Å., Sjøborg, K., Kalnes, R. S. & Tingleff, T. (2020). Initialbehandling av premature. https://www.legeforeningen.no/foreningsledd/fagmed/norsk-gynekologisk-forening/veiledere/veileder-i-fodselshjelp/preterm-fodsel/

Modi, N., Lewis, H., Al-Naqeeb, N., Ajayi-Obe, M., Doré, C. J. & Rutherford, M. (2001). The effects of repeated antenatal glucocorticoid therapy on the developing brain. *Pediatric research*, *50*(5), 581–585.

Moreira, R. S., Magalhães, L. C. & Alves, C. R. (2014). Effect of preterm birth on motor development, behavior, and school performance of school-age children: A systematic review. *Jornal de pediatria*, *90*(2), 119–134.

Mu, S. C., Tsou, K. S., Hsu, C. H., Fang, L. J., Jeng, S. F., Chang, C. H. & Tsou, K. I. (2008). Cognitive development at age 8 years in very low birth weight children in Taiwan. *Journal of the Formosan Medical Association*, *107*(12), 915–920. https://doi.org/10.1016/S0929-6646(09)60014-0

Mulder, H., Pitchford, N. J. & Marlow, N. (2011). Processing speed mediates executive function difficulties in very preterm children in middle childhood. *Journal of the International Neuropsychological Society*, *17*(3), 445–454. https://doi.org/10.1017/s1355617711000373

Nagy, A., Kalmár, M., Beke, A. M., Gráf, R. & Horváth, E. (2022). Intelligence and executive function of school-age preterm children in function of birth weight and perinatal complication. *Applied Neuropsychology: Child*, *11*(3), 400–411.

Ng, E., Taddio, A. & Ohlsson, A. (2017). Intravenous midazolam infusion for sedation of infants in the neonatal intensive care unit. *Cochrane Database of Systematic Reviews*, (1).

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F. & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American psychologist*, *67*(2), 130.

Norman, M., Hallberg, B., Abrahamsson, T., Björklund, L. J., Domellöf, M., Farooqi, A., Bruun, C. F., Gadsbøll, C., Hellström-Westas, L., Ingemansson, F. et al. (2019). Association between year of birth and 1-year survival among extremely preterm infants in sweden during 2004-2007 and 2014-2016. *Jama*, *321*(12), 1188–1199.

Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*, *5*, 1–10.

Oza, S., Lawn, J. E., Hogan, D. R., Mathers, C. & Cousens, S. N. (2014). Neonatal cause-of-death estimates for the early and late neonatal periods for 194 countries: 2000–2013. *Bulletin of the World Health Organization*, *93*, 19–28.

Perin, J., Mulick, A., Yeung, D., Villavicencio, F., Lopez, G., Strong, K. L., Prieto-Merino, D., Cousens, S., Black, R. E. & Liu, L. (2022). Global, regional, and national causes of under-5 mortality in 2000–19: An updated systematic analysis with implications

for the sustainable development goals. *The Lancet Child & Adolescent Health*, *6*(2), 106–115.

Petrill, S. A. & Wilkerson, B. (2000). Intelligence and achievement: A behavioral genetic perspective. *Educational Psychology Review*, *12*, 185–199.

Piaget, J. (2005). *The psychology of intelligence*. Routledge.

Pinto-Martin, J., Whitaker, A., Feldman, J., Cnaan, A., Zhao, H., Bloch, J. R., McCulloch, D. & Paneth, N. (2004). Special education services and school performance in a regional cohort of low-birthweight infants at age nine. *1*(2), 120–9.

Potharst, E. S., Houtzager, B. A., Van Sonderen, L., Tamminga, P., Kok, J. H., Last, B. F. & Van Wassenaer, A. G. (2012). Prediction of cognitive abilities at the age of 5 years using developmental follow-up assessments at the age of 2 and 3 years in very preterm children. *Developmental Medicine & Child Neurology*, *54*(3), 240–246.

Pritchard, V. E., Bora, S., Austin, N. C., Levin, K. J. & Woodward, L. J. (2014). Identifying very preterm children at educational risk using a school readiness framework. *Pediatrics*, *134*(3), e825–e832.

Rees, P., Callan, C., Chadda, K. R., Diviney, J., Harnden, F., Gardiner, J., Battersby, C., Gale, C. & Sutcliffe, A. (2023). Childhood outcomes after low-grade intraventricular haemorrhage: A systematic review and meta-analysis. *Developmental Medicine & Child Neurology*.

Rickards, A. L., Kelly, E. A., Doyle, L. W. & Callanan, C. (2001). Cognition, academic progress, behavior and self-concept at 14 years of very low birth weight children. *J Dev Behav Pediatr*, *22*(1), 11–8. https://doi.org/10.1097/00004703-200102000-00002

Rimol, L. M., Botellero, V. L., Bjuland, K. J., Løhaugen, G. C., Lydersen, S., Evensen, K. A. I., Brubakk, A.-M., Eikenes, L., Indredavik, M. S., Martinussen, M. et al. (2019). Reduced white matter fractional anisotropy mediates cortical thickening in adults born preterm with very low birthweight. *Neuroimage*, *188*, 217–227.

Roberts, D., Brown, J., Medley, N. & Dalziel, S. R. (2017). Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. *Cochrane database of systematic reviews*, (3).

Roberts, G., Lim, J., Doyle, L. W. & Anderson, P. J. (2011). High rates of school readiness difficulties at 5 years of age in very preterm infants compared with term controls. *Journal of developmental & behavioral pediatrics*, *32*(2), 117–124.

Roberts, R. M., George, W. M., Cole, C., Marshall, P., Ellison, V. & Fabel, H. (2013). The effect of age-correction on IQ scores among school-aged children born preterm. *Australian Journal of Educational & Developmental Psychology*, *13*, 1–15.

Rose, S. A., Feldman, J. F., Jankowski, J. J. & Van Rossem, R. (2011). Basic information processing abilities at 11 years account for deficits in IQ associated with preterm birth. *Intelligence*, *39*(4), 198–209. https://doi.org/10.1016/j.intell.2011.03.003

Ruiz-Peláez, J. G., Charpak, N. & Cuervo, L. G. (2004). Kangaroo mother care, an example to follow from developing countries. *Bmj*, *329*(7475), 1179–1181.

Saccone, G., Gragnano, E., Ilardi, B., Marrone, V., Strina, I., Venturella, R., Berghella, V. & Zullo, F. (2022). Maternal and perinatal complications according to maternal age: A systematic review and meta-analysis. *International Journal of Gynecology & Obstetrics*, *159*(1), 43–55.

Saigal, S. & Doyle, L. W. (2008). An overview of mortality and sequelae of preterm birth from infancy to adulthood. *The Lancet*, *371*(9608), 261–269.

Samuelsson, S., Finnström, O., Flodmark, O., Gäddlin, P.-O., Leijon, I. & Wadsby, M. (2006). A longitudinal study of reading skills among very-low-birthweight children: Is there a catch-up? *Journal of Pediatric Psychology*, *31*(9), 967–977.

Sarda, S. P., Sarri, G. & Siffel, C. (2021). Global prevalence of long-term neurodevelopmental impairment following extremely preterm birth: A systematic literature review. *Journal of International Medical Research*, *49*(7).

Schneider, W., Niklas, F. & Schmiedeler, S. (2014). Intellectual development from early childhood to early adulthood: The impact of early IQ differences on stability and change over time. *Learning and Individual Differences*, *32*, 156–162.

Seashore, H., Wesman, A. & Doppelt, J. (1950). The standardization of the Wechsler Intelligence Scale for Children. *Journal of Consulting Psychology*, *14*(2).

Sentenac, M., Chaimani, A., Twilhaar, S., Benhammou, V., Johnson, S., Morgan, A. & Zeitlin, J. (2022). The challenges of heterogeneity in gestational age and birthweight inclusion criteria for research synthesis on very preterm birth and childhood cognition: An umbrella review and meta-regression analysis. *Paediatric and perinatal epidemiology*, *36*(5), 717–725.

Seri, I. & Evans, J. (2008). Limits of viability: Definition of the gray zone. *Journal of Perinatology*, *28*(1), S4–S8.

Shah, N. R. & Bracken, M. B. (2000). A systematic review and meta-analysis of prospective studies on the association between maternal cigarette smoking and preterm delivery. *American journal of obstetrics and gynecology*, *182*(2), 465–472.

Simons, S. H., van Dijk, M., Anand, K. S., Roofthooft, D., van Lingen, R. A. & Tibboel, D. (2003). Do we still hurt newborn babies?: A prospective study of procedural pain and analgesia in neonates. *Archives of pediatrics & adolescent medicine*, *157*(11), 1058–1064.

Sistiaga, A., Garmendia, J., Aliri, J., Marti, I. & Labayru, G. (2021). A validated WISC-V short-form to estimate intellectual functioning in very preterm children at early school age. *Frontiers in Psychology*, *12*, 789124.

Sivanandan, S. & Sankar, M. J. (2023). Kangaroo mother care for preterm or low birth weight infants: A systematic review and meta-analysis. *BMJ Global Health*, *8*(6).

Smith, G. C., Gutovich, J., Smyser, C., Pineda, R., Newnham, C., Tjoeng, T. H., Vavasseur, C., Wallendorf, M., Neil, J. & Inder, T. (2011). Neonatal intensive care unit stress is associated with brain development in preterm infants. *Annals of neurology*, *70*(4), 541–549.

Sølsnes, A. E., Grunewaldt, K. H., Bjuland, K. J., Stavnes, E. M., Bastholm, I. A., Aanes, S., Østgård, H. F., Håberg, A., Løhaugen, G. C., Skranes, J. et al. (2015). Cortical morphometry and IQ in vlbw children without cerebral palsy born in 2003–2007. *NeuroImage: Clinical*, *8*, 193–201.

Squarza, C., Picciolini, O., Gardon, L., Ravasi, M., Gianni, M. L., Porro, M., Bonzini, M., Gangi, S. & Mosca, F. (2017). Seven years cognitive functioning and early assessment in extremely low birth weight children. *Front Psychol*, *8*, 1257. https://doi.org/10.3389/fpsyg.2017.01257

Sterne, J. A., Egger, M. & Smith, G. D. (2001). Investigating and dealing with publication and other biases in meta-analysis. *Bmj*, *323*(7304), 101–105.

Stoll, B. J., Hansen, N. I., Bell, E. F., Walsh, M. C., Carlo, W. A., Shankaran, S., Laptook, A. R., Sánchez, P. J., Van Meurs, K. P., Wyckoff, M. et al. (2015). Trends in care practices, morbidity, and mortality of extremely preterm neonates, 1993-2012. *Jama*, *314*(10), 1039–1051.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, *35*(5), 401–426.

Sweet, D. G., Carnielli, V. P., Greisen, G., Hallman, M., Klebermass-Schrehof, K., Ozek, E., Te Pas, A., Plavka, R., Roehr, C. C., Saugstad, O. D. et al. (2023). European consensus guidelines on the management of respiratory distress syndrome: 2022 update. *Neonatology*, *120*(1), 3–23.

Takayanagi, T., Egashira, M., Yamaguchi, T., Murata, N., Yokota, G., Matsuo, K., Ogata, T., Egashira, T., Iwanaga, M. & Mizukami, T. (2013). Cognitive outcome of very-low-birthweight infants at 6 years of age. *Pediatrics international*, *55*(5), 594–598.

Tam, E. W., Chau, V., Ferriero, D. M., Barkovich, A. J., Poskitt, K. J., Studholme, C., Fok, E. D.-Y., Grunau, R. E., Glidden, D. V. & Miller, S. P. (2011). Preterm cerebellar growth impairment after postnatal exposure to glucocorticoids. *Science Translational Medicine*, *3*(105).

Tinelli, F., Anobile, G., Gori, M., Aagten-Murphy, D., Bartoli, M., Burr, D. C., Cioni, G. & Concetta Morrone, M. (2015). Time, number and attention in very low birth weight children. *Neuropsychologia*, *73*, 60–9. https://doi.org/10.1016/j.neuropsychologia.2015.04.016

Tommiska, V., Lano, A., Kleemola, P., Klenberg, L., Lehtonen, L., Lopponen, T., Olsen, P., Tammela, O. & Fellman, V. (2020). Analysis of neurodevelopmental outcomes of preadolescents born with extremely low weight revealed impairments in multiple

developmental domains despite absence of cognitive impairment. *Health Sci Rep*, *3*(3), e180. https://doi.org/10.1002/hsr2.180

Tosello, B., Méziane, S., Resseguier, N., Marret, S., Cambonie, G., Zahed, M., Brévaut-Malaty, V., Beltran Anzola, A. & Gire, C. (2021). The neurobehavioral phenotype of school-aged, very prematurely born children with no serious neurological sequelae: A quality of life predictor. *Children*, *8*(11), 943.

Twilhaar, E. S., Wade, R. M., De Kieviet, J. F., Van Goudoever, J. B., Van Elburg, R. M. & Oosterlaan, J. (2018). Cognitive outcomes of children born extremely or very preterm since the 1990s and associated risk factors: A meta-analysis and meta-regression. *JAMA pediatrics*, *172*(4), 361–367.

Upadhyay, R. P., Naik, G., Choudhary, T. S., Chowdhury, R., Taneja, S., Bhandari, N., Martines, J. C., Bahl, R. & Bhan, M. K. (2019). Cognitive and motor outcomes in children born low birth weight: A systematic review and meta-analysis of studies from south asia. *BMC pediatrics*, *19*, 1–15.

van Noort-van der Spek, I. L., Franken, M.-C. J. & Weisglas-Kuperus, N. (2012). Language functions in preterm-born children: A systematic review and meta-analysis. *Pediatrics*, *129*(4), 745–754.

van Ool, J. S., Hurks, P. P., Snoeijen-Schouwenaars, F. M., Tan, I. Y., Schelhaas, H. J., Klinkenberg, S., Aldenkamp, A. P. & Hendriksen, J. G. (2018). Accuracy of WISC-III and WAIS-IV short forms in patients with neurological disorders. *Developmental Neurorehabilitation*, *21*(2), 101–107.

van't Westende, C., Peeters-Scholte, C., Jansen, L., van Egmond-van Dam, J. C., Tannemaat, M. R., de Bruine, F. T., van den Berg-Huysmans, A. A., Geraedts, V. J., Gouw, A. A., Steggerda, S. J., Stam, C. J. & van de Pol, L. A. (2020). The degree of prematurity affects functional brain activity in preterm born children at school-age: An EEG study. *Early Hum Dev*, *148*, 105096. https://doi.org/10.1016/j.earlhumdev.2020.105096

van Veen, S., van Wassenaer-Leemhuis, A. G., Oosterlaan, J., van Kaam, A. H. & Aarnoudse-Moens, C. S. H. (2020). Eight-year-old very and extremely preterm children showed more difficulties in performance intelligence than verbal intelligence. *Acta Paediatr*, *109*(6), 1175–1183. https://doi.org/10.1111/apa.15095

van Veen, S., Aarnoudse-Moens, C. S., van Kaam, A. H., Oosterlaan, J. & van Wassenaer-Leemhuis, A. G. (2016). Consequences of correcting intelligence quotient for prematurity at age 5 years. *The Journal of pediatrics*, *173*, 90–95.

Vermeulen, K., van Beek, P. E., van der Horst, I. E., Pop, V. J. M., van Dam, M., Vugs, B. & Andriessen, P. (2022). Toddler motor performance and intelligence at school age in preterm born children: A longitudinal cohort study. *Early Hum Dev*, *166*, 105549. https://doi.org/10.1016/j.earlhumdev.2022.105549

Volpe, J. J. (2009). Brain injury in premature infants: A complex amalgam of destructive and developmental disturbances. *The Lancet Neurology*, *8*(1), 110–124.

Wechsler, D. (1991). *Wechsler intelligence scale for children third edition*. Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence scale for children, fourth edition*. Pearson.

Wechsler, D. (2014). *Wechsler intelligence scale for children, fifth edition*. Pearson.

Wendel, K., Klingenberg, C., Moen, A., Bentsen, M., Støen, R. & Songstad, N. T. (2023). 5.6 Bronkopulmonal dysplasi. https://www.helsebiblioteket.no/innhold/retningslinjer/pediatri/nyfodtmedisin-veiledende-prosedyrer-fra-norsk-barnelegeforening/5-lunge-og-respirasjon/5.6-bronkopulmonal-dysplasi#-helsebiblioteket-innhold-retningslinjer-pediatri-nyfodtmedisin-veiledende-prosedyrer-fra-norsk-barnelegeforening-5-lunge-og-respirasjon-56-bronkopulmonal-dysplasi

Wilcox, A. J. (2001). On the importance—and the unimportance—of birthweight. *International journal of epidemiology*, *30*(6), 1233–1241.

Wilson-Ching, M., Pascoe, L., Doyle, L. W. & Anderson, P. J. (2014). Effects of correcting for prematurity on cognitive test scores in childhood. *Journal of paediatrics and child health*, *50*(3), 182–188.

Wolke, D., Söhne, B., Ohrt, B. & Riegel, K. (1995). Follow-up of preterm children: Important to document dropouts. *The Lancet*, *345*(8947), 447.

Wong, H. S., Santhakumaran, S., Cowan, F. M. & Modi, N. (2016). Developmental assessments in preterm children: A meta-analysis. *Pediatrics*, *138*(2).

World Health Organization. (2003). *Kangaroo mother care: A practical guide*. World Health Organization. https://www.who.int/publications/i/item/9241590351

World Health Organization. (2021). Immediate "kangaroo mother care" and survival of infants with low birth weight. *New England Journal of Medicine*, *384*(21), 2028–2038. https://doi.org/10.1056/NEJMoa2026486

World Health Organization. (2022). *WHO recommendations for care of the preterm or low birth weight infant*. World Health Organization.

World Health Organization. (2023a). *Born too soon: Decade of action on preterm birth*. World Health Organization.

World Health Organization. (2023b). *Preterm birth*. Retrieved 6th September 2023, from https://www.who.int/news-room/fact-sheets/detail/preterm-birth

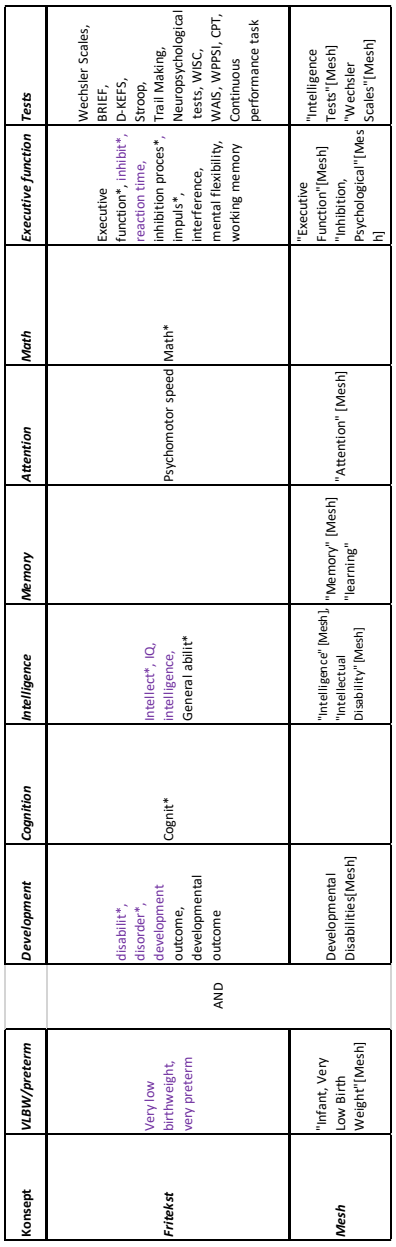
Zeng, Y., Ge, G., Lei, C. & Zhang, M. (2022). Beyond fetal immunity: A systematic review and meta-analysis of the association between antenatal corticosteroids and retinopathy of prematurity. *Frontiers in Pharmacology*, *13*, 759–742.

Appendix

A Table of search concepts

Figure A.1

Table of concepts included in the search. Words in purple were limited to title or abstract



Konsept	VLBW/preterm		Development	Cognition	Intelligence	Memory	Attention	Math	Executive function	Tests
Fritekst	Very low birthweight, very preterm	AND	disabilit*, disorder*, development outcome, developmental outcome	Cognit*	Intellect*, IQ, intelligence, General abilit*		Psychomotor speed	Math*	Executive function*, inhibit*, reaction time, inhibition proces*, impuls*, interference, mental flexibility, working memory	Wechsler Scales, BRIEF, D-KEFS, Stroop, Trail Making, Neuropsychological tests, WISC, WAIS, WPPSI, CPT, Continuous performance task
Mesh	"Infant, Very Low Birth Weight"[Mesh]		Developmental Disabilities[Mesh]		"Intelligence" [Mesh], "Intellectual Disability" [Mesh]	"Memory" [Mesh], "learning"	"Attention" [Mesh]		"Executive Function"[Mesh] "Inhibition, Psychological"[Mesh]	"Intelligence Tests"[Mesh] "Wechsler Scales"[Mesh]

B Search log

Premature og kognitiv funksjon – Søk utført 28-09-22

Til stede: Ingrid Remøy, Magnus Rom Jensen, Lisbeth Jahren

Resultater søk:

PsycInfo (Ovid): 8020
Medline (Ovid): 4014
Web of Science: 7362

Samlet resultat fra søk: 19396

P-Populasjon	**AND**	1-Development	2-Cognition	3-Intelligence	4-Memory	5-Attention	6-Math/Acad	7-Executive function	8-Tests

PsycInfo Ovid Søkeoppsett [ti=title, ab=abstract, id=key concepts]

1. ("Low birth weight" or "<28 weeks" or "<32 weeks" or "<37 weeks" or LBW or VLBW or ELBW or EPT or VPT).ti,ab,id.
2. ((Premature or Preterm) adj10 (Bab* or Infant* or Birth or Child* or Born)).ti,ab,id.
3. Premature Birth/
4. or/1-3
5. (Disabilit* or Disorder* or "Development* outcome").ti,ab,id.
6. Developmental Disabilities/
7. Cogniti*.ti,ab,id.
8. Cognitive Ability/
9. Cognition/
10. Cognitive Impairment/
11. (Intelligence or Intellect* or IQ or "General abilit*").ti,ab,id.
12. Intelligence/
13. exp Intellectual Development Disorder/
14. Intellectual Development/
15. Memory.ti,ab,id.
16. Learning.ti,ab,id.
17. Memory/
18. Learning/

19. Learning Ability/
20. Learning Disabilities/
21. Learning Disorders/
22. Psychomotor*.ti,ab,id.
23. Attention/
24. Perceptual Motor Processes/
25. Psychomotor Development/
26. Reaction Time/
27. (Math* or Academic*).ti,ab,id.
28. Mathematical Ability/
29. Mathematics Achievement/
30. Academic Achievement/
31. ("Executive function*" or Inhibit* or "Reaction time" or Impuls* or Interference* or Flexibilit*).ti,ab,id.
32. Executive Function/
33. Short Term Memory/

34. Response Inhibition/
35. (Wechsler adj3 (Scale* or Test*)).ti,ab,id.
36. "Neuropsychological test*".ti,ab,id.
37. ("Behavior Rating Inventory of Executive Function" or "BRIEF-2" or "BRIEF 2" or "BRIEF-P" or "BRIEF P" or D-KEFS or Stroop or "Trail Making" or WISC* or WAIS* or WPPSI* or CPT* or "Continuous performance*").ti,ab,id.
38. Intelligence Measures/
39. Neuropsychological Assessment/
40. Wechsler Adult Intelligence Scale/
41. Wechsler Intelligence Scale for Children/
42. Wechsler Memory Scale/
43. or/5-42
44. 4 and 43
45. limit 44 to (peer reviewed journal and yr="2000 - 2023")

Medline Ovid Søkeoppsett [ti=title, ab=abstract, kw=keyword]

1. ("Low birth weight" or "<28 weeks" or "<32 weeks" or "<37 weeks" or LBW or VLBW or ELBW or EPT or VPT).ti,ab,kw.
2. ((Premature or Preterm) adj10 (Bab* or Infant* or Birth or Child* or Born)).ti,ab,kw.
3. Premature Birth/
4. exp Infant, Premature/

5. exp Infant, Very Low Birth Weight/
6. or/1-5
7. ((Disabilit* or Disorder* or Development*) adj3 (Intellect* or Cogniti* or Neuropsychologic*)).ti,ab,kw.
8. Cogniti*.ti,ab,kw.
9. Cognition Disorders/
10. Cognitive Dysfunction/
11. (Intelligence or Intellect* or IQ or "General abilit*").ti,ab,kw.
12. Intelligence/
13. Intellectual Disability/
14. ("Working memory" or "Short Term Memory").ti,ab,kw.
15. (Memory adj5 (Abilit* or Disabilit* or Disorder* or Problem* or Difficult* or Defici*)).ti,ab,kw.
16. (Learning adj3 (Abilit* or Disabilit* or Disorder* or Problem* or Difficult* or Defici*)).ti,ab,kw.
17. Memory/
18. Learning/
19. Learning Disabilities/
20. Psychomotor*.ti,ab,kw.
21. Attention/
22. Psychomotor Performance/
23. Reaction Time/
24. (Math* or Academic*).ti,ab,kw.
25. Mathematics/
26. exp Academic Performance/
27. ("Executive function*" or "Response inhibit*" or "Reaction time" or Impuls* or Interference* or Flexibilit*).ti,ab,kw.
28. Executive Function/
29. Memory, Short-Term/
30. exp Inhibition, Psychological/
31. (Wechsler adj3 (Scale* or Test*)).ti,ab,kw.
32. (Neuropsychological adj (test* or assessment*)).ti,ab,kw.
33. ("Behavior Rating Inventory of Executive Function" or "BRIEF-2" or "BRIEF 2" or "BRIEF-P" or "BRIEF P" or D-KEFS or Stroop or "Trail Making" or WISC* or WAIS* or WPPSI* or CPT* or "Continuous performance*").ti,ab,kw.
34. Intelligence Tests/
35. Neuropsychological Tests/
36. Wechsler Scales/
37. Wechsler Memory Scale/

38. Stroop Test/
39. Trail Making Test/
40. or/7-39
41. 6 and 40
42. limit 41 to yr="2000 - 2023"

Web of Science Søkestreng

((TS=("Low birth weight" or "<28 weeks" or "<32 weeks" or "<37 weeks" or LBW or VLBW or ELBW or EPT or VPT or ((Premature or Preterm) NEAR/10 (Bab* or Infant* or Birth or Child* or Born)))) AND TS=((Disabilit* or Disorder* or Development*) NEAR/1 (Intellect* or Cogniti* or Neuropsychologic*) or Cognit* NEAR/2 (Abilit* or Function* or Dysfunction* or Impairment* or Disabilit* or Disorder* or Problem* or Difficult* or Defici*) or Intelligence or Intellect* or IQ or Memory NEAR/5 (Abilit* or Disabilit* or Disorder* or Problem* or Difficult* or Defici*) or Learning NEAR/3 (Abilit* or Disabilit* or Disorder* or Problem* or Difficult* or Defici*) or Psychomotor* or (Mathemetic* or Academic*) NEAR/3 (Abilit* or Aachievement*) or "Executive function*" or "response inhibit*" or "Reaction time" or Impuls* or Interference* or Flexibilit* or "Working memory" or "Short Term Memory" or (Wechsler NEAR/3 (Scale* or Test*)) or (Neuropsychological NEAR/3 (Test* or Assessment*)) or BRIEF* or "D-KEFS*" or Stroop* or "Trail Making*" or WISC* or WAIS* or WPPSI* or CPT* or "Continuous performance*")) AND PY=(2000-2023)

C Funnel Plots

Figure C.1

Funnel plot for <1000g

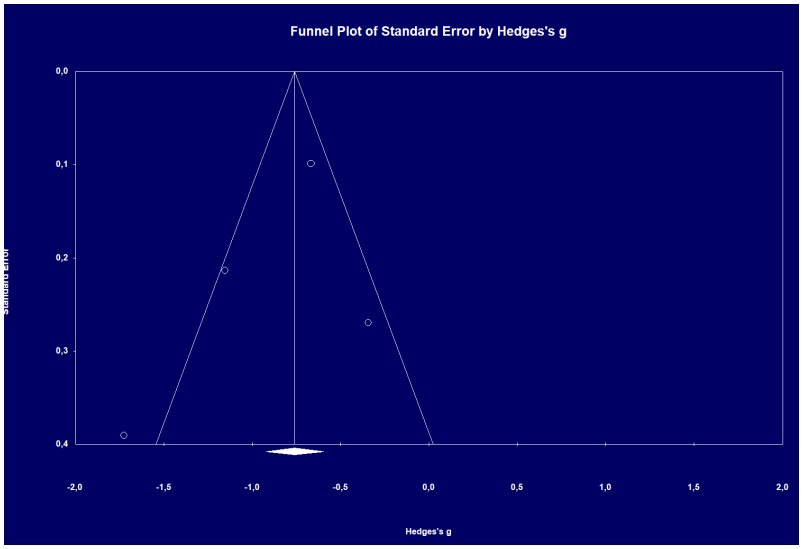


Figure C.2

Funnel plot for <1500g

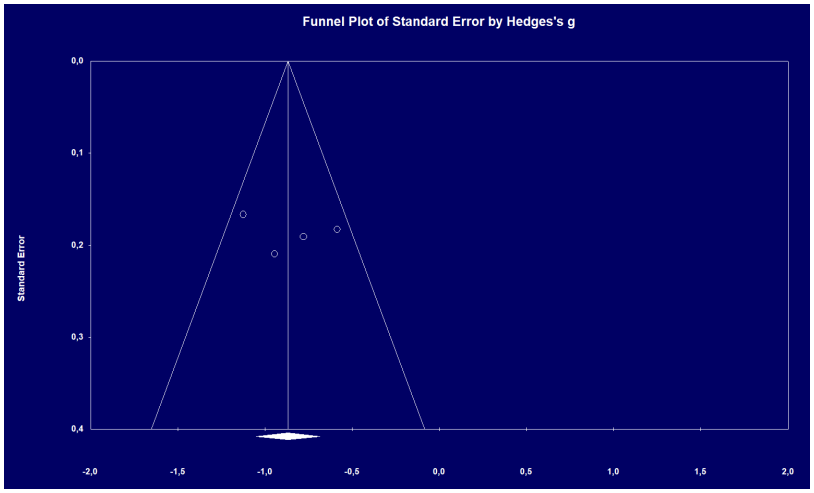
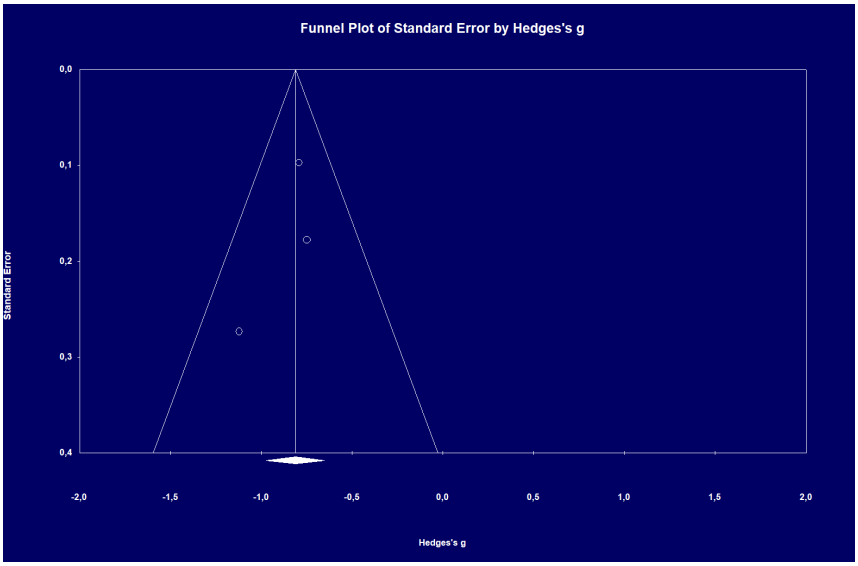
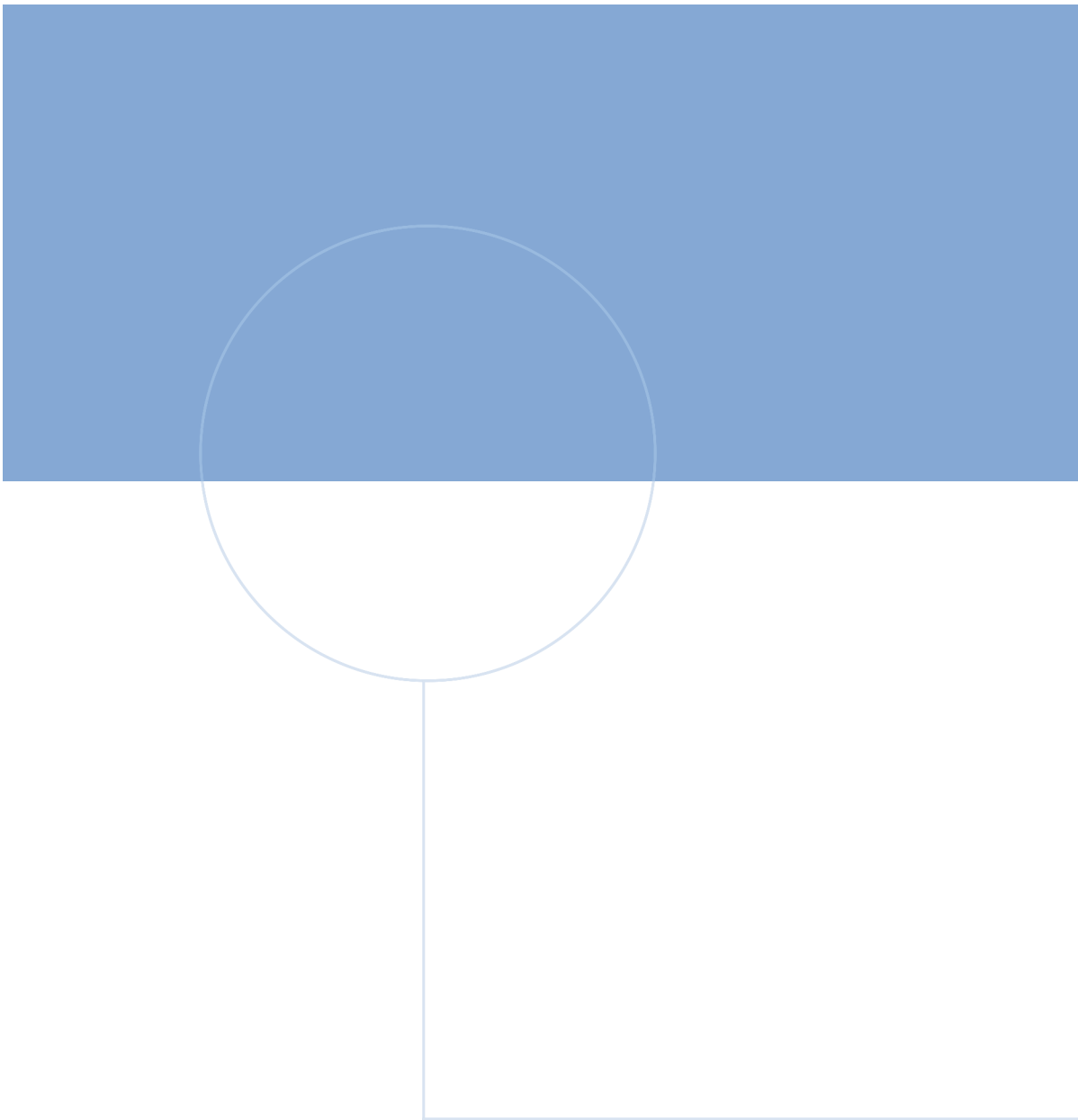


Figure C.3

Funnel plot for GA <32 weeks





NTNU

Kunnskap for en bedre verden