National Institute for Health and Care Excellence

V 2.0

Developmental follow-up of children and young people born preterm

Full Guideline

NICE Guideline

Methods, evidence and recommendations

February 2017

Draft for Consultation

Developed by the National Guideline Alliance, hosted by the Royal College of Obstetricians and Gynaecologists

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Contents

Introduction

This guideline focuses on the specialist developmental support and surveillance needed for the early identification of developmental problems and disorders in children born preterm.

The proportion of babies born preterm in the UK, defined as birth before 37 weeks' gestation, has remained steady for several years at 7.4%. In 2014 this amounted to 48,985 from a total of 656,957 live births, of which 2438 (5% of preterm births and 0.4% of all births) were before 28 weeks' gestation.

Preterm birth is associated with an increased risk of developmental problems and disorders. These include developmental challenges, physical, sensory, cognitive and learning disorders, and emotional and behavioural problems. These may extend into adolescence and, in some cases, be lifelong. In particular, the risk and prevalence of impairments that affect educational attainment rise sharply in children born before 28 weeks' gestation. Although most major disorders are detectable in the first 2 years of life, several developmental disorders and problems, particularly those that have an impact on the child's ability to participate and on their educational attainment, may not be apparent until they are older.

Identifying developmental problems and disorders in all children (born preterm or at term) is currently through the Healthy Child Programme, which incorporates nationally approved population screening programmes recommended by Public Health England. This guideline aims to improve the identification of developmental problems and disorders in children born preterm by setting standards for follow-up. This is expected to improve outcomes for these children by reducing variation in follow-up and enabling benchmarking of neonatal care. Developmental surveillance up to and at 2 years (corrected age) is recommended for identifying major problems and disorders. A later developmental assessment for children at high risk aims to identify problems that are more apparent at school age, with a view to supporting education plans for the child.

1 Guideline summary

1.1 Group membership and NGA staff

Table 1: Committee members

Name	Role
Gillian Baird	Chair, Consultant Developmental Paediatrician - Guy's and St Thomas' NHS trust
Jennifer Baulcomb	Educational Psychologist - Evalina London Children's hospital, Guy's and St Thomas' NHS trust
Joe Fawke	Consultant Neonatologist - University Hospitals of Leicester NHS trust
Joanna Goodman	Lay member
Celia Harding	Speech and Language Therapist - Royal Free Hospital and Barnet Hospital
Sarra Hoy	Lay member
Betty Hutchon	Head Paediatric Occupational Therapist - Royal Free Hospital
Sally Jary	Clinical Specialist Paediatric Physiotherapist - Bristol Royal Hospital
Samantha Johnson	Developmental Psychologist and Senior Research Fellow - Leicester University
Nashwa Matta	Associate specialist in neonatology and child development - Princess Royal maternity hospital
Nicola O'Connor	Lay member
Tilly Pillay	Neonatal Consultant - Staffordshire, Shropshire and Black Country Newborn Network
Claire Rohan	Consultant Paediatrician - Chase Farm and Barnet NHS Trust
Co-opted members	
Kristie Hill	Health Visitor/ Parenting Practitioner - Somerset Partnership NHS Trust
Neil Marlow	Professor of Neonatal Medicine - UCL EGA Institute for Women's Health
Jugnoo Rahi	Professor of Ophthalmic Epidemiology - Institute of Child Health and Institute of Ophthalmology UCL and Honorary Consultant Ophthalmologist - Great Ormond Street Hospital NHS Foundation Trust

Table 2: NGA staff

Name	Role
Alex Bates	Health Economist (until September 2016)
Vanessa Delgado Nunes	Guideline Lead (until November 2016)
Annabel Flint	Project Manager (until September 2016)
Yelan Guo	Systematic Reviewer (until October 2016)
Sadia Janjua	Systematic Reviewer
Maija Kallionen	Systematic Reviewer
Taryn Krause	Guideline Lead (from November 2016)
Stephen Murphy	Clinical Advisor (child health)
Sabrina Naqvi	Project Manager (from August 2016)
Matthew Prettyjohns	Health Economist (from September 2016)
Timothy Reeves	Information Scientist
Victoria Rowlands	Project Manager (from October 2016)

1.2.1 Developmental support and surveillance algorithm

Time	Developmental support and surveillance			
	Enhanced			Routine
Time	Children born before 28+0 weeks' gestation	Children born between 28^{+0} and 30^{+0} weeks' gestation	Children born between 30^{+0} and 36^{+6} weeks' gestation who are at increased risk of developmental problems or disorders[1]	All Children born before 37^{+0} weeks
Birth through 2 years (corrected age)	Enhanced developmental support from a single point of contact within the neonatal service, whom parents and carers can contact after discharge Tailored support provided using a range of approaches which may include face-to-face meetings, in the clinic or home, a telephone helpline, or electronic communication Minimum of 2 face-to-face follow-up visits to review any developmental concerns			Routine postnatal care and support as described in NICE guideline on postnatal care up to 8 weeks after birth Surveillance from the Healthy Child Programme
Age 2 (corrected age)	Developmental assessment: -checks for any developmental problems or disorders (including check for global developmental delay using the PARCA-R) -ensure checks of vision and hearing have been carried out in line with national recommendations			
Age 4	Developmental assessment should: -be conducted by professionals with appropriate skills -take into account information provided by parent or carers -include a review of previous assessments and information from all other relevant sources -include checks for developmental problems and disorders use: - the Strengths and Difficulties Questionaire (SDQ) to check for social, attentional, emotional and behavioural problems - as a minimum, the Wechsler Preschool and Primary Scales of Intelligence 4th Edition (WPPSI) test, including subscales for verbal comprehension, visual spatial skills, fluid reasoning, working memory and processing speed: - if the WPPSI is not suitable (for example, because of sensory or motor impairment), use a suitable alternative. -ensure that children born preterm who are having a 4-year developmental assessment have been offered orthoptic vision screening as recommended by the National Screening Committee	Surveillance from the Healthy Child Programme		

[1] Risk factors include: a brain lesion on neuroimaging likely to be associated with developmental problems or disorders (for example, grade 3 or 4 intraventricular haemorrhage or cystic periventricular leukomalacia), grade 2 or 3 hypoxic ischaemic encephalopathy, neonatal bacterial meningitis, herpes simplex encephalitis, . Consider providing enhanced developemental support for children who do not have any of the above risk factors but who are thought, using clinical judgement, to be at risk, taking into account the presence and severity of risk factors.

1.3 Recommendations

1. Be aware that children born preterm are at increased risk of developmental problems and disorders.
2. Be aware that for recommendations in this section:
    - for some developmental problems and disorders there was an absence of evidence about overall risk and prevalence in children born preterm, and some papers included specific gestational ages at birth from which the committee was unable to extrapolate to other gestational ages
    - for some developmental problems and disorders the evidence was underpowered to detect an effect
    - other gestational ages and other factors not listed here might also be associated with increased risk of developmental problems and disorders.
3. Be aware that children born preterm are at increased risk of cerebral palsy, and that:
    - the following are independent risk factors:
        - grade 3 or 4 intraventricular haemorrhage
        - cystic periventricular leukomalacia
        - neonatal sepsis
        - bronchopulmonary dysplasia for which mechanical ventilation was still needed at 36 weeks' postmenstrual age
        - antenatal steroids not given
        - postnatal steroids given to babies born before 32+0 weeks' gestation
    - prevalence increases with decreasing gestational age.

    See also the NICE guideline on cerebral palsy in children and young people under 25.
4. Be aware that children born preterm are at increased risk of motor problems, and that the following are independent risk factors:
    - brain lesions (for example, grade 3 or 4 intraventricular haemorrhage, periventricular leukomalacia, infarct)
    - necrotising enterocolitis that needed surgery
    - neonatal sepsis
    - severe retinopathy of prematurity.
5. Be aware that there is increased prevalence of developmental coordination disorder in children born preterm compared with the general population.
6. Be aware that children born preterm are at increased risk of intellectual disability, and that:
    - the following are independent risk factors:
        - grade 3 or 4 intraventricular haemorrhage

    - cystic periventricular leukomalacia
    - neonatal sepsis in babies born before 28+0 weeks' gestation
    - necrotising enterocolitis that needed surgery in babies born before 33+0 weeks' gestation
    - bronchopulmonary dysplasia for which mechanical ventilation was still needed at 36 weeks' postmenstrual age in babies born before 28+0 weeks' gestation
    - severe retinopathy of prematurity in babies born before 28+0 weeks' gestation
    - small for gestational age
    - postnatal steroids given to babies born before 33+0 weeks' gestation
    - mother from a low-income or disadvantaged background
  - prevalence increases with decreasing gestational age.

7. Be aware that children born preterm are at increased risk of having special educational needs, and that the following are independent risk factors:
   - brain lesions detected by ultrasound
   - male sex.
8. Be aware that children born preterm are at increased risk of low educational attainment at the end of the early years foundation stage and at key stage 1, and that:
   - prevalence of low educational attainment increases with decreasing gestational age
   - there is increased risk of low attainment for reading and numeracy in children born before 26+0 weeks' gestation
   - the following are independent risk factors for delayed numeracy in children born before 32+0 weeks' gestation:
     - intracranial haemorrhage
     - bronchopulmonary dysplasia for which mechanical ventilation was still needed at 36 weeks' postmenstrual age.
9. Be aware that children born before 33+0 weeks' gestation are at increased risk of symptoms of hyperactivity, impulsivity and particularly inattention at preschool and school ages.
10. Be aware that children born before 28+0 weeks' gestation are at increased risk of ADHD, and that male sex is an independent risk factor.
11. Be aware that children born before 28+0 weeks' gestation are at increased risk of symptoms of social communication impairment, which may suggest a problem in the autism spectrum.
12. Be aware that children born preterm are at increased risk of autism spectrum disorder, and that:
    - the following are independent risk factors:
      - intracranial haemorrhage in babies born before 34+0 weeks' gestation
      - male sex

- prevalence increases with decreasing gestational age.

13. Be aware that children born preterm are at increased risk of emotional and behavioural problems, particularly internalising behaviours and passivity, at preschool and primary school ages, and that the following are independent risk factors:
    - major brain lesions (for example, periventricular leukomalacia, parenchymal lesions)
    - mother with mental health problems
    - mother younger than 25 years
    - mother from a low-income or disadvantaged background.
14. Be aware that children born preterm are at increased risk of speech, language and communication problems and disorders, and that the following are independent risk factors for language disorder :
    - grade 3 or 4 intraventricular haemorrhage
    - cystic periventricular leukomalacia
    - male sex.
15. Be aware that children born preterm are at increased risk of oro-motor feeding problems, and that this increased risk persists until at least 6 years of age in children born before 26+0 weeks.
16. Be aware that children born preterm are at increased risk of sleep apnoea up to 6 years of age.
17. Be aware that the prevalence of visual impairment increases with decreasing gestational age in children born preterm, and that the following are independent risk factors:
    - grade 3 or 4 intraventricular haemorrhage with a shunt
    - neonatal sepsis in babies born before 33+0 weeks' gestation
    - retinopathy of prematurity requiring treatment.
18. Be aware that the prevalence of hearing impairment increases with decreasing gestational age in children born preterm, and that neonatal sepsis is an independent risk factor in babies born before 28+0 weeks' gestation.
19. Be aware that children born before 32+0 weeks' gestation are at increased risk of executive function problems at preschool and school ages.
20. Be aware that children born preterm are at increased risk of developmental problems, and that the following are independent risk factors:
    - small for gestational age
    - male sex
    - mother from a low-income or disadvantaged background
    - black, Asian or other minority ethnic group
    - multiple pregnancy.
21. Provide information about the risk and prevalence of developmental problems and disorders to parents or carers of preterm babies, and discuss this with them.

22. Provide information to parents or carers of preterm babies that is tailored to their individual circumstances, taking into account:
    - their child's potential developmental needs
    - their level of education
    - any social care needs they have
    - any cultural, spiritual or religious beliefs.
    - the need for consistency in information sharing among healthcare professionals
23. Follow the principles in the NICE guideline on patient experience in NHS services in relation to communication (including different formats and languages), information and shared decision-making.
24. Provide emotional and psychological support as needed to parents or carers of preterm babies.
25. Provide information to parents or carers of preterm babies about opportunities for peer support.
26. Before discharging a preterm baby:
    - agree a discharge plan with the parents or carers
    - ensure that the discharge plan includes clear information about any antenatal and perinatal risk factors for developmental problems and disorders (see section 4.3)
    - share the discharge plan with parents or carers and with primary and secondary healthcare teams.
27. Help parents or carers to gain the knowledge, skills and confidence they need to look after their baby at home and support the baby's developmental needs, taking into account that they are likely to be anxious about managing their baby's care after discharge. This may relate to:
    - interaction with the baby
    - managing feeding
    - patterns of sleeping
    - impact on day-to-day living, such as social isolation because of fear of infection.
28. Involve the social support networks (which may include partners, grandparents or other family members) of parents and carers of a baby born preterm when planning discharge and during follow-up.
29. Explain to parents and carers at the time of discharge that their child's developmental (corrected) age, which is calculated from their original due date (and not the date they were born), will be used for the first 2 years when assessing their functional and developmental skills (such as walking and talking).
30. Inform parents or carers of all preterm babies about the Healthy Child Programme, which includes national recommendations for all children about screening (for example, newborn hearing screening) and surveillance (including social, emotional, behavioural and language development).

31. Inform parents or carers about the routine postnatal care and support available as described in the NICE guideline on postnatal care up to 8 weeks after birth.

32. Healthcare professionals providing postnatal care and support in the community for babies born preterm should have the skills and knowledge to recognise and manage problems in these babies, including:
    - providing feeding support
    - addressing concerns about sleeping
    - facilitating interaction between the parents or carers and the baby.

33. Provide enhanced developmental support and surveillance by a multidisciplinary team (see section 5.2.3) up to 2 years (corrected age) for children born preterm who have, or are at increased risk of, developmental disorders or problems, based on the following criteria:
    - born before 30+0 weeks' gestation **or**
    - born between 30+0 and 36+6 weeks' gestation and has or had 1 or more of the following risk factors:
        - a brain lesion on neuroimaging likely to be associated with developmental problems or disorders (for example, grade 3 or 4 intraventricular haemorrhage or cystic periventricular leukomalacia)
        - grade 2 or 3 hypoxic ischaemic encephalopathy
        - neonatal bacterial meningitis
        - herpes simplex encephalitis in the neonatal period

34. Consider providing enhanced developmental support and surveillance by a multidisciplinary team (see section 5.2.3) up to 2 years (corrected age) for children born between 30^{+0} and 36^{+6} weeks' gestation who do not have any of the risk factors listed in section 5.2.1 but are thought, using clinical judgement, to be at increased risk of developmental problems or disorders in the first 2 years of life and taking into account the presence and severity of risk factors (see recommendations 3-20)

35. Inform parents or carers of preterm babies who meet the defined criteria about the arrangements for enhanced developmental support and surveillance for their child.

36. Provide parents or carers of a preterm baby having enhanced developmental support with a single point of contact within the neonatal service for outreach care after discharge.

37. Use a range of approaches when providing enhanced developmental support and tailor the support to take account of individual preferences and needs. Approaches may include:
    - face-to-face meetings, in clinics or in the home
    - a telephone helpline
    - electronic communication, for example by text message or email.

38. For all children born preterm who are having enhanced developmental surveillance, provide:
    - a minimum of 2 face-to-face follow-up visits in the first 2 years of life **and**

- a developmental assessment at 2 years (corrected age) (see recommendation 43).

39. At each visit for a child born preterm who is having enhanced developmental surveillance:
    - ensure that this is conducted by professionals with appropriate skills (see recommendation 50 and 51)
    - ask parents or carers whether they have any concerns about their child's development
    - include checks for developmental problems and disorders (See recommendation 40)
    - carefully assess and review any developmental concerns arising either from parent or carer report or at the visit itself
    - correct for gestational age up to 2 years (corrected) when assessing development
    - discuss any concerns with parents or carers
    - consider further investigation or referral if a developmental problem or disorder is suspected or present
    - refer the child to the appropriate local pathway if needed.
40. At each visit for a child born preterm who is having enhanced developmental surveillance up to 2 years (corrected age), and at the 4-year assessment (for children born before 28+0 weeks; see recommendation 46), check for signs and symptoms of problems and disorders as appropriate, such as:
    - cerebral palsy (see recommendation 3)
    - global developmental delay
    - autism spectrum disorder (See recommendation 12)
    - visual impairment
    - hearing impairment
    - feeding problems
    - sleep problems
    - speech, language and communication problems
    - motor problems
    - attention, impulsivity and hyperactivity
    - emotional and behavioural problems
    - executive functional problems
    - special educational needs
41. Recognise the following as possible early motor signs of cerebral palsy:
    - delayed motor milestones, such as late sitting, crawling or walking (correcting for gestational age)
    - unusual fidgety movements or other abnormalities of movement, including asymmetry or paucity of movement
    - abnormalities of tone, including hypotonia (floppiness) or spasticity (stiffness)

- persisting feeding difficulties.

42. For guidance on recognising signs and symptoms of possible autism spectrum disorder, see the NICE guideline on autism spectrum disorder in under 19s: recognition, referral and diagnosis.

43. Provide a developmental assessment at 2-years (corrected age) for children born preterm who are having enhanced developmental surveillance. This assessment should include:
    - all aspects listed in recommendation 39
    - at a minimum, use the Parent Report of Children's Abilities - Revised (PARCA-R) to identify if the child is at risk of global developmental delay, early intellectual disability or language problems:
    - if the PARCA-R is not suitable (for example, because of poor English language comprehension or the child being outside the validated age range of 22 to 26 months), use a suitable alternative.
    - ensuring that checks of vision and hearing have been carried out in line with national recommendations.

44. If findings from the developmental assessment at 2 years (corrected age) or 4 years (see recommendation 46), taking into account all of its components, suggest any developmental problems or disorders:
    - refer the child to an appropriate local pathway, which may involve child health and education services
    - share information with:
    - o parents or carers
    - o primary and secondary healthcare teams
    - ask parents or carers for permission to share information with:
    - o education services
    - o social care services as appropriate.

45. After the developmental assessment at 2 years (corrected age):
    - advise parents or carers of all children that their child should continue to be followed-up in the healthy child programme **and**
    - advise parents or carers of children born before 28+0 weeks' gestation that the child will also be offered a further developmental assessment at 4 years .

46. Provide a developmental assessment at 4 years for all children born before 28+0 weeks' gestation. This assessment should:
    - be conducted by professionals with appropriate skills (see recommendations 49 and 50)
    - take into account information provided by parent or carers (see recommendation 39)
    - include a review of previous assessments and information from all other relevant sources
    - include checks for developmental problems and disorders (see recommendation 40)
    - use:

- the Strengths and Difficulties Questionaire (SDQ) to check for social, attentional, emotional and behavioural problems
- as a minimum, the Wechsler Preschool and Primary Scales of Intelligence 4th Edition (WPPSI) test, including subscales for verbal comprehension, visual spatial skills, fluid reasoning, working memory and processing speed:
- if the WPPSI is not suitable (for example, because of sensory or motor impairment), use a suitable alternative.
- include ensuring that the child has been offered orthoptic vision screening as recommended by the National Screening Committee.

47. Provide a comprehensive summary of the child's strengths and difficulties, including any developmental problems and disorders, after the 4-year assessment that:
    - is in a format that is accessible to parents and carers
    - if needed, informs the development of a plan for intervention and support, including educational support.

    See also recommendation 44 about referral and information sharing

48. Enhanced developmental support and surveillance for children born preterm who meet the defined criteria (see recommendations 33,34,45) should:
    - be provided as an integral part of a neonatal service working together with local health services
    - empower parents and carers to be involved in decisions about their child's care
    - be delivered by a multidisciplinary team with the necessary skills (see recommendation 49)
    - record outcomes at specified time points for national audit (see section 5.2.4)
    - be monitored by checking adherence to the recommendations in this guideline, including follow-up rates and outcomes, as part of the routine provision of neonatal care by neonatal operational delivery networks and commissioners

49. Multidisciplinary teams delivering enhanced developmental support and surveillance for children born preterm should include professionals with knowledge and expertise in the following areas:
    - neonatal care
    - development of children born preterm, including developmental problems and disorders (see recommendation 40)
    - providing support in the community, for example for feeding problems
    - administering and interpreting results from questionnaires and standardised tests (such as the PARCA-R, SDQ and WPPSI)
    - collating information from a range of sources to facilitate decision making and writing reports
    - local care pathways, including Early Years education.

50. Multidisciplinary teams delivering enhanced developmental support and surveillance for children born preterm should include the following professionals:
    - for follow-up to 2 years (corrected age):
        - neonatologist or paediatrician with expertise in neonatal care
        - occupational therapist or physiotherapist
        - outreach nurse or nurse with expertise in neonatal care
    - for the assessment at 4 years (see recommendation 46):
        - clinical or educational psychologist
        - paediatrician with expertise in neurodevelopment.
51. Multidisciplinary teams delivering enhanced developmental support and surveillance for children born preterm should have access to the following professionals:
    - community nurse
    - occupational therapist
    - physiotherapist
    - paediatric neurologist
    - paediatrician with expertise in neurodevelopment
    - dietitian
    - speech and language therapist.
52. Record the following information, as applicable, in the National Neonatal Research Database for every child born preterm who has enhanced developmental surveillance:
    - whether the child had specialist neonatal care and details of discharge
    - at the assessment at 2 years (corrected age) (see recommendation 4343)
        - diagnosis of cerebral palsy
        - Gross Motor Function Classification System (GMFCS) score if cerebral palsy is present
        - PARCA-R score
        - epilepsy that is currently being treated
        - impairments of hearing, vision, speech and language, and motor skills
    - at the assessment at 4 years (see recommendation 46)
        - diagnosis of cerebral palsy
        - GMFCS score if cerebral palsy is present
        - WPPSI full scale IQ score, and subscale scores for verbal comprehension, visual spatial skills, fluid reasoning, working memory and processing speed
        - SDQ total difficulty score, subscale scores and impact score
        - any formal clinical diagnoses of a developmental disorder (for example, autism spectrum disorder)

- epilepsy that is currently being treated
- the presence of a hearing impairment, defined as profound deafness or impairment severe enough to need hearing aids or cochlear implant
- results of national orthoptic vision screening (see recommendation 46).

53. Record routine educational measures at key stage 2 (including special educational needs and disability [SEND]) on an operational delivery network-wide basis, to allow educational outcomes at 11 years to be linked to neonatal information.

1.4 Research recommendations

1. What support do parents and carers report was helpful to them in the care of children who were born preterm at the time of transfer to education services?
2. What is the accuracy of the parent-completed Parent Report of Children's Abilities-Revised (PARCA-R) questionnaire for predicting intellectual disability, language impairment and special educational needs at age 4 years for children born preterm?
3. What is the accuracy of the parent-completed Ages and Stages Questionnaire, 3rd edition (ASQ-3) for detecting concurrent intellectual disability and motor impairment between the ages of 2 years (corrected) and 4 years in children born preterm?
4. What is the accuracy of the parent-completed Strengths and Difficulties Questionnaire (SDQ) for predicting social, attentional, emotional and behavioural problems in children born before 28+0 weeks' gestation?
5. What is the accuracy of the Preschool Language Scales 5th edition (PLS-5), completed by parents together with a speech and language therapist, for detecting speech and language problems at 2 years (corrected age) in children born preterm?
6. What is the accuracy of a Wechsler Preschool and Primary Scale of Intelligence 4th Edition (WPPSI-IV) assessment at age 4 years for predicting later educational difficulties in children of primary school age who were born before 28^{+0} weeks' gestation?
7. Does enhanced developmental support and surveillance improve outcomes for the parents and carers of children born preterm?

1.5 Other versions of the guideline

NICE produce a number of versions of this guideline:

- The 'short guideline' lists the recommendations, context and recommendations for research.
- NICE Pathways brings together all connected NICE guidance.

1.6 Schedule for updating the guideline

Following publication, NICE will undertake a reviews at specified times to determine whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update. The review for update process is presented and in accordance with the NICE guidelines manual 2014.

2 Development of the guideline

2.1 What is a NICE guideline?

National Institute for Health and Care Excellence (NICE) guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the NHS – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of healthcare. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific review questions.

NICE clinical guidelines can:

- provide recommendations for the treatment and care of people by healthcare professionals
- be used to develop standards to assess the clinical practice of individual healthcare professionals
- be used in the education and training of healthcare professionals
- help patients to make informed decisions
- improve communication between patients and healthcare professionals.

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

- The guideline topic is referred to NICE from the Department of Health.
- Stakeholders register an interest in the guideline and are consulted throughout the development process.
- The scope is prepared by the National Guideline Alliance (NGA).
- The NGA establishes a Guideline Committee.
- A draft guideline is produced after the group assesses the available evidence and makes recommendations.
- There is a consultation on the draft guideline.
- The final guideline is produced.

2.2 Remit

NICE received the remit for this guideline from the Department of Health. It commissioned the NGA to produce the guideline.

The remit for this guideline is to develop a clinical guideline on the developmental follow-up of preterm babies.

2.3 Who developed this guideline?

A multidisciplinary guideline Committee comprising healthcare professionals and researchers as well as lay members developed this guideline (see Table 1).

The Committee met every 4 to 6 weeks during the development of the guideline. At the start of the development process all group members were required to declare interests including consultancies, fee-paid work, shareholdings, fellowships and support from the healthcare industry in accordance with NICE's policy on Conflicts of Interest. At all subsequent group meetings, members declared all subsequent potential conflicts of interest.

Members were either required to withdraw completely or for part of the discussion if their declared interest made it appropriate. The details of declared interests and the actions taken are shown in Appendix C:.

Staff from the NGA provided methodological support and guidance for the development process. The team working on the guideline included a guideline lead, project manager, systematic reviewers, health economists and information scientists. They undertook systematic searches of the literature, appraised the evidence, conducted data analysis and cost-effectiveness analysis (where appropriate) and drafted the guideline in collaboration with the Committee.

2.4 What the guideline covers

2.4.1 Groups that will be covered

This guideline covers the following groups:

- Babies, children and young people under 18 years who were born preterm (less than 37 weeks of pregnancy).

2.4.2 Key clinical issues that will be covered

The following clinical issues are covered in this guideline:

1. The risk of developmental problems (such as feeding difficulties) and developmental disorders (such as cerebral palsy or autism) in relation to gestational age at birth for babies, children and young people who were born preterm, and other factors that might affect their risk.
2. Identifying developmental problems and disorders in babies, children and young people who were born preterm.
3. Providing information about the development of preterm babies for parents and carers and children and young people who were born preterm.
4. Providing support (for example, help with feeding difficulties, including continuing breastfeeding if appropriate, and with parent child interaction) for babies, children and young people who were born preterm and their parents and carers.
5. Service delivery for developmental follow-up after preterm birth.

For further details please refer to the scope in Appendix A: and review questions in Table 4.

2.4.3 Clinical issues that will not be covered

This guideline does not cover:

1. Diagnosing and managing developmental disorders such as autism and cerebral palsy. These areas are covered by existing NICE guidance on autism diagnosis in children and young people and autism: the management and support of children and young people on the autism spectrum, and in guidance being developed on the diagnosis and management of cerebral palsy in children and young people.
2. Reducing the risk of preterm birth.

2.5 Relationship between the guidance and other NICE guidance

- Preterm labour and birth (2015). NICE guideline 25.

- Postnatal care up to 8 weeks after birth (2006). NICE guideline 37.
- Autism spectrum disorder in under 19s: recognition, referral and diagnosis (2011). NICE guideline 128.
- Spasticity in under 19s: management (2012). NICE guideline 145
- Mental health problems in people with learning disabilities: prevention, assessment and management (2016). NICE guideline 53
- Cerebral palsy in under 25s: diagnosis and management. NICE guideline. Publication expected January 2017
- Intrapartum care for high risk women. NICE guideline. Publication expected November 2017.
- Faltering growth - recognition and management of faltering growth in children. Publication expected October 2017.
- Social and emotional wellbeing in secondary education (2009). Public health guideline 20.
- Social and emotional wellbeing in primary education (2008). Public health guideline 12.
- Patient experience in adult NHS services: improving the experience of care for people using adult NHS services (2012). Clinical guideline 138.
- Sepsis: recognition, diagnosis and early management. NICE guideline 51
- Neonatal infection (early onset): antibiotics for prevention and management (2012). Clinical guideline 149

3 Guideline development methodology

This guideline was developed in accordance with the methods outlined in the NICE guidelines manual 2012 until the beginning of development phase and thereafter in accordance with the updated NICE guidelines manual 2014 (Table 3).

Table 3: Summary of manuals used during the guideline development

Phase of development	Manual
Scoping phase	2012 NICE Manual
Development phase	2014 NICE Manual
Consultation phase	
Validation phase	

3.1 Developing the review questions and protocols

The review questions were drafted by the NGA technical team, then refined and validated by the Committee. The questions were based on the key clinical areas identified in the scope (Appendix A:). Literature searches, critical appraisal and synthesis of the evidence was conducted for each review question.

The review framework was determined by the type of question:

- prognostic reviews – population, risk factors and outcomes
- prevalence reviews –population, outcomes/conditions of interest and context
- reviews of diagnostic test accuracy –population, index tests, reference standard and target condition
- qualitative reviews –population, area of interest and outcomes.

A total of 9 review questions were identified (Table 4).

Table 4: Review questions

Question	
1	What is the risk of developmental problems in babies, children and young people born preterm at different gestational ages? How do the following factors influence the risk of developmental problems in babies, children and young people born preterm: • biological factors • neonatal factors • socioeconomic, maternal and environmental factors • postnatal factors?
2	What is the risk of developmental disorders in babies, children and young people born preterm at different gestational ages? How do the following factors influence the risk of developmental disorders in babies, children and young people born preterm: • biological factors • neonatal factors • socioeconomic, maternal and environmental factors • postnatal factors?
3	What is the prevalence of developmental problems in babies, children and young people born preterm?
4	What is the prevalence of developmental disorders in babies, children and young people born preterm?

Question	
5	What information about development and follow-up arrangements should be provided to parents and carers of preterm babies and to children and young people who were born preterm?
6	What support do parents and carers report was or would have been helpful to them in the care of infants who were born preterm both at discharge and during subsequent follow-up?
7	What is the usefulness of the following screening strategies in the identification of children and young people born preterm with intellectual disability, speech and language disorder, specific learning difficulty, social, emotional and mental health, and developmental co-ordination disorder: • healthy child programme (including plus/enhanced) • parental observation/concern • teachers observation/concern • formal screening tests?
8	What is the most effective setting and staffing model for the follow-up for the identification of developmental problems and disorders and support of babies, children and young people born preterm?
9	What information should be shared between those delivering NHS commissioned care and also between the NHS and the educational sector on the developmental follow-up of babies, children and young people born preterm?

3.2 Searching for evidence

3.2.1 Clinical literature searches

Systematic literature searches were undertaken to identify all published clinical evidence relevant to each review question.

Databases were searched using medical subject headings, free-text terms and study type filters where appropriate. Where possible, searches were restricted to retrieve articles published in English. All searches were limited by date to 1990 onwards because the change in the use of surfactants at this time significantly altered outcomes in areas covered by the guideline. All searches were conducted in the MEDLINE, Embase and Health Technology Assessments (HTA) databases as well as various databases that form parts of The Cochrane Library. All searches were updated on 20th October 2016. Any studies added to the databases after this date (including those published prior to this date but not yet indexed) were not considered relevant for inclusion.

Search strategies were quality assured by cross-checking reference lists of relevant papers, analysing search strategies from other systematic reviews and asking Guideline Committee members to highlight key studies. All search strategies were also quality assured by an Information Scientist who was not involved in the development of the search. Details of the search strategies, including study type filters that were applied and databases that were searched, can be found in Appendix E:.

All references suggested by stakeholders at the time of the scope consultation were considered for inclusion. During the scoping stage, searches were conducted for guidelines, health technology assessments, systematic reviews, economic evaluations and reports on biomedical databases and websites of organisations relevant to the topic. Formal searching for grey literature, unpublished literature and electronic, ahead-of-print publications was not routinely undertaken.

3.2.2 Health economic literature searches

Systematic literature searches were also undertaken to identify relevant published health economic evidence. A broad search was conducted to identify evidence relating to developmental follow-up of preterm babies in the following databases: NHS Economic Evaluation Database (NHS EED), Health Technology Assessment (HTA), Medline, Cochrane Central Register of Controlled Trials (CCTR) and Embase with an economic search filter applied. Where possible, the search was restricted to articles published in English and studies published in languages other than English were not eligible for inclusion.

The search strategies for the health economic literature search are included in Appendix E:. All searches were updated on 20th October 2016. Any studies added to the databases after this date (including those published prior to this date but not yet indexed) were not included unless specifically stated in the text.

3.3 Reviewing and synthesising the evidence

The process for reviewing and synthesising the evidence was as follows:

- The titles and abstracts of records retrieved by the literature searches were sifted for relevance, and potentially relevant publications were obtained in full text.
- Full papers were reviewed against inclusion and exclusion criteria in order to identify relevant studies (review protocols are included in Appendix D:).
- Relevant studies were critically appraised using the appropriate checklist as specified in the NICE guidelines manual 2014. For diagnostic questions the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist was used. For prognostic (risk factors) reviews, the quality of the evidence was assessed using the checklist developed and published by Hayden et al. 2013. For prevalence questions, the quality of the evidence was assessed by using the tool developed and published by The Joanna Briggs Institute (The Joanna Briggs Institute, 2014; Munn et al. 2014). For qualitative reviews, a checklist for qualitative based on the Critical Appraisal Skills Programme (CASP) checklist for qualitative studies (http://www.casp-uk.net/casp-tools-checklists) was used.
- Key information was extracted on the study's methods, PICO factors and results. This is presented in summary tables within each chapter of the guideline and evidence tables (in Appendix K:Appendix J:).
- Summaries of evidence by outcome were generated and then presented to the Committee for discussion:
    - Prognostic (risk) studies – data were presented as measures of association (odds ratios, risk ratios, hazard ratios and adjusted hazard ratios); the decision about whether meta-analysis could be conducted was based on the appraisal of heterogeneity between the studies. In all cases meta-analysis was not considered appropriate.
    - Prevalence studies – data were presented as measures of prevalence or incidence during a period of time (proportions with their 95% confidence intervals); the decision about whether meta-analysis could be conducted was based on the appraisal of heterogeneity between the studies. In all cases meta-analysis was not considered appropriate.
    - Diagnostic/predictive accuracy studies – presented as measures of diagnostic/predictive test accuracy (sensitivity, specificity, positive and negative likelihood ratio); the decision about whether meta-analysis could be conducted was based on the appraisal of heterogeneity between the studies. In all cases meta-analysis was not considered appropriate.
    - Qualitative studies – the themes of the studies were organised in a modified version of a GRADE profile, where possible, along with quality assessment otherwise presented in a narrative form.

  - Delivering enhanced support and surveillance review – narrative summaries of the included literature (including grey literature) were presented.
- Double-sifting was done by a second reviewer for a 5% sample of the abstract list for searches prioritised for health economic modelling and those for complex reviews. If discrepancies were observed, they were solved on a one-by-one basis.
- Double-data extraction was done by a second reviewer for a 5% sample for a review question that were considered complex in order to assure the quality of the data extraction and minimise potential risk of reviewer bias or error.

3.3.1 Type of studies

The type(s) of study design considered optimal for inclusion depended on the review question being asked.

- For clinical prediction (risk) and diagnostic and prognostic reviews, prospective observational studies of N>50 participants were prioritised for inclusion. This is based on the requirements proposed by Green (1991) which is a sample size greater than or equal to 50 participants plus a minimum of 8 variables or predictors.
- For prevalence reviews, the Committee prioritised cross-sectional studies and prospective cohort studies (national registries were preferred) with sample sizes greater than 250 participants. The larger sample size was required for precision.
- For qualitative reviews: the Committee prioritised studies that have collected and analysed data qualitatively (for example using interviews, focus groups, surveys and thematic analysis). Studies that only reported quantitative descriptive data were not prioritised for this type of review.
- For the review about delivering enhanced support and surveillance, the Committee prioritised randomised controlled trials and observational studies. However, they agreed that in the absence of such evidence, grey literature, including expert opinion papers and published developmental follow-up models should be considered.

Sample size cut-offs were agreed with the Committee at the time of protocol development, due to the methodological considerations outlined below and their knowledge of the published evidence base for each topic.

Please refer to Appendix D: for full details on the study design of studies selected for each review question.

3.3.2 Data synthesis

3.3.2.1 Prognostic (risk) and prevalence reviews

Study results were presented according to the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-analyses) (see Appendix F:). Risk factors that were assessed in a multi-variated regression analysis model with adjustment for important confounders were reported. To assist with the ease of interpretation, only results from studies where outcomes were assessed dichotomously were included and reported. Prevalence estimates (proportions) with their 95% confidence intervals were reported or calculated where sufficient data were available. Odds ratios that were adjusted in multivariate analyses for the prespecified confounders were considered the preferred measure.

Studies were categorised according to type of outcome and where data were available, results were reported by subgroups pre-specified in the review protocol. As GRADE is not suitable for this type of review the overall confidence in quality of the evidence was made using the methods described in section 3.3.3.1.

The appropriateness of meta-analysis was assessed by considering whether there was clinical variation and/or methodological heterogeneity across studies. Specifically, the following factors were considered:

- inclusion/exclusion criteria of participants
- age of participants at time of assessment
- whether confounders and risk factors were adjusted for in multivariate analysis models
- whether studies adjusted for the same confounders and risk factors in multivariate analyses
- how outcomes are defined
- measurement tools and scales for the assessment of outcomes
- consistency of results.

Risk factors were also presented graphically in forest plots (Appendix J:). The forest plots displayed all the evidence assessing the association between a risk factor and an outcome as odds ratios.

Prevalence estimates were also presented graphically by outcomes in forest plots (Appendix J:). The forest plots displayed all studies that assessed the prevalence and an estimate of the prevalence of that outcome in the sample is presented as a percentage with 95% confidence intervals. The forest plots for prevalence were presented in a non-logarithmic scale for better visual presentation.

The forest plots for both risk and prevalence evidence were organised by outcome where evidence allowed and in presence of a lot of evidence for an outcome also by gestational age group specified in the review protocols. The forest plots were generated using the statistical software STATA.

3.3.2.2 Diagnostic test accuracy reviews

For studies assessing the diagnostic accuracy of screening tools (index test) compared to diagnostic tests (reference standard) the following outcomes were considered:

- sensitivity
- specificity
- positive likelihood ratio (LR+)
- negative likelihood ratio (LR-).

These diagnostic accuracy parameters (with 95% CI) were obtained from the studies or calculated by the technical team using data from the studies (Table 5).

The following definitions were used when summarising the levels of sensitivity or specificity for the Committee:

- High: 90% and above
- Moderate: 75% to 89%
- Low: 74% or below

The following definitions were used when summarising the likelihood ratios for the Committee:

- Very useful test: LR+ higher than 10, LR- lower than 0.1
- Moderately useful test: LR+ 5 to 10, LR- 0.1 to 0.2
- Not a useful test: LR+ lower than 5, LR- higher than 0.2

Table 5: '2 x 2' table for calculation of diagnostic accuracy parameters

	Reference standard positive	Reference standard negative	Total
Index test result positive	True positive (TP)	False positive (FP)	TP+FP (Total number of subjects with positive result in screening tool)
Index test result negative	False negative (FN)	True negative (TN)	FN+TN (Total number of subjects with negative results in screening tool)
Total	TP+FN (Total number of subjects with diagnosis)	FP+TN (Total number of subjects without diagnosis)	TP+FP+FN+Tn=N (Total number of subjects in study)
Sensitivity=TP/(TP+FN) Specificity=TN/(TN+FP) Positive likelihood ratio=sensitivity/(1-specificity) Negative likelihood ratio=(1-sensitivity)/specificity			

3.3.2.3 Qualitative reviews

A thematic approach was used to identify concepts across qualitative studies. Where possible, a meta-synthesis was conducted to combine results. Themes or new perspectives of a particular topic from the studies were extracted and the characteristics summarised. Common concepts were categorised and tabulated including how many studies contributed to an overarching theme. Sampling of studies continued until no new relevant qualitative data emerged known as 'theoretical saturation' (Dixon-Wood 2005). A final selection of included studies was agreed between two reviewers. Themes from the individual studies were categorised into overarching categories of themes with sub-themes. Themes were derived from direct quotes from individual studies by those who were interviewed. A thematic map was then developed to demonstrate the relationship between themes and subthemes.

3.3.3 Appraising the quality of evidence

3.3.3.1 Prognostic outcomes

Quality of prognostic studies and evidence was assessed using the checklist created by Hayden et al. (2013).

This risk of bias for each risk factor across studies was derived by assessing the risk of bias across 6 domains for each study: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, and statistical analysis and reporting, with the last 4 domains being assessed for each outcome. More details about the quality assessment for prognostic studies are shown in Table 6. The assessment of the overall quality of the evidence was based on the reviewer's judgment considering the assessment of all the 6 domains. For example, if there was a high risk of bias in any domain, the evidence was considered to be of low quality; if there was moderate risk of bias as defined by Hayden et al. (2013) in some of the domains, the evidence was considered to be moderate quality;and if there was low risk of bias in all domains, the evidence was considered to be of high quality.

Table 6: Assessment of risk of bias for prognostic factor studies based on Hayden et al. (2013)

Risk of bias	Explanation
Study participation	Assessment of whether or not there was adequate participation in the study by eligible individuals; if the population and sample were described; if the recruitment and sampling were described and considered appropriate; if inclusion and exclusion criteria were adequately described.
Study attrition	Assessment of whether there was an adequate follow-up rate for study participants; reasons for losses to follow-up were described; the individuals lost to follow-up were adequately described; assessment was done whether the ones lost to follow-up differed from the ones who completed the follow-up.
Prognostic factor measurement	Assessment of whether or not a clear description of the prognostic (risk) factor is provided; the method of assessing or measuring the prognostic factor is valid and reliable; and is the same for every participant.
Outcome measurement	Assessment of whether or not a clear definition of the outcome was provided; the measurement or assessment of outcome is valid and reliable; the method and setting of outcome measurement is the same for every participant.
Study confounding	Assessment of whether or not important confounders were adequately measured, described and adjusted for in the analyses.
Statistical analysis and reporting	Assessment of whether or not there is sufficient presentation of data to assess the adequacy of the analytical strategy; the statistical model is adequate; the reporting of results is adequate, clear and not selective.

3.3.3.2 Prevalence outcomes

Quality of prevalence outcomes was assessed using the checklist created by The Joanna Briggs Institute (The Joanna Briggs Institute, 2014; Munn et al., 2014).

The quality was assessed based on answering ‘yes’, ‘no’, ‘unclear', or “not applicable” to the following questions:

- Was the sample representative of the target population?
- Were the study participants recruited in an appropriate way?
- Was the sample size adequate?
- Were the study subjects and setting described in detail?
- Is the data analysis conducted with sufficient coverage of the identified sample?
- Were objective, standard criteria used for measurement of the condition?
- Was the condition measured reliably?
- Was there appropriate statistical analysis?
- Are all important confounding factors/ subgroups/differences identified and accounted for?
- Were subpopulations identified using objective criteria?

The assessment of the overall quality of the evidence was based on the reviewer’s judgment considering the answers to the questions above. For example, if there were several “no” and “unclear” answers, the quality of the evidence was considered to be low or very low; if there were some “unclear” answers the quality of the evidence was considered to be moderate; and if all answers for the above questions were “yes” or did not raise concern, the evidence was considered to be of high quality.

3.3.3.3 Diagnostic outcomes

For diagnostic accuracy studies, the Quality Assessment of Diagnostic Accuracy Studies version 2 (QUADAS-2) checklist was used to assess risk of bias and applicability of the evidence (Whiting et al., 2011). The assessment of risk of bias and applicability of patient selection, index test, reference standard and flow and timing were done. More details of the QUADAS-2 is given in Table 7.

Table 7: Summary of assessment of risk of bias and applicability of diagnostic accuracy evidence according to QUADAS-2

Domain	Patient selection	Index test	Reference standard	Flow and timing
Description	Describe methods of patient selection: Describe included patients (prior testing, presentation, intended use of index test and setting):	Describe the index test and how it was conducted and interpreted:	Describe the reference standard and how it was conducted and interpreted:	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded from the 2x2 table (refer to flow diagram): Describe the time interval and any interventions between index test(s) and reference standard:
Signalling questions (yes/no/unclear)	Was a consecutive or random sample of patients enrolled? Was a case-control design avoided? Did the study avoid inappropriate exclusions?	Were the index test results interpreted without knowledge of the results of the reference standard? If a threshold was used, was it pre-specified?	Is the reference standard likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of the results of the index test?	Was there an appropriate interval between index test(s) and reference standard? Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?
Risk of bias: High/low/unclear	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?
Concerns regarding applicability: High/low/unclear	Are there concerns that the included patients do not match the review question?	Are there concerns that the index test, its conduct, or interpretation differ from the review question?	Are there concerns that the target condition as defined by the reference standard does	

Domain	Patient selection	Index test	Reference standard	Flow and timing
			not match the review question?	

From http://www.bristol.ac.uk/social-community-medicine/projects/quadas/quadas-2/

For the assessment of the overall quality of the diagnostic accuracy evidence, adapted GRADE methodology was used. At the time of writing, the GRADE methodology, as developed by the international GRADE working group, was available for RCTs and observational studies only. We adapted the quality assessment elements and outcome presentation for diagnostic accuracy studies. GRADE methodology takes into account the assessment of 5 different domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias. Note that publication bias was not systematically considered in this guideline. Table 8 gives more details of the different domains. The assessment of risk of bias and indirectness were based on the QUADAS-2 assessment described above.

The overall quality of the diagnostic accuracy evidence was based on the sum of the grading of the different domains of GRADE. Inconsistency was not considered applicable when no meta-analysis was performed. The reasons or criteria used for downgrading were specified in the footnotes of the adapted GRADE tables.

Table 8: Summary of the adapted GRADE methodology to assess the quality of diagnostic accuracy evidence

Quality element	Description
Risk of bias (study limitations)	Defined as anything that causes a consistent deviation from the truth. Bias can be perceived as a systematic error; for example, if a study was carried out several times and there was a consistently wrong answer, the results would be inaccurate. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect. A study with a poor methodological design does not automatically imply high risk of bias; the bias is considered individually for each outcome and it is assessed whether this poor design will impact on the estimation of the intervention effect. Based on the assessment using QUADAS-2 checklist.
Inconsistency	Inconsistency refers to an unexplained heterogeneity of results. Only applicable when meta-analysis is performed.
Indirectness	Indirectness refers to differences in, for example, study population, index test, and comparator (reference standard) between the available evidence and the review question. Based on the assessment using QUADAS-2 checklist.
Imprecision	Results were considered imprecise when the estimates have wide confidence intervals based on visual inspection.

3.3.3.4 Qualitative studies

The main quality assessment domains are organised across the definition of population included, the appropriateness of methods used and the completeness of data analysis and the overall relevance of the study participants to the population of interest for the guideline.

Individual studies were assessed for methodological limitations using an adapted Critical Appraisal Skills Programme (CASP, 2013) checklist for qualitative studies, where items in the original CASP checklist were adapted and fitted into 5 main quality appraisal areas according to the following criteria:

- aim (description of aims and appropriateness of the study design)

- sample (clear description, role of the researcher, data saturation, critical review of the researchers' influence on the data collection)
- rigour of data selection (method of selection, independence of participants from the researchers, appropriateness of participants)
- data collection analysis (clear description, how are categories or themes derived, sufficiency of presented findings, saturation in terms of analysis, the role of the researcher in the analysis, validation)
- results / findings (clearly described, applicable and comprehensible, theory production)
- An adapted GRADE-CERQual (Lewin 2015) approach was used to present and summarise qualitative findings across studies. This approach considers the quality of evidence by themes. Themes may have originated from an individual study or been identified through a number of individual themes or components of themes from a number of included studies. Quality is assessed in the domains described in Table 9.

Table 9: Domains of quality considered in qualitative studies

Quality element	Description
Risk of bias ('Study limitations')	Limitations in the study design and implementation may bias the interpretation of the qualitative themes that are identified. High risk of bias for the majority of the evidence decreases confidence in the estimate of the effect. Qualitative studies are not usually randomised and therefore would not be downgraded for study design from the outset and start as high level evidence.
Coherence of findings	The extent to which different individual themes or components of themes from studies fit into a wider network of overarching themes. For example, many components (relationship and rapport, clinical experience, information provision) can contribute to an overarching theme of healthcare professional factors in shared decision-making. Even though each individual study may not mention each factor the overall theme is coherent.
Applicability (or relevance) of evidence	The extent to which the evidence supporting the review finding is applicable to the context specified in the review question. In the case of this guideline qualitative evidence from the UK was prioritised over and above data from other contexts.
Theme saturation / sufficiency	Theme saturation or sufficiency refers to a similar concept in qualitative research. This refers to whether a theoretical point of theme saturation was achieved at which point no further citations or observations would provide more insight or suggest a different interpretation of this theme. Individual studies that may have contributed to a theme or subtheme may have been conducted in a manner that by design would have not reached theoretical saturation on an individual study level.

3.3.4 Evidence statements

Evidence statements are statements that summarise the key features of the clinical evidence presented. The wording of the evidence statements reflects the amount of certainty in the estimate of effect. They are presented by comparison (for interventional reviews) or by description of outcome where appropriate and encompass the following key features of the evidence:

- the number of studies and the number of participants for a particular outcome
- a brief description of the participants
- an indication of the direction of effect (if 1 treatment is beneficial or harmful compared with the other, or whether there is no difference between the 2 tested treatments)
- a description of the overall quality of evidence.

3.3.5 Evidence of cost effectiveness

The health economic evidence presented in the guideline aims to inform the Committee about potential economic issues and ensure that the recommendations represent a cost-effective use of healthcare resources. Health economic evaluations aim to integrate data on benefits (ideally in terms of quality adjusted life years [QALYs]), harms and costs of different care options.

3.3.5.1 Literature review

The Health Economist assessed the titles and abstracts of publications identified by the literature searches using the pre-defined eligibility criteria specified in Table 10.

Table 10: Inclusion and exclusion criteria for the systematic reviews of economic evaluations

Inclusion criteria
• intervention or comparators according to the scope • study population according to the scope • full economic evaluations (cost-utility, cost-effectiveness, cost-benefit or cost-consequence analyses) that assess both costs and outcomes associated with the interventions of interest
Exclusion criteria
• abstracts with insufficient methodological details • conference papers published before January 2014

Once the screening of titles and abstracts was complete, full versions of the selected papers were obtained for assessment. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) for this search on economic evaluations is presented in Appendix F:.

As well as reviewing the published economic literature, as described above, new economic analysis was undertaken in selected areas prioritised by the Committee in conjunction with the health economist. Topics were prioritised on the basis of the following criteria, in accordance with the NICE guidelines manual:

- the overall importance of the recommendation, which may be a function of the number of patients affected and the potential impact on costs and health outcomes per patient
- the current extent of uncertainty over cost effectiveness, and the likelihood that economic analysis will reduce this uncertainty
- the feasibility of building an economic model

The following priority areas for de novo economic analysis were agreed by the Committee after formation of the review questions and consideration of the available health economic evidence:

- screening strategies for the identification of children and young people born preterm with intellectual disability, speech and language disorder and specific leaning difficulty
- delivery of enhanced support and surveillance

The methods and results of de novo economic analyses are reported in Appendix H:. When new economic analysis was not prioritised, the Committee made a qualitative judgement regarding cost effectiveness by considering expected differences in resource and cost use between options, alongside clinical effectiveness evidence identified from the clinical evidence review.

3.3.5.2 Cost effectiveness criteria

NICE's report Social value judgements: principles for the development of NICE guidance sets out the principles that Committees should consider when judging whether an intervention offers good value for money. In general, an intervention was considered to be cost effective if either of the following criteria applied (given that the estimate was considered plausible):

- the intervention dominated other relevant strategies (that is, it was both less costly in terms of resource use and more clinically effective compared with all the other relevant alternative strategies), or;
- the intervention cost less than £20,000 per QALY gained compared with the next best strategy, or;
- the intervention provided clinically significant benefits at an acceptable additional cost when compared with the next best strategy.

The Committee's considerations of cost-effectiveness are discussed explicitly in the 'Consideration of economic benefits and harms' section for each topic. .

3.4 Developing recommendations

Over the course of the guideline development process, the Committee was presented with:

- evidence tables of the clinical and economic evidence reviewed from the literature (see Appendix H:, Appendix I:, Appendix K:)
- summary of clinical and economic evidence and quality assessment
- forest plots (Appendix J:)
- a description of the methods and results of the cost-effectiveness analysis undertaken for the guideline (Appendix H:, Appendix I:).

Recommendations were drafted on the basis of the Committee's interpretation of the available evidence, taking into account the balance of benefits, harms and costs between different courses of action. Firstly, the net benefit over harm (clinical effectiveness) was considered, focusing on the critical outcomes, although most of the reviews in the guideline were outcome driven. The Committee took into account the clinical benefits and harms when one intervention was compared with another. The assessment of net benefit was moderated by the importance placed on the outcomes (the Committee's values and preferences), and the confidence the Committee had in the evidence (evidence quality). Secondly, the Committee assessed whether the net benefit justified any differences in costs.

When clinical and economic evidence was of poor quality, conflicting or absent, the Committee drafted recommendations based on their expert opinion. The considerations for making consensus-based recommendations include the balance between potential harms and benefits, the economic costs or implications compared with the economic benefits, current practices, recommendations made in other relevant guidelines, patient preferences and equality issues. The Committee also considered whether the uncertainty was sufficient to justify delaying making a recommendation to await further research, taking into account the potential harm of failing to make a clear recommendation.

The wording of recommendations was agreed by the Committee and focused on the following factors:

- the actions healthcare professionals need to take
- the information readers need to know
- the strength of the recommendation (for example the word 'offer' was used for strong recommendations and 'consider' for weak recommendations)
- the involvement of parents, carers and families in decisions about treatment and care

- consistency with NICE's standard advice on recommendations about drugs, waiting times and ineffective intervention.

The main considerations specific to each recommendation are outlined in the 'Recommendations and link to evidence' sections within each section.

3.4.1 Research recommendations

When areas were identified for which good evidence was lacking, the Committee considered making recommendations for future research. Decisions about inclusion were based on factors such as:

- the importance to patients or the population
- national priorities
- potential impact on the NHS and future NICE guidance
- ethical and technical feasibility.

3.4.2 Validation process

This guidance is subject to a 6-week public consultation and feedback as part of the quality assurance and peer review of the document. All comments received from registered stakeholders receive individual responses that are posted on the NICE website when the pre-publication check of the full guideline occurs.

3.4.3 Disclaimer

Healthcare providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by practitioners in light of individual patient circumstances, the wishes of the patient, clinical expertise and available resources.

The National Guideline Alliance (NGA) disclaims any responsibility for damages arising out of the use or non-use of these guidelines and the literature used in support of these guidelines.

3.4.4 Funding

The NGA was commissioned by the National Institute for Health and Care Excellence (NICE) to undertake the work on this guideline.

4 Risk and prevalence of developmental problems and disorders

4.1 Introduction

Children born preterm are thought to be at increased risk of a range of developmental problems and disorders that may have a short or long term, and often cumulative, impact on a child's health, development and well-being.

Developmental problems and disorders typically present on a continuum, with disorders considered to represent the severe end of the spectrum. Although a child may not meet the diagnostic criteria for a developmental disorder they may still experience substantial developmental difficulties that impact on their everyday life. The prevalence of these conditions is thought to be associated with the degree of prematurity at birth.

Developmental problems may include functional issues with feeding, sleeping and toileting, excessive crying or irritability during infancy, delayed motor or language development during the early years, sensory difficulties, behavioural, social and emotional problems, deficits in executive functions and special educational needs throughout childhood and adolescence. They may present independently or co-exist with other developmental problems or disorders.

Developmental disorders may include intellectual disability or global developmental delay, cerebral palsy, speech and/or language disorders, attention-deficit/hyperactivity disorder, developmental coordination disorder, specific learning disorders, autism spectrum disorder, other mental and behavioural disorders and sensory impairments such as hearing and visual impairments.

Information about the potential risk and prevalence of developmental problems and disorders can be used to support the early identification of difficulties as they arise so that appropriate support and therapeutic intervention is provided. This information can, in turn, be used to guide service planning inclusion the provision of health, education and social care and requirements for developmental surveillance.

4.1.1 Risk of developmental problems

Review question:

What is the risk of developmental problems in babies, children and young people born preterm at different gestational ages?

How do the following factors influence the risk of developmental problems in babies, children and young people born preterm:

- **biological factors**
- **neonatal factors**
- **socioeconomic, maternal and environmental factors**
- **postnatal factors?**

4.1.1.1 Description of clinical evidence

The aim of this review was to identify different factors (gestational age at birth; biological factors; neonatal factors; maternal, social or environmental; and postnatal factors) that can affect the risk of developmental problems in babies, children and young people born preterm. Developmental problems considered as outcomes included sensory sensitivity; functional problems with feeding, sleeping or toileting; motor, developmental and language delay;

executive function; problems specific to infancy (excessive crying, irritability, poor self-regulation); behavioural, social, emotional, attention problems; and special educational needs.

Studies were included if they:

1. were prospective cohort studies (in addition, two retrospective population-based studies were included for special educational needs outcome where evidence is otherwise scarce)
2. were multi-centre or national population-based studies;
3. included only participants born after 1990 (two exceptions where small number of participants were born before 1990);
4. confounders were adjusted for in the analyses.

For full details see review protocol in Appendix D:.

In total, fifty-one publications were included in the review (Adams-Chapman 2008; Allred 2014; Brown 2014; Carlo 2011; Chan 2014; de Jong 2015; Delobel-Ayoub 2006; Delobel-Ayoub 2009; Farooqi 2016; Farooqi 2013; Farooqi 2007; Fevang 2016; Guellec 2011; Gurka 2010; Higa Diez 2016; Hintz 2005; Hornman 2016; Johnson 2016; Johnson 2015a; Johnson 2015b; Johnson 2011; Kerstjens 2013; Kerstjens 2012; Kerstjens 2011; Larroque 2011; Laughon 2009; MacKay 2010; MacKay 2013; Martin 2010; Migraine 2013; Odd 2016; Odd 2013a; Odd 2013b; O'Shea 2008; Peacock 2012; Potijk 2015; Quigley 2012; Rautava 2010; Raynes-Greenow 2012; Reijneveld 2006; Samara 2010; Schendel 1997; Shah 2012; Shankaran 2004; Singer 2001; Stene-Larsen 2014; Stoll 2004; Sullivan 2015; Vohr 2005; Vohr 2000; Woythaler 2011). The sample sizes ranged from 169 (Farooqi 2013) to 407503 (MacKay 2013; MacKay 2010).

Seventeen publications are from the United States (Adams-Chapman 2008; Allred 2014; Carlo 2011; Gurka 2010; Hintz 2005; Laughon 2009; Martin 2010; O'Shea 2008; Schendel 1997; Shah 2012; Shankaran 2004; Singer 2001; Stoll 2004; Vohr 2005; Vohr 2000; Woythaler 2011). Elevenpublications are from the UK (Chan 2014; Johnson 2016; Johnson 2015a; Johnson 2015b; MacKay 2010; MacKay 2013; Odd 2016; Odd 2013a; Odd 2013b; Peacock 2012; Quigley 2012; Sullivan). Two publications are from the UK and Ireland (Samara 2010; Johnson 2011). Seven publications are from the Netherlands (de Jong 2015; Hornman 2016; Kerstjens 2013; Kerstjens 2012; Kerstjens 2011; Potijk 2015; Reijneveld 2006) and five publications are from France (Delobel-Ayoub 2006; Delobel-Ayoub 2009; Guellec 2011; Larroque 2011; Migraine 2013). Threepublications from Sweden (Farooqi 2016; Farooqi 2013; Farooqi 2007) and two from Norway (Fevang 2016; Stene-Larsen 2014) One publication comes from the following countries: Australia (Raynes-Greenow 2012); Canada (Brown 2014); Finland (Rautava 2010); and Japan (Higa Diez 2016).

Forty-nine publications used data from population-based, multicentre or regional prospective cohort studies (Adams-Chapman 2008; Allred 2014; Brown 2014; Carlo 2011; Chan 2014; de Jong 2015; Delobel-Ayoub 2006; Delobel-Ayoub 2009; Farooqi 2016; Farooqi 2013; Farooqi 2007; Fevang 2016; Guellec 2011; Gurka 2010; Higa Diez 2016; Hintz 2005; Hornman 2016; Johnson 2016; Johnson 2015a; Johnson 2015b; Johnson 2011; Kerstjens 2013; Kerstjens 2012; Kerstjens 2011; Larroque 2011; Laughon 2009; Martin 2010; Migraine 2013; Odd 2106; Odd 2013a; Odd 2013b; O'Shea 2008; Peacock 2012; Potijk 2015; Quigley 2012; Rautava 2010; Raynes-Greenow 2012; Reijneveld 2006; Samara 2010; Schendel 1997; Shah 2012; Shankaran 2004; Singer 2001; Stene-Larsen 2014; Stoll 2004; Sullivan 2015; Vohr 2005; Vohr 2000; Woythaler 2011). Two publications used data from retrospective cohort studies using population-based data (MacKay 2010; MacKay 2013).

The fifty-one publications included in this review come from twenty-three different studies. Eight publications from the United States derive from the work of the Eunice Kennedy Shriver National Institute of Child Health and Human Development's (NICHD) Neonatal Research Network (NRN) (Adams-Chapman 2008; Carlo 2011; Hintz 2005; Shah 2012; Shankaran 2004; Stoll 2004; Vohr 2000, Vohr 2005). These publications include cohorts born at different time spans between 1993 and 2011, therefore, the cohort included in each study differ across the publications. Four publications are from the Extremely Low Gestational Age Newborns (ELGAN) study from the United States (Allred 2014; Laughon 2009; Martin 2010; O'Shea 2008). Another four publications come from the French study called Etude Epidemiologique sur les Petits Ages Gestationnels (EPIPAGE) (Delobel-Ayoub 2006; Delobel-Ayoub 2009; Guellec 2011; Larroque 2011). Five publications are from the Longitudinal Preterm Outcome Project (Lollipop) in the Netherlands (Hornman 2016; Kerstjens 2013; Kerstjens 2012; Kertsjens 2011; Potijk 2015). Five publications derive from the Avon Longitudinal Study of Parents and Children (ALSPAC) from the United Kingdom (Odd 2016; Odd 2013a; Odd 2013b; Peacock 2012; Sullivan 2015). Three publications are from the Late to Moderately Preterm Birth Study (LAMBS) in the UK (Johnson 2016; Johnson 2015a; Johnson 2015b).Two publications are from the EPICure Study (Johnson 2011; Samara 2010). Another two publications use data from the same school census from Scotland (MacKay 2010; MacKay 2013). The different publications within the same studies examine different risk factors and/or different outcomes or assess the children at different age. The rest of the included studies had one publication from the cohort studied.

In relation to gestational age, in total thirty-four publications were included in the review (Brown 2014; Chan 2014; de Jong 2015; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Farooqi 2016; Farooqi 2013; Farooqi 2007; Fevang 2016; Gurka 2010; Higa Diez 2016; Hornman 2016; Johnson 2016; Johnson 2015a; Johnson 2015b; Kerstjens 2011; Kerstjens 2012; Larroque 2011; MacKay 2010; MacKay 2013; Migraine 2013; Odd 2013a; Odd 2013b; Peacock 2012; Potijk 2015; Quigley 2012; Rautava 2010; Raynes-Greenow 2012; Reijneveld 2006; Samara 2010; Schendel 1997; Stene-Larsen 2014; Sullivan 2015; Woythaler 2011).Six publications reported on functional problems (de Jong 2015; Johnson 2016; Migraine 2013; Raynes-Greenow 2012; Samara 2010; Sullivan 2015); ten publications reported on motor, developmental or language problems (Brown 2014; de Jong 2015; Johnson 2015a; Kerstjens 2012; Kerstjens 2011; Odd 2013b; Rautava 2010; Schendel 1997; Stene-Larsen 2014; Woythaler 2011); three publications reported on executive function (Farooqi 2016; Farooqi 2013; Rautava 2010); fourteen publications reported on behavioural, social, emotional or attention problems (de Jong 2015; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Farooqi 2013; Farooqi 2007; Fevang 2016; Gurka 2010; Higa Diez 2016; Hornman 2016; Johnson 2015b; Potijk 2015; Rautava 2010; Reijneveld 2006; Schendel 1997); and seven publications reported on special educational needs (Chan 2014; Larroque 2011; MacKay 2013; MacKay 2010; Odd 2013a; Peacock 2012; Quigley 2012). No evidence on sensory sensitivity was found.

In relation to biological factors (sex of the child, being born small for gestational age, and ethnicity or race), ten publications were included (Delobel-Ayoub 2009; Delobel-Ayoub 2006; Guellec 2011; Johnson 2016; Johnson 2015a; Johnson 2015b; Johnson 2011; Kerstjens 2013; Shankaran 2004; Vohr 2000). Two publications reported on functional problems (Johnson 2016; Vohr 2000); four publications reported on motor, developmental or language problems (Johnson 2015a; Kerstjens 2013; Shankaran 2004; Vohr 2000); four publications reported on behavioural, social, emotional, or attention problems (Delobel-Ayoub 2009; Delobel-Ayoub 2006; Guellec 2011; Johnson 2015b); two publications reported on special educational needs (Guellec 2011; Johnson 2011). No evidence on sensory sensitivity or executive function in relation to biological risk factors.

In relation to neonatal factors (brain abnormalities, sepsis, retinopathy of prematurity, necrotising enterocolitis, exposure to antenatal steroids, exposure to postnatal steroids, bronchopulmory dysplasia), eighteen publications were included in the review (Adams-Chapman 2008; Allred 2014; Carlo 2011; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Hintz 2005; Johnson 2015b; Johnson 2011; Kerstjens 2013; Kerstjens 2012; Laughon 2009; Martin 2010; O'Shea 2008; Shah 2012; Shankaran 2004; Stoll 2004; Vohr 2005; Vohr 2000). One publication reported on functional problems (Vohr 2000); Fourteen publications reported on motor, developmental or language problems (Adams-Chapman 2008; Allred 2014; Carlo 2011; Hintz 2005; Kerstjens 2013; Kerstjens 2012; Laughon 2009; Martin 2010; O'Shea

2008; Shah 2012; Shankaran 2004; Stoll 2004; Vohr 2005; Vohr 2000); and three publications reported on behavioural, social, emotional or attention problems (Delobel-Ayoub 2009; Delobel-Ayoub 2006; Johnson 2015b). One publication reported on special educational needs (Johnson 2011). No evidence on sensory sensitivity or executive function in relation to different neonatal factors.

In relation to different social, environmental or maternal factors (socioeconomic status, maternal substance abuse, maternal alcohol abuse, multiple pregnancy, chorioamnionitis, neglect, maternal age and maternal mental health disorder), ten publications were included (Delobel-Ayoub 2009; Delobel-Ayoub 2006; Johnson 2016; Johnson 2015a; Johnson 2015b; Johnson 2011; Kerstjens 2013; Potijk 2015; Shankaran 2004; Singer 2001). One publication reported on functional problems (Johnson 2016). Four publications reported on motor, developmental or language problems (Johnson 2015a; Kerstjans 2013; Shankaran 2004; Singer 2001); and four publications reported on behavioural, social, emotional or attention problems (Delobel-Ayoub 2009; Delobel-Ayoub 2006; Johnson 2015b; Potijk 2015). One publication reported on special educational needs (Johnson 2011). No evidence on sensory sensitivity, functional problems, or executive function in relation to different maternal, social or environmental factors.

The feasibility of combining study data using meta-analysis was assessed. Due to the following differences between studies, it was not considered appropriate to pool the results:

- the inclusion/exclusion criteria for participants
- ages of participants at the time of assessment
- confounders adjusted for in multivariate analysis models
- outcome definitions and measurement tools
- consistency of results.

4.1.1.2 Summary of included studies

Table 11: Summary of included studies in relation to gestational age

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Functional problems with feeding/sleeping/toileting						
de Jong 2015 (The Netherlands)	Multicentre prospective cohort	n=116 moderately preterm children (32-36 weeks gestation) n=99 term children (37-41 weeks gestation)	Analyses were adjusted for maternal education level and maternal age at birth.	Behavioural problems were assessed with the CBCL. For total problems and broadband scales, scores of 60 or above were considered abnormal. For the subscales, scores of 65 or above were considered abnormal.	At 24 months (corrected age) Sleep problems Term: Reference 32-36 weeks: OR 0.53 (0.06-4.43)	High
Johnson 2016 (UK)	Prospective population-based cohort study	n=628 late and moderately preterm (LMPT) children (32-36 weeks) n=759 term controls (>=37 weeks)	The analyses between term and LMPT group were adjusted for sex, SGA, SES index score, and nasogastric tube feeding >2 weeks. The analyses within the LMPT group included the following variables: behaviour problems, delayed social competence, SGA and nasogastric tube feeding.	A validated eating behaviour questionnaire (4) was used to assess the presence of eating difficulties in the 4 domains of refusal/picky eating (e.g., poor appetite, food refusal, selective eating), oral motor problems (e.g., problems biting, chewing, or swallowing; gagging; or choking on food), oral hypersensitivity (e.g., aversion to being touched around	At 2 years (corrected age) Total feeding problems Term: Reference 32-36 weeks: RR 1.44 (1.01-2.03) Refusal/picky eating Term: Reference 32-36 weeks: RR 1.30 (0.84-1.98) Oral motor problems Term: Reference 32-36 weeks: RR 1.65 (1.05-2.58) Oral hypersensitivity Term: Reference 32-36 weeks: RR 1.22 (0.69-2.13)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				the mouth or having things put in the mouth), and eating behaviour problems (e.g., has tantrums or makes a mess during meals). >90th percentile of the term control group were used to identify children with clinically significant eating difficulties.	Eating behaviour problems Term: Reference 32-36 weeks: RR 0.88 (0.53-1.45)	

Migraine 2013 (France)	Multicentre prospective cohort study	n=234 children born <33 weeks GA
		(n=54 children 32 weeks GA;
		n=78 children 30-31 weeks GA;
		n=54 children 28-29 weeks GA;
		n=48 children <28 weeks GA)
		
		n=245 term controls (>37 weeks)
			Maternal age, maternal BMI, maternal education level, breastfeeding, gestational age, birth-weight z score and gender.	The Children's Eating Difficulties Questionnaire was completed by parents. 2 domains of low drive to eat and narrow food repertoire were generated. Subjects scoring in the highest quintile for these outcomes were defined as having eating difficulties.	At 24 months of age (corrected)
					Low drive to eat
					>37 weeks: Reference
					32 weeks: OR 1.33 (0.59-2.98)
					30-31 weeks: OR 1.17 (0.54-2.55)
					28-29 weeks: OR 2.01 (0.89-4.56)
					<28 weeks: OR 1.63 (0.69-3.81)
					Low food variety
					>37 weeks: Reference
					32 weeks: OR 0.87 (0.39-1.94)
					30-31 weeks: OR 1.10 (0.55-2.21)
					28-29 weeks: OR 0.97 (0.42-2.24)
						Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					<28 weeks: OR 0.75 (0.31-1.82)	
Raynes-Greenow 2012 (Australia)	Prospective cohort study (using record linked population health data)	n=3115 children born at <32 weeks; n=22039 children born at 32-36 weeks; n=377952 children born at >36 weeks	Sex, maternal age, caesarean section, pregnancy hypertension, number of previous pregnancies, any neonatal resuscitation, and neonatal morbidity (admitted to the special care nursery and/or the neonatal intensive care unit).	Data from births from 2000–2004 were obtained via the NSW Midwives Data Collection, a legislated population-based surveillance system that includes information on all babies born at ≥ 20 weeks gestation or weighing ≥ 400 g. The primary outcome was sleep apnoea diagnosis in childhood, first diagnosed between 1 and 6 years of age. Children with sleep apnoea were identified from those hospital records with the ICD-10 code G47.3: sleep apnoea, central or obstructive.	At 2.5 to 6 years if age Sleep apnea diagnosis >36 weeks: Reference 32-36 weeks: OR 1.19 (1.03-1.34) <32 weeks: 2.74 (2.16-3.49)	Moderate
Samara 2010 (UK and Ireland)	Population based prospective cohort study (EPICure)	n=223 preterm children (<26 weeks') n=148 full-term controls	Cognitive, neuromotor and pervasive behavioural difficulties.	Parents completed a specially developed eating questionnaire. Items were grouped into four categories: refusal-faddy eating problems, oral motor problems, oral hypersensitivity	At age 6 years (assumed to be chronological) Total eating difficulties Controls: Reference Preterm: OR 2.5 (1.3-4.8) Oral motor problems Controls: Reference Preterm: OR 2.7 (1.3-5.7)	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				problems and behavioural problems around meals. A total eating problems score was also constructed. Higher scores on each scale indicate more problems. To derive clinical categories, each scale was dichotomised into normal versus clinical (scores above the 90th centile or near according to the comparison group).	Refusal-faddy eating problems Controls: Reference Preterm: OR 1.6 (0.8-3.3) Behavioural problems around meals Controls: Reference Preterm: OR 1.6 (0.7-3.6) Oral hypersensitivity problems Controls: Reference Preterm: OR 1.9 (0.8-4.7)	
Sullivan 2016 (UK)	Regional prospective cohort study (ALSPAC)	N=13, 973 children alive at 12 months N=8769 children with 3 or more bedwetting measures	Adjusted for the confounders including gender and socioeconomic status (family adversity)	At ages 4.5, 5.5, 6.5, 7.5 and 9.5 years (4-9 years), parents were asked about how often their child wets their bed. The frequency of bedwetting was further divided into three categories: no current bed wetting, infrequent bedwetting (< once or about once a week), and frequent bedwetting (2-5 times a week, nearly every night, or more than once a week). Frequent	At 4 to 9 years age Risk of frequent persistent bedwetting <37 weeks GA: OR 0.82 (95%CI 0.40-1.70)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				bedwetting corresponds to the frequency of bedwetting required for a DSM-V diagnosis of nocturnal enuresis.		
Motor, developmental and language delay						
Brown 2014 (Canada)	Population based prospective cohort	n=15099 at 2-3 years n=12302 at 4-5 years	Adjusted for alcohol during pregnancy, smoking during pregnancy, placental ischaemia, delivery mode, other biological determinants (not described further), delivery mode, gestational age, partnership status, number of siblings, family income adequacy, maternal education, maternal age at birth of child, maternal health, maternal mental health, family functioning, parenting interactions, parenting effectiveness and parenting consistency.	Developmental delay was measured at 2-3 years using the Motor and Social Development Scale. Scores were standardised by 1-month age groups and children scoring ≥1 SD below the mean were classified as having a delay. Receptive vocabulary delay was measured at 4-5 years using the PPVT-R. The number of correct responses is computed and an age-standardised score is based on 1-month age groups. Children scoring ≥1 SD below the mean were classified as having a delay.	At 2-3 years(assumed to be chronological age) Risk of developmental delay 39-41 weeks: Reference 34-36 weeks: RR 1.13 (0.90-1.42) At 4-5 years (assumed to be chronological age) Risk of receptive vocabulary delay 39-41 weeks: Reference 34-36 weeks: RR 1.06 (0.79-1.43)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
de Jong 2015 (The Netherlands)	Multicentre prospective cohort	n=116 moderately preterm children (32-36 weeks gestation) n=99 term children (37-41 weeks gestation)	Analyses were adjusted for maternal education level and maternal age at birth.	Developmental delay was assessed with the Bayley III scales. Scores of <7 were defined as mild developmental delay for each of the subscales.	At 24 months (corrected for gestation) Cognitive developmental delay Term: Reference 32-36wks: OR 0.89 (0.19-4.15) Fine motor developmental delay Term: Reference 32-36wks: OR 0.48 (0.04-6.36) Gross motor developmental delay Term: Reference 32-36wks: OR 1.61 (0.69-3.73) Receptive communication developmental delay Term: Reference 32-36wks: OR 2.07 (0.37-11.56) Expressive communication developmental delay Term: Reference 32-36wks: OR 0.48 (0.13-1.75) At 24 months (chronological age) Cognitive developmental delay Term: Reference	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					32-36wks: OR 2.19 (0.56-8.63) Fine motor developmental delay Term: Reference 32-36wks: OR 2.13 (0.40-11.44) Gross motor developmental delay Term: Reference 32-36wks: OR 2.30 (1.03-5.13) Receptive communication developmental delay Term: Reference 32-36wks: OR 3.52 (0.69-17.82) Expressive communication developmental delay Term: Reference 32-36wks: OR 1.03 (0.33-3.17)	
Johnson 2015 (UK)	Prospective cohort study	n=638 late/moderately preterm infants n=765 term controls	Sex, SES-index and SGA.	Cognitive impairment was assessed using the Parent Report of Children's Abilities-Revised (PARCA-R). Scores for non-verbal cognition and expressive language were combined to give a total parent report composite. These scores are	At 2 years (corrected age) Risk of cognitive impairment Term: Reference 32-36 weeks: RR 2.09 (1.19-3.64)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				strongly correlated with scores on gold standard developmental tests. Moderate/severe cognitive impairment was identified as a score corresponding to with PRC scores < 2.5th percentile in the term reference group.		
Kerstjens 2012 (The Netherlands)	Population based prospective cohort study	n=832 moderately preterm children (32 to 35+6 weeks)	Variables included in the final model were: birth asphyxia, tertiary NICU admission, hypoglycaemia, hyperbilirubinaemia, SGA and gender.	Parents completed the Dutch version of the 48 months ASQ. The scores on each domain add up to an ASQ total problems score. A score of >2SDs below the mean for the Dutch reference group was considered to indicate developmental delay.	At 43-49 months (assumed to be chronological age) Risk of abnormal ASQ total problems score Low gestational age 34 to 35+6 weeks: Reference 32 to 33+6 weeks: not significant on univariate analysis	Moderate
Kerstjens 2011 (The Netherlands)	Population based prospective cohort study	n=1983 total sample n=512 children born at <32 weeks of gestation n=927 children born at 32-35	Maternal age, mother's birth country, parental education, single-parent family, sex, multiple birth and SGA.	The Dutch version of the age 48 month form of the Ages and Stages questionnaire was used to assess development. The ASQ covers five domains: communication, fine motor function, gross motor	At 4 years (assumed to be chronological age) Risk of developmental delay (ASQ total score <2SD below the mean) Term: Reference <32 weeks: OR 3.2 (1.88-5.37) 32-35 weeks: OR 1.5 (0.89-2.52) 32-33 weeks: OR 1.5 (0.81-2.92)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		weeks of gestation n=544 children born at 38-41 weeks of gestation		function,personal-social functioning and problem solving. The total score was calculated by adding all the domain scores and dividing by five. The individual domain scores, and the total score were dichotomized at 2SD below the mean score of the Dutch reference group as normal/abnormal.	34-35 weeks: OR 1.5 (0.84-2.52) Risk of fine motor impairment (ASQ Fine motor score <2SD below the mean) Term: Reference <32 weeks: OR 3.6 (2.02-6.38) 32-35 weeks: OR 2.0 (1.17-3.54) 32-33 weeks: OR 2.5 (1.32-4.87) 34-35 weeks: OR 1.8 (1.01-3.22) Risk of gross motor impairment (ASQ Gross motor score <2SD below the mean) Term: Reference <32 weeks: OR 3.5 (2.04-5.94) 32-35 weeks: OR 1.3 (0.75-2.21) 32-33 weeks: OR 1.0 (0.46-2.06) 34-35 weeks: OR 1.4 (0.81-2.50)	
Odd 2013b (UK)	Regional prospective cohort study	Overall: n=741 moderate/late preterm infants n=13102 term infants	Ethnicity, housing, crowding and maternal education, socioeconomic group, car ownership, maternal age, gender, parity, weight, length and head circumference	3 of the 8 subtests of the MABC were used. These subtests were selected to test the three realms of coordination: manual dexterity (placing pegs task), ball skills (throwing bean bag	At age 7-8 years (assumed to be chronological) Abnormal heel-to-toe score Term: Reference Moderate/late preterm: OR 1.27 (0.98-1.63) Abnormal bean-bag score Term: Reference	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		With data on abnormal heel-to-toe score: n=331 preterm n=6501 full-term With data on abnormal bean-bag score: n=332 preterm n=6512 full-term With data on abnormal peg-score and abnormal coordination summary score: n=328 preterm n=6414 full-term	at birth, mode of delivery, maternal hypertension, pyrexia and need for resuscitation at birth.	into box) and balance (heel-toe walking). A summary score of all three tests was derived (range 0-15). The top 5th centile was used to define severe motor coordination difficulties.	Moderate/late preterm: OR 1.17 (0.91-1.50) Abnormal peg score Term: Reference Moderate/late preterm: OR 1.40 (1.08-1.81) Abnormal coordination summary score Term: Reference Moderate/late preterm: OR 1.39 (1.12-1.72)	
Rautava 2010 (Finland)	Population based cohort study	n=588 preterm (<32 weeks gestation) and/or VLBW (≤1500g) children n=176 term controls (38-42 weeks gestation)	Sex, family structure and the mother's and father's years of education and employment status.	The FTF was used to assess behavioural outcomes. Results are presented as rate ratios comparing mean scores in preterm/VLBW children to controls.	At 5 years of age (chronological) Motor skills Term: Reference Preterm: RR 2.22 (1.83-2.69) Gross motor skills Term: Reference Preterm: RR 2.89 (2.16-3.86) Fine motor skills Term: Reference Preterm: RR 1.91 (1.59-2.30) Language	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Term: Reference Preterm: RR 1.64 (1.33-2.01) Comprehension Term: Reference Preterm: RR 1.61 (1.25-2.07) Expressive language skills Term: Reference Preterm: RR 1.65 (1.31-2.07) Communication Term: Reference Preterm: RR 1.76 (1.30-2.38)	
Schendel 1997 (USA)	Regional prospective cohort study	n=367 VLBW children (<1500g) n=555 NBW children (≥2500g) n=524 MLBW children (1500-2499g) Note that small number of participants were born prior to 1990 (study dates 12/1989-03/1991).	Adjusted for gender, maternal age, maternal education, maternal race, marital status, Medicaid use, maternal residence, maternal smoking and alcohol intake.	The Denver II was used to screen children for possible developmental delay. Nine outcomes were used in this analysis. Eight of the outcomes were based on two measures of performance in each of four domains: personal-social, language, fine motor adaptive skills and gross motor skills. One of the two domain specific measures was whether the child failed a task in each	At 9-34 months Risk of questionable overall performance (>=2 cautions) NBW: Reference VLBW: OR 2.74 (1.74-4.31) MLBW: Reference VLBW: OR 1.66 (1.09-2.51) Risk of abnormal overall performance (>=2 delays) NBW: Reference VLBW: OR 4.81 (2.51-9.23) MLBW: Reference VLBW: OR 2.02 (1.18-3.45) Risk of ≥ 1 caution in language outcomes NBW: Reference VLBW: OR 2.16 (1.39-3.37)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				domain for which 75-90% of children of the same (adjusted) age would pass. This was denoted as receiving a caution score in a given domain. The other measure was whether a child failed on or more tasks in each domain for which at least 90% of children of the same age would be expected to pass (denoted as receiving a delay score in that domain).	MLBW: Reference VLBW: OR 1.41 (0.93-2.12) Risk of ≥ 1 delay in language outcomes NBW: Reference VLBW: OR 2.97 (1.61-5.47) MLBW: Reference VLBW: OR 1.79 (1.04-3.09) Risk of ≥ 1 caution in fine motor-adaptive outcomes NBW: Reference VLBW: OR 2.10 (1.26-3.50) MLBW: Reference VLBW: OR 1.42 (0.88-2.28) Risk of ≥ 1 delay in fine motor-adaptive outcomes NBW: Reference VLBW: OR 4.88 (2.34-10.20) MLBW: Reference VLBW: OR 1.6 (0.9-2.84) Risk of ≥ 1 caution in gross motor outcomes NBW: Reference VLBW: OR 4.95 (2.89-8.47) MLBW: Reference VLBW: OR 2.16 (1.39-3.34) Risk of ≥ 1 delay in gross motor outcomes NBW: Reference VLBW: OR 6.26 (2.87-13.65)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					MLBW: Reference VLBW: OR 2.54 (1.38-4.68) Risk of >=1 caution in personal-social outcomes NBW: Reference VLBW: OR 2.12 (1.38-3.24) MLBW: Reference VLBW: OR 1.64 (1.09-2.48) Risk of >=1 delay in personal-social outcomes NBW: Reference VLBW: OR 3.21 (1.51-6.68) MLBW: Reference VLBW: OR 2.74 (1.36-5.53)	
Stene-Larsen 2014 (Norway)	Prospective population based cohort study.	Sample recruited n=101624 (Original sample in Mother and Birth Cohort Study) Sample analysed after exclusions n=32314 children (n=1673 children born at 34-36 weeks; n=30641 children born at 39-41 weeks)	Emergency Caesarean delivery, maternal gestational diabetes, preeclampsia/HELLP syndrome, multiple gestation, small for gestational age, 5 minute Apgar score ≤6, diagnosis of respiratory distress or intracranial bleeding and use of mechanical ventilation after birth.	Child communication impairments at the age of 18 months were measured using 3 specifically selected items from the Ages and Stages Questionnaire (ASQ), as rated by the child's mother. Two of these assess receptive communication skills and the other assesses expressive communication skills. To identify children at risk for clinically significant communication impairments, a cutoff	At 18 months of age Communication impairments Term: Reference 34-36 weeks: OR 1.74 (1.41-2.14) At 36 months of age Communication impairments Term: Reference 34-36 weeks: OR 1.19 (0.96-1.47) Expressive language impairments Term: Reference 34-36 weeks: OR 1.37 (1.09-1.73)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				of 2SD above the cohort mean was set. Communication impairments at 36 months were assessed using 6 items from the ASQ measuring expressive (3 items) and receptive (3 items) communication skills, as rated by the child's mother. A cut off of 2SD above the cohort mean was set to identify children at risk. Expressive communication impairment was measured using the parent-based assessment of grammar abilities (Dale 2003). Mothers are asked to select which category best describes how their child talks: (1) not yet talking, (2) talking, but not understandably, (3) talking in single word utterances, such as "milk", (4) child is talking in 2-3 word phrases, such as "me		

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				got ball", (5) child is talking in fairly complete sentences, such as "can I go outside?" and (6) child is talking in long and complicated sentences, such as "when I went to the park, I went on the swings". The measure was dichotomised so that a score of ≥5 was coded 0 and a score of ≤4 was coded 1.		
Woythaler 2011 (US)	Prospective national cohort study.	n=1200 late preterm babies n=6300 term babies	Gestational age, plurality, maternal race, education, marital status, depression, prenatal care, primary language, infant gender, poverty level, delivery type, fetal growth and any breast milk feeding.	Psychomotor development index (PDI) using the Bayley Short Form Research edition (BSF-R). This was administered in the child's home by trained personnel. Each administrator's testing and scoring were validate through in person quality control visits and videotaped interviews. Score of <70 considered as a delay.	At 24 months (chronological age) Risk of severe psychomotor developmental delay (PDI score <70) Term: Reference Late preterm: OR 1.56 (1.29-1.88) Risk of mild psychomotor developmental delay (PDI score 70-84) Term: Reference Late preterm: OR 1.58 (1.37-1.83)	Moderate
Executive function						

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Farooqi 2016 (Sweden)	Regional prospective cohort study	N=134 extremely preterm infants (<26 weeks') N=103 term infants	Adjusted for sex, composite social risk, and mother's country of origin.	Executive function (cognitive function and behavioural function) was measured using the following tests: Six core subsets were selected from Wechsler Intelligence Scale for Children (WISC-III-R) to assess general intelligence (full scale IQ), cognitive assessment (inhibition, working memory and shifting strategy) related to executive function Tower test of Delis-Kaplan Executive Function Scale (D-KEFS) was used to visual attention and visual spatial skills (spatial planning, rule learning, Inhibition, establishing and maintaining cognitive set/problem solving) To assess behavioural parameters related to executive function, parts of the Five to Fifteen (FTF) were	At 10 to 15 years (chronological age) Executive function (EPT (23-25 weeks GA, total) vs control, in total population, scoring <-2SD on WISC-III-R): Verbal working memory (digit span): OR 12.8 (95%CI 3-56) Non-verbal memory (coding): OR 10.0 (95%CI 2.9-35.0) Spatial conceptualisation (block design): OR 18.0 (95%CI 4-77) Visual reasoning (picture arrangement): OR 4.7 (95%CI 1.8-12.7) Planning ability (Tower test): OR 26.0 (95%CI 3.4-192) Executive function (EPT (23-25 weeks GA) vs control, in those children who did not have NSI and had FSIQ >70) (scoring <-2SD on WISC-III-R) Verbal working memory (digit span): OR 3.6 (95%CI 0.7-19) Non-verbal memory (coding): OR 5.5 (95%CI 1.1-27)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				used to assess attention, hyperactivity/impulsivity, hypoactivity, planning/organisation, and working memory. The domains of the parent and teacher FTF were collapsed into a primary Executive Function Composite Score (EFCS) domain) The learning skills domain from the FTF was used to assess learning skills (teacher and parent reported) in school subjects (maths, reading and writing, as well as coping in learning). Impairments in the inattention individual domains of executive function and learning skills were defined as 2 SD (>95th percentile) greater than the normative mean in the parent FTF or 2SD above the mean z scores for controls in the teacher FTF,	Memory, attention, distractibility (Arithmetic): OR 7.9 (95%CI 1.7-37) Visual reasoning (picture arrangement): OR 2.1 (95%CI 0.6-7.3) Planning ability (Tower test): P 0.007 Spatial conceptualisation (block design): P <0.001 Behavioural assessment (EPT (23-25 weeks GA) vs control, in total population, scoring >2SD on FTF) Executive function composite score (parent): OR 16.1 (95%CI 2.1-122.1) Executive function composite score (teacher): OR 5.7 (95%CI 2.1-15.4) Attention (parent): OR 13.5 (95%CI 1.8-104.0) Attention (teacher): OR 5.6 (95%CI 2.2-14.0) Hyperactivity/impulsivity (parent): P <0.001 Hyperactivity/impulsivity (teacher): OR 2.6 (95%CI 0.95-67.0) Hypoactivity (parent): OR 4.4 (95%CI 1.2-15.7)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				corresponding to significant difficulties	Hypoactivity (teacher): OR 5.0 (95%CI 1.8-13.8) Planning/organisation (parent): OR 4.6 (95%CI 1.9-10.9) Planning/organisation (teacher): OR 8.6 (95%CI 2.9-25.4) Working memory (parent): OR 5.6 (95%CI 1.9-16.8) Working memory (teacher): OR 9.6 (95%CI 3.3-28.6) Behavioural assessment (EPT (23-25 weeks GA) vs control, in those children who did not have NSI and had FSIQ>70, scoring >2SD above mean on FTF) Executive function composite score (parent): P= 0.003 Executive function composite score (teacher): OR 5.8 (95%CI 1.6-21.1) Attention (parent): P= 0.002 Attention (teacher): OR 4.2 (95%CI 1.5-11.9) Hyperactivity/impulsivity (parent): P=0.007 Hyperactivity/impulsivity (teacher):	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					OR 1.8 (95%CI 0.85-6.0), P=0.35 Hypoactivity (parent): OR 10.7 (95%CI 1.3-89.9) Hypoactivity (teacher): OR 6.3 (95%CI 1.8-22.4) Planning/organisation (parent): OR 3.3 (95%CI 1.2-9.6) Planning/organisation (teacher): OR 6.7 (95%CI 1.8-24.2) Working memory (parent): OR 10.2 (95%CI 1.3-83.2) Working memory (teacher): OR 9.9 (95%CI 2.1-45.0) Learning skills (EPT (23-25 weeks GA) vs control, in those children who did not have NSI and had FSIQ >70, scoring >2SD on FTF) Reading/writing (parent): OR 12.5 (95%CI 1.6-99.1) Reading/writing (teacher): OR 3.6 (95%CI 1.3-9.7) Mathematics (parent): OR 21.4 (95%CI 2.8-165.2) Mathematics (teacher): OR 8.8 (95%CI 3.5-22.2) General learning (parent): P <0.001 General learning (teacher): OR 18.2 (95%CI 2.3-142.6)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Farooqi 2013 (Sweden)	Population based prospective cohort study	n=83 preterm children (<26 weeks') n=86 term controls	Gender, social risk and family function.	The FTF questionnaire was used to assess aspects of executive function and attention/hyperactivity. Scores of >2SD above the mean were considered problem scores.	At 11 years of age (assumed to be chronological age) Total population Hypoactivity problems Term: Reference Preterm: OR 1.5 (0.5-4.5)† Preterm: OR 3.8 (1.2-12.2)‡ Planning/Organising problems Term: Reference Preterm: OR 5.9 (2.1-16.9) † Preterm: OR 4.7 (1.6-13.4)‡ Working memory problems Term: Reference Preterm: OR 8.6 (1.8-39.7)† Preterm: OR 5.5 (2.1-14.5)‡ Population after excluding those with neurosensory impairment Hypoactivity problems Term: Reference Preterm: OR 1.6 (0.47-5.3)† Preterm: OR 5.1 (1.3-19.1)‡ Planning/Organising problems Term: Reference Preterm: OR 5.03 (1.6-16.2) † Preterm: OR 5.9 (1.8-18.8)‡ Working memory problems	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Term: Reference Preterm: OR 14.2 (1.7-116.2)† Preterm: OR 6.6 (2.4-18.8)‡ † as rated by parents ‡ as rated by teachers	
Rautava 2010 (Finland)	Population based cohort study	n=588 preterm (<32 weeks gestation) and/or VLBW (≤1500g) children n=176 term controls (38-42 weeks gestation)	Sex, family structure and the mother's and father's years of education and employment status.	The FTF was used to assess behavioural outcomes. Results are presented as rate ratios comparing mean scores in preterm/VLBW children to controls.	At 5 years of age (chronological) Planning/Organising problems Term: Reference Preterm: RR 1.34 (1.07-1.68) Memory problems Term: Reference Preterm: RR 1.26 (1.01-1.58)	Moderate
Behavioural, social, emotional or attention problems						
de Jong 2015 (The Netherlands)	Multicentre prospective cohort	n=116 moderately preterm children (32-36 weeks gestation) n=99 term children (37-41 weeks gestation)	Analyses were adjusted for maternal education level and maternal age at birth.	Behavioural problems were assessed with the CBCL. For total problems and broadband scales, scores of 60 or above were considered abnormal. For the subscales, scores of 65 or above were considered abnormal.	At 24 months (corrected age) Total behavioural problems Term: Reference 32-36wks: OR 1.37 (0.31-6.02) Internalising problems Term: Reference 32-36wks: OR 3.70 (0.41-33.09) Externalising problems Term: Reference 32-36wks: OR 1.88 (0.54-6.54)	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Emotionally reactive Term: Reference 32-36wks: OR 3.70 (0.40-34.22) Somatic complaints Term: Reference 32-36wks: OR 2.26 (0.58-8.83) Withdrawn Term: Reference 32-36wks: OR 0.76 (0.04-15.14) Attention problems Term: Reference 32-36wks: OR 1.06 (0.28-4.04)	
Delobel-Ayoub 2006 (France)	Population based prospective cohort study (EPIPAGE)	n=1228 preterm babies born at 22-32 weeks n=447 term controls born at 39-40 weeks	For the comparison of term and preterm children, OR were adjusted for gender, maternal age at birth, birth order, maternal education, marital status of the mother, hospitalization during the last year, neurodevelopmental delay and the health of the child (assessed by the parents) at 3 years of age. For the analyses based on preterm children only	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At 3 years of age (assumed chronological) Total difficulties score Term: Reference Preterm: OR 1.9 (1.3-2.8) Hyperactivity Term: Reference Preterm: OR 1.7 (1.2-2.5) Conduct problems Term: Reference Preterm: OR 1.6 (1.1-2.3) Emotional symptoms Term: Reference Preterm: OR 1.4 (1.0-2.1) Peer problems	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			OR were also adjusted for gestational age, cerebral lesions and hospitalization in NICU ≥13 weeks.		Term: Reference Preterm: OR 1.5 (1.0-2.3) Within the preterm cohort only Gestational age Total difficulties score 31-32 weeks: Reference 29-30 weeks: OR 0.9 (0.6-1.3) 24-28 weeks: OR 1.4 (0.9-2.2)	
Delobel-Ayoub 2009 (France)	Population based prospective cohort study (EPIPAGE)	n=1102 preterm babies born at 22-32 weeks n=375 term controls born at 39-40 weeks	All outcomes adjusted for cognitive performance, maternal age at birth, development of the child (assessed by the parents), hospitalisations between birth and 5 years and mental wellbeing of the mother during the previous month. For the analyses comparing preterm and term children, OR were also adjusted for the health of the child.	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At age 5 years (assumed chronological age) Total difficulties score Term: Reference Preterm: OR 1.8 (1.2-2.8) Within the preterm cohort Total difficulties score Gestational age (24-26 weeks, 27-28 weeks, 29-30 weeks, 31-32 weeks (ref)) Not significant on univariate analysis	Moderate
Farooqi 2013 (Sweden)	Population based prospective cohort study	n=83 preterm children (<26 weeks')	Gender, social risk and family function.	The FTF questionnaire was used to assess aspects of executive	At 11 years of age Total population Attention problems	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		n=86 term controls		function and attention/hyperactivity. Scores of >2SD above the mean were considered problem scores.	Term: Reference Preterm: OR 2.8 (0.81-9.6)† Preterm: OR 4.2 (1.3-13.5)‡ Hyperactivity/impulsivity problems Term: Reference Preterm: OR 2.3 (0.72-7.2)† Preterm: OR 2.7 (0.7-10.9)‡ Population after excluding those with neurosensory impairment Attention problems Term: Reference Preterm: OR 2.5 (0.6-11.2)† Preterm: OR 5.2 (1.4-19.7)‡ Hyperactivity/impulsivity problems Term: Reference Preterm: OR 1.8 (0.48-6.9)† Preterm: OR 2.0 (0.5-9.1)‡ † as rated by parents ‡ as rated by teachers	
Farooqi 2007 (Sweden)	Nationally-representative population-based cohort study	n=169 total sample n=83 extremely immature (EI) children born before 26 completed weeks of gestation n=86 control children with normal birth	Sex, social risk, family function, maternal mental health risk score, and presence of a chronic medical condition.	Parents completed the Child Behavior Checklist (CBCL) for ages 4 to 18 years and the teachers completed the analogous Teacher Report Form (TRF). Both forms include 118 items for scoring particular	At 11 years Anxious/depressed Term: Reference <26 week: OR 2.56 (1.06-6.18) † <26 week: OR 3.54 (1.39-9.03) ‡ Withdrawn Term: Reference <26 week: OR 2.9 (1.27-6.63) †	High

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		weight born at term at the same hospital, of the same gender and nearest in birth date (7 days) to the extremely immature child.		behaviour/emotional problems, plus 2 open-ended problem items. Principal-component analyses reveal 8 sets of behaviours: withdrawn, somatic complaints, anxious or depressed, social problems, thought problems, attention problems, delinquent behaviour, and aggressive behaviour. Principal-factor analyses of the 8 categories produce 2 broad groupings, namely, internalizing, derived from the sum of the items in the first 3 sets, and externalizing, derived from the last 2 (delinquent behaviour and aggressive behaviour). The remaining 3 categories (social, thought, and attention problems) represent problems that fit either broad grouping. Scores above the 90th percentile for the	<26 week: OR 3.15 (1.25–8.0) ‡ Somatic complaints Term: Reference <26 week: OR 1.26 (0.42-3.72) † <26 week: OR 3.94 (1.37–11.32) ‡ Social problems Term: Reference <26 week: OR 1.92 (0.79-4.63) † <26 week: OR 2.86 (1.08–7.58) ‡ Thought problems Term: Reference <26 week: OR 1.78 (0.71-4.5) † <26 week: OR 5.04 (1.87–13.61) ‡ Attention problems Term: Reference <26 week: OR 3.46 (1.40-8.54) † <26 week: OR 3.43 (1.26–9.35) ‡ Aggressive behaviour Term: Reference <26 week: OR 0.99 (0.36-2.73) † <26 week: OR 1.33 (0.53–3.33) ‡ Delinquent behaviours Term: Reference <26 week: OR 0.87 (0.31-	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				control subjects of the same gender were classified as being in the abnormal range. Children completed a self-report with a depression self-rating scale (DSRS).32 The DSRS is an 18-item self-report questionnaire composed of a psychiatric symptom checklist that measures anxiety and depression. Scores above the 90th percentile for the control subjects of the same gender were classified as being in the abnormal range.	2.49) † <26 week: OR 2.20 (0.89–5.45) ‡ Internalising behaviours Term: Reference <26 week: OR 3.35 (1.38-8.11) † <26 week: OR 3.51 (1.41–8.78) ‡ Externalising behaviours Term: Reference <26 week: OR 0.76 (0.22-2.61) † <26 week: OR 1.76 (0.65–4.76) ‡ Total problems Term: Reference <26 week: OR 2.86 (1.17-7.0) † <26 week: OR 3.1 (1.19–8.07) ‡ † as rated by parents ‡ as rated by teachers	
Fevang 2016 (Norway)	National prospective cohort study	n=216 extremely preterm/extremely low birth weight (EP/ELBW) children (born at <28 weeks of gestation or with birth weight <1000 g) n=1767 reference children with	Father's educational status.	The Autism Spectrum Screening Questionnaire (ASSQ) consists of 27 items reflecting symptoms of ASD. The Swanson, Noland, and Pelham Questionnaire, Revision IV (SNAP-IV) is a screening tool for ADHD.	Assessed at 11 years Autism spectrum disorder symptoms (ASSQ >=95th percentile) Parent report Term: reference EP/ELBW: OR 2.3 (1.4-3.8) Teacher report Term: reference EO/ELBW: OR 6.6 (4.3-10)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		parental reported data and n=1880 reference children with teacher reported data		A 5-item parental version of SCARED to assess anxiety symptoms. Five unvalidated OCD questions derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition and International Classification of Diseases, 10th Edition guidelines were used. The Strength and Difficulties Questionnaire (SDQ) is a general behavioural screening. These items are collapsed to form the total difficulties score. The Screen for Child Anxiety Related Emotional Disorders (SCARED) and the Symptoms of Obsessive-Compulsive Disorder questionnaires were completed by parents, and the other questionnaires	Inattention symptoms (SNAP-IV) Parent report Term: reference EP/ELBE: OR 4.8 (3.2-7.6) Teacher report Term: reference EP/ELBE: OR 5.6 (3.6-8.7) Hyperactivity/impulsivity symptoms (SNAP-IV) Parent report Term: reference EP/ELBE: OR 3.3 (2.1-5.2) Teacher report Term:reference EP/ELBW: OR 2.7 (1.6-4.6) Anxiety symptoms (SCARED) Parent report Term: reference EP/ELBW: OR 2.3 (1.4-3.7) OCD symptoms Parent report Term: reference EP/ELBW: OR 2.6 (1.6-4.3) SDQ total difficulties Parent report Term: reference EP/ELBW: OR 3.1 (2.1-4.6) Teacher report Term: reference EP/ELBW: OR 4.0 (2.7-5.8)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				by both parents and teachers. A scale score ≥95th percentile for the reference group was classified as a high score for all the questionnaires except for the Strengths and Difficulties Questionnaire (SDQ), for which the total difficulties score ≥90th percentile (TDS90) is accepted as a high score.		
Gurka 2010 (US)	Prospective cohort study	n=1298 (of which n=53 born at 34-36 weeks of gestation, the rest at term)	Child race (white vs non-white), maternal age (in years), maternal education (in years), whether the mother experiences health problems during the pregnancy, delivery type (vaginal vs caesarean), mean Home Observation for Measurement of the Environment scores during the first 3 years of life (a measure of the quality of the home environment), mean	Behavioural and emotional problems: externalising behaviours; internalising behaviours; aggressive behaviours; anxiety/depression, assessed with the Child Behaviour Checklist (CBCL) completed by parents. The CBCL has been age-standardized on large samples of children in the US and abroad. Each of the 118	From 4 to 15 years of age (full-term vs late-preterm): External behaviours: No significant difference between the groups over time. Internal behaviours: No significant difference between the groups over time. Aggressive behaviours: No significant difference between the groups over time. Anxiety/depression: No significant difference between the groups over time.	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			maternal depression scores (Center for Epidemiological Studies-Depression Scales) during the first 3 years of the child's life, and the mother's verbal ability, assessed using the Peabody Picture Vocabulary Test-Revised.	problem items is scored on a Likert scale based on the preceding 6 months. Scores on each item are summed to give a raw total problem score, which is then converted to a T-score (mean [SD]=50 [10]). Higher scores indicare more behavioral and emotional problems. Four of the scales in the study were used in the study to examine behavioural and emotional functioning.		
Higa Diez 2016 (Japan)	Population-based national longitudinal cohort study (Longitudinal Survey of Babies in the 21st Century)	n=34163 (total sample) n=356 children born at <34 weeks n=1287 children born at 34-36 weeks n=children born at 37-38 weeks (results not presented) n=children born at 39-41 weeks (reference group)	Sex, singleton or not, maternal age at delivery, maternal education attainment and maternal smoking status.	Parents filled in the Child Behaviour Checklist (CBCL) 4-18 for Japan. A total of seven behavioural outcomes were assessed, three in relation to attention problems: interrupting people; inability for the child to wait for his/her turn during play; failure to pay attention to the surrounding area when crossing a	At 8 years Attentional problems: Interrupting people 39-41 weeks: Reference 34-36 weeks: OR 1.05 (0.93-1.19) <34 weeks: OR 1.10 (0.89-1.38) Inability to wait his/her turn: 39-41 weeks: Reference 34-36 weeks: OR 1.28 (1.03-1.59) <34 weeks: OR 1.72 (1.22-2.43)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				street, and four in relation to delinquent/aggressive behaviours: lying; destroying toys and/or books; hurting other people; causing disturbance in public.	Failure to pay attention crossing street: 39-41 weeks: Reference 34-36 weeks: OR 0.98 (0.85-1.14) <34 weeks: OR 1.09 (0.84-1.42) Subjects who presented adverse outcomes for all attentional problems: 39-41 weeks: Reference 34-36 weeks: OR 1.43 (0.98-2.09) <34 weeks: OR 2.21 (1.24-3 95) Delinquent/aggressive behaviours: Lying 39-41 weeks: Reference 34-36 weeks: OR 1.10 (0.96-1.26) <34 weeks: OR 1.15 (0.96-1.46) Destroying toys/books 39-41 weeks: Reference 34-36 weeks: OR 1.15 (0.95-1.39) <34 weeks: OR 1.46 (1.07-1.99) Hurting other people 39-41 weeks: Reference 34-36 weeks: OR 1.08 (0.90-1.29)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					<34 weeks: OR 1.23 (0.90-1.69) Disturbance in public 39-41 weeks: Reference 34-36 weeks: OR 1.20 (1.04-1.38) <34 weeks: OR 1.14 (0.89-1.48) Subjects who presented adverse outcomes for all delinquent/aggressive behaviours 39-41 weeks: Reference 34-36 weeks: OR 1.02 (0.63-1.65) <34 weeks: OR 1.46 (0.71-3.00)	
Hornman 2016 (The Netherlands)	Multicentre prospective cohort study (Lollipop)	n=1054 preterm children (<36 weeks) (n=653 moderately preterm children [32-35 weeks] n=401 early preterm children [25-31 weeks]) n=389 term children as comparisons	Gender, SGA, smoking during pregnancy, being part of a multiple pregnancy, multiparity, low education level of parents, and 1-parent family.	Emotional and behavioural problems were assessed with the validated Dutch version of the Child Behaviour Checklist (CBCL), applicable for ages 1.5-5 years. The CBCL consists of 99 problem items, each item can be rated by the parents as not true (0), somewhat/sometimes true (1), or very/often true (2). From these ratings, the total,	At age 4 and 5 years Total emotional/behavioural problems (CBCL >=84th percentile) Emerging problems (normal score at 4 y, abnormal at 5 y) Term: Reference <36 weeks: OR 1.58 (0.71-3.49) 32-35 weeks: OR 1.42 (0.62-3.27) 25-31 weeks: OR 1.88 (0.78-4.52)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				internalising, and externalising problem scales were constructed. >=84th percentile of the scale was considered subclinical or clinical. The dichotomised CBCL outcomes at ages 4 and 5 years were combined, resulting in 4 categories: consistently normal (normal score at both 4 and 5 years), emerging problems (normal score at 4 years, abnormal score at 5 years), resolving problems (abnormal score at 4 years, normal score at 5 years), and persistent problems (abnormal score at both 4 and 5 years).	Resolving problems (abnormal score at 4 y, normal score at 5 y) Term: Reference <36 weeks: OR 2.71 (1.43-5.15) 32-35 weeks: OR 3.10 (1.61-5.96) 25-31 weeks: OR 1.94 (0.92-4.12) Persistent problems (abnormal score at both 4 and 5 y) Term: Reference <36 weeks: OR 2.02 (1.07-3.81) 32-35 weeks: OR 1.93 (0.99-3.74) 25-31 weeks: OR 2.17 (1.07-4.41) Internalising problems (CBCL >=84th percentile) Emerging problems (normal score at 4 y, abnormal at 5 y) Term: Reference <36 weeks: OR 1.23 (0.72-2.09) 32-35 weeks: OR 1.17 (0.67-2.05) 25-31 weeks: OR 1.34 (0.73-2.49)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Resolving problems (abnormal score at 4 y, normal score at 5 y) Term: Reference <36 weeks: OR 2.18 (1.16-4.09) 32-35 weeks: OR 2.16 (1.13-4.15) 25-31 weeks: OR 2.22 (1.09-4.51) Persistent problems (abnormal score at both 4 and 5 y) Term: Reference <36 weeks: OR 2.04 (1.21-3.45) 32-35 weeks: OR 1.90 (1.10-3.29) 25-31 weeks: OR 2.31 (1.28-4.17) Externalising problems (CBCL >=84th percentile) Emerging problems (normal score at 4 y, abnormal at 5 y) Term: Reference <36 weeks: OR 2.54 (1.21-5.32) 32-35 weeks: OR 2.63 (1.23-5.63) 25-31 weeks: OR 2.37 (1.03-5.47)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Resolving problems (abnormal score at 4 y, normal score at 5 y) Term: Reference <36 weeks: OR 1.59 (0.90-2.81) 32-35 weeks: OR 1.85 (1.03-3.32) 25-31 weeks: OR 1.07 (0.53-2.17) Persistent problems (abnormal score at both 4 and 5 y) Term: Reference <36 weeks: OR 2.25 (1.26-4.03) 32-35 weeks: OR 2.31 (1.26-4.23) 25-31 weeks: OR 2.14 (1.10-4.15)	
Johnson 2015b (UK)	Prospective population-based cohort study	n=625 late and moderately preterm (LMPT, 32-36 weeks) n=760 term controls	Age, sex, SES-index category, SGA, infant cognitive impairment.	To assess behavioural outcomes, parents completed the Brief Infant Toddler Social Emotional Assessment (BITSEA). The BITSEA “problem scale” comprises 31 items that assess behaviour problems in the areas of externalizing problems,	At 2 years (corrected age) Behaviour problem Term: Reference 32-36 weeks: RR 1.13 (0.8-1.42) Delayed competence Term: Reference 32-36 weeks: RR 1.28 (1.03-1.58) Problem or delay Term: Reference 32-36 weeks: RR 1.17 (1.00-1.38)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				internalizing difficulties, dysregulation, maladaptive behaviours, and atypical behaviours. The BITSEA "competence scale" comprises 11 items that assess areas of attention, compliance, mastery motivation, prosocial peer relations, empathy, imitation/play skills, and social relatedness and is designed to identify children who have delays or deficits in the acquisition of social-emotional competencies (irrespective of whether behaviour problems are present).	Problem and delay Term: Reference 32-36 weeks: RR 1.34 (0.91-1.97)	
Potijk 2015 (The Netherlands)	Multicentre prospective cohort study	n=915 moderately preterm children (32-35+6 weeks gestation) n=543 term children (38-	Socioeconomic status, gestational age, gender, number of siblings and maternal age.	The Dutch version of the CBCL was used to identify behavioural problems. The authors state that "American cut-offs"	At age 4 years (assumed to be chronological) Total behavioural problems GA: OR 1.24 (1.00-1.56) Externalising problems GA: OR 1.31 (1.05-1.63) Internalising problems	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		41+6 weeks gestation)		were used to identify problem scores.	GA: OR 1.41 (1.13-1.73) OR represent the risk per SD decrease in GA.	
Rautava 2010 (Finland)	Population based cohort study	n=588 preterm (<32 weeks gestation) and/or VLBW (≤1500g) children n=176 term controls (38-42 weeks gestation)	Sex, family structure and the mother's and father's years of education and employment status.	The FTF was used to assess behavioural outcomes. Results are presented as rate ratios comparing mean scores in preterm/VLBW children to controls.	At 5 years of age (chronological) Hyperactive/impulsive Term: Reference Preterm: RR 1.28 (1.07-1.53) Attention Term: Reference Preterm: RR 1.81 (1.47-2.23) Emotional/behavioural problems Term: Reference Preterm: RR 1.49 (1.20-1.84) Internalising Term: Reference Preterm: RR 1.56 (1.19-2.05) Externalising Term: Reference Preterm: RR 1.39 (1.09-1.78) Obsessive compulsive Term: Reference Preterm: RR 1.79 (1.22-2.62)	Moderate
Reijneveld 2006	Population based cohort study	n=402 preterm (<32 weeks) and/or VLBW	Adjustment was performed for gender, family composition,	The CBCL was used to assess behavioural outcomes. Results	At 5 years of age (assumed to be chronological)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
(The Netherlands)		(<1500g) children n=6007 reference children from the general population	number of siblings and maternal educational level. The authors state that no important differences were noted, therefore unadjusted results are reported.	were dichotomised into clinical ranges at the 97th percentile for the individual syndrome scales, and at the 90th percentile for the total problems score and internalising/ externalising scales.	Total problems General population: Reference Preterm/VLBW:OR 1.60 (1.18-2.17) Internalising problems General population: Reference Preterm/VLBW: OR 1.06 (0.71-1.57) Externalising problems General population: Reference Preterm/VLBW: OR 1.48 (1.08-2.03) Withdrawn General population: Reference Preterm/VLBW:OR 1.72 (0.82-3.60) Somatic complaints General population: Reference Preterm/VLBW:OR 1.90 (1.10-3.28) Anxious/depressed General population: Reference Preterm/VLBW:OR 1.15 (0.41-3.20) Social problems General population: Reference	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Preterm/VLBW:OR 2.62 (1.38-5.16) Thought problems General population: Reference Preterm/VLBW:OR 2.72 (1.49-4.94) Attention problems General population: Reference Preterm/VLBW:OR 3.45 (2.02-5.89) Delinquent behaviour General population: Reference Preterm/VLBW:OR 2.65 (1.39-5.08) Aggressive behaviour General population: Reference Preterm/VLBW:OR 1.58 (0.90-2.77) Sex problems General population: Reference Preterm/VLBW:OR 1.48 (0.68-3.24)	
Special educational needs						
Chan 2014 (UK)	A nationally representative longitudinal study (The Millennium	n=6031	Sex, child's age in school year taking into account premature children who if born at full	School performance was investigated using the statutory Key Stage 1 (KS1) teacher assessments	At 7 years of age KS1 overall Term (39-40 weeks): Reference	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
	Cohort Study (MCS))		term would have been placed in the year below, multiple birth, firstborn status, mother's age, mother's education, mother's social class, marital status, smoking during pregnancy.	performed in the third school year in England. At KS1, children generally perform between level 1 (below expected level) to level 3 (considerably above the expected level), with adequate performance categorised as achieving level 2 or above. KS1 results were obtained from the Department of Education's National Pupil Database.	<32 weeks: OR 1.78 (1.24-2.54) 32-33wks: OR 1.71 (1.15-2.54) 34-36 weeks: OR 1.36 (1.09-1.68) KS1 reading Term (39-40 weeks): Reference <32 weeks: OR 1.84 (1.12-3.05) 32-33 weeks: OR 1.82 (1.12-2.98) 34-36 weeks: OR 1.55 (1.2-2) KS1 writing Term (39-40 weeks): Reference <32 weeks: OR 1.82 (1.24-2.68) 32-33 weeks: OR 1.69 (1.14-2.5) 34-36 weeks: OR 1.35 (1.07-1.71) KS1 speaking and listening Term (39-40 weeks): Reference <32 weeks: OR 2.48 (1.63-3.78) 32-33 weeks: OR 1.58 (0.79-3.17) 34-36 weeks: OR 1.36 (0.96-1.94) KS1 mathematics	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Term (39-40 weeks): Reference <32 weeks: OR 1.89 (0.92-3.64) 32-33 weeks: OR 1.96 (0.97-3.99) 34-36 weeks: OR 1.03 (0.66-1.59) KS1 science Term (39-40 weeks): Reference <32 weeks: OR 1.87 (0.93-3.74) 32-33 weeks: OR 2.25 (1.16-4.38) 34-36 weeks: OR 1.33 (0.91-1.94)	
Larroque 2011 (France)	Population based prospective cohort (EPIPAGE)	n=1439 preterm children (22-32 weeks) n=327 term controls (39-40 weeks)	Maternal age at childbirth, parity, maternal level of education, maternal birth place, SES and sex.	Parental questionnaire was used to identify whether the child attended special schooling or had additional support at school.	At 8 years (assumed to be chronological) Risk of being in an institution or special school/class Term: Reference Preterm: OR 3.0 (0.9-9.8) Risk of being in a mainstream class with the year repeated Term: Reference Preterm: OR 4.4 (2.3-8.2) Risk of needing special care and/or support at school Term: Reference Preterm: OR 2.0 (1.5-2.6)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
MacKay 2010 (UK)	Retrospective cohort using national registry data	n=21959 preterm (24-36 weeks) n=130798 term controls (40 weeks) n=407503 total sample of the study (including 37-39 GA and >40 GA) Note that some participants were born prior to 1990 (participants aged 5 to 18 years were assessed in 2005).	Infant sex, maternal age and height, marital status, parity, birth weight centile, induction of labour, mode of delivery, year of delivery, previous spontaneous and therapeutic abortions and 5 minute Apgar score.	The 2005 school census was used to identify children with reported special educational needs.	At 5-18 years of age (assumed to be chronological) Risk of SEN according to gestational age 40 weeks : Reference 33-36 weeks : OR 1.53 (1.43-1.63) 28-32 weeks : OR 2.66 (2.38-2.97) 24-27 weeks : OR 6.92 (5.58-8.58)	Moderate
MacKay 2013 (UK)	Retrospective cohort using national registry data	n=21959 preterm (24-36 weeks) n=215935 term controls (40 - 41 weeks) Note that some participants were born prior to 1990 (participants aged 5 to 18 years were assessed in 2005).	Infant sex, maternal age and height, marital status, parity, induction of labour, mode of delivery, year of delivery, previous spontaneous and therapeutic abortions, and the 5 minute Apgar score.	The 2005 school census was used to identify children with reported special educational needs.	At 5-18 years of age(assumed chronological) Risk of sensory SEN according to gestational age 40-41wks: Reference 33-36wks: OR 1.73 (1.18-2.52) 28-32wks: OR 4.44 (2.56-7.71) 24-27wks: OR 23.64 (12.03-46.45) Risk of physical or motor SEN according to gestational age 40-41wks: Reference	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					33-36wks: OR 2.99 (2.27-3.95) 28-32wks: OR 16.01 (11.78-21.75) 24-27wks: OR 29.69 (17.49-50.40) Risk of language SEN according to gestational age 40-41wks: Reference 33-36wks: OR 1.03 (0.72-1.48) 28-32wks: OR 1.88 (0.99-3.55) 24-27wks: OR 1.64 (0.22-12.02) Risk of social, emotional or behavioural SEN according to gestational age 40-41wks: Reference 33-36wks: OR 1.34 (1.12-1.61) 28-32wks: OR 1.24 (0.80-1.92) 24-27wks: OR 1.90 (0.60-6.07) Risk of specific learning difficulties SEN according to gestational age 40-41wks: Reference 33-36wks: OR 1.26 (1.09-1.46) 28-32wks: OR 1.54 (1.13-2.12)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					24-27wks: OR 3.56 (1.80-7.05) Risk of intellectual SEN according to gestational age 40-41wks: Reference 33-36wks: OR 1.93 (1.74-2.14) 28-32wks: OR 3.11 (2.56-3.77) 24-27wks: OR 11.67 (8.46-16.10) Risk of ASD SEN according to gestational age 40-41wks: Reference 33-36wks: OR 0.93 (0.72-1.21) 28-32wks: OR 1.95 (1.29-2.96) 24-27wks: OR 2.56 (0.80-8.20) Risk of unspecified SEN according to gestational age 40-41wks: Reference 33-36wks: OR 1.56 (1.26-1.94) 28-32wks: OR 2.42 (1.60-3.65) 24-27wks: OR 5.01 (2.16-11.64)	
Odd 2016 (UK)	Regional prospective cohort study (ALSPAC)	N=775 children born at <37 weeks of gestation	Adjusted for ethnicity, maternal education, socio-economic group, age, gender, maternal parity,	Mandatory UK educational assessments done at 4 stages, the stages are Key Stage (KS) 1	At 5-7 years Low score at KS1 Matched for date of birth	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			weight at birth, length and birth, head circumference at birth, mode of birth, maternal hypertension.	at 5-7 years, KS2 at 7-11 years, KS3 at 11-14 years, and KS4 at 14-16 years. The test is done at the end of each stage. Governmental standards set the minimum standard expected at each stage of the first 3 stages and this was used as the cut-off for a low score. At the end of KS4 children take their school exams and an a-priori cut-off of 5 General Certificates of Secondary Education (GCSE) or equivalent at A* to C level was used to define a normal score at this age. At KS4, <5 passes at A* to C level was considered as poor/low attainment at KS4. Children identified as having special educational needs (SEN) in KS4 were identified from the Pupil Level Annual	Term (37-42 weeks): Reference Preterm (<37 weeks): aOR 1.44 (95% CI 1.17-1.77) At 7-11 years Low score at KS2 Matched for date of birth Term (37-42 weeks): Reference Preterm (<37 weeks): aOR 1.20 (95% CI 0.99-1.46) At 11-14 years Low score at KS3 Matched for date of birth Term (37-42 weeks): Reference Preterm (<37 weeks): aOR 1.11 (95% CI 0.91-1.35) At 14-16 years Low score at KS4 Matched for date of birth Term (37-42 weeks): Reference Preterm (<37 weeks): aOR 1.10 (95% CI 0.91-1.34) At 14-16 years SEN Matched for date of birth Term (37-42 weeks): Reference Preterm (<37 weeks): aOR 1.39 (95% CI 1.14-1.68)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				School Census (PLASC).		
Odd 2013a (UK)	Regional prospective cohort study (ALSPAC)	n=722 preterm infants (<37 weeks) n=11268 term infants (37-42 weeks) Note that these numbers represent the full cohort, but data on Low KS1 score was obtained for 11169 children and data on special educational needs was obtained for 6174 children. Numbers in different GA group not reported by outcome.	Adjusted for ethnicity, housing, crowding and maternal education, socioeconomic group, car ownership, age, gender, parity, weight, length and head circumference at birth, mode of delivery, maternal hypertension and pyrexia.	Teachers were asked to report whether the child had ever had special educational needs provision.	At 8 years of chronological age Risk of special education needs Term: Reference < 37 weeks: OR 1.57 (1.19-2.07) 32-36 weeks: OR 1.53 (1.15-2.03) < 32 weeks: OR 1.98 (0.82-4.82) At 8 years of adjusted age Risk of special education needs Term: Reference < 37 weeks: OR 1.59(1.20-2.11) 32-36 weeks: OR 1.51 (1.13-2.03) < 32 weeks: OR 2.36 (0.98-5.67)	High
Peacock 2012 (UK)	Population-based longitudinal study	n=10279 children in total (n=9683 childen born at 37-41 weeks and n=596 born at 32-36 weeks)	Sex, age at testing, birth weight z score for gestational age and gender, pregnancy size, maternal age, mode of delivery, parity, maternal smoking, maternal education	Data on Key Stage 1 assessments were obtained from local education authorities. The results for the three assessment domains (reading, writing and mathematics) were	At 5-7 years Success in KS1 overall assessment (at least level 2 in reading, writing and mathematics) Term (37-41 weeks): Reference	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			and social class, ethnicity, housing tenure and crowding, car use, family income and single parenthood.	dichotomized, with success defined as achieving at least level 2, the expected level of attainment. Overall KS1 score defined as having at least level 2 in all three domains.	Preterm (32-36 weeks): OR 0.74 (0.59-0.92) Success in KS1 reading assessment (at least level 2) Term (37-41 weeks): Reference Preterm (32-36 weeks): OR 0.74 (0.58-0.94) Success in KS1 writing assessment (at least level 2) Term (37-41 weeks): Reference Preterm (32-36 weeks): OR 0.74 (0.59-0.94) Success om KS1 mathematics assessment (at least level 2) Term (37-41 weeks): Reference Preterm (32-36 weeks): OR 0.62 (0.48-0.80)	
Quigley 2012 (UK)	Population-based cohort study	n=7650 total n=84 <32 weeks; very preterm n=92 32–33 weeks; moderately preterm n=471 34–36 weeks; late preterm n=1596 37–38 weeks; early term;	Sex, ethnicity, whether firstborn, multiple birth, breastfeeding duration, month of birth (age within the school year) and mother's age, marital status, education, social class and whether languages other than English were spoken at home.	Foundation stage profile (FSP) records the child's achievement as measured by their teacher at the end of their first school year. Teachers are trained in how to conduct the assessments, which are based on observations during the whole year. The FSP captures the	At 5 years Not good level of overall achievement 23-31 weeks: RR 1.19 (1.00-1.42) 32-33 weeks: RR 1.19 (0.98-1.45) 34-36 weeks: RR 1.12 (1.04, 1.22) 39-41 weeks: Reference Not working securely in all three scales of personal, social and emotional	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		n=5407 39–41 weeks; full term		'Early Learning Goals' as a set of 13 assessment scales across six areas of learning: 1) personal, social and emotional development, 2) communication, language and literacy, 3) mathematical development, 4) Knowledge and understanding of the world, 5) Physical development, and 6) Creative development. Also, the following categories were assessed: working securely in all the six above-mentioned areas of learning; good level of overall achievement.	development 23-31 weeks: RR 1.53 (1.16, 2.00) 32-33 weeks: RR 1.25 (0.92, 1.72) 34-36 weeks: RR 1.14 (0.99, 1.32) 39-41 weeks: Reference Not working securely in all four scales of communication, language and literacy 23-31 weeks: RR 1.17 (0.99, 1.39) 32-33 weeks: RR 1.21 (0.98, 1.48) 34-36 weeks: RR 1.11 (1.02, 1.22) 39-41 weeks: Reference Not working securely in all three scales of mathematical development 23-31 weeks: RR 1.56 (1.21, 2.01) 32-33 weeks: RR 1.35 (1.02, 1.8) 34-36 weeks: RR 1.16 (1, 1.34) 39-41 weeks: Reference Not working securely in the 'knowledge and understanding of the world' scale	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					23-31 weeks: RR 1.32 (0.9, 1.93) 32-33 weeks: RR 1.47 (0.93, 2.33) 34-36 weeks: RR 1.30 (1.08, 1.56) 39-41 weeks: Reference Not working securely in the ‘physical development’ scale 23-31 weeks: RR 1.82 (1.12, 2.96) 32-33 weeks: RR 1.64 (0.99, 2.73) 34-36 weeks: RR 1.27 (0.92, 1.74) 39-41 weeks: Reference Not working securely in the ‘creative development’ 23-31 weeks: RR 1.77 (1.3, 2.41) 32-33 weeks: RR 1.46 (0.94, 2.27) 34-36 weeks: RR 1.22 (1.02, 1.46) 39-41 weeks: Reference	

Abbreviations:AGA-appropriate for gestational age; ASD-autism spectrum disorder; ASQ-Ages and Stages Questionnaire; BMI-body mass index; BRIEF-Behaviour Rating Inventory of Executive Function; CBCL-Child Behaviour Checklist; ELBW-extremely low birth weight; FTF-Five to Fifteen questionnaire; GA-gestational age K-ABC-Kaufman Assessment Battery for Children; MABC-Movement Assessment Battery for Children; MPC-Mental Processing Composite; NBW-normal birth weight; NICU-neonatal intensive care unit; OR-odds ratio; PPVT-R- Peabody Picture Vocabulary Test-Revised; RR-relative risk; SD-standard deviation; SDQ- Strengths and Difficulties Questionnaire; SEN-special educational needs; SES-socioeconomic status; SGA-small for gestational age; VLBW-very low birth weight

Table 12: Summary of included studies on biological factors

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Functional problems with feeding/sleeping/toileting						
Johnson 2016 (UK)	Prospective population-based cohort study	n=628 late and moderately preterm (LMPT) children (32-36 weeks) n=759 term controls (>=37 weeks)	Behaviour problems, delayed social competence, SGA and nasogastric tube feeding.	A validated eating behaviour questionnaire (4) was used to assess the presence of eating difficulties in the 4 domains of refusal/picky eating (e.g., poor appetite, food refusal, selective eating), oral motor problems (e.g., problems biting, chewing, or swallowing; gagging; or choking on food), oral hypersensitivity (e.g., aversion to being touched around the mouth or having things put in the mouth), and eating behaviour problems (e.g., has tantrums or makes a mess during meals). >90th percentile of the term control group were used to identify children with clinically	At 2 years (corrected age) Total feeding problems AGA: Reference SGA: RR 1.57 (0.99-2.49)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				significant eating difficulties.		
Vohr 2000 (US)	Multicentre prospective cohort study	n=1151	Outborn status, maternal hypertension, antenatal steroids, maternal education, race, caesarean section, birth weight, surfactant, early-onset sepsis, late-onset sepsis, grades 3 and 4 IVH/PVL, chronic lung disease (oxygen requirement at 36 weeks), postnatal steroids, small for gestational age, gender, and adjusted age at time of testing.	No independent feeding, not clear how assessed but they report that a basic, functional, gross motor skills were assessed derived from the work of Russell et al. and Palisano et al.	At 18-22 months of age (corrected) No independent feeding Male (vs female): Not significant (OR (95% CI) not reported numerically) SGA (vs AGA): Not significant (OR (95% CI) not reported numerically) Race white (vs non-white): Not significant (OR (95% CI) not reported numerically)	Low
Motor, developmental and language delay						
Johnson 2015 (UK)	Prospective cohort study	n=638 late/moderately preterm infants	SES, preeclampsia, sex, breast milk at discharge.	At 2 years (corrected age), cognitive impairment was assessed using the Parent Report of Children's Abilities-Revised (PARCA-R).	At 2 years of age (corrected) Moderate/severe cognitive impairment (<2.5th percentile PARCA-R) White ethnic group: Reference Non-white ethnic group: RR 2.06 (1.10-3.83) Female: Reference Male: RR 7.04 (2.52-19.67)	Moderate
Kerstjens 2013 (The Netherlands)	Population based prospective cohort study	n=834 moderately	Maternal somatic illness, maternal mental illness,	Parents completed the Dutch version of the 48 months ASQ.	At 43-49 months (chronological age)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		preterm children (32-35 weeks)	maternal pre pregnancy obesity, in vitro fertilization, SGA, sex, multiple pregnancy, breech presentation, foetal and maternal induction of birth, Caesarean delivery, assisted delivery, SES and parity	The scores on each domain add up to an ASQ total problems score. A score of >2SDs below the mean for the Dutch reference group was considered to indicate developmental delay.	Abnormal ASQ total problems score SGA: OR 2.75 (1.25-6.08) Male sex: OR 4.20 (2.09-8.46)	
Shankaran 2004 (US)	Prospective cohort study	n=246	Neonatal brain lesions, antenatal steroid exposure, sex, ethnicity/race, household income, BPD, surfactant administration, steroids for BPD, Medicaid, no high school degree, 2-parent household.	The Bayley Scales of Infant Development (BSID-II) was used to assess Psychomotor Developmental Index (PDI). A delay in psychomotor development was considered with a PDI score <70. BSID-II was administered by clinical psychologists or psychometricians trained to reliability.	At 18-22 months of age (corrected) PDI <70 (BSID-II) Female: Reference Male: OR 1.3 (0.7-2.6) Non-black: Reference Black: OR 1.2 (0.6-2.5)	Low
Vohr 2000 (US)	Multicentre prospective cohort study	n=1151	Out born status, maternal hypertension, antenatal steroids, maternal education, race, caesarean section, birth weight, surfactant, early-onset sepsis, late-onset sepsis, grades	No independent walking, not clear how assessed but they report that a basic, functional, gross motor skills were assessed derived from the work of Russell et al.d Palisano et al.	At 18-22 months of age (corrected) No independent walking Male (vs female): Not significant (OR (95% CI) not reported numerically) SGA (vs AGA): Not significant (OR (95% CI) not reported numerically)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			3 and 4 IVH/PVL, chronic lung disease (oxygen requirement at 36 weeks), postnatal steroids, small for gestational age, gender, and adjusted age at time of testing.	Psychomotor Developmental Index (PDI) score <70, assessed with Bayley Scale of Infant Development II (BSID-II)	Race white (vs non-white): Not significant (OR (95% CI) not reported numerically) PDI <70 (Bayley-II) Male (vs female): Not significant (OR (95% CI) not reported numerically) SGA (vs AGA): Not significant (OR (95% CI) not reported numerically) Race white (vs non-white): Not significant (OR (95% CI) not reported numerically)	
Behavioural, social, emotional and attention problems						
Delobel-Ayoub 2006 (France)	Population based prospective cohort study (EPIPAGE)	n=1228 preterm babies born at 22-32 weeks	For the comparison of term and preterm children, OR were adjusted for gender, maternal age at birth, birth order, maternal education, marital status of the mother, hospitalization during the last year, neurodevelopmental delay, the health of the child (assessed by the parents) at 3 years of age, gestational age, cerebral lesions and hospitalization in NICU ≥13 weeks.	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At 3 years of age (assumed chronological) Gender Total difficulties score Female: Reference Male: OR 1.3 (0.9-1.7) SGA status Total difficulties score Not a significant predictor on univariate analysis	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Delobel-Ayoub 2009 (France)	Population based prospective cohort study (EPIPAGE)	n=1102 preterm babies born at 22-32 weeks	All outcomes adjusted for cognitive performance, maternal age at birth, development of the child (assessed by the parents), hospitalisations between birth and 5 years, health of the child and mental wellbeing of the mother during the previous month.	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At age 5 years (assumed chronological age) Gender Not significant on multivariate analysis	Moderate
Guellec 2011 (France)	Population based prospective cohort study (EPIPAGE)	n=1677 preterm babies born at 24-32 weeks	All outcomes adjusted for GA, gender, social class of the family, type of pregnancy (single versus multiple), antenatal corticosteroids, maternal age, nationality and parity.	Behavioural problems were assessed using the French version of the SDQ which was completed by the parents. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At 5 years of age (assumed chronological) 24-28 week preterm infants Inattention-hyperactivity symptoms AGA: Reference SGA: OR 1.29 (0.37-4.46) Total behavioural difficulties AGA: Reference SGA: OR 2.30 (0.82-6.48) 29-32 week preterm infants Inattention-hyperactivity symptoms AGA: Reference SGA: OR 1.78 (1.10-2.89) Total behavioural difficulties AGA: Reference SGA: OR 0.98 (0.59-1.63)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Johnson 2015b (UK)	Prospective population-based cohort study	n=625 late and moderately preterm (LMPT, 32-36 weeks) n=760 term controls	Not clearly reported. Variables that were significant (p<.05) in univariable analyses were all entered into the model. Variables that were not significant in this model were dropped in turn until only those variables significant at p <.05 were included in the final model. Variables that had been dropped were entered back into this final model one at a time to assess their significance.	Parents completed the Brief Infant Toddler Social Emotional Assessment (BITSEA). The BITSEA "competence scale" comprises 11 items that assess areas of attention, compliance, mastery motivation, prosocial peer relations, empathy, imitation/play skills, and social relatedness and is designed to identify children who have delays or deficits in the acquisition of social-emotional competencies (irrespective of whether behavior problems are present). Infants were identified as having delayed social competence if their total competence score was <15th percentile of children of the same age and sex in the BITSEA	At 2 years (corrected age) Delayed socioemotional competence Ethnicity White: Reference Non-white: RR 1.68 (1.26-2.24) Sex Female: Reference Male: RR 1.27 0.96-1.67) SGA AGA: Reference SGA: NS	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				standardization sample.		
Special educational needs						
Guellec 2011 (France)	Population based prospective cohort study (EPIPAGE)	n=1439 preterm babies born at 24-32 weeks	Adjusted for GA, gender, social class of the family, maternal age and parity.	School difficulties were defined by special schooling (institution or special school, special class in mainstream school, mainstream class) or low grades.	At age 8 years 24-28 week preterm infants School difficulties AGA: Reference SGA: OR 1.39 (0.47-4.14) 29-32 week preterm infants School difficulties AGA: Reference SGA: OR 1.74 (1.07-2.82)	Low
Johnson 2011 (UK & Ireland)	*<Insert Note here>* Population-based cohort study (EPICure Study)	n=219	Sex, gestational age, birth weight, maternal ethnicity, maternal age, maternal education, SES, antenatal steroids, preterm premature rupture of membranes, vaginal breech delivery, chorioamnionitis, fetal heart rate >100 bpm at 5 minutes, admission temperature <35c, CRIB score, NEC, postnatal steroids for chronic lung disease, any breast milk given, duration of NICU admission.	Teachers completed a questionnaire about if special educational needs (SEN) provision was utilized by the child.	At age 11 years SEN provision Female: Reference Male: OR 3.08 (1.48-6.40)	Low

Abbreviations: AGA-appropriate for gestational age; ASQ-Ages and Stages Questionnaire; GA-gestational age; K-ABC-Kaufman Assessment Battery for Children; MPC-Mental Processing Composite; NICU-neonatal intensive care unit; OR-odds ratio; SD-standard deviation; SDQ-Strengths and Difficulties Questionnaire; SGA-small for gestational age

Table 13: Summary of included studies on neonatal factors

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Functional problems in feeding/sleeping/toileting						
Vohr 2000 (US)	Multicentre prospective cohort study	n=1151	Outborn status, maternal hypertension, antenatal steroids, maternal education, race, caesarean section, birth weight, surfactant, early-onset sepsis, late-onset sepsis, grades 3 and 4 IVH/PVL, chronic lung disease (oxygen requirement at 36 weeks), postnatal steroids, small for gestational age, gender, and adjusted age at time of testing.	No independent feeding, not clear how assessed but they report that a basic, functional, gross motor skills were assessed derived from the work of Russell et al. and Palisano et al.	At 18-22 months of age (corrected) No independent feeding IVH/PVL grade III-IV: Significantly increased odds (OR (95% CI) not reported numerically) Postnatal steroids : Not significant (OR (95% CI) not reported numerically) NEC: Not significant (OR (95% CI) not reported numerically) BPD at 36 weeks: Significantly increased odds (OR (95% CI) not reported numerically) Late-onset sepsis: Not significant (OR (95% CI) not reported numerically) Early-onset sepsis: Not significant (OR (95% CI) not reported numerically) Antenatal steroids: Not significant (OR (95% CI) not reported numerically)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Motor, developmental and language delay						
Adams-Chapman 2008 (US)	19 centres of the National Institute of Child Health and Human Development Neonatal Research Network, neonatal data obtained from the Generic Database of the research network, follow-up examinations done prospectively.	n=6161 children with severe IVH or no IVH studied in depth in this study, and classified into 5 groups: 1) no IVH/no shunt n=5163 2) IVH grade 3/no shunt n=459 3) IVH grade 3/shunt n=103 4) IVH grade 4/no shunt n=311 5) IVH grade 4/shunt n=125	Study center, gestational age, birth weight, gender, race, caesarean section delivery, multiple birth, antenatal steroid exposure, postnatal steroid exposure, surfactant use, respiratory distress syndrome, bronchopulmonary dysplacia (BPD), patent ductus arteriosus, periventricular leukomalacia (PVL), infection group, caregivers' education.	Psychomotor Development Index (PDI) <70, assessed by Bayley Scales of Infant Development IIR, administered by certified examiners).	At 18-22 months of age (corrected) PDI <70 IVH 3/no shunt: Reference IVH 3/shunt: OR 1.61 (1.32-1.96) No IVH/no shunt: Reference IVH 3/shunt: OR 2.45 (2.06-2.91) IVH 4/no shunt: Reference IVH 4/shunt: OR 1.94 (1.61-2.34) No IVH/no shunt: Reference IVH 4/shunt: OR 2.90 (2.45-3.43)	Moderate
Allred 2014 (US)	Prospective cohort study in 14 participating institutions in the Extremely Low Gestational Age Newborn (ELGAN) Study	n=1085	Gestational age, birth weight z-score categories, hyperoxemia (a PaO2 in the highest quartile on 2 of the first 3 postnatal days), Score of Neonatal Acute Physiology-II (SNAP-II) in the highest quartile, culture-proven bacteraemia in the first 28 days,	Psychomotor Development Index (PDI), assessed by Bayley Scales of Infant Development (2nd edition) by certified examiners. PDI <70 was considered as a delay in psychomotor development.	At 24 months PDI <55 No ROP stage 3+: Reference ROP stage 3+: OR 1.6 (1.03-2.4) No ROP plus disease: Reference ROP plus disease: OR 1.8 (1.1-3.1) No ROP zone 1: Reference	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			mechanical or high frequency on 14 or more days, and growth velocity in the lowest quartile.		ROP zone 1: OR 1.1 (0.6-2.2) No ROP threshold: Reference ROP threshold: OR 1.8 (0.6-5.0) No ROP prethreshold: Reference ROP prethreshold: OR 1.9 (1.1-3.1) PDI 56-69 No ROP stage 3+: Reference ROP stage 3+: OR 1.6 (1.03-2.5) No ROP plus disease: Reference ROP plus disease: OR 1.4 (0.7-2.6) No ROP zone 1: Reference ROP zone 1: OR 2.2 (1.2-4.2) No ROP threshold: Reference ROP threshold: OR 2.1 (0.7-6.6) No ROP prethreshold: Reference ROP prethreshold: OR 1.6 (0.9-2.9)	
Carlo 2011 (US)	Cohort study in 23 National Institute of Child	n=4924 total sample (children born at 22-25	Maternal variables (age, marital status, race, diabetes,	Bayley II Psychomotor Development index	At 18-22 months of age (corrected) PDI <70 (Bayley)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
	Health and Human Development Neonatal Research Network centres	weeks of gestation) n=72 children born at 22 weeks of gestation n=553 children born at 23 weeks of gestation n=1755 children born at 24 weeks of gestation n=2544 children born at 25 weeks of gestation	hypertension/preeclampsia, rupture of membranes >24h, antepartum haemorrhage, and delivery mode), multiple birth, gender, and centre, unless otherwise stated.	(PDI), a score <70 considered a delay.	22-25 weeks of gestation No antenatal corticosteroids: Reference Antenatal corticosteroids: OR 0.79 (0.65-0.96) 22 weeks of gestation No antenatal corticosteroids: Reference Antenatal corticosteroids: OR 1.47 (0.48-4.50)* 23 weeks of gestation No antenatal corticosteroids: Reference Antenatal corticosteroids: OR 0.93 (0.58-1.50) 24 weeks of gestation No antenatal corticosteroids: Reference Antenatal corticosteroids: OR 0.69 (0.49-0.95) 25 weeks of gestation No antenatal corticosteroids: Reference Antenatal corticosteroids: OR 0.82 (0.60-1.11) *Only adjusted for gender due to convergence problems because of low outcome prevalence.	
Hintz 2005 (US)	Multicentre cohort study using data from the National Institute of Child	n=2948	Network centre, use of antenatal glucocorticoids, rupture of membranes >24h,	Psychomotor development index (PDI), assessed through the Bayley Scales of Infant	At 18-22 months of age (corrected) PDI <70 (BSID-II)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
	Health and Human Development Neonatal Research Network Very Low Birth Weight Registry		out born status, estimated gestational age, gender, race, birth weight, small for gestational age, surfactant therapy, intraventricular haemorrhage grade 3 or 4 or cystic periventricular leukomalacia, sepsis, postnatal steroid treatment, bronchopulmonary dysplasia, and highest level of education attained by the primary caregiver.	Development-II (BSID-II). A score of <70 considered as a delay.	No NEC: Reference Surgical NEC: OR 1.95 (1.25-3.04) No NEC: Reference Medical NEC: OR 1.08 (0.66-1.80)	
Kerstjens 2012 (The Netherlands)	Population based prospective cohort study	n=832 moderately preterm children (32 to 35+6 weeks)	Variables included in the final model were: birth asphyxia, tertiary NICU admission, hypoglycaemia, hyperbilirubinaemia, SGA and gender.	Parents completed the Dutch version of the 48 months ASQ. The scores on each domain add up to an ASQ total problems score. A score of >2SDs below the mean for the Dutch reference group was considered to indicate developmental delay.	At 43-49 months (assumed to be chronological age) Risk of abnormal ASQ total problems score Septicaemia (both clinical symptoms and at least one positive blood culture result): Not significant on univariate analysis	Moderate
Kerstjens 2013 (The Netherlands)	Population based prospective cohort study	n=834 moderately preterm children (32-35 weeks)	SES and parity	Parents completed the Dutch version of the 48 months ASQ. The scores on each domain add up to an ASQ total problems	At 43-49 months (chronological age) Abnormal ASQ total problems score	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
				score. A score of >2SDs below the mean for the Dutch reference group was considered to indicate developmental delay.	Antenatal steroids: OR not significant in the univariate regression	
Laughon 2009 (US)	Prospective cohort study in 14 institutions in the Extremely low gestational age new born (ELGAN) study	n=915	Gestational age, single mother, complete course of antenatal steroids, caesarean delivery, delivery for preeclampsia or foetal indications, SNAP-II in the top quartile, Pao2 missing (week 1), transfusions (packed red blood cells), pulmonary deterioration, early and persistent pulmonary dysfunction, ventriculomegaly, echolucent lesion, echodense lesion, NEC stage II or worse, methylxanthine, patent ductus arteriosus, patent ductus arteriosus ligation, chronic lung disease without mechanical	Psychomotor Developmental Index (PDI) assessed by the Bayley Scales of Infant Development-2nd Edition (BSID-II). Score of <55 was considered a considerable delay.	At 2 years PDI <55 (BSID-II) No BPD: Reference BPD without mechanical ventilation: OR 1.1 (0.6–2.0) BPD with mechanical ventilation: OR 1.9 (0.97–3.9) No complete course of antenatal steroids: Reference Complete course of antenatal steroids: OR 2.4 (1.5-3.8)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			ventilation at 36 weeks, chronic lung disease with mechanical ventilation.			
Martin 2010 (USA)	Multicentre prospective cohort	n=1155 preterm infants (23-27+6 weeks)	All models are adjusted for public insurance, maternal or foetal initiator for delivery, GA (23-24, 25-26, 27 weeks), birth weight Z score 1 and thrombosis of the foetal stem vessels of the placenta and include a random effect cluster term for birth hospital.	The Bayley Scales of Infant Development-Second Edition was administered by examiners unaware of the infant's medical history. A score of < 70 (more than 2SD below the mean) was taken to represent significant psychomotor delay (PDI).	At 2 years of age (corrected) PDI <70 (Bayley-II) No NEC or late bacteraemia: Reference Medical NEC: OR 0.8 (0.3-1.9) Surgical NEC: OR 2.7 (1.2-6.4) Late bacteraemia: OR 1.3 (0.9-1.9)	High
O'Shea 2008 (US)	Prospective cohort study in 14 hospitals in 11 cities in 5 states in the US.	n=1017	Gestational age (23-24, 25-26, or 27 weeks), receipt of a complete course of antenatal corticosteroid, caesarean delivery, and Medicaid insurance at 2 years' corrected age.	Psychomotor Development Index (PDI) assessed using Bayley Scales of Infant Development - Second Edition (BSID-II). A score of <70 considered delayed psychomotor development.	At 24 months of age (corrected) PDI <70 (BSID-II) No IVH: Reference IVH: RR 2.10 (95% CI 1.50-2.90) No early PVL: Reference Early PVL: RR 2.10 (95% CI 1.40-3.20) No cystic PVL: Reference Cystic PVL: RR 4.30 (95% CI 2.30-8.10) No PIVH: Reference PIVH: RR 4.00 (95% CI 2.20-7.00)	Moderate
Shah 2012 (US)	Population-based cohort study	n=865	Birth weight, race, gender, multiple	Bayley Scales of Infant Development-II	At 18-22 months of age PDI <70 (Bayley)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
	utilizing data from the National Institute of Child Health Neonatal Research Network registry and the Cincinnati Collaborative Outreach Program Database.		births, antenatal steroids, surfactant, bronchopulmonary dysplasia, sepsis, and any intraventricular haemorrhage.	(BSID-II) (for infant born before 2006) and Bayley Scales of Infant Development-III (BSID-III) (for infants born after 1/1/2006) was used to obtain psychomotor developmental index (PDI). A score of <70 was considered an impaired psychomotor development.	No NEC: Reference NEC: OR 2.64 (1.18-5.91)	
Shankaran 2004 (US)	Prospective cohort study	n=246	Neonatal brain lesions, antenatal steroid exposure, sex, ethnicity/race, household income, BPD, surfactant administration, steroids for BPD, Medicaid, no high school degree, 2-parent household.	The Bayley Scales of Infant Development (BSID-II) was used to assess Psychomotor Developmental Index (PDI). A delay in psychomotor development was considered with a PDI score <70. BSID-II was administered by clinical psychologists or psychometricians trained to reliability.	At 18-22 months of age (corrected) PDI <70 (BSID-II) ICH grade 3-4: OR 1.1 (0.6-2.3) PVL: OR 3.1 (1.1-9.4) Any antenatal steroids: OR 0.9 (0.5-1.7) BPD: Not significant	Low
Stoll 2004 (US)	Multicentre cohort study using data from the National Institute of Child Health and Human	n=6314	Study centre, gestational age, birth weight, sex, race/ethnicity, rupture of membranes >24 h, CS, multiple birth, antenatal antibiotics,	Psychomotor developmental index (PDI), assessed with Bayley Scales of Infant Development II (BSID-II). A score of	At 18-22 months of age (corrected) PDI <70 (BSID-II) No infection: Reference Sepsis alone: OR 1.5 (1.2-1.9)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
	development (NICHD) Neonatal Research Network registry.		antenatal steroids, postnatal steroids, surfactant use, respiratory distress syndrome, bronchopulmonary dysplasia, patent ductus arteriosus, intraventricular haemorrhage grade 3-4, periventricular leukomalacia, maternal age at time of delivery, caregiver's level of education.	<70 considered a delay.	Sepsis + NEC: OR 2.4 (1.7-3.4) Meningitis with or without sepsis: OR 1.7 (1.1-2.5)	
Vohr 2005 (US)	Multicentre cohort study using data from 12 different centres of the National Institute of Child Health and Human Development Neonatal Research Network.	n=3785	Epoch, gestational age group, birth weight; gender, small for gestational age, multiple births, surfactant, grades 3 to 4 IVH, PVL, sepsis, oxygen requirement at 36 weeks, white vs. non-white race, out born vs. inborn status, caesarean section vs. vaginal delivery, maternal education <12 years vs. >=12 years, private health insurance vs. public, conventional ventilation vs. none,	Psychomotor Development Index, assessed by Bayley Scales of Infant Development II (BSID-II) or a gross motor assessment (not defined). A score of <70 was considered a delay in psychomotor development.	At 18-22 months of age (corrected) PDI <70 (Bayley) No PVL: Reference PVL: Significantly increased odds (OR and 95% CI not reported numerically) No grade 3-4 IVH: Reference Grade 3-4 IVH: Significantly increased odds (OR and 95% CI not reported numerically) No postnatal steroids: Reference Postnatal steroids : OR 1.99 (1.56-2.55)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			adjusted age at the time of assessment, centre, and the 4 interventions of interest: antenatal steroids (yes, no), high-frequency ventilation vs. none; days to regain birth weight, and postnatal steroids (yes, no).		No BPD: Reference BPD: Significantly increased odds (OR and 95% CI not reported numerically) No sepsis: Reference Sepsis: Not significant (OR and 95% CI not reported numerically) No antenatal steroids: Reference Antenatal steroids: OR 0.66 (0.52-0.84)	
Vohr 2000 (US)	Multicentre prospective cohort study	n=1151	Out born status, maternal hypertension, antenatal steroids, maternal education, race, caesarean section, birth weight, surfactant, early-onset sepsis, late-onset sepsis, grades 3 and 4 IVH/PVL, chronic lung disease (oxygen requirement at 36 weeks), postnatal steroids, small for gestational age, gender, and adjusted age at time of testing.	No independent walking, not clear how assessed but they report that a basic, functional, gross motor skills were assessed derived from the work of Russell et al.d Palisano et al. Psychomotor Developmental Index (PDI) score <70, assessed with Bayley Scale of Infant Development II (BSID-II)	At 18-22 months of age (corrected) No independent walking IVH/PVL grade III-IV: Significantly increased odds (OR (95% CI) not reported numerically) Postnatal steroids : Significantly increased odds (OR (95% CI) not reported numerically) NEC: Not significant (OR (95% CI) not reported numerically) BPD at 36 weeks: Significantly increased odds (OR (95% CI) not reported numerically)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
					Late-onset sepsis: Not significant (OR (95% CI) not reported numerically) Early-onset sepsis: Not significant (OR (95% CI) not reported numerically) Antenatal steroids: Not significant (OR (95% CI) not reported numerically) PDI <70 (Bayley-II) IVH/PVL grade III-IV: Significantly increased odds (OR (95% CI) not reported numerically) Postnatal steroids : Significantly increased odds (OR (95% CI) not reported numerically) NEC: Significantly increased odds (OR (95% CI) not reported numerically) BPD at 36 weeks: Significantly increased odds (OR (95% CI) not reported numerically) Late-onset sepsis: Not significant (OR (95% CI) not reported numerically) Early-onset sepsis: Not significant (OR (95% CI) not reported numerically) Antenatal steroids: Not significant (OR (95% CI) not reported numerically)	

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Behavioural, social, emotional and attention problems						
Delobel-Ayoub 2006 (France)	Population based prospective cohort study (EPIPAGE)	n=1228 preterm babies born at 22-32 weeks	For the comparison of term and preterm children, OR were adjusted for gender, maternal age at birth, birth order, maternal education, marital status of the mother, hospitalization during the last year, neurodevelopmental delay, the health of the child (assessed by the parents) at 3 years of age, gestational age, cerebral lesions and hospitalization in NICU ≥13 weeks.	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At 3 years of age (assumed chronological) Total difficulties score (SDQ 10th percentile) Cerebral lesions No lesion: Reference Minor lesion: OR 1.3 (0.9-2.0) Moderate lesion: OR 0.9 (0.6-1.5) Major lesions: OR 2.4 (1.1-5.2) BPD Total difficulties score Not a significant predictor on univariate analysis	Moderate
Delobel-Ayoub 2009 (France)	Population based prospective cohort study (EPIPAGE)	n=1102 preterm babies born at 22-32 weeks	All outcomes adjusted for cognitive performance, maternal age at birth, development of the child (assessed by the parents), hospitalisations between birth and 5 years, health of the child and mental wellbeing of the	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At age 5 years (assumed chronological age) Total difficulties score (SDQ 10th percentile) Cerebral lesions Not significant on univariate analysis	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			mother during the previous month.			
Johnson 2015b (UK)	Prospective population-based cohort study	n=625 late and moderately preterm (LMPT, 32-36 weeks) n=760 term controls	Not clearly reported. Variables that were significant (p<.05) in univariable analyses were all entered into the model. Variables that were not significant in this model were dropped in turn until only those variables significant at p <.05 were included in the final model. Variables that had been dropped were entered back into this final model one at a time to assess their significance.	Parents completed the Brief Infant Toddler Social Emotional Assessment (BITSEA). The BITSEA "competence scale" comprises 11 items that assess areas of attention, compliance, mastery motivation, prosocial peer relations, empathy, imitation/play skills, and social relatedness and is designed to identify children who have delays or deficits in the acquisition of social-emotional competencies (irrespective of whether behaviour problems are present).	At 2 years (corrected age) Delayed socioemotional competence Antenatal steroids not given: reference Antenatal steroid given: NS	Low
Special educational needs						
Johnson 2011 (UK & Ireland)	Population-based cohort study (EPICure Study)	n=219	Sex, gestational age, birth weight, maternal ethnicity, maternal age, maternal education, SES, antenatal steroids,	Teachers completed a questionnaire about if special educational needs (SEN) provision was utilized by the child.	At age 11 years SEN provision Abnormal last cerebral ultrasound: OR 3.72 (1.16-11.91)	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			preterm premature rupture of membranes, vaginal breech delivery, chorioamnionitis, foetal heart rate >100 bpm at 5 minutes, admission temperature <35c, CRIB score, NEC, postnatal steroids for chronic lung disease, any breast milk given, duration of NICU admission.		NEC: not significant (not reported) Any antenatal steroids: not significant (not reported) Any postnatal steroids for chronic lung disease: not significant (not reported)	

Abbreviations: ASQ-Ages and Stages Questionnaire; BPD-bronchopulmonary dysplasia; GA-gestational age; GMFCS-Gross Motor Functional Classification System; MDI-Mental Development Index; NEC-necrotising enterocolitis; NICU-neonatal intensive care unit; OR-odds ratio; PDI-Psychomotor Development Index; SD-standard deviation; SDQ-Strengths and Difficulties Questionnaire; SGA-small for gestational age; NEC-necrotising enterocolitis; SEN-special educational needs

Table 14: Summary of included publications on social, environmental and maternal factors

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Functional problems						
Johnson 2016 (UK)	Prospective population-based cohort study	n=628 late and moderately preterm (LMPT) children (32-36 weeks) n=759 term controls (>=37 weeks)	The analyses between term and LMPT group were adjusted for sex, SGA, SES index score, and nasogastric tube feeding >2 weeks. The analyses within the LMPT group	A validated eating behaviour questionnaire (4) was used to assess the presence of eating difficulties in the 4 domains of refusal/picky eating (e.g., poor appetite, food refusal, selective eating),	At 2 years (corrected age) Total feeding problems SES-index Low risk: Reference Medium risk: NS in univariate analysis High risk: NS in univariate analysis	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			included the following variables: behaviour problems, delayed social competence, SGA and nasogastric tube feeding.	oral motor problems (e.g., problems biting, chewing, or swallowing; gagging; or choking on food), oral hypersensitivity (e.g., aversion to being touched around the mouth or having things put in the mouth), and eating behaviour problems (e.g., has tantrums or makes a mess during meals). >90th percentile of the term control group were used to identify children with clinically significant eating difficulties.		
Motor, developmental and language delay						
Johnson 2015 (UK)	Prospective cohort study	n=638 late/moderately preterm infants	Ethnicity, sex, preeclampsia, any breast milk at discharge.	At 2 years (corrected age), cognitive impairment was assessed using the Parent Report of Children's Abilities-Revised (PARCA-R).	At 2 years of age (corrected) Moderate/severe cognitive impairment (<2.5th percentile PARCA-R) Socioeconomic status index Low risk: Reference Medium risk: RR 2.86 (1.24-6.57) High risk: RR 2.36 (1.02-5.48)	Moderate
Kerstjens 2013 (The Netherlands)	Population based prospective cohort study	n=834 moderately	SES and parity	Parents completed the Dutch version of the 48 months ASQ.	At 43-49 months (chronological age)	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
		preterm children (32-35 weeks)		The scores on each domain add up to an ASQ total problems score. A score of >2SDs below the mean for the Dutch reference group was considered to indicate developmental delay.	Abnormal ASQ total problems score Maternal pre-existing mental illness (depression, psychosis, other): OR 1.32 (0.14-12.3) Maternal age <20 years: not significant in the univariate regression Multiple pregnancy: OR 1.86 (1.02-3.42)	
Shankaran 2004 (US)	Prospective cohort study	n=246	Neonatal brain lesions, antenatal steroid exposure, sex, ethnicity/race, household income, BPD, surfactant administration, steroids for BPD, Medicaid, no high school degree, 2-parent household.	The Bayley Scales of Infant Development (BSID-II) was used to assess Psychomotor Developmental Index (PDI). A delay in psychomotor development was considered with a PDI score <70. BSID-II was administered by clinical psychologists or psychometricians trained to reliability.	At 18-22 months of age (corrected) PDI <70 (BSID-II) Socioeconomic status Household income >=$20 000: Reference Household income <$20 000: OR 1.5 (0.7-3.2)	Low
Singer 2001 (US)	Prospective cohort study	n=69 very low birth weight infants	Not clearly reported: “When the baseline differences […the effects of IVH, the only neonatal neurologic complication which differed between the groups…] were	The Bayley Scales of Infant Development that is described as widely used assessment toll of infant development. The psychomotor index (PDI) measures gross and fine motor	At 3 years PDI <70 (BSID) Maternal cocaine use When baseline differences were controlled, the effects of cocaine on intellectual disability remained significant	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			controlled, the effects of cocaine on these developmental outcomes remained significant"	control and coordination.		
Behavioural, social, emotional or attention problems						
Delobel-Ayoub 2006 (France)	Population based prospective cohort study (EPIPAGE)	n=1228 preterm babies born at 22-32 weeks	For the comparison of term and preterm children, OR were adjusted for gender, maternal age at birth, birth order, maternal education, marital status of the mother, hospitalization during the last year, neurodevelopmental delay, the health of the child (assessed by the parents) at 3 years of age, gestational age, cerebral lesions and hospitalization in NICU ≥13 weeks.	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At 3 years of age (assumed chronological) Total difficulties score Maternal age at birth 25-34 years: Reference <25 years: OR 2.5 (1.7-3.7) ≥35 years: OR 0.9 (0.5-1.4)	Moderate
Delobel-Ayoub 2009 (France)	Population based prospective cohort study (EPIPAGE)	n=1102 preterm babies born at 22-32 weeks	All outcomes adjusted for cognitive performance, maternal age at birth, development of the child (assessed by the parents), hospitalisations between birth and 5 years, health of the child and mental	The SDQ was used to assess behavioural problems. Cut-offs were defined so that 10% of the term control group were considered to have a behavioural problem.	At age 5 years (assumed chronological age) Total difficulties score Socioeconomic status Not significant on multivariate analysis Mental wellbeing of the mother during the previous month Very well: Reference	Moderate

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
			wellbeing of the mother during the previous month.		Fairly well: OR 1.8 (1.2-2.7) Fairly or very poor: OR 3.4 (1.9-6.3) Maternal age at birth 25-34 yrs: Reference <25 yrs: OR 1.6 (1.0-2.4) ≥35 yrs: OR 0.6 (0.4-1.0)	
Johnson 2015b (UK)	Prospective population-based cohort study	n=625 late and moderately preterm (LMPT, 32-36 weeks) n=760 term controls	Not clearly reported. Variables that were significant (p<.05) in univariable analyses were all entered into the model. Variables that were not significant in this model were dropped in turn until only those variables significant at p <.05 were included in the final model. Variables that had been dropped were entered back into this final model one at a time to assess their significance.	Parents completed the Brief Infant Toddler Social Emotional Assessment (BITSEA). The BITSEA "competence scale" comprises 11 items that assess areas of attention, compliance, mastery motivation, prosocial peer relations, empathy, imitation/play skills, and social relatedness and is designed to identify children who have delays or deficits in the acquisition of social-emotional competencies (irrespective of whether behaviour problems are present).	At 2 years (corrected age) Delayed socioemotional competence SES-index Low risk: reference Medium risk: RR 1.60 (1.14-2.24) High risk: RR 1.98 (1.41-2.75) Maternal substance abuse Non-drug user: reference Recreational drugs use during pregnancy: RR 1.70 (1.03-2.82) Multiple pregnancy Singleton: reference Multiple pregnancy: NS	Low

Study	Data Source	Sample and Population studied	Adjustment	Measures of Outcomes	Prognostic outcomes/Results	Study Quality
Potijk 2015 (The Netherlands)	Multicentre prospective cohort study	n=915 moderately preterm children (32-35+6 weeks gestation) n=543 term children (38-41+6 weeks gestation)	Socioeconomic status, gestational age, gender, number of siblings and maternal age.	The Dutch version of the CBCL was used to identify behavioural problems. The authors state that “American cut-offs” were used to identify problem scores.	At age 4 years (assumed to be chronological) Socioeconomic status Total behavioural problems SES: OR 1.42 (1.14-1.77) Externalising problems SES: OR 1.21 (0.99-1.50) Internalising problems SES: OR 1.26 (1.03-1.54) OR represent the risk per SD decrease in SES.	High
Special educational needs						
Johnson 2011 (UK & Ireland)	Population-based cohort study (EPICure Study)	n=219	Sex, gestational age, birth weight, maternal ethnicity, maternal age, maternal education, SES, antenatal steroids, preterm premature rupture of membranes, vaginal breech delivery, chorioamnionitis, foetal heart rate >100 bpm at 5 minutes, admission temperature <35c, CRIB score, NEC, postnatal steroids for chronic lung disease, any breast milk given, duration of NICU admission.	Teachers completed a questionnaire about if special educational needs (SEN) provision was utilized by the child.	At age 11 years SEN provision Maternal age (per 10 years): not significant (not reported) SES: not significant (not reported) Chorioamnionitis (suspected or proven): not significant (not reported)	Low

Abbreviations: CBCL-Child Behaviour Checklist; OR-odds ratio; SD-standard deviation; SES-socioeconomic status; SDQ-Strengths and Difficulties Questionnaire; NEC-nectotising enterocolitis; NICU-neonatal intensive care unit; SEN-special educational needs

4.1.1.3 Economic evidence

No health economic search was undertaken for this review question and consequently no evidence was found. This question focused on the risk of various developmental problems rather than whether any strategy for the management of these problems represents a cost-effective use of resources. Therefore, this question is not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review.

4.1.1.4 Evidence statements

4.1.1.4.1 Feeding problems

In relation to gestational age

Moderate to low quality evidence from three studies on feeding problems was mixed when comparing preterm infants to term controls. Moderate evidence from one study (n=479) showed no difference in the risk of a low drive to eat or low food variety at the age of 2 years (corrected age) among those born at <28 weeks, 28-29 weeks, 30-31 weeks or 32 weeks of gestation (Migraine 2013). Another study (n=371) also showed no difference in the risk of food refusal/faddy eating problems, behavioural problems around eating or oral hypersensitivity problems, but did find an increased risk of overall eating difficulties and oral motor problems at 6 years among children born extremely preterm (<26 weeks) (moderate quality evidence, Samara 2010). Another low quality study (n=1323) also found an increased risk of overall eating difficulties and oral motor problems at 2 years (corrected age) among children born at 32-36 weeks of gestation (Johnson 2016).

In relation to biological factors

Sex of the child

Low quality evidence from two studies found no association between sex of the child and feeding problems. One study (n=1151) examined the association between sex and no independent feeding at 18-22 months corrected age among children born with birth weight <1000 g (Vohr 2000). Another study (n=584) found no association between sex of the child and feeding difficulties at 2 years (corrected age) among moderate to late children born preterm (32-36 weeks) (Johnson 2016).

Small for gestational age

Low quality evidence from two studies show somewhat mixed results. One low quality study (n=1151) examined the association between being preterm and small for gestational age and no independent feeding at 18-22 months corrected age among children born with birth weight <1000 g (Vohr 2000). No significant association was found. Another low quality study (n=584) found a borderline significant increased risk of feeding difficulties at 2 years of corrected age among children born small for gestational age at 32-36 weeks of gestation (Johnson 2016).

Ethnicity or race

Low quality evidence from one study (n=1151) examined the association between the ethnicity or race of the preterm child and no independent feeding at 18-22 months corrected age (Vohr 2000). No significant association was found.

In relation to neonatal factors

Brain abnormalities

Low quality evidence from one study (n=1151) among children born with birth weight <1000g found an increased odds of lack of independent feeding at 18-22 months corrected age with neonatal intraventricular haemorrhage (IVH) grade III-IV (Vohr 2000).

Sepsis

Low quality evidence from one study (n=1151) among children born with birth weight <1000g found no association between neonatal culture-proven sepsis (neither early-onset nor late-onset) and lack of independent feeding at 18-22 months of corrected age (Vohr 2000).

Retinopathy of prematurity

No evidence was identified on the relationship between ROP and functional problems with feeding.

Necrotising enterocolitis

Low quality evidence from one study (n=1151) among children born with birth weight <1000g found no association between NEC and lack of independent feeding at 18-22 months of corrected age (Vohr 2000).

Antenatal exposure to steroids

Low quality evidence from one study (n=1151) among children born with birth weight <1000g found no association between antenatal exposure to steroids and lack of independent feeding at 18-22 months of corrected age (Vohr 2000).

Postnatal exposure to steroids

Low quality evidence from one study (n=1151) among children born with birth weight <1000g found no association between postnatal exposure to steroids and lack of independent feeding at 18-22 months of corrected age (Vohr 2000).

Bronchopulmonary dysplasia

Low quality evidence from one study (n=1151) showed an increased odds of lack of independent feeding at 18-22 months of corrected age with bronchopulmonary dysplasia at 36 weeks among children born with birth weight <1000 g (Vohr 2000).

In relation to social, environmental or maternal factors

Socioeconomic status

Low quality evidence from one study (n=584) found no association between socioeconomic status and feeding difficulties at 2 years (corrected age) among children born at 32-36 weeks of gestation (Johnson 2016).

No evidence was identified on the relationship between other maternal, social and environmental factors and functional problems with feeding.

4.1.1.4.2 *Sleeping problems*

In relation to gestational age

Moderate quality evidence from two studies on sleeping problems in relation to gestational age at birth showed was available. One publication (n=215) found no significant difference in sleeping problems between preterm children and term controls at the age of 2 years (de Jong 2015). However, another publication (n=398961) found a significantly increased odds of sleep apnoea diagnosis among children born preterm compared to children born full term (increased odds was found among children born at <32 weeks of gestation and among children born at 32-36 weeks of gestation, Raynes-Greenow 2012).

In relation to biological factors

No evidence was identified on the relationship between biological factors and functional problems with sleeping.

In relation to neonatal factors

No evidence was identified on the relationship between neonatal factors and functional problems with sleeping.

In relation to social, environmental or maternal factors

No evidence was identified on the relationship between maternal, social and environmental factors and functional problems with sleeping.

4.1.1.4.3 *Toileting problems*

In relation to gestational age Moderate quality evidence from one study (n=8769) found no association between gestational age and frequent bedwetting at 4 to 9 years age among children born at <37 weeks of gestation (Sullivan 2015).

In relation to biological factors

No evidence was identified on the relationship between biological risk factors and functional problems with toileting.

In relation to neonatal factors

No evidence was identified on the relationship between neonatal risk factors and functional problems with toileting.

In relation to social, environmental or maternal factors

No evidence was identified on the relationship between social, environmental or maternal factors and functional problems with toileting.

4.1.1.4.4 *Motor delay*

In relation to gestational age

Six publications of moderate to high quality provided evidence on the association of gestational age at birth and motor delay. Sample sizes ranged from 215 to 13843.

Moderate quality evidence from four studies provided mixed evidence on fine motor delay in relation to gestational age. One study (n=215) found no significant effect of being born at 32-36 weeks of gestation compared with term on fine motor skills when using the Dutch version of the Bayley Scales of Infant Development 3rd edition (BSID-III) at 24 months of age, both corrected and uncorrected (de Jong 2012). However, the three other studies found an increased odds of fine motor delay among children born preterm. One study (n=1983) used the Ages and Stages Questionnaire (ASQ) for children aged 4 years and found an increased odds of fine motor delay among children born at <32 weeks, 32-33, 34-35 and 32-35 weeks of gestation (Kerstjens 2011). One study (n=764) assessed children at 5 years of age with the Five to Fifteen (FTF) questionnaire and found an increased odds of fine motor skills problems among children born at <32 weeks of gestation compared to full term children (Rautava 2010). Another study (n=1356) assessed children between ages 9 to 34 months with the Denver II tool and found increased odds of one or more fine motor-adaptive cautions as well and one or more fine motor-adaptive delays among very low birth weight (mean gestational age of 28.4 weeks) compared with normal birth weight children (Schendel 1997). The same publication did not find a significant effect on either outcome when comparing the very low birth weight children with moderately low birth weight children (mean gestational age of 35.6 weeks).

Moderate quality evidence from the same four studies on gross motor delay in relation to gestational age is mixed. One study (n=215) found no significant effect of being born at 32-36 weeks of gestation on gross motor skills assessed with the Dutch version of the BSID-III at 24 months corrected age but found an increased odds when children were assessed at 24 months uncorrected age (de Jong 2015). Another study (n=1983) using the ASQ assessed children at 4 years and found an increased odds of gross motor delay among children born <32 weeks of gestation (compared with children born at full term) but not among children born at 32-33, 34-35, or 32-35 weeks of gestation (Kerstjens 2011). In another study (n=764), children born before 32 weeks of gestation were found to have a significantly increased odds of gross motor delay at 5 years assessed by FTF questionnaire (Rautava 2010). This study also looked at combined motor skills and found a significant effect. The study using Denver II tool (n=1356) found an increased odds of one or more gross motor cautions and one or more gross motor delays among very low birth children (mean gestational age of 28.4 weeks) compared to normal birth weight children and compared to moderately low birth weight children (mean gestational age of 35.6 weeks) (Schendel 1997).

High quality evidence from one study (n=13843) looked at specific motor delays using Movement Assessment Battery for Children (MABC) and found and increased odds of abnormal peg score (assessing manual dexterity) and abnormal coordination summary score (including balance, ball skills and peg scores) among children born at 32-35 weeks of gestation compared with full term born children assessed at 7 to 8 years (Odd 2013b). No significant effect was found on heel-to-toe score (assessing balance) or bean-bag score (assessing ball skills). Moderate quality evidence from another study (n=7500) used Bayley Short Form Research edition (BSF-R) to assess psychomotor development of children born at 34-36 weeks of gestation (compared to children born at full term) at 2 years of age and found and increased odds of psychomotor developmental index (PDI) of <70 and PDI 70-84 (Woythaler 2011).

In relation to biological factors

Sex of the child

Low quality evidence from two studies (n=246 and n=1151) found no associations between the sex of the child and motor delay (PDI <70 and lack of independent walking) among preterm babies (born at <25 weeks of gestation or with birth weight of 401-1000 g), assessed at 18-22 months of corrected age (Shankaran 2004; Vohr 2000).

Small for gestational age

Low quality evidence from one study (n=1151) found no association between being born small for gestational age and PDI score of <70 and lack of independent walking at 18-22 months of corrected age among children born with birth weight 401-1000 g (Vohr 2000).

Ethnicity or race

Low quality evidence from two studies (n=246; n=1151) on the relationship between ethnicity/race and motor delay among children born preterm show no association among preterm children (born at <25 weeks of gestation or with birth weight of 401-1000 g), on PDI <70 (Shankaran 2004; Vohr 2000) and lack of independent walking (Vohr 2000) between black and non-black children (Shankaran 2004) and between white and non-white children (Vohr 2000) assessed at 18-22 months of corrected age with BSID.

In relation to neonatal factors

Brain abnormalities

Low to moderate quality evidence from four studies (sample sizes ranging from 246 to 6161) was available on the relationship between neonatal brain lesions among children born preterm (born at <28 weeks of gestation or with birth weight <1000 g) and motor delay at 18-24 months corrected age (Adams-Chapman 2008; O'Shea 2008; Shankaran 2004; Vohr 2000). All studies found increased odds of PDI <70 with different types of brain lesions (intraventricular haemorrhage [IVH], IVH grade III-IV, IVH III with shunt, IVH IV with shunt, periventricular leukomalacia [PVL], cystic PVL, early PVL, periventricular haemorrhagic infarction). One study (n=1151) also found an association with IVH or PVL grade III-IV and lack of independent walking (Vohr 2000). One publication (n=246) found no association between intracranial haemorrhage (ICH) grade III-IV and PDI <70 (Shankaran 2004).

Sepsis

Low to high quality evidence from four studies (sample sizes ranging from 1151 to 6314) on the relationship between neonatal sepsis and motor delay show mixed results (Martin 2010; Stoll 2004; Vohr 2005; Vohr 2000). High quality evidence from a study (n=1155) found no association between culture-proven late-onset neonatal sepsis and abnormal PDI at 2 years of age (Martin 2010). Moderate quality evidence from another study found an increased odds of abnormal PDI score at 18-22 months corrected age among preterm children (with birth weight 1000 g or less) that had had neonatal culture-proven sepsis with antibiotic therapy for more than five days, that had had neonatal sepsis with NEC, and that had had neonatal meningitis with or without sepsis (Stoll 2004). Low to moderate quality evidence from two publications of the same study project examining cohorts born at different times (n=3785 and n=1151) found no association between sepsis and abnormal PDI score at 18-22 months corrected age (Vohr 2005; Vohr 2000). The latter also did not fund an association between sepsis and lack of independent walking.

Retinopathy of prematurity

Moderate quality evidence from one study on the association between different severities of ROP (vs no ROP) and abnormal PDI score (either <55 or 55-69) show mixed findings (Allred 2014). The evidence shows a general tendency of increased odds of abnormal PDI score for all severities of ROP, however, not all of them reached statistical significance. ROP stage 3+, however, showed significantly increased odds of PDI <55 and PDI 55-69. The children were born earlier than 28 weeks of gestation and they were assessed at 24 months of age.

Necrotising enterocolitis

Low to high quality evidence from four studies (sample sizes ranging from 865 to 2948) on the association between necrotising enterocolitis (NEC) and psychomotor development (assessed by BSID) show somewhat mixed results (Hintz 2005; Martin 2010: Shah 2012; Vohr 2000). High quality evidence from one study (n=1155) and moderate quality evidence from another study (n=2948) showed a significant increase in the odds of an abnormal PDI for preterm infants (23 to 27+6 weeks of gestation or birth weight of 401-1000 g) who had NEC requiring surgery but not for ones with medically managed NEC (Hintz 2005; Martin 2010). Moderate quality evidence from one study (n=865) showed an increased odds of abnormal PDI score with NEC grade II or higher and low quality evidence from another study (n=1151) showed an increased odds of abnormal PDI score with NEC (unspecified) (Shah 2004; Vohr 2000). The same low quality publication also reported that there was no association between NEC and lack of independent walking (Vohr 2000). All outcomes were assessed at around 2 years of age.

Antenatal steroid exposure

Low to moderate quality evidence from five studies on the association between antenatal steroid exposure and motor delay (assessed by BSID) show mixed results (Carlo 2011; Laughon 2009; Shankaran 2004; Vohr 2005; Vohr 2000). Moderate quality evidence from two studies (n=4924; n=3785) found reduced odds of PDI score <70 at 18-22 months of corrected age among preterm children (born 22-32 weeks of gestation) with exposure to antenatal steroids (Carlo 2011; Vohr 2005). The first study also performed stratified analysis for each week of gestation (from 22 to 25 weeks), the findings are mixed but largely did not reach statistical significance. Low quality evidence from two other studies (n=246; n=1151) found no association between antenatal steroids and PDI <70 at 18-22 months of corrected age among extremely preterm children (<25 weeks of gestation or with birth weight 401-1000 g) (Shankaran 2004; Vohr 2000). The latter publication also found no association on lack of independent walking. Moderate quality evidence from one study (n=915) found an increased odds of PDI score <55 among preterm children (born <28 weeks of gestation) at 24 months of age (Laughon 2009).

Postnatal steroid exposure

Low to moderate quality evidence from two studies (=3785 and n=1151, respectively) on the relationship between postnatal exposure to steroids and motor delay found an increased odds of PDI score <70 (Vohr 2005; Vohr 2000). The latter publication also found an increased odds of lack of independent walking. The children were born at 22-32 weeks of gestation or with birth weight 401-1000 g and assessed at 18-22 months of corrected age.

Bronchopulmonary dysplasia

Low to moderate quality evidence from four studies (sample sizes ranging from 246 to 3785) on the association between bronchopulmonary dysplasia (BPD, need of additional oxygen at 36 weeks) and motor delay show mixed results (Laughon 2009; Shankaran 2004; Vohr 2005; Vohr 2000). Moderate quality evidence from one study (n=915) found no association with PDI score of <55 when looking at BPD without mechanical ventilation and a near-significant association when looking at BPD with mechanical ventilation among children born <28 weeks of gestation and assessed at 24 months of age (Laughon 2009). Low to moderate quality evidence from two publications from one large study project (n=3785 and n=1151, respectively) found an increased odds of PDI <70 at 18-22 months of age with BPD among children were born at 22-32 weeks of gestation or with birth weight 401-1000 g (Vohr 2005; Vohr 2000). The latter publication also found an association with lack of independent walking. Low quality evidence from one study (n=246) found no association among children born <25 weeks of gestation and assessed at 18-22 months of corrected age (Shankaran 2004).

In relation to social, environmental or maternal factors

Socioeconomic status

Low quality evidence from one study (n=246) found no association between socioeconomic status (household income <$20000/year vs >=€20000) and PDI <70 (assessed by BSID) among children born at <25 weeks of gestation and assessed at 18-22 months of corrected age (Shankaran 2004).

Substance abuse

Low quality evidence from one study (n=82) found a significant association between maternal cocaine use and abnormal psychomotor developmental index score (BSID) at three years of age among children born with birth weight <1500 g (Singer 2001).

No other evidence was identified on the relationship between other social, environmental or maternal factors and motor delay.

4.1.1.4.5 *Language delay*

In relation to gestational age

Moderate quality evidence from five studies (sample sizes ranging from 215 to 32314) on the association between gestational age and language problems show mixed findings (Brown 2014; de Jong 2015; Rautava 2010; Stene-Larsen 2014; Schendel 1997). One study (n=12302) found no association among children 34-36 weeks of gestation (versus term) and receptive vocabulary delay (assessed with Peabody Picture Vocabulary Test-Revised, PPVT-R) at 4-5 years of age (Brown 2014). Another study (n=215) found no association between gestational age (32-36 weeks versus term) and receptive communication delay or expressive communication delay (assessed with the Dutch version of the BSID-III at 24 months of age (corrected and uncorrected) (de Jong 2015). Another study (n=764) found an increased odds of language problems, expressive language skills problem and communication problem (assessed with the FTF questionnaire) at 5 years of age among children born <32 weeks of gestation (Rautava 2010). One study (n=32314) found an increased risk of communication problems (assessed with 3 items from the ASQ) at 18 months of age among children born at 34-36 weeks of gestation (compared to term) (Stene-Larsen 2014). The same children were assessed at 36 months of age and the association was no longer significant (assessed with 6 items from the ASQ). However, there was an increased odds of expressive language impairments at 36 weeks months of age. Finally, one study (n=1356) found an increased odds of language cautions and language delays (assessed with Denver-II tool) among children born with very low birth weight (mean gestational weeks 28.4) compared with children born with normal birth weight (mean gestational weeks 39.4) (Schendel 1997). The children were assessed between ages 9 to 34 months corrected age. The same study compared children born with very low birth weight (mean gestational weeks 28.4) with children born with moderately low birth weight (mean gestational weeks 35.6) and found an increased odds of language delays, however, language cautions did not reach statistical significance.

In relation to biological factors

No evidence was identified on the relationship between biological factors and language delay.

In relation to neonatal factors

No evidence was identified on the relationship between neonatal factors and language delay.

In relation to social, environmental or maternal factors

No evidence was identified on the relationship between social, environmental or maternal factors and language delay.

4.1.1.4.6 *Other developmental delay*

In relation to gestational age

Moderate quality evidence on the relationship between gestational age and other developmental delays from six studies (sample sizes ranging from 764 to 15099) show mixed results (Brown 2014; Johnson 2015a; Kerstjens 2011; Kerstjens 2012; Rautava 2010; Schendel 1997). One study (n=15099) found no association between developmental delay (assessed with Motor and Social Development Scale) and gestational age among children born at 34-36 weeks of gestation and assessed at 2-3 years of age (Brown 2014). Another study (n=1983) found no association between gestational age and developmental delay (ASQ total score <2SD) at 4 years of age among children born at 32-35 weeks of gestation (compared to term) but found a significantly increased odds developmental delay among children born <32 weeks of gestation (Kerstjens 2011). Another publication of the same study (n=832) compared children born at 32-33 gestational weeks to children born at 34-35 gestational weeks and found no association with developmental delay between the two preterm groups (Kerstjens 2012). One study (n=764) found an increased odds of comprehension problem (assessed with the FTF questionnaire) at 5 years among children born at <32 weeks of gestation (Rautava). Another study (n=1403) found an increased odds of moderate to severe cognitive impairment (assessed with PARCA-R) at 2 years of corrected age among children born at 32-36 weeks of gestation (Johnson 2015a). Finally, one study (n=1356) used Denver-II tool to assess developmental delay at 9-34 months of age and found an increased odds of questionable overall performance and abnormal overall performance in the Denver-II test among children born with very low birth weight (mean gestational weeks 28.4) compared to normal birth weight children (mean gestational weeks 39.4) and compared to moderately low birth weight children (mean gestational weeks 35.6) (Schendel 1997).

In relation to biological factors

Sex of the child

Moderate quality evidence from two studies (n=638; n=834) showed increased odds of developmental delay for male preterm children as compared to females (Johnson 2015a; Kerstjens 2013). Developmental problems were assessed by ASQ in the first publication; and moderate to severe cognitive impairment was assessed by PARCA-R screening tool in the second publication. These children were born at 32 to 36 weeks and were assessed at 2 years of corrected age in the first study and at 43 to 49 months of age in the second study.

Small for gestational age

Moderate quality evidence from one study (n=834) showed an increase in the risk of developmental delay (assessed by ASQ) for SGA preterm children, when compared to those preterm children born appropriate for gestational age (Kerstjens 2013). The children were assessed at between 43 and 49 months of age, and were born at 32 to 35 weeks.

Ethnicity or race

Moderate quality evidence from one study (n=1403) found an increased odds of moderate to severe cognitive impairment (assessed by PARCA-R screening tool) among non-white children compared with white children (born at 32-36 weeks of gestation) assessed at 2 years of corrected age even after adjusting for socioeconomic status (Johnson 2015).

In relation to neonatal factors

Sepsis

Moderate quality evidence from one study (n=832) found no association between neonatal sepsis (defined as clinical symptoms and at least one positive blood culture) and developmental problems (ASQ total problems <2SD) among children born at 32-35 weeks of gestation and assessed at 43-49 months of age (Kerstjens 2012a).

Antenatal steroids

Moderate quality evidence from one study (n=834) found no association between antenatal exposure to steroids and developmental problems (ASQ total problems <2SD) among children born at 32-35 weeks of gestation and assessed at 43-49 months of age (Kerstjens 2013).

No evidence was identified on the relationship between the other neonatal factors and other developmental delay.

In relation to social, environmental or maternal factors

Socioeconomic status

Moderate quality evidence from one study (n=1403) on the association between socioeconomic status and moderate to severe cognitive impairment show that lower socioeconomic status was associated with increased odds of cognitive impairment (Johnson 2015a). This study included children born at 32-36 weeks of gestation and they were assessed at 2 years of corrected age using PARCA-R screening tool.

Maternal age

Moderate quality evidence from one study (n=834) found no association between maternal age under 20 years and developmental problems (ASQ total problems <2SD) among children born at 32-35 weeks of gestation and assessed at 43-49 months of age (Kerstjens 2013).

Maternal mental health

Moderate quality evidence from one study (n=834) found no association between maternal mental illness and developmental problems (ASQ total problems <2SD) among children born at 32-35 weeks of gestation and assessed at 43-49 months of age (Kerstjens 2013).

Multiple pregnancy

Moderate quality evidence from one study (n=834) shows an association between multiple pregnancy and developmental problems (ASQ total problems <2Sd) among children born at 32-35 weeks of gestation and assessed at 43-49 months of age (Kerstjens 2013).

No other evidence was identified on the relationship between other social, environmental and maternal factors and other developmental delays.

4.1.1.4.7 Executive function

In relation to gestational age

Low to high quality evidence from three studies (n=134; n=169; n=764) on executive function in preterm children as compared to term controls show somewhat mixed findings (Farooqi

2016; Farooqi 2013; Rautava 2010). Children in these studies were all born at <32 weeks and/or ≤1500g and the children were assessed between 5 and 16 years of age. One study (n=764) found an increased odds of planning or organising problems and memory problems at 5 years among children born at <32 weeks of gestation or with birth weight of <1500 g assessed with the FTF questionnaire (Rautava 2010). Similarly, another study (n=169) found an increased odds of problems with planning or organisation and working memory reported by both parents and teachers among children born at <26 weeks of gestation compared to term children at 11 years (assessed with the FTF questionnaire) (Farooqi 2013). In another study of low quality, preterm children born at <26 weeks of gestation (as compared to term controls) who were assessed between 10 and 15 years of age were found to have increased odds of problems with verbal, non-verbal working memory, spatial conceptualisation visual reasoning, and planning ability (assessed with the WISC III-R questionnaire domains for executive function, and Tower test D-KEFS). In the same study, children were found to have increased odds of behavioural problems with attention, hypoactivity, planning and organisation, working memory, (reported by parents and teachers, assessed with the FTF questionnaire domains for executive function) (Farooqi 2016).In relation to biological factors

No evidence was identified on the relationship between biological factors and executive function.

In relation to neonatal factors

No evidence was identified on the relationship between neonatal factors and executive function.

In relation to social, environmental or maternal factors

No evidence was identified on the relationship between social, environmental or maternal factors and executive function.

4.1.1.4.8 Behavioural, social, emotional and attention problems

In relation to gestational age

Low to high quality evidence from fourteen studies examine the relationship between gestational age (preterm compared to term) and different behavioural, social, emotional and attention problems.

Low to high quality evidence from eight studies (sample sizes ranging from 169 to 6409) examined the relationship between gestational age and total behavioural problems assessed with either the Strengths and Difficulties Questionnaire (SDQ) or the Child Behaviour Checklist (CBCL) (or the equivalent for teachers Teacher Report Form [TRF]) (de Jong 2014; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Farooqi 2007; Fevang 2016; Hornman 2016; Johnson 2015b; Potijk 2015; Reijneveld 2006). The findings are somewhat mixed.

Two studies used the SDQ. Moderate quality evidence from one study (n=1675) found an increased odds of total behavioural difficulties at 3 years of age among children born at 22-32 weeks of gestation (Delobel-Ayoub 2006). The effect remained when these children were assessed again at 5 years of age (n=1477, Delobel-Ayoub 2009). When comparing the total behavioural problems between preterm children born at different gestational ages, no significant differences were observed when assessed at 3 and 5 years of age (Delobel-Ayoub 2009; Delobel-Ayoub 2006). Low quality evidence from another study (n=2098) found a significantly increased odds of total behavioural problems at 11 years of age among children born at <28 weeks of gestation or with birth weight <1000 g (Fevang 2016).

Five studies used the CBCL to assess total behavioural problems among children born preterm. Moderate to high quality evidence from two studies publications (n=6409; n=169) show an increased risk of total behavioural problems at 5 years and at 11 years of age among children born at less than 32 gestational weeks or with a birth weight or less than 1500 g (Reijneveld 2006; Farooqi 2007). Moderate quality evidence from another study (n=1458) shows a borderline significant association with total behavioural problems at 4 years of age among children born at 32-35 weeks of gestation (Potijk 2015). Moderate quality evidence from one study (n=215) among moderate and late children born preterm (32-36 weeks) shows no significant association with total behavioural problems at 24 months of corrected age (de Jong 2015). One publication (n=1443) with moderate quality evidence on total behavioural problems assessed at four years and at five years looked if the abnormal CBCL total score was present at either four or five years of age, or both, categorising outcome of total problems into emerging (normal score at four years but abnormal score at five years), resolving (abnormal score at four years but normal score at 5 years) and persistent (abnormal score at both 4 and 5 years) problems (Hornman 2016). The study found no difference in emerging problems among children born at <36 weeks of gestation, or at 32-35 weeks of gestation, or at 25-31 weeks of gestation compared to term born children. The study showed an increased odds of resolving problems among the children born at <36 weeks and children born at 32-35 weeks but not among children born at 25-31 weeks. There was an increased odds of persistent total problems among children born at <36 weeks and children born at 25-31 weeks and a borderline significant increased odds among children born at 32-35 weeks of gestation.

Additionally, low quality evidence from one study (n=1385) show no association between gestational age and behaviour problems among moderate to late children born preterm when using the Brief Infant Toddler Social Emotional Assessment (BITSEA) at 2 years (corrected age) (Johnson 2015b). The same study reports an increased odds of delayed socioemotional competence among the children.

Low to high quality evidence from four studies (sample sizes ranging from 169 to 1675) on the association between gestational age and hyperactivity show mixed findings (Delobel-Ayoub 2006; Farooqi 2013; Fevang 2016; Rautava 2010). High quality evidence from one study (n=169) found no association among children born at <26 weeks of gestation and assessed at 11 years of age using the FTF questionnaire with both parental report and teacher report (Farooqi 2013). No association was found even after excluding the ones with neurosensory impairment. Moderate quality evidence from one study (n=1675) found an increased odds of hyperactivity (assessed by parents with SDQ) among children born at 22-32 weeks of gestation and assessed at 3 years of age (Delobel-Ayoub 2006). Moderate quality evidence from another study (n=764) also found an increased odds of hyperactivity or impulsivity among children born at <32 weeks of gestation or with a birth weight of <1500 g (Rautava 2010). The children were assessed at 5 years of age through parental report on the FTF questionnaire. Low quality evidence from one study (n=2098) found increased odds of hyperactivity/impulsivity at 11 years among children born <28 weeks of gestation or with birth weight <1000 g (assessed with Swanson, Noland, and Pelham Questionnaire, Revision IV [SNAP-IV]) (Fevang 2016).

Moderate to high quality evidence from two studies show mixed findings on the association between gestational age and hypoactivity (Farooqi 2013; Rautava 2010). High quality evidence from one study (n=169) found no significant association between being born <26 weeks of gestation (versus term) and hypoactivity (assessed with the FTF questionnaire) when using parental report (Farooqi 2013). When teacher report was used, an increased odds of hypoactivity was observed. The results remained even when excluding children with neurosensory impairment. The children were assessed at 11 years of age. Moderate quality evidence from another study (n=764) found a significantly increased odds of hypoactivity (parental report through the FTF questionnaire) at 5 years of age among children born <32 weeks of gestation or with birth weight <1500 g (Rautava 2010).

Low to high quality evidence from seven studies (sample sizes ranging from 169 to 34163) on the relationship between gestational age and attention problems show mixed findings (de Jong 2014; Farooqi 2013; Farooqi 2007; Fevang 2016; Higa Diez 2016; Rautava 2010;

Reijneveld 2006). Three studies used the Child Behaviour Checklist (CBCL) and two studies used the FTF questionnaire. One study used the SNAP-IV. The children were assessed between 24 months corrected age and 11 years chronological age and the prematurity of the children ranged from <26 weeks of gestation to 36 weeks of gestation. High quality evidence from one study (n=169) show an increased odds of attention problems among children born <26 weeks of gestation and assessed at 11 years through FTF questionnaire filled in by teachers (Farooqi 2013). However, no significant association was among the same population when FTF questionnaire was filled in by parents. The results remained the same after excluding the children with neurosensory impairment. Moderate quality evidence from one study (n=764) show an increased odds of attention problems among children born at <32 weeks of gestation or with birth weight ≤1500 g when assessed at 5 years of age with FTF questionnaire using parental report (Rautava 2010). Moderate quality evidence from another study (n=6409) show an increased risk of attention problems among preterm children (born at <32 weeks of gestation or with birth weight <1500 g) at 5 years of age assessed with the CBCL (Reijneveld 2006). Moderate quality evidence from one study (n=215) found no association to attention problems at 24 months of corrected age among children born 32-36 weeks of gestation and assessed with the CBCLde Jong 2015). Moderate quality evidence from a nationally representative study from Japan (n=34163) using the CBCL (parental report) to assess different types of attention problems among children born preterm compared to their term peers at 8 years of age found children born preterm (at <34 weeks or at 34-36 weeks of gestation) being more likely to have problems waiting for their turn during play. However, no difference between term and preterm children were observed in the attention problem domains of “interrupting people” and “failure to pay attention when crossing the street”. When looking at children who presented problems in all of the above mentioned attention domains, there was a significant association among children born at <34 weeks of gestation. The association among children born at 34-36 weeks of gestation did not reach statistical significance. Low quality evidence from one study (n=2098) found an association between being born at <28 weeks of gestation or with birth weight <1000 g and inattention problems (assessed with SNAP-IV) at 11 years of age (Fevang 2016).

Moderate to high quality evidence from seven studies (sample sizes ranging from 169 to 6409) show mixed results on the association between gestational age and internalising behaviours among preterm children (versus term children) (de Jong 2015; Farooqi 2007; Gurka 2010; Hornman 2016; Potijk 2015; Rautava 2010; Reihneveld 2006). The children were assessed aged between 24 months (corrected) and 11 years of age using either the CBCL or the FTF questionnaire. Moderate quality evidence from two different studies (n=764; n=6409) that both examined children born at <32 weeks of gestation or with birth weight of <1500 g show mixed findings (Rautava 2010; Reijneveld 2006). The first study found an increased risk of internalising problems at 5 years of age using the FTF questionnaire, while the other publication found no association using the CBCL. Evidence from a third study (n=1458) shows an increased odds of internalising problems among children born at 32-35 weeks of gestation who were assessed at 4 years of age with the CBCL (Potijk 2015), however, evidence from another study (n=215) show no association among children born at 32-36 weeks of gestation at 24 months of corrected age using the CBCL (de Jong 2015). Low quality evidence from another study (n=1298) observing children born late-preterm (34-36 weeks) and their full-term born peers from ages 4 until 15 years show no significant difference in internalising behaviours between the groups (Gurka 2010). A high quality evidence from a study (n=169) show an association between being born at <26 weeks of gestation and internalising problems at 11 years when assessed by both parents (CBCL) and teachers (Teacher Report Form [TRF], parallel form of CBCL for teachers) (Farooqi 2007). One publication (n=1443) with moderate quality evidence on internalising problems assessed at four years and at five years looked if the abnormal score was present at either four or five years of age, or both, categorising outcome of internalising problems into emerging (normal score at four years but abnormal score at five years), resolving (abnormal score at four years but normal score at 5 years) and persistent (abnormal score at both 4 and 5 years) problems (Hornman 2016). The study found no difference in emerging internalising problems among children born at <36 weeks of gestation, or at 32-35 weeks of gestation, or at 25-31 weeks of gestation compared to term born children. The study found an increased odds of resolving internalising problems and persistent internalising problems among the children born at <36 weeks, children born at 32-35 weeks and children born at 25-31 weeks.

Low to high quality evidence from five studies (sample sizes ranging from 169 to 6409) that observed specific internalising behaviours using the CBCL show mixed findings (de Jong 2015; Farooqi 2007; Fevang 2016; Gurka 2010; Reijneveld 2006). The populations in these studies vary as well as the age at assessment. Three different studies (sample sizes ranging from 169 to 6409) presenting moderate to high quality evidence report mixed findings on withdrawn behaviour (de Jong 2015; Farooqi 2007; Reijneveld 2006). Two studies found no association between gestational age and withdrawn behaviour at 24 months of corrected age among children born at 32-36 weeks) (de Jong 2015)) and at 5 years of age among children born at <32 weeks or with birth weight <1500 g (Reijneveld 2006). However, the third study found an increased odds of withdrawn behaviour at 11 years of age among children born extremely preterm (<26 weeks) when assessed by both parents and teachers (Farooqi 2007). The same three studies with moderate to high evidence report mixed findings on somatic complaints as well. Moderate quality evidence from one study (n=215) show no association with somatic complaints at 24 months corrected age among children born a 32-36 weeks of gestation (de Jong 2015). Moderate quality evidence from another study among children with lower gestational age (<32 weeks or birth weight or <1500 g), however, show an increased odds of somatic complaints at 5 years (Reijneveld 2006). High quality evidence from a third study show an association between extreme prematurity (<26 weeks) and somatic complaints at 11 years of age when children were assessed by teachers but not when they were assessed by parents (Farooqi 2007).

Moderate quality evidence from three studies (samples sizes ranging from 169 to 6409) on the association between prematurity and depression or anxiety symptoms show mixed findings (Farooqi 2007; Fevang 2016; Reijneveld 2006). Moderate quality evidence from one study (n=6409) using the CBCL found no association between being born at <32 weeks of gestation (or with birth weight <1500 g) and anxious/depressed behaviours at 5 years of age (Reijneveld 2006). However, high quality evidence from another study (n=169) using the CBCL (and TRF) found a significantly increased odds of anxious/depressed behaviours at 11 years of age among extremely children born preterm (<26 weeks) when the child was assessed by both parents and teachers (Farooqi 2007). However, the latter study used a less strict cut-off (90th percentile) than the first study (97th percentile). The latter study, however, did not find an association between being born extremely premature and child self-reported depression symptoms (depression self-rating scale [DSRS], Farooqi 2007). Low quality evidence from another study (n=2098) show an association between being born at <28 weeks or with birth weight <1000 g and anxiety symptoms (assessed with the Screen for Child Anxiety Related Emotional Disorders [SCARED], Fevang 2016).

Moderate to high quality evidence from seven studies (sample sizes ranging from 169 to 6409) on the relationship between gestational age and externalising behaviours show mixed findings. High quality evidence from one study (n=169) among children born extremely preterm (<26 weeks) show no association between gestational age and externalising behaviours at 11 years of age (CBCL/TRF) (Farooqi 2007). Moderate quality evidence from another study (n=215) among children born at 32-36 weeks of gestation show no association with gestational age and externalising behaviour (CBCL) at 24 months (corrected) (de Jong 2015). Low quality evidence from one study (n=1298) that assessed children from 4 to 15 years of age show no difference in externalising behaviours between children born preterm (34-36 weeks) and full-term born children. However, moderate quality evidence from three studies (sample sizes ranging from 764 to 6409) show preterm children (<36 weeks of gestation) to be more likely to present externalising behaviours than term children at 4 and 5 years of age (assessed with FTF questionnaire and the CBCL) (Potijk 2015; Rautava 2010;

Reijneveld 2006). One publication (n=1443) with moderate quality evidence on externalising problems assessed at four years and at five years looked if the abnormal score was present at either four or five years of age, or both, categorising outcome of externalising problems into emerging (normal score at four years but abnormal score at five years), resolving (abnormal score at four years but normal score at 5 years) and persistent (abnormal score at both 4 and 5 years) problems (Hornman 2016). The study found an increased odds of emerging externalising problems among children born at <36 weeks of gestation, or at 32-35 weeks of gestation, or at 25-31 weeks of gestation compared to term born children. The study found an increased odds of resolving externalising problems among children born at 32-35 weeks of gestation but not among children born at <36 weeks or 25-31 weeks of gestation. The study found an increased odds of persistent internalising problems among the children born at <36 weeks, children born at 32-35 weeks and children born at 25-31 weeks.

High quality evidence from a population-based study (n=169) show no association between being born extremely preterm (<26 weeks) and aggressive or delinquent behaviours at 11 years of age (assessed by parents and teachers with CBCL/TRF) (Farooqi 2007). Moderate quality evidence from another population-based study (n=34163) from Japan on the association between prematurity and delinquent or aggressive behaviours at 8 years of age show no association with gestational age and lying behaviour and hurting other people (Higa Diez 2016). However, children born at <34 weeks of gestation were more likely to destroy toys or books compared to their term peers (not significant among children born at 34-36 weeks) and children born at 34-36 weeks of gestation were more likely to cause disturbances in public (not significant among children born at <34 weeks). When looking at children with problems in all the above mentioned delinquency/aggressive behaviour domains, no significant association was found between preterm and term born children in this study. Moderate quality evidence from another study (n=6409) found an association with delinquent behaviour at 5 years of age among children born <32 gestational weeks or with birth weight <1500 g (Reijneveld 2006). The same study did not find a significant association for aggressive behaviour. Similarly, low quality evidence from one study (n=1298) did not show a difference in aggressive behaviours (assessed with CBCL) in preterm (34-36 weeks) and full-term born children from age 4 to 15 years of age (Gurka 2010).

Moderate quality evidence from a study (n=1675) show an association with gestational age 22-32 weeks (versus term) and conduct problems when assessed at 3 years of age with the SDQ (Delobel-Ayoub 2006). The same study found a borderline significant association with peer problems and emotional symptoms. Moderate quality evidence from another study (n=215) show no association between being born at 32-36 weeks of gestation and being abnormally emotionally reactive at 24 months of corrected age (assessed with the CBCL) (de Jong 2015). Moderate quality evidence from one study (n=6409) show a significantly increased odds of social problems and thought problems at 5 years of age among children born at <32 weeks of gestation (assessed with the CBCL) (Reijneveld 2006). No association was found between gestational age and sex problems at 5 years in the same study. High quality evidence from another study (n=169) show an increased odds of social problems and thought problems among children born extremely preterm (<26 weeks) at 11 years when assessed by teachers (TRF) but not when assessed by parents (CBCL) (Farooqi 2007). Moderate quality evidence from one study (n=1356) that examined the association between gestational age and personal-social problems show an increased risk of one or more personal-social cautions and personal-social delays among children born with very low birth weight (mean gestational weeks 28.4) compared with children born with normal birth weight (mean gestational weeks 39.4) and compared with children born with moderately low birth weight (mean gestational weeks 35.6) when assessed with Denver-II tool between ages 9 to 34 months (corrected) (Schendel 1997). Moderate quality evidence from one study (n=764) show an increased risk of emotional or behavioural problems and obsessive compulsive behaviour at 5 years among children born at <32 weeks of gestation (assessed with the FTF questionnaire) (Rautava 2010). Low quality evidence from one study (n=2098) show an association between being born extremely preterm (<28 weeks or with birth weight <1000 g) and symptoms of obsessive compulsive disorder at 11 years (Fevang 2016). The same study found an association between gestational age and both parent- and teacher-reported symptoms of autism spectrum disorder (assessed by Autism Spectrum Screening Questionnaire [ASSQ]) at 11 years.

In relation to biological factors

Sex of the child

Low to moderate quality evidence from two studies (three publications, sample sizes ranging from 625 to 1228) shows no association between child's sex and behavioural problems among children born preterm (Delobel-Ayoub 2009; Delobel-Ayoub 2006; Johnson 2015b). The first study assessed children born <33 weeks of gestation at 3 and 5 years of age with the SDQ (Delobel-Ayoub 2006; Delobel-Ayoub 2009) and the second study assessed moderate to late preterm (32-36 weeks) children at two years corrected age on delayed socioemotional competence (assessed with BITSEA) (Johnson 2015b).

Small for gestational age

Moderate quality evidence from two studies (n=1228; n=1277) showed no difference in total behavioural difficulties for SGA preterm infants as compared to those who were appropriate for gestational age. Children were assessed at 3 to 5 years of age and were born at 22-32 weeks. However, one of these studies did observe an increase in the risk of inattention-hyperactivity symptoms for SGA preterm infants born at 29-32 weeks (Delobel-Ayoub 2006; Guellec 2011). In addition, low quality evidence from one study (n=625) found no association between being born SGA and delayed socioemotional competence (assessed with BITSEA) at 2 years (corrected age) among moderate to late children born preterm (Johnson 2015b).

Ethnicity or race

Low quality evidence from one study (n=625) show an association between being non-white and delayed socioemotional competence (assessed with BITSEA) at 2 years (corrected age) among moderate to late children born preterm (Johnson 2015b).

In relation to neonatal factors

Brain abnormalities

Moderate quality evidence from one study show an increase in the risk of behavioural difficulties (assessed with the SDQ) for preterm infants with major cerebral lesions when assessed at the age of 3 years (Delobel-Ayoub 2006). The children were born at 22-32 weeks, and 1228 children were included. The same study (different publication, n=1102) conducted further follow-up at 5 years of age and found no association between brain lesions (level of severity not considered) and behavioural problems (Delobel-Ayoub 2009).

Antenatal steroids

Low quality evidence from one study (n=625) show no association between exposure to antenatal steroids and delayed socioemotional competence (assessed with BITSEA) at 2 years (corrected age) among moderate to late children born preterm (Johnson 2015b).

Bronchopulmonary dysplasia

Moderate quality evidence from one study (n=1228) did not show any difference in the risk of behavioural problems for preterm infants who had bronchopulmonary dysplasia, as compared to those who did not (Delobel-Ayoub 2006). The children were born at 22-32 weeks and followed up at 3 years of age.

No evidence was identified on the relationship between other neonatal factors and behavioural, social, emotional and attention problems.

In relation to social, environmental or maternal factors

Socioeconomic status

Low to moderate quality evidence from three studies show mixed results on behavioural outcomes in relation to socioeconomic status. Moderate quality evidence from one study (n=1102) show no association between socioeconomic status and behavioural problems (assessed with the SDQ) in very preterm (22-32 weeks) at 5 years (Delobel-Ayoub 2009). Moderate quality evidence from another study (n=1458) show an increase in the odds of behavioural problems and internalising problems (assessed with the CBCL) for children born to families with lower socioeconomic status (Potijk 2015). Increased odds of externalising problems was borderline significant. This study included children born between 32 and 41 weeks of gestation and followed up at 4 years. Low quality evidence from a third study (n=625) found an association between lower socioeconomic status and delayed socioemotional competence (assessed with BITSEA) at 2 years of age (corrected) among moderate to late children born preterm (Johnson 2015b).

Maternal substance abuse

Low quality evidence from one study (n=625) show an association between recreational use of drugs during pregnancy and delayed socioemotional competence (assessed with BITSEA) at 2 years (corrected age) among moderate to late children born preterm (Johnson 2015b).

Multiple pregnancy

Low quality evidence from one study (n=625) show no association between multiple pregnancy and delayed socioemotional competence (assessed with BITSEA) at 2 years (corrected age) among moderate to late children born preterm (Johnson 2015b).

Maternal age

Moderate quality evidence from one study (n=1228) show an increase in the risk of behavioural problems (assessed by SDQ) for preterm infants (born at 22-32 weeks gestation and followed up at 3 years of age) born to mothers less than 25 years (compared with mothers 25-34 years) (Delobel-Ayoub 2006). Maternal age of 35 years or more was not associated with behavioural problems in this study. When the children were followed up at 5 years of age (n=1102), the increased odds of behavioural problems of preterm children of mothers younger than 25 years at the time of birth remained borderline significant (Delobel-Ayoub 2009). The association between maternal age 35 years or older and behavioural problems also became borderline significant (borderline reduced odds of behavioural problems compared with maternal age of 25-34 years).

Maternal mental health

Moderate quality evidence from one study (n=1102) show an increase in the risk of behavioural problems (assessed by the SDQ) at 5 years of age for preterm children (born at 22-32 weeks) born to mothers with poorer self-reported mental health (Delobel-Ayoub 2009).

4.1.1.4.9 Special educational needs

In relation to gestational age

Low to high quality evidence from five different studies (eight publications, sample sizes ranging from 6031 to 407503) on the relationship between gestational age and special education needs (SEN) show somewhat mixed findings. SEN were defined differently across the studies, similarly the sample sizes and age at assessment varied between studies.

Moderate quality evidence from one study (n=1766) show children born at <33 weeks of gestation (compared to term) to be at an increased risk of needing special care and/or support at school and repeating a year when in a mainstream class (Larroque 2011). The children were assessed at 8 years of age. Being in an institution or special class or school did not reach statistical significance. High quality evidence from another study (n=6174) that looked at teacher-reported SEN (through a questionnaire) of children born preterm at different gestational ages against their term peers matched by either chronological age, corrected age, or corrected age and year of schooling show mixed findings (Odd 2013a). When matched by chronological age (i.e. uncorrected age) or by corrected age, there was an increased odds of special educational needs among children born premature in all gestational groups (<37 weeks, 32-36 weeks and <32 weeks of gestation), however, due to a small number in the <32 weeks group, it did not reach statistical significance. When matched by corrected age and year of schooling, no statistically significant association was found in either gestational age group. The children were assessed at age 8 years. Moderate quality evidence from one study (n=12586) showed increased odds of SEN in children born at <37 weeks of gestation compared to the term group at 14 to 16 years age (Odd 2016). Moderate quality evidence from one study (n=407503) using school census data among 5- to 18-year-old school children show an increased odds of learning disability or physical disability that impact learning among children born preterm compared with term born children, the effect size increasing as gestational age decreases (MacKay 2010). The same study (different publication) also looked at specific types of SEN at 5-18 years (MacKay 2013). Increased odds of sensory SEN, physical or motor SEN, specific learning difficulty SEN, intellectual SEN, and unspecified SEN were observed with increasing effect estimate as gestational age decreases. However, language SEN, social, emotional or behavioural SEN and autistic spectrum disorder SEN showed mixed findings that mainly did not reach statistical significance.

Low to high quality evidence from four studies (sample sizes ranging from 6031 to 12586) mostly show an association between low gestational age and poor performance in Key Stages 1 (KS1) score (Chan 2014; Odd 2016; Odd 2013a; Peacock 2012). High quality evidence from one study (n=11169) that examined the overall KS1 score at 8 years in children born at different gestational ages against their term peers matched by either chronological age, corrected age, or corrected age and year of schooling show slightly mixed findings (Odd 2013a). When matched by chronological age and corrected age, all preterm children (<32, 32-36, and <37 weeks) had an increased odds of low KS1 score compared to their term peers. However, when matched by corrected age and the year of schooling, the association was no longer statistically significant in either gestational age group. Low quality evidence from another study (n=6031) show an increased odds of low overall KS1 at 7 years of age among children born preterm compared with term (<32, 32-33, and 34-36 weeks of gestation) (Chan 2014). This study also looked at KS1 scores on specific domains and found an increased odds of low KS1 reading score and low KS1 writing score among all preterm children regardless of their gestational age at birth. Low KS1 speaking and listening score was only significant among the children born at <32 weeks of gestation. There was no statistical difference between children born preterm and term born children on low KS1 mathematics score. Low KS1 science score was only significant among the children born at 32-33 weeks of gestation. Moderate quality evidence from one study (n=10279) show a decreased odds of success in KS1 overall assessment among children born at 32-36 weeks of gestation compared to their full-term born peers (Peacock 2012). Children born preterm were also less likely to succeed in KS1 reading, writing and mathematics assessments compared to the term children. Finally one study of moderate quality showed an increased odds of low KS1 score among children born at <37 weeks of gestation compared to full term children at age 5-7 years (Odd 2016).

Moderate quality evidence from one study (n=7650) that reports teacher assessment of the Foundation Stage Profile (FSP) of children at 5 years of age, comparing term born children to children born preterm (23-31 weeks; 32-33 weeks; and 34-36 weeks) was available (Quigley 2012). A significant or borderline significant association with not good level of overall achievement was found among all gestational age groups compared to full-term born children. The children born at 23-31 weeks of gestation had an increased odds of performing poorly in personal, social and emotional development scales. Children born at 34-36 weeks of gestation had a borderline significant increased odds. All gestational age groups had a borderline significant increased odds of performing poorly in communication, language and literacy. All preterm children had increased odds of performing poorly in mathematical development scales, the association among late children born preterm (34-36 weeks) was borderline significant.

In relation to biological factors

Small for gestational age

Low quality evidence from one study (n=1439) that examined the association between being born small for gestational age and having school difficulties (defined as needing special schooling or having low grades, reported by parents) among children born preterm at eight years of age was available (Guellec 2011). No association was found among small for gestational age children born at 24-28 weeks of gestation but an increased odds of school difficulties was found among small for gestational age children born at 29-32 weeks of gestation.

No evidence was identified on the relationship between other biological factors and special educational needs.

Sex

Low quality evidence from one study (n=219) show male children born at <26 weeks of gestation to be more likely to be provided special educational needs support at 11 years compared to their female peers (Johnson 2011).

Ethnicity or race

Low quality evidence from one study (n=219) show no association between maternal ethnicity and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

In relation to neonatal factors

Brain abnormalities

Low quality evidence from one study (n=219) show a significant association between abnormal cerebral ultrasound scan and special educational needs at 11 years among children born at <26 gestational weeks (Johnson 2011).

Necrotising enterocolitis

Low quality evidence from one study (n=219) show no association between necrotising enterocolitis and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

Antenatal steroids

Low quality evidence from one study (n=219) show no association between any exposure to antenatal steroids and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

Postnatal steroids

Low quality evidence from one study (n=219) show no association between any exposure to postnatal steroid for chronic lung disease and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

No evidence was identified on the relationship between other neonatal factors and special educational needs.

In relation to social, environmental or maternal factors

Maternal age

Low quality evidence from one study (n=219) show no association between maternal age and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

Socioeconomic status

Low quality evidence from one study (n=219) show no association between socioeconomic status and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

Chorioamnionitis

Low quality evidence from one study (n=219) show no association between chorioamnionitis (suspected or proven) and special educational needs among extremely children born preterm at 11 years of age (born at <26 gestational weeks) (Johnson 2011).

No evidence was identified on the relationship between other social, environmental or maternal factors and special educational needs.

4.1.2 Risk of developmental disorders

Review question:

What is the risk of developmental disorders in babies, children and young people born preterm at different gestational ages?

How do the following factors influence the risk of developmental disorders in babies, children and young people born preterm:

- **biological factors**
- **neonatal factors**
- **socioeconomic, maternal and environmental factors**
- **postnatal factors?**

4.1.2.1 Description of clinical evidence

The aim of this review was to identify different factors (gestational age at birth; biological factors; neonatal factors; social, environmental or maternal factors; and postnatal factors) that can affect the risk of developmental disorders in babies, children and young people born preterm. Biological factors that were considered included sex of the child, being born small for gestational age, and ethnicity or race. Neonatal factors included brain abnormalities, sepsis, retinopathy of prematurity, necrotising enterocolitis, exposure to antenatal steroids, exposure to postnatal steroids, and bronchopulmory dysplasia. Social, maternal or environmental factors included socioeconomic status, maternal substance abuse, maternal alcohol abuse, multiple pregnancy, chorioamnionitis, neglect, maternal age and maternal mental health problems. Postnatal factors included epilepsy and age at establishing oral feeding.

Developmental disorders considered as outcomes included cerebral palsy (CP), intellectual disability, specific learning impairment, speech and language impairment, attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), mental and behavioural disorders, developmental co-ordination disorder and hearing and visual impairments. In addition, composite neurodevelopmental or neurosensory outcomes were considered. Composite neurodevelopmental outcome was defined as the child having one or more of the following disorders: moderate to severe intellectual disability, CP or motor delay, vision impairment or hearing impairment. Composite neurosensory outcome was defined as having one or more of the following disorders: CP or motor delay, vision impairment or hearing impairment.

Studies were included if they: 1) were prospective or retrospective population-based or multi-centre cohort studies; 2) included only participants born after 1990; 3) confounders were adjusted for in the analyses. For full details see review protocol in Appendix D:.

In total, 64 publications were included in the review (Adams-Chapman 2008; Allred 2014; Ambalavanan 2012; Andrews 2008; Beaino 2010; Beaino 2011; Bolisetty 2014; Burnett 2014; Carlo 2011; Davis 2007; DeJesus 2013; Foix-L'Helias 2008; Goldstein 2013; Guellec 2011; Hansen 2004; Helderman 2012; Herber-Jonat 2014; Hillemeier 2011; Hintz 2005; Hirvonen 2014; Hoffman 2015; Hwang 2013; Johnson 2010; Johnson 2011; Kallen 2015; Kent 2012; Kiechl-Kohlendorfer 2013; Kuzniewicz 2014; Larroque 2008; Laughon 2009; Leversen 2010; Marret 2007; Merhar 2012; O'Shea 2008; Mikkola 2005; Miyazaki 2016; Moore 2012; Natarajan 2012; Odd 2013; Pappas 2014; Payne 2013; Perrott 2003; Petrini 2009; Rabie 2015; Rogers 2013; Serenius 2013; Shah 2012; Shankaran 2004; Singer 2001; Singh 2013; Stoll 2004; Sukhov 2012; Tommiska 2003; Toome 2013; VanMarter 2011; Victorian Infant Collaborative Study Group 2000; Vincer 2006; Vohr 2000; Vohr 2005; Walsh 2005; Wolke 2008; Wong 2014; Wood 2005; Woythaler 2011).

Thirty-three of the studies came from the United States (Adams-Chapman 2008; Allred 2014; Ambalavanan 2012; Andrews 2008; Carlo 2011; DeJesus 2013; Goldstein 2013; Helderman 2012; Hillemeier 2011; Hintz 2005; Hoffman 2015; Kuzniewicz 2014; Laughon 2009; Merhar 2012; O'Shea 2008; Moore 2012; Natarajan 2012; Pappas 2014; Payne 2013; Petrini 2009; Rabie 2015; Rogers 2013; Shah 2012; Shankaran 2004; Singer 2001; Singh 2013; Stoll 2004; Sukhov 2012; VanMarter 2011; Vohr 2000; Vohr 2005; Walsh 2005; Woythaler 2011). Six studies came from both Australia (Bolisetty 2014; Burnett 2014; Davis 2007; Kent 2012; Victorian Infant 2000; Wong 2014) and France (Beaino 2010; Beaino 2011; Foix-L'Helias 2008; Guellec 2011; Larroque 2008; Marret 2007). Four studies came from the United Kingdom and Ireland (Johnson 2010; Johnson 2011; Wolke 2008; Wood 2005) and 1 study came from the United Kingdom (Odd 2013). Three studies came from Finland (Hirvonen 2014; Mikkola 2005; Tommiska 2003). Two studies came from Canada (Perrott 2003; Vincer 2006) and Sweden (Kallen 2015; Serenius 2013). One study came from each of the following countries: Austria (Kiechl-Kohlendorfer 2013), Denmark (Hansen 2004), Estonia (Toome 2013), Germany (Herber-Jonat 2014), Japan (Miyazaki 2016), Norway (Leversen 2010), and Taiwan (Hwang 2013).

Fifty-three studies were population-based or multi-centre prospective cohort studies (Adams-Chapman 2008; Allred 2014; Ambalavanan 2012; Andrews 2008; Beaino 2010; Beaino 2011; Bolisetty 2014; Burnett 2014; Carlo 2011; Davis 2007; Foix-L'Helias 2008; Guellec 2011; Hansen 2004; Helderman 2012; Herber-Jonat 2014; Hillemeier 2011; Hwang 2013; Johnson 2010; Johnson 2011; Kallen 2015; Kent 2012; Kiechl-Kohlendorfer 2013; Kuzniewicz 2014; Larroque 2008; Leversen 2010; Marret 2007; Merhar 2012; O'Shea 2008; Mikkola 2005; Natarajan 2012; Odd 2013; Payne 2013; Perrott 2003; Petrini 2009; Rabie 2015; Rogers 2013; Serenius 2013; Shah 2012; Shankaran 2004; Singer 2001; Singh 2013; Stoll 2004; Tommiska 2003; Toome 2013; VanMarter 2011; Victorian Infant 2000; Vincer 2006; Vohr 2000; Vohr 2005; Walsh 2005; Wolke 2008; Wood 2005; Woythaler 2011). Nine studies were retrospective cohort studies (DeJesus 2013; Goldstein 2013; Hintz 2005; Hoffman 2015; Laughon 2009; Miyazaki 2016; Moore 2012; Pappas 2014; Wong 2014) and two studies used population-based registry data (HGirvonen 2014; Sukhov 2012).

Seventeen publications stemmed from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN). Twelve publications came from the Extremely Low Gestational Age Newborns (ELGAN) study from the US (Allred 2014; Helderman 2012; Hillemeier 2011; Kuzniewicz 2014; Laughon 2009; O'Shea 2008; Petrini 2009; Rabie 2015; Rogers 2013; Singh 2013; VanMarter 2011; Woythaler 2011). Six publications came from the French Etude Epidemiologique sur les Petits Ages Gestationnels (EPIPAGE) study (Beaino 2010; Beaino 2011; Foix-L'Helias 2008; Guellec 2011; Larroque 2008; Marret 2007). Four publications came from the EPICure study from the UK and Ireland (Johnson 2010; Johnson 2011; Wolke 2008; Wood 2005). Three publications came from an Australian cohort of extremely preterm infants admitted to any of the 10 NICUs within New South Wales (NSW) and the Australian Capital Territory (Bolisetty 2014; Kent 2012; Wong 2014) and three publications came from the Victorian Infant Collaborative Study Group (Burnett 2014; Davis 2007; Victorian Infant Collaborative Study Group 2000). The rest of the included publications were the only publications from their respective cohort studies.

Gestational age as a risk for developmental disorders

Nineteen studies studied the association between gestational age (preterm versus term) and different developmental disorders (Burnett 2014; Helderman 2012; Hillemeier 2011; Hirvonen 2014; Johnson 2010; Johnson 2011; Kent 2012; Kuzniewicz 2014; Larroque 2008; Odd 2013; Petrini 2009; Rabie 2015; Rogers 2013; Serenius 2013; Singh 2013; Sukhov 2012; Toome 2013; Wolke 2008; Woythaler 2011). Five of these studies looked at the association between gestational age and CP (Hirvonen 2014; Odd 2013; Petrini 2009; Sukhov 2012; Toome 2013). Eight studies looked at the association between gestational age and intellectual disability (Burnett 2014Helderman 2012; Hillemeier 2011; Larroque 2008; Petrini

2009; Serenius 2013; Singh 2013; Toome 2013; Woythaler 2011). Four studies looked at the association between gestational age and speech, language and communication delay (Rabie 2015; Serenius 2013; Toome 2013; Wolke 2008). Four studies looked at the association between gestational age and mental and behavioural disorders Burnett 2014; Johnson 2010; Rogers 2013; Singh 2013). Two studies looked at the association between gestational age and autism spectrum disorder (Kuzniewicz 2014; Singh 2013) and attention deficit hyperactivity disorder (Rabie 2015; Singh 2013). Two studies looked at the association between gestystional age and neurosensory or neurodevelopmental composite outcome (Kent 2012; Toome 2013). One study looked at the association between gestational age and specific learning difficulties (Johnson 2010).

No evidence was found on the association between gestational age and developmental co-ordination disorder among children born preterm. No evidence was identified on the association between gestational age and hearing or visual impairment, although these outcomes were included in a composite outcome measure in 2 studies (Kent 2012; Toome 2013).

Biological factors as risk for developmental disorders

Twenty-four publications studied the association between biological factors (sex of the child, being born small for gestational age, and ethnicity or race) and developmental disorders among children born preterm (Ambalavanan 2012; Andrews 2008; Beaino 2011; Bolisetty 2014; Davis 2007; De Jesus 2013; Guellec 2011; Hansen 2004; Helderman 2012; Hirvonen 2014; Hoffman 2015; Hwang 2013; Kent 2012; Kuzniewicz 2014;Leversen 2010; Marret 2007; Moore 2012; Natarajan 2012; Shankaran 2004; Singh 2013; Tommiska 2003; Toome 2013; Vohr 2000; Walsh 2005). Ten of these studies reported on the association between biological factors and CP (Andrews 2008; Beaino 2011; Guellec 2011; Hansen 2004; Hirvonen 2014; Marret 2007; Shankaran 2004; Tommiska 2003; Toome 2013; Vohr 2000). Twelve studies reported on the association between biological factors and intellectual disability (Ambalavanan 2012; Andrews 2008; Beaino 2011; Hansen 2004; Helderman 2012; Hoffman 2015; Marret 2007; Natarajan 2012; Shankaran 2004; Singh 2013; Toome 2013; Vohr 2000) and two studies on speech, language or communication impairment (Hoffman 2015; Toome 2013). One study reported on the association between biological factors and mental or behavioural disorders (Singh 2013) and four studies on ASD (Hwang 2013; Kuzniewicz 2014; Moore 2012; Singh 2013), and one study on ADHD (Singh 2013). One study reported on the association between biological factors and vision impairment and hearing impairment (DeJesus 2013). Six studies looked at the association between different biological factors and composite neurodevelopmental or neurosensory outcome (Bolisetty 2014; Kent 2012; Leversen 2010; Shankaran 2004; Toome 2013; Walsh 2005).

No evidence was found on the association between different biological factors and developmental co-ordination disorder or specific learning impairment among children born preterm.

Neonatal factors as risk for developmental disorders

Forty publications reported on the association between different neonatal factors (brain abnormalities, sepsis, retinopathy of prematurity, necrotising enterocolitis, exposure to antenatal steroids, exposure to postnatal steroids, bronchopulmory dysplasia) and developmental disorders amonf children born preterm (Adams-Chapman 2008; Allred 2014; Andrews 2008; Beaino 2010; Beaino 2011; Bolisetty 2014; Carlo 2011; Foix-L'Helias 2008; Goldstein 2013; Hansen 2004; Herber-Jonat 2014; Hintz 2005; Hirvonen 2014; Hoffman 2015; Hwang 2013; Johnson 2010; Kallen 2015; Kiechl-Kohlendorfer 2013; Kuzniewicz 2014; Laughon 2009; Leversen 2010; Merhar 2012; O'Shea 2008; Mikkola 2005; Natarajan 2012; Payne 2013; Perrott 2003; Shah 2012; Shankaran 2004; Stoll 2004; Tommiska 2003; Toome 2013; VanMarter 2011; Victorian Infant Collaborative Study Group 2000; Vincer 2006; Vohr 2000; Vohr 2005; Walsh 2005; Wong 2014; Wood 2005). Of these studies, 22 reported on the association between different neonatal factors and CP (Adams-Chapman 2008; Allred 2014; Andrews 2008; Beaino 2010; Beaino 2011; Carlo 2011; Foix-L'Helias 2008; Hansen 2004; Hintz 2005; Hirvonen 2014; Mikkola 2005; Payne 2013; Shankaran 2004; Stoll 2004; Tommiska 2003; Toome 2013; VanMarter 2011; Victorian Infant Collaborative Study Group 2000; Vincer 2006; Vohr 2000; Vohr 2005; Wood 2005), and 22 reported on intellectual disability (Adams-Chapman 2008; Allred 2014; Andrews 2008; Beaino 2010; Beaino 2011; Carlo 2011; Foix-L'Helias 2008; Hansen 2004; Hintz 2005; Hoffman 2015; Kallen 2015; Laughon 2009; O'Shea 2008; Mikkola 2005; Natarajan 2012; Payne 2013; Shah 2012; Shankaran 2004; Stoll 2004; Toome 2013; Vohr 2000; Vohr 2005). One study reported on the association between neonatal factors and specific learning impairment (Kiechl-Kohlendorfer 2013), and three studies reported on speech, language or communication impairment (Hoffman 2015; Payne 2013; Toome 2013). One study reported on the association between neonatal factors and mental dirorders (Johnson 2010), and 2 studies on ASD (Hwang 2013; Kuzniewicz 2014). Four studies reported on the association between neonatal factors and vision impairment (Adams-Chapman 2008; Carlo 2011; Mikkola 2005; Stoll 2004), and three studies on hearing impairment (Adams-Chapman 2008; Carlo 2011; Stoll 2004). Nineteen studies reported on the association between neonatal factors and composite neurodevelopmental or neurosensory outcome (Adams-Chapman 2008; Bolisetty 2014; Carlo 2011; Goldstein 2013; Herber-Jonat 2014; Hintz 2005; Kallen 2015; Leversen 2010; Merhar 2012; Payne 2013; Perrott 2003; Shah 2012; Shankaran 2004; Stoll 2004; Toome 2013; Victorian Infant Collaborative Study Group 2000; Vohr 2005; Walsh 2005; Wong 2014).

No evidence was found on the association between neonatal factors and developmental co-ordination disorder and ADHD among children born preterm.

Social, environmental and maternal factors as risk for developmental disorders

Fourteen publications studied the association between different social, environmental and maternal factors (socioeconomic status, maternal substance abuse, maternal alcohol abuse, multiple pregnancy, chorioamnionitis, neglect, maternal age and maternal mental health problems) and developmental disorders among children born preterm (Beaino 2010; Beaino 2011; Hirvonen 2014; Hoffman 2015; Kallen 2015; Leversen 2010; Marret 2007; Miyazaki 2016; Pappas 2014; Shankaran 2004; Singer 2001; Tommiska 2003; Toome 2013; Wood 2005). Ten of these studies reported on the risk of CP (Beaino 2010; Beaino 2011; Hirvonen 2014; Marret 2007; Miyazaki 2016; Pappas 2014; Shankaran 2004; Tommiska 2003; Toome 2013; Wood 2005), and ten on intellectual disability (Beaino 2010; Beaino 2011; Hoffman 2015; Kallen 2015; Marret 2007; Miyazaki 2016; Pappas 2014; Shankaran 2004; Singer 2001; Toome 2013). Two studies reported on speech, language or communication impairment (Hoffman 2015; Toome 2013) and one on vision impairment and hearing impairment (Miyazaki 2016). Six studies reported on the association between different social, environmental or maternal factors on composite neurodevelopmental or neurosensory outcome (Kallen 2015; Leversen 2010; Pappas 2014; Shankaran 2004; Singer 2001; Toome 2013).

No evidence was found on the association between social, environmental and maternal factors and developmental co-ordination disorder, specific learning impairment, mental disorders, ASD, or ADHD among children born preterm.

No evidence was found on the association between postnatal factors and developmental disorders among children born preterm.

The feasibility of combining study data using meta-analysis was assessed. Due to the following differences between studies, it was not considered appropriate to pool the results:

- the inclusion/exclusion criteria for participants
- ages of participants at the time of assessment

- confounders adjusted for in multivariate analysis models
- outcome definitions and measurement tools
- consistency of results.

4.1.2.2 Summary of included studies

Table 15: Summary of studies on the association between gestational age and developmental disorders

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Cerebral palsy						
Odd 2013	Prospective regional cohort study	N=13,843 Analysis compares moderate to late preterm infants (32-36 weeks) to full term (37-42 weeks)	Infants with cerebral palsy were identified from hospital and community health service records, and the diagnosis confirmed at age 4 using the Standard Recording of Motor Deficit	Ethnicity, housing, crowding, maternal education, socio-economic group, car ownership, age, gender, parity, weight, length, head circumference at birth, mode of delivery, maternal hypertension, pyrexia, need for resuscitation at birth	Cerebral palsy at 7 years age: Term: reference 32-36 weeks: OR 6.38 (2.28-17.76)	Moderate
Hirvonen 2014	Population based retrospective cohort using national registry data	Overall sample: n=1039263 Sample size after exclusions: n=1018302 (included for comparisons of cerebral palsy risk at different gestational ages)	All inpatient and outpatient visits due to a CP diagnosis in public hospitals were registered. The diagnosis of CP in Finland is based on medical history, ultrasound and MRI data, and multidisciplinary	Maternal age, maternal smoking status, primiparous, previous C-section, maternal diabetes, multiple pregnancy, order of fetuses,	By the age of 7 years Cerebral palsy Gestational age Term: Reference <32 weeks: OR 9.37 (7.34-11.96) 32+0 to 33+6 weeks: OR 5.12 (4.13-6.34) 34+0 to 36+6 weeks: OR 2.35 (1.99 to 2.77)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
		n=53078	evaluations in the paediatric neurology units of 20 secondary level central hospitals and 5 tertiary level university hospitals. The diagnosis is included in the database as soon as it has been established	assisted reproductive technology, cervical cerclage, chorionic villus sampling, PROM, preeclampsia, time of birth, antenatal steroid use, place of birth, mode of delivery, gender, gestational weight, birth weight <1500g, Apgar score, umbilical artery pH, admission to neonatal unit, ventilator, resuscitation at birth, phototherapy, antibiotic therapy, RDS, sepsis, intracranial haemorrhage, convulsions and hyperbilirubinaemia.		
Petrini 2009	Regional retrospective cohort study	n=141321 Analysis compares	ICD 9 codes of patient diagnoses in electronic medical	Maternal ethnicity, sex, multiple	During follow-up time of up to 5.5 years Cerebral palsy	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
		preterm infants to full term (37-41 weeks)	records were used to identify cases of cerebral palsy and developmental delay/mental retardation.	pregnancy and size for gestational age.	Term: Reference 34-36 weeks: HR 3.39 (2.54-4.52) 30-33 weeks: HR 7.87 (5.38-11.51)	
Sukhov 2012	Retrospective cohort study using population registry data	n=6,145,357 Analysis compares different groups of preterm infants to term (≥37 weeks)	Cerebral palsy was identified through an administrative database from 21 non-profit regional centres which provide therapy services to people with developmental disabilities including CP.	Maternal age, parity, maternal education, payer-source, ethnicity, timing of initiation of prenatal care, number of prenatal visits, gestational age, birthweight, multiple pregnancy, gender, placental abruption, fetal distress, mild to severe birth asphyxia, birth defects, birth trauma, meningitis and cord prolapse.	At between 5 and 15 years Cerebral palsy Term: Reference 32-36 weeks: OR 2.20 (2.05-2.36) 28-31 weeks: OR 8.83 (8.04-9.70) < 28 weeks: OR 18.21 (16.70-19.86)	Moderate
Intellectual disability						
Woythaler 2011	Population based prospective cohort study	n=1200 preterm infants (34-36+6 weeks)	The mental development index (MDI)) of the Bayley Short Form Research edition (BSF-R) were	Gestational age, plurality, maternal race, education, marital status,	At 2 years chronological age Severe developmental delay Term: Reference 34-36+6 weeks: OR 1.51 (1.26-1.82)	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
		n=6300 term infants (≥37 weeks)	used to identify developmental delay and psychomotor developmental delay. Abnormal scores were identified as mild abnormality (between 1SD and 2SD below the mean score) and severe abnormality (<2SD below the mean score).	depression, prenatal care, primary language, infant gender, poverty level, delivery type, fetal growth and any breast milk feeding.	Mild developmental delay Term: Reference 34-36+6 weeks: OR 1.43 (1.22-1.67)	
Serenius 2013	Population based prospective cohort study (EXPRESS)	n=456 preterm infants (<27 weeks) n=701 full term controls (37-41 weeks)	Cognitive, language and motor development were all assessed with the Bayley- Scales of Infant and Toddler Development (Bayley-III). Cognitive, language and motor development was considered normal if the composite score on the respective Bayley-III scale was within 1 SD of the norm, mildly impaired if the score was between 1 and 2SD below the norm, moderately impaired if the score was between 2 and 3 SD below the norm, and	Maternal country of birth (Nordic/non-Nordic), maternal and paternal educational level	At 2.5 years corrected age Mild cognitive impairment Term: Reference <27 weeks: OR 4.3 (2.3-7.9) Mild mental developmental delay Term: Reference <27 weeks: OR 3.0 (1.8-5.0) Moderate mental developmental delay Term: Reference <27 weeks: OR 6.4 (2.4-17.1)	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			severely impaired if the score was < 3SD below the norm.			
Larroque 2008	Population based prospective cohort study (EPIPAGE)	n=1534 preterm children born at 22 to 32 completed weeks gestation n=320 term controls born at 39-40 weeks	Mental Processing Composite (MPC) of the Kaufmann Assessment Battery for Children (K-ABC) was used to assess intellectual disability. Scores of <2SD below the mean were taken as abnormal.	Maternal age, parity, maternal education, maternal birthplace and socioeconomic status.	At age 5 years Intellectual disability (MPC score 55-69) Term: Reference 22-32 weeks: OR 3.4 (1.8-6.4)	High
Petrini 2009	Regional retrospective cohort study	n=141321 Analysis compares preterm infants to full term (37-41 weeks)	ICD 9 codes of patient diagnoses in electronic medical records were used to identify cases of cerebral palsy and developmental delay/mental retardation.	Maternal ethnicity, sex, multiple pregnancy and size for gestational age.	During follow-up time of up to 5.5 years For the outcome of Developmental delay/mental retardation Term: Reference 34-36 weeks: HR 1.25 (1.01-1.54) 30-33 weeks: HR 1.90 (1.34-2.71)	Moderate
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (≥37 weeks)	Parents were asked to self- report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	Household composition, place of residence and highest household/parental education.	During follow-up period of between 2 and 17 years Intellectual disability/mental retardation Term: Reference <37 weeks: OR 2.74 (2.02-3.73)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Helderman 2012	Multicentre Prospective cohort study	Sample recruited: n=1506 Sample eligible for assessment: n=1200 Sample analysed after exclusions:n=921	The assessment of developmental delays (determined by cognitive impairment Mental Development Index [MDI]) at 24-months adjusted age at 24-months included the Bayley Scales of Infant Development-2nd Edition (BSID-II). Cognitive impairment was defined as an MDI <70. An MDI <55 was considered severe cognitive impairment.	Single mother, BMI>30, vaginal/cervical infection, caesarean delivery, BWZ <-2, mother's education <12 years or >16 years, Hospital cluster	Intellectual disability (developmental delay - Mental Developmental Index [MDI]) Gestational age 23–24 week - (RR [95% CIs]) Referent group is infants with MDI <70 MDI < 55: 1.9 (0.97, 3.6) MDI = 55–69: 1.0 (0.5, 1.9) Gestational age 25–26 week - (RR [95% CIs]) Referent group is infants with MDI <70 MDI < 55: 1.2 (0.7, 2.1) MDI = 55–69: 0.8 (0.5, 1.3)	Moderate (the study was downgraded for risk of bias because the confounders for adjustment were not reported clearly)
Hillemeier 2011	National longitudinal cohort study	n=7,200	Cognitive delay was assessed at 24 and 48 months age using the Bayley Short Form-Research Edition (BSF-R). Children scoring the lowest 10% of the scale were considered to have cognitive delay. At 48 months, Bayley assessment was not possible due to age, therefore a standardised assessment developed for other large studies of child development. Children	Adjustment for sex, age, race/ethnicity, socioeconomic variables, characteristics of gestation and infant status at birth	At 24 months: Cognitive delay Gestational age Term Ref Moderately preterm (33-36 weeks) OR 1.07 (NS, 95% CI not presented) Very preterm (<=32 weeks) 1.52 (NS) The model adjusted for sex, age, race/ethnicity, socioeconomic variables, characteristics of gestation and infant status at birth. At 48 months: Cognitive delay Gestational age	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			scoring lowest 10% were considered to have cognitive delay		Term Ref Moderately preterm (33-36 weeks) 1.10 (NS) Very preterm (<=32 weeks) 1.86 (NS) The model adjusted for sex, age, race/ethnicity, socioeconomic variables, characteristics of gestation and infant status at birth.	
Speech and/or language disorder						
Serenius 2013	Population based prospective cohort study (EXPRESS)	n=456 preterm infants (<27 weeks) n=701 full term controls (37-41 weeks)	Cognitive, language and motor development were all assessed with the Bayley- Scales of Infant and Toddler Development (Bayley-III). Cognitive, language and motor development was considered normal if the composite score on the respective Bayley-III scale was within 1 SD of the norm, mildly impaired if the score was between 1 and 2SD below the norm, moderately impaired if the score was between 2 and 3 SD below the norm, and	Maternal country of birth (Nordic/non-Nordic), maternal and paternal educational level	Mild language impairment at 2.5 years corrected age Term: Reference <27 weeks: OR 3.5 (1.9-6.4) Moderate language impairment Term: Reference <27 weeks: OR 5.1 (1.9-13.8)	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			severely impaired if the score was < 3SD below the norm.			
Rabie 2015	Retrospective cohort study using population registry data	n=38802 Analysis compares late preterm infants to full term (39-41+6 weeks)	ICD-9 codes from Medicaid files were used to identify children with ADHD and developmental speech and/or language delay.	Birth weight, SGA and LGA, gender, ethnicity, hospital characteristics and maternal medical comorbidities (diabetes, hypertension, anaemia, chronic lung disease, herpes, neurologic disorder, coagulation disorder, obesity, depression).	At age 3-5 years. Developmental speech and/or language delay Term: Reference 34-36+6: HR 1.36 (1.23-1.50)	Low
Wolke 2008	National cohort study	n=308 children born <=25 gestational weeks n=241 children survived to follow-up n=160 full-term born children as comparison group, matched by age and sex	Serious impairment in receptive and expressive language ability, evaluated using the Preschool Language Scale-3 (UK) (PLS-3) which comprises Auditory Comprehension and Expressive Communication scales. ◦Total score	Adjusting for MPC score (cognitive ability)	Outcomes assessed at median age of 6 years and 4 months: Serious impairment in language abilities Total score: Full-term Extremely preterm Ref 1.3 (0.3-5.3) Auditory comprehension: Full-term Extremely preterm Ref 1.6 (0.3-9.8)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			◦Auditory comprehension ◦Expressive communication ◦Articulation screener Outcome were dichotomized a priori using a cut-off of 2 SD or the 10th/90th percentiles as appropriate (not specified which one was used for this outcome).		Expressive communication: Full-term Extremely preterm Ref 1.2 (0.2-6.5) Articulation screener: Full-term Extremely preterm Ref 1.1 (0.3-4) Model adjusted for cognitive impairment score (MPC score).	
Attention deficit hyperactivity disorder						
Burnett 2014	Prospective geographical cohort study	n=215 early preterm/extremely low birth weight infants n=157 normal birth weight (>2499 g) controls n=372 in total	Standardized face-to-face clinical interview and questionnaires were used to assess the mental health status in late adolescence ADHD, any type (All ADHD types assessed with the ADHD module of the Children's Interview for Psychiatric Syndromes (ChIPS)) ADHD, combined type ADHD, inattentive type ADHD, hyperactive/impulsive type	Adjusting for sex, parental education and childhood SES.	At age 18 years: ADHD, any type Normal BW EP/ELBW Reference 2.67 (1.08-6.58) ADHD, combined type Normal BW EP/ELBW Reference 4.9 (0.56-43.24) ADHD, hyperactive/impulsive type Normal BW EP/ELBW Reference NR (0 cases in the control group	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Rogers 2013	Cross sectional survey	n=39 preterm (34-36 weeks) n=154 full term (40-41 weeks)	The Preschool Age Psychiatric Assessment (PAPA) was used to establish DSM-IV Axis 1 diagnoses. It was administered by bachelor's or master's level clinicians and final diagnoses were derived using computerised algorithms.	Sex, family income, IQ and ethnicity.	At age 3-6 years Risk of ADHD Term: Reference 34-36 weeks: OR 0.81 (0.29-2.29) ADHD-inattentive Term: Reference 34-36 weeks: OR 1.21 (0.11-13.22)	Low
Rabie 2015	Retrospective cohort study using population registry data	n=38802 Analysis compares late preterm infants to full term (39-41+6 weeks)	ICD-9 codes from Medicaid files were used to identify children with ADHD and developmental speech and/or language delay.	Birth weight, SGA and LGA, gender, ethnicity, hospital characteristics and maternal medical comorbidities (diabetes, hypertension, anaemia, chronic lung disease, herpes, neurologic disorder, coagulation disorder, obesity, depression). OR are unadjusted, as	At age 3-5 years. ADHD Term: Reference 34-36+6 weeks: HR 1.21 (0.98-1.49)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
				adjustment for sex and socioeconomic status did not affect the results significantly.		
Johnson 2010	Population based prospective cohort study (EPICure)	n=219 preterm children born at <26 weeks n=152 term controls (exact gestation not described)	The Development and Wellbeing Assessment was administered via a telephone interview with parents. Potential cases were identified using computer based scoring algorithms, and final DSM-IV diagnoses were assigned by two child and adolescent psychiatrists on review of summary sheets and clinical transcripts	OR are unadjusted, as adjustment for sex and socioeconomic status did not affect the results significantly.	At age 11 years ADHD Term: Reference <26 weeks: OR 4.3 (1.5-13.0) ADHD inattentive subtype Term: Reference <26 weeks: OR 10.5 (1.4-81.1) ADHD combined type Term: Reference <26 weeks: OR 2.1 (0.5-7.9)	Moderate
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (≥37 weeks)	Parents were asked to self- report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	Household composition, place of residence and highest household/parental education.	During follow-up period of between 2 and 17 years ADHD Term: Reference <37 weeks: OR 1.49 (1.29-1.73)	Low
Autism spectrum disorder						
Kuzniewicz 2014	Regional prospective cohort study	n=195021 Analysis compares preterm infants to term (≥37 weeks)	Cases of autistic spectrum disorder identified through a regional autism registry.	Gestational age, sex, maternal age, maternal education and	During follow-up period of age 2-11 Austism spectrum disorder Term: Reference	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			Cases were defined as children with at least one diagnosis of ASD made at an ASD evaluation centre, or by a clinical specialist, or by a general paediatrician.	small for gestational age.	34-36 weeks: HR 1.3 (1.1-1.4) 27-33 weeks: HR 1.4 (1.1-1.8) 24-26 weeks: HR 2.7 (1.5-5.0)	
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (≥37 weeks)	Parents were asked to self- report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	Household composition, place of residence and highest household/parental education.	During follow-up period of between 2 and 17 years Autism spectrum disorder Term: Reference <37 weeks: OR 2.26 (1.69-3.03)	Low
Specific learning difficulty						
Johnson 2011	Population based prospective cohort study (EPICure)	n=219 preterm children born at < 26 weeks n=153 term controls (exact gestation not described)	Wechsler Individual Achievement Test to measure mathematics and reading ability. Scores of <2SD below the mean were taken as abnormal.	OR are unadjusted, as adjustment for maternal education and socioeconomic status did not affect the results significantly.	At age 11 years Reading impairment Term: Reference < 26 weeks: OR 21.6 (6.6-70.4) Mathematics impairment Term: Reference < 26 weeks: OR 58.7 (14.2-242.9)	Moderate
Mental and behavioural disorders						
Burnett 2014	Prospective geographical cohort study	n=215 early preterm/extremely low birth weight infants n=157 normal birth weight (>2499 g) controls	•Any anxiety or mood disorder (All DSM-IV Axis I disorders (mood, anxiety, substance use, psychotic, eating and adjustment disorders) assessed with the	Adjusting for sex, parental education and childhood SES.	At age 18 years: Any anxiety or mood disorder Normal BW EP/ELBW Reference 1.08 (0.61-1.91) Any mood disorder Normal BW EP/ELBW Reference 0.96 (0.51-1.84)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
		n=372 in total	Structured Clinical Interview dor DSM-IV Disorders, Axis 1 Non-Patient version (SCIP-I/NP), Assessments supplemented by questionnaires examining recent anxiety and depression symptoms: the Beck Anxiety Inventory (BAI) and the Center for Epidemiologic Studies Depression Scale - Revised (CESD-R).) •Any mood disorder •Any anxiety disorder •Co-morbid anxiety and mood disorder		Any anxiety disorder Normal BW EP/ELBW Reference 1.11 (0.53-2.33) Co-morbid anxiety and mood disorder Normal BW EP/ELBW Reference 0.90 (0.34-2.41) Any SCID-I/NP diagnosis Normal BW EP/ELBW Reference 1.16 (0.67-2.04)	
Rogers 2013	Cross sectional survey	n=39 preterm (34-36 weeks) n=154 full term (40-41 weeks)	The Preschool Age Psychiatric Assessment (PAPA) was used to establish DSM-IV Axis 1 diagnoses. It was administered by bachelor's or master's level clinicians and final diagnoses were derived using computerised algorithms.	Sex, family income, IQ and ethnicity.	At age 3-6 years Oppositional Defiant Disorder Term: Reference 34-36 weeks: OR 2.30 (0.98-5.40) Conduct Disorder Term: Reference 34-36 weeks: OR 1.60 (0.55-4.66) Any anxiety diagnosis Term: Reference	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
					34-36 weeks: OR 3.74 (1.59-8.78)	
Johnson 2010	Population based prospective cohort study (EPICure)	n=219 preterm children born at <26 weeks n=152 term controls (exact gestation not described)	The Development and Wellbeing Assessment was administered via a telephone interview with parents. Potential cases were identified using computer based scoring algorithms, and final DSM-IV diagnoses were assigned by two child and adolescent psychiatrists on review of summary sheets and clinical transcripts.	OR are unadjusted, as adjustment for sex and socioeconomic status did not affect the results significantly.	At age 11 years Major depression Term: Reference <26 weeks: OR 2.2 (0.2-21.0) Conduct disorder Term: Reference <26 weeks: OR 0.9 (0.4-2.2) Oppositional defiant disorder Term: Reference <26 weeks: OR 1.0 (0.4-2.4)	Moderate
Singh 2013	Cross sectional survey	n=85,535 Separated into premature children (born at <37 weeks) and term children (≥37 weeks)	Parents were asked to self- report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	Household composition, place of residence and highest household/parental education.	During follow-up period of between 2 and 17 years Conduct disorder (including oppositional defiant disorder) Term: Reference <37 weeks: OR 1.50 (1.21-1.86) Anxiety Term: Reference <37 weeks: 1.58 (1.31-1.91) Depression Term: Reference <37 weeks: 1.33 (1.01-1.74)	Low
Composite outcomes						

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Kent 2012	Population based longitudinal cohort study	Sample size N=2701 Followed up at 2-3 years: n=1473	Assessment of outcome involved examination of 4 domains: developmental, neurologic, vision, and hearing Developmental assessment used the Griffiths Mental Developmental Scales or Bayley Scales of Infant Development II Neurologic assessment included evaluation of muscle tone, primitive reflexes, automatic reactions, and volitional movement Cerebral palsy was diagnosed if the child had non-progressive motor impairment characterised by abnormal muscle tone and a decreased range or decreased control of movements, accompanied by neurologic signs Moderate to severe functional disability was defined as one or more of the following: developmental delay	Multiple regression analysis adjusted for male versus female, gestational age, birth weight percentiles, antepartum haemorrhage, pregnancy-induced hypertension, foetal stress, emergency caesarean delivery, Apgar score < 7 at 5 min, outborn versus inborn	At 2-3 years corrected age Gestational age: 27-28 weeks GA: reference 22-26 weeks GA: OR 2.444 (1.831-3.263)	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			(<2SD below the mean for adjusted age determined by the Griffiths Mental Developmental Scales or BSID-II, cerebral palsy (unable to walk without aids), bilateral blindness (visual acuity <6/60 in better eye), or bilateral deafness (requiring bilateral hearing aids or cochlear implants)			
Toome 2013	Population based prospective cohort study	n=155 preterm infants (<32 weeks) n=153 full term controls (≥37 weeks)	Cerebral palsy was defined according to the guidelines of the Surveillance of Cerebral Palsy in Europe collaborative group. The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills. A composite outcome measure of neurodevelopmental impairment was used. This includes any one (or more) of the following criteria: CP with GMFCS level 2,3,4 or 5; cognitive	Antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge,	At corrected age of 2 years Moderate/Severe neurodevelopmental disability (CP with GMFCS level 2,3,4 or 5; cognitive and/or language composite scores of ≤-2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.) OR 0.7 (0.6-0.9) per additional week of gestational age	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			and/or language composite scores of ≤-2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.	maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family		

1 **Table 16: Summary of studies on the association between different biological factors and developmental disorders**

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
Cerebral palsy						
De Jesus 2013	Retrospective cohort study	N=2971 - Infants born between 23 0/7 and 26 6/7 weeks GA	SGA -adjusted for: Random effects variable, male, sex, multiple birth, GA, antenatal corticosteroid use, hypertension, and maternal education	Moderate or severe cerebral palsy (CP) based on presence of bilateral hearing loss (with or without amplification) or bilateral blindness (vision <20/200).	CP assessed at 18-22 months corrected age among children born between 23 and 26 weeks' GA: moderate or severe CP: SGA: OR 2.55, 95%CI 1.69-3.86	Moderate
Shankaran 2004	Prospective cohort study	N=246	Male gender Black race Risk factors adjusted for each other plus surfactant administration, steroids for BPD, Medicaid, no high school degree, 2-parent household	CP was defined as none-progressive central nervous system disorder characterised by abnormal muscle tone in at least one extremity and abnormal control of movement and posture	CP assessed at 18-22 months corrected age among children born at 23.6 weeks GA; Male: 1.2 (0.6-2.4) Black: 1.0 (0.5-2.2)	Low

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
Tommiska 2003	Population based prospective cohort study	N= 208 Infants with a birth weight below 1000g and gestational age of at least 22 full weeks.	Male gender: -Adjusted for: multiparity, pre-eclampsia, premature rupture of membranes, maternal infection, antenatal steroid treatment, hyperstimulation or in vitro fertilisation, maternal age below 20 or above 40, smoking, marital status, social class 1-4, birth in secondary level hospital, catchment area for the different hospitals, vaginal delivery, birth weight (100g groups), intrauterine growth restriction, gestational age, male gender, multiple birth, anomalies, respiratory distress syndrome, septicaemia, necrotising enterocolitis with perforation and intraventricular haemorrhage grades 2-4.		CP assessed at age 18 months corrected age among children born ≥ 22 weeks' GA: Male gender Not a significant independent predictor on multivariate analysis	High
Toome 2013	Prospective population based cohort.	N=187	Male gender SGA	Cerebral palsy was defined according to the guidelines of the Surveillance of Cerebral	Assessed at corrected age 2 years.	High

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
		Born at mean 28.8 (28.4-29.1) weeks gestation	-Adjusted for: Gestational age SGA Maternal age Low income of the family Multiple births Antenatal steroids Postnatal steroids BPD (defined as oxygen dependency at 36 weeks) ROP stage 3-5 with laser therapy Positive blood culture sepsis NEC stage 2-3	Palsy in Europe collaborative group, and the Gross Motor Function Classification System (GMFCS) was used to quantify motor function in infants with CP.	Risk of cerebral palsy Male gender SGA Not found to be significant predictors	
Hansen 2004	Prospective cohort	N=252 Children born at < 28 weeks' GA	Male gender: -Adjusted for: IVH, NEC, CRIB-score (high), chronic lung disease, and mechanic ventilation during neonatal course	Cerebral palsy was diagnosed in accordance with the criteria as defined in the Surveillance of cerebral palsy in Europe	CP assessed at age 5 years among children born < 28 weeks' GA: Sex/boy: 0.5 (0.2-1.6)	Moderate
Marret 2007	Population based prospective cohort	n=2457 children born at 30-34 weeks gestation	Male gender -Adjusted for: Cerebral palsy was defined as at least two of: abnormal posture or movement, increased tone and hyperreflexia. When the diagnosis of	Cerebral palsy was defined as at least two of: abnormal posture or movement, increased tone and hyperreflexia. When the diagnosis of cerebral palsy was in doubt, a panel of trained paediatricians met to discuss the case.	CP assessed at age 5 years age among children born at 30-34 weeks gestation: Female: Reference Male: OR 1.5 (0.9-2.5)	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
			cerebral palsy was in doubt, a panel of trained paediatricians met to discuss the case.	.		
Andrews 2008	Prospective cohort study	N=375 Children born between 23 and 32 weeks' GA	African American ethnicity; -Adjusted for: gestational age and ethnicity. The study did not clearly report on how many multiple regression models were run for the results reported.	Cerebral palsy was defined as an abnormal muscle tone in at least one extremity and abnormal control of movement and posture	CP assessed at age 6 years among children born between 23 and < 32 weeks' GA: African American ethnicity: OR 0.1, 95% C.I. 0.01 – 0.6	High
Hirvonen 2014	Population based retrospective cohort using national registry data.	N= 53,078 Children born at between 32 and 34-36+6 Weeks GA	Male SGA -Adjusted for: period of study (1991-1995, 1996-2001 or 2002-2008), maternal age, maternal smoking status, primiparous, previous C-section, maternal diabetes, multiple pregnancy, order of fetuses, assisted reproductive technology, cervical cerclage, chorionic villus sampling, PROM, preeclampsia, time of birth, antenatal steroid use, place of birth,	The diagnosis of CP in Finland is based on medical history, ultrasound and MRI data, and multidisciplinary evaluations in the paediatric neurology units of 20 secondary level central hospitals and 5 tertiary level university hospitals	CP assessed at age 7: Within very preterm infants, <32 weeks gestation Sex Female: Reference Male: OR 1.34 (1.11-1.61) SGA Appropriate for gestational age*: Reference Small for gestational age: OR 0.75 (0.57-0.99) Within moderately preterm infants, 32+0	Low

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
			mode of delivery, gender, gestational weight, birth weight <1500g, Apgar score, umbilical artery pH, admission to neonatal unit, ventilator, resuscitation at birth, phototherapy, antibiotic therapy, RDS, sepsis, intracranial haemorrhage, convulsions and hyperbilirubinaemia.		to 33+6 weeks gestation: Sex Female: Reference Male: OR 1.11 (0.80-1.55) SGA Appropriate for gestational age*: Reference Small for gestational age: OR 1.10 (0.57-2.13) Within late preterm infants, 34+0 to 36+6 weeks gestation Sex Female: Reference Male: OR 0.98 (0.75-1.28) SGA Appropriate for gestational age*: Reference Small for gestational age: OR 1.85 (1.25-2.75)	
Guellec 2011	Prospective cohort study	N=2846 n=1822 children with follow-up at 5 years on CP and	Small for gestational age (SGA) (vs appropriate for gestational age AGA)	Cerebral palsy (CP), defined according to the European CP Network definition, children were classified as having CP if they had abnormal posture or	Outcome(s) at age Outcomes assessed at 5 years of age: Cerebral palsy (CP)	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
		cognitive outcome (disorders)	Adjusted for gestational age, gender, special class of the family, type of pregnancy (single vs multiple).	movement, increased tone or hyperreflexia (spastic CP), involuntary movements (dyskinetic CP), or loss of coordination (ataxic CP). Detailed medical and neurologic examintion in which tone, reflexes, postures and movements were assessed. Trained paediatricians reviewed data for children with abnormal results on neurologic examination to validate the diagnosis of CP and assess the severity.	1) Infants born at 24-28 weeks of gestation: AGA (>=20th centile):reference; SGA (<10th centile): 1.73 (0.54-5.60) 2) Infants born at 29-32 weeks of gestation: AGA (>=20th centile): reference; SGA (<10th centile): 0.39 (0.14-1.08)	
Intellectual disability						
Natarajan 2012	Prospective cohort study	N=963 Born at < 27 weeks gestation	Male gender SGA -Adjusted for: gestational age status, surgical NEC, severe IVH or cystic PVL, bloodstream infection, and antenatal steroids	Results of a structured neurologic examination by trained examiners and language and cognitive scores on Bayley Scales of Infant Development III at 18-22 months corrected age Cognitive composite score < 70 was defined as cognitive impairment	Assessed at 18 to 22 months corrected age among children born at < 27 weeks' GA: Cognitive impairment (composite score) (<70): OR (95%CI): Male: 1.39 (0.86-2.24) SGA: 2.60 (1.23-5.50)	Moderate
Amabalavanan 2012	Multicentre prospective cohort study	Sample recruited - n=14147	Male;	Intellectual disability was assessed by the Mental Developmental Index <70 on Bayley Scales of Infant Development-II,	At 18-22 months corrected age Intellectual disability (developmental delay: Mental Developmental Index [MDI <70]] Sex, Male gender - (OR [95% CIs])	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
					Referent group is not reported (assume is MDI>70) 1.62 (1.42–1.86)	
Guellec 2011	Prospective cohort study	N=2846 n=1822 children with follow-up at 5 years on CP and cognitive outcome (disorders)	Small for gestational age (SGA) (vs appropriate for gestational age AGA) Adjusted for gestational age, gender, special class of the family, type of pregnancy (single vs multiple).	Cognitive deficiency, defined by a Mental processing Composite (MPC) <85 (-1SD) assessed by the French version of the Kaufman Assessment Battery for Children, administered by trained psychologist.	At age 5 years: Cognitive deficiency SGA Infants born at 24-28 weeks GA: AGA (>=20th centile): reference SGA (<10th centile): 1.05 (0.34-3.19) Infants born at 29-32 weeks GA: AGA (>=20th centile): reference SGA (<10th centile): 1.73 (1.12-2.69)	
Helderman 2012	Multicentre prospective cohort study	Sample recruited: n=1506 Sample eligible for assessment: n=1200 Sample analysed after exclusions:n=921	Gender Ethnicity; Neonatal data were collected from the newborn's medical record.	The assessment of developmental delays (determined by cognitive impairment Mental Development Index [MDI]) at 24-months adjusted age at 24-months included the Bayley Scales of Infant Development-2nd Edition (BSID-II). Cognitive impairment was defined as an MDI <70. An MDI <55 was considered severe cognitive impairment.	At 24 months corrected age Intellectual disability (developmental delay MDI Male gender: (RR [95% CIs]) Referent group is infants with MDI <70 MDI MDI < 55: 2.5 (1.6, 4.1) MDI = 55–69: 2.0 (1.3, 3.2)	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
					Ethnicity (non- white race): (RR [95% CIs]) Referent group is infants with MDI <70 MDI MDI < 55: 2.3 (1.4, 3.8) MDI = 55–69: 2.1 (1.3, 3.5)	
Hoffman 2015	Retrospective cohort study	Sample recruited - n=3790	Ethnicity;	The primary study outcomes were BSID-III composite cognitive and language scores.	At 18-22 months corrected age (intellectual disability) Cognitive Composite <70 - (RR [95% CIs]) Referent group is not reported 0.79 (0.56–1.12)	Moderate
Vohr 2000	Multicentre cohort study	N=1151	Male; Ethnicity; SGA	Mental development index (MDI) <70, assessed by Bayley Scales of Infant Development-II (BSID-II)	At 18-22 months corrected age: MDI <70: Not significant (only reported graphically)	Low
Shankaran 2004	Multicentre prospective cohort study	n=246 preterm infants ≤24 weeks' gestation and ≤750g	Male; Ethnicity; Adjusted for: -risk factors were adjusted for each other, plus surfactant administration, steriods for BPD, Medicaid, No	The Bayley Scales of Infant Development (BSID-II), including the Mental Developmental Index (MDI) .	At 18-22 months' corrected age among those born ≤24 weeks' GA; Cognitive impairment (MDI < 70): OR (95%CI) Male: 2.1 (1.1-4.0)	Low

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
			high school degree, 2-parent household		Black: 1.9 (0.9- 3.8)	
Singh 2013	Cohort study	N=85,535	Male gender Adjusted for: age, sex, race/ethnicity, household composition, place of residence, and household education and income levels	Self-reported development problems; For the outcome of behavioural/emotional problems, it was measured as a composite, global mental health indicator which include depression, anxiety, or behavioural or conduct problems in the child. For disorders, parents were asked whether they were told by a doctor that their child had a disorder between age 2 to 17 years;	Intellectual disability/mental retardation, AOR (95%CI) at 2 to 17 years: Gender: Female: Reference Male: 1.70 (1.25-2.31) Race/ethnicity: Non-Hispanic white: reference Hispanic: 0.65 (0.36-1.19) Non-Hispanic black: 0.87 (0.60) Non-Hispanic mixed race: 1.00 (0.61-1.64) Other: 0.41 (0.23-0.76)	Low
Toome 2013	Prospective population based cohort.	N=187 Born at mean 28.8 (28.4-29.1) weeks gestation	Male gender SGA -Adjusted for: Gestational age SGA Maternal age Low income of the family Multiple births Antenatal steroids Postnatal steroids	The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below the mean for cognitive and language composite scores.	Assessed at corrected age 2 years among children born mean 28.8 weeks' GA. Risk of cognitive composite score <-2SD Male gender SGA not found to be significant predictors	HIgh

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
			BPD (defined as oxygen dependency at 36 weeks) ROP stage 3-5 with laser therapy Positive blood culture sepsis NEC stage 2-3			
Beaino 2011	Population based prospective cohort	N= 1503 Children born between 24-32wk's GA	Gender SGA -Adjusted for: neonatal cerebral lesions, gestational age of 28 weeks or less, gender, small for gestational age, Apgar score below 7 at one minute, NEC, BPD at 36 weeks, acute anaemia, late-onset anaemia and postnatal corticosteroid), social factors (parental socioeconomic status, number of siblings) and breast feeding.	The assessment used the Kaufman Assessment Battery for Children (K-ABC) test. Overall cognitive ability was evaluated by the Mental Processing Composite score, which was available for 1503 infants. Cognitive deficiency was classified as severe when the MPC score was below 70 (-2SD below the norm).	Severe cognitive deficiency assessed at age 5 years, among children born between 24 and 32 weeks' GA: Male: OR 1.08 (0.74-1.57) SGA: OR 2.49 (1.41-4.40)	Moderate
Hansen 2004	Prospective cohort	N=252 Children born at < 28 weeks' GA	Male -Adjusted for: IVH, NEC, CRIB-score (high), chronic lung disease, and mechanic ventilation during neonatal course	Intelligence test: Wechsler's Preschool and Primary Scale of Intelligence-Revised, WPPSI-R, was used as an intelligence test. Intellectual disability: An IQ score below -2 SD from the mean of a reference group	Assessed at age 5 years among children born < 28 weeks' GA: For the outcome of IQ score below 2 -SD of the mean: Sex/boy: 1.0 (0.5-2.0)	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
				classified children with intellectual disability.		
Speech and/or language disorder						
Hoffman 2015	Retrospective cohort study	N=3790 Sample recruited - n=3790 infants (456 born to adolescent mothers + 3364 born to adult mothers)	Ethnicity Regression models were used to compare relative risk (RR) of adverse outcomes at 18 to 22 months, controlling for infant and maternal characteristics that varied significantly between groups When control variables were highly related or overlapped, only 1 control variable was included to avoid overestimation problems due to multicollinearity		At 18-22 months corrected age Intellectual disability (Language Composite <70 and <85) Language Composite <70 - (RR [95% CIs]) Referent group is not reported1.10 (0.83–1.46)	Moderate
Toome 2013	Prospective population based cohort.	N=187 Born at mean 28.8 (28.4-29.1) weeks gestation	Male gender SGA -Adjusted for: Gestational age SGA Maternal age Low income of the family Multiple births Antenatal steroids Postnatal steroids	The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below the mean for cognitive and language composite scores.	Assessed at corrected age 2 years among children born mean 28.8 weeks' GA. Risk of language composite score <-2SD Male gender No: Reference Yes: OR 4.9 (1.1-21.8) SGA not found to be a significant predictor	High

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
			BPD (defined as oxygen dependency at 36 weeks) ROP stage 3-5 with laser therapy Positive blood culture sepsis NEC stage 2-3			
Autism spectrum disorder (ASD)						
Kuzniewicz 2014	Retrospective cohort study using population registry data	N= 235,198preterm children born at 24-34 weeks' GA	SGA -Adjusted for: gestational age, gender, maternal age, maternal education and SGA.	ASD: Children with a diagnosis of austism, Asperger syndrome or pervasive developmental disorder not otherwise specified were identified. The minimum age of children in the cohort was 3 years of age at the time the registry was assessed. ASD cases were defined as children with at least 1 diagnosis of ASD made at an ASD evaluation centre, or by a clinical specialist (psychiatrist, psychologist or developmental paediatrician) outside of the evaluation centre, or by a general paediatrician.	Diagnosis of ASD at age 2 to 11 years among children born at 24-34 weeks' GA: SGA: No: Reference Yes: HR 3.0 (1.4-6.3)	High
Hwang 2013	National prospective study	N= 1078 preterm children born at early preterm (GA<28 weeks), later preterm (GA 28-36 weeks), full term (≥37 weeks GA) weeks' GA	Male -Adjusted for: BPD, birth weight, and cerebral dysfunction	Infantile autism: children with autism were diagnosed and coded by their doctors based on ICD-9-CM definitions.	Infantile Autism assessed at age 8 to 11 years among children born preterm/extremely low birth weight (750g-1499g)weeks' GA: OR (95% CI)	Low

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
					Male: 4.1 (3.1-5.3)	
Moore 2012	Retrospective cohort	n=21717 children with autism, of which a proportion were children born preterm	SGA -Adjusted for: maternal age, race, hypertension, preeclampsia, diabetes, birth order, twin gestation, and months since last live birth.	Cases of autism were identified by: 1. An autistic level of one on any Client Development Evaluation Report or 2. An International Classification of Diseases 9th edition (ICD-9) code of 299.0 (autistic disorder), 299.8 or 299.9	Autism assessed at age 11 years: SGA 5-10 % (stratified by gestational age groups): Reference: AGA>10 to <90%=1.00 Among those 23-31 weeks GA: SGA: OR 1.36 95%CI 0.91-2.02 32-33 weeks GA: SGA: OR 1.00 95%CI 0.57-1.78 34-36 weeks GA: SGA: OR 1.12 95%CI 0.91-1.38	High
Singh 2013	Cross sectional survey	N=85, 535 Separated into premature children (born at <37 weeks) and term children (≥37 weeks)	Gender Ethnicity Household composition, place of residence and highest household/parental education	Parents were asked to self-report whether their child had been diagnosed with one of the disorders by a doctor or health care provider.	At age 2 to 17 years Autism spectrum disorder, AOR (95%CI): Gender: Female: Reference Male:4.49 (3.48-5.80) Race/ethnicity: Non-Hispanic white: reference Hispanic: 0.85 (0.53-1.36) Non-Hispanic black: 0.61 (0.41-0.92)	Low

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
					Non-Hispanic mixed race: 1.07 (0.75-1.55) Other: 0.60 (0.40-0.89) ADD/ADHD, AOR (95%CI): Gender: Female: Reference Male:2.43 (2.15-2.75) Race/ethnicity: Non-Hispanic white: reference Hispanic: 0.42 (0.33-0.54) Non-Hispanic black: 0.64 (0.53-0.77) Non-Hispanic mixed race: 0.91 (0.74-1.11) Other: 0.33 (0.25-0.43)	
Hearing impairment						
De Jesus 2013	Retrospective cohort study	N=2971 - Infants born between 23 0/7 and 26 6/7 weeks GA	SGA -adjusted for: Random effects variable, male, sex, multiple birth, GA, antenatal corticosteroid use, hypertension, and maternal education	Neurodevelopmental impairment was defined as presence of at least one of the following: 1. A composite score <70 on the cognitive component of the Bayley Scales of Infant and Toddler Development (BSID-III); 2. Moderate or severe cerebral palsy (CP) based on presence of bilateral hearing loss (with or without amplification) or bilateral blindness (vision <20/200).	Assessment at 18-22 months corrected age among children born between 23 and 26 weeks' GA: For the outcome of hearing loss with or without amplification: SGA: OR 1.38, 95%CI 0.44-4.36 (P=0.58)	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
Vision impairment						
De Jesus 2013	Retrospective cohort study	N=2971 - Infants born between 23 0/7 and 26 6/7 weeks GA	SGA -adjusted for: Random effects variable, male, sex, multiple birth, GA, antenatal corticosteroid use, hypertension, and maternal education	Neurodevelopmental impairment was defined as presence of at least one of the following: 1. A composite score <70 on the cognitive component of the Bayley Scales of Infant and Toddler Development (BSID-III); 2. Moderate or severe cerebral palsy (CP) based on presence of bilateral hearing loss (with or without amplification) or bilateral blindness (vision <20/200).	Assessment at 18-22 months corrected age among children born between 23 and 26 weeks' GA: For the outcome of blindness (<20/200 vision bilaterally): SGA: OR 10.9, 95%CI 2.15-55.5	Moderate
Developmental coordination disorder (DCD)						
Davis 2007	Prospective cohort study	N=298 consecutive preterms N=262 randomly selected infants	Male;	Fine and gross motor abilities were assessed using the Movement Assessment Battery for Children (MABC), age band 2 for 7 to 8 year olds Cut off of the 5th centile was used to denote children with DCD Full scale IQ was sued as a measure of general cognitive ability Parents and teachers completed the Behaviour Assessment System for Children	Outcome at age: Developmental Coordination Disorder at 8 and 9 years age After adjustment for all other perinatal variables, only male sex increased the risk of a child having developmental coordination disorder, with P value 0.017	Low
Composite outcomes						
Shankaran 2004	Multicentre prospective cohort study	n=246 preterm infants ≤24 weeks'	Male; Ethnicity; Adjusted for:	Neurodevelopmental impairment (NDI) was defined as CP, MDI or PDI < 70,	At 18-22 months' corrected age among	High

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
		gestation and ≤750g	-risk factors were adjusted for each other, plus surfactant administration, steroids for BPD, Medicaid, No high school degree, 2-parent household.	bilateral blindness, or hearing impaired with amplification.	those born ≤24 weeks' GA; NDI: OR (95%CI) Male: 1.4 (0.7-2.6) Black: 1.1 (0.6-2.2)	
Walsh 2005	Prospective cohort study	n=3041 children born at 25.8 ±2.23 weeks postmenstrual age.	Male SGA Ethnicity Adjusted for: male, SGA, ethnicity, PLV, Grade III-IV IVH, Postnatal steroids, Antental steriods	The Bayley Scales of Infant Development - II, including the mental scale, psychomotor scale, and the behavior rating scale, were administered by developmental specailists trained. BSID-II scores of 100 ± 15 represent the mean ± 1 standard deviation The neurologic examination is based on the Amiel-Tison neurologic assessment. Infants were scored as normal if no abnormalities were observed in the examination.	Outcomes assessed at 18-22 months Postmenstrual age, among children born at 25.8 ±2.23 weeks postmenstrual age. NDI: Male gender: 1.62 (1.32-1.93) SGA was not found to be a significant predictor	Moderate
Bolisetty 2014	Retrospective multicentre cohort study	N= 1472 Born between 23 and 28+6 weeks'	Male gender; SGA (<10th percentile and <3rd percentile)	Moderate neurosensory impairment was defined as the presence of developmental delay (Griffiths Mental Developmental Scale General Quotient or Bayley Scales of Infant Development MDI between 2 and 3 SD below the mean), moderate cerebral palsy (able to walk with the assistance of aids) or deafness (requiring amplification with bilateral hearing aids or	At 2-3 years' corrected age among children born between 23 and 28 weeks' GA: Moderate to severe neurosensory impairment Male gender No: Reference Yes: OR 1.81 (1.32-2.47)	Moderate

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
				unilateral/bilateral cochlear implant). Severe neurosensory impairment was defined as developmental delay (GMDS-GQ or MDI less than 3 SD below the mean), severe cerebral palsy (unable to walk with the assistance of aids) or bilateral blindness (visual acuity <6/60 in the better eye).	SGA <10th percentile No: Reference Yes: OR 1.94 (1.09-3.46)	
Toome 2013	Prospective population based cohort.	N=187 Born at mean 28.8 (28.4-29.1) weeks gestation	Male gender SGA -Adjusted for: Gestational age SGA Maternal age Low income of the family Multiple births Antenatal steroids Postnatal steroids BPD (defined as oxygen dependency at 36 weeks) ROP stage 3-5 with laser therapy Positive blood culture sepsis NEC stage 2-3	A composite outcome measure of neurodevelopmental impairment was also used. This includes any one (or more) of the following criteria: CP with GMFCS level 2,3,4 or 5; cognitive and/or language composite scores of ≤-2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.	Assessed at corrected age 2 years among children born mean 28.8 weeks' GA. Risk of neurodevelopmental impairment Male gender SGA Both not found to be significant predictors Risk of cognitive composite score <-2SD Male gender SGA not found to be significant predictors	High
Kent 2012	Population based longitudinal cohort study	N=2701 N=1473 followed up at 2-3 years	Male gender SGA	Moderate to severe functional disability was defined as one or more of the following: developmental delay (<2SD	Moderate to severe disability among male and female infants at	High

Study	Data Source	Sample and Population studied	Risk factors and adjustment	Measures of Outcomes	Prognostic outcomes	Study Quality
				below the mean for adjusted age determined by the Griffiths Mental Developmental Scales or BSID-II, cerebral palsy (unable to walk without aids), bilateral blindness (visual acuity <6/60 in better eye), or bilateral deafness (requiring bilateral hearing aids or cochlear implants)	2 to 3 years corrected age Gender: Female: reference Male: OR 1.877 (1.398-2.521) SGA: AGA: reference SGA: OR 2.077 (1.376-3.136)	
Leversen 2010	Prospective population based cohort study	n=376 preterm babies discharged home alive	Gender Small for gestational age Adjusted for gestational age, gender, multiple pregnancy, chorioamnionitis, preeclampsia, antenatal steroids, PROM, Caesarean section, SGA, illness severity score (a score of the lowest and highest FiO2 requirements and the largest base deficit during the first 12 hours of life), septicaemia, BPD, patent ductus arteriosus, NEC, postnatal steroids, cranial ultrasound findings and retinopathy of prematurity	The outcome reported was a composite finding of "major neurosensory disabilities". This includes cerebral palsy, blindness (classified as legally blind) or complete deafness.	Major neurosensory disability at 2 years Gender Female: Reference Male: OR 1.3 (0.5-3.8) Small for gestational age No: Reference Yes: OR 3.0 (0.5-19.9)	Moderate

1

1 **Table 17: Summary of studies on the association between different neonatal factors and developmental disorders**

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Cerebral palsy						
Hintz 2005 (USA)	Retrospective cohort study	N= 2948 extremely low birth weight infants, mean GA not reported;	NEC -adjusted for: network centre, use of antenatal glucocorticoids, rupture of membranes >24h, outborn status, estimated gestational age, gender, race, birth weight, small for gestational age, surfactant therapy, intraventricular haemmorrhage grade 3 or 4 or cystic periventricular leukomalacia, sepsis, postnatal steroid treatment, bronchopulmonary dysplasia, and highest level of education attained by the primary caregiver	Cerebral palsy (CP) was defined as a non-progressive central nervous system disorder characterized by abnormal muscle tone in at least 1 extremity and abnormal control of movement and posture	CP assessed at 18-22 months corrected age among children born extremely low birth weight: NEC surgical: OR 1.31 (0.8-2.14) NEC medical: OR 0.68 (0.38-1.29)	Moderate
Vincer 2006 (Canada)	Prospective cohort study	N= 672 children born at < 31wks GA	Antenatal corticosteroids	CP was defined as a disorder of control of movement or posture	CP assessed at age 24 months: Antenatal corticosteroids: OR 0.53 (0.27 – 1)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			Postnatal dexamethasone use IVH grade III and IV -adjusted for: gestational age <28 weeks vs >28 weeks to 30 weeks; postnatal dexamethasone use; patent ductus artriosus; severe hyaline membrane disease; resuscitation in the delivery room; IVH grades 3 and 4; antenatal corticosteroid use. Other variables that were considered and tested for in the stepwise backward manner were: Maternal age at delivery; maternal substance use; pregnancy-induced hypertension; chlorioamnionitis; funisitis;	secondary to a non-progressive brain lesion.	Postnatal dexamethasone use: OR 2.245 (1.24 -4.06) IVH grade III and IV : OR 7.78 (3.43 -18.34)	

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			oligohydramnios; polyhydramnios; multiple birth; major anomaly; hydrops fetalis; SGA; maternal analgesic use; maternal anaesthetics; premature rupture of membranes; birth depression, 5-min Apgar score; cardiopulmonary resuscitation; indomethacin use; hypernatremia, hyponetremia; unconjugated bilirubin; hypoglycemia; gender of the infant.			
Payne 2013 (USA)	Prospective cohort study	N= 1472 children born at < 27 weeks' GA	Low grade PIVH Severe PIVH Antenatal steroids Sepsis Postnatal steroids -adjusted for: PIVH severity (3 levels), gestational age, sex,	Any cerebral palsy (CP), defined as abnormal tone or reflexes in at least one extremity and abnormal control of movement or posture to a degree that interferes with age-appropriate activity assessed with the Amiel-Tison neurologic assessment	CP assessed at 18-22 months corrected age: Low grade PIVH versus no PIVH: OR 1 (0.61-1.64) Severe PIVH no PIVH: OR 3.43 (2.24-5.27) Severe PIVH versus low grade PIVH: OR 3.44 (1.96-5.98) Antenatal steroids: OR 0.69 (0.42-1.14)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			race/ethnicity, maternal education, chorioamnionitis, sepsis, antenatal steroid exposure, postnatal steroid exposure, high frequency ventilation and patent ductus arteriosus	and Palisano's Gross Motor Function Classification System (GMFCS).	Sepsis: OR 1.48 (1.03-2.11) Postnatal steroids: OR 1.44 (0.92-2.26)	
Vohr 2005 (USA)	Prospective cohort study	N= 3785 children born at 22 to 32 weeks' GA	PVL; IVH grade III-IV: BPD: Sepsis: Antenatal steroids: -adjusted for: gestational age group; birth weight; gender; small for gestational age; multiple births; surfactant; grades 3 to 4 IVH; PVL; sepsis; oxygen requirement at 36 weeks; white vs. non-white race; outborn vs. inborn status caesarean section vs. vaginal delivery;	CP, defined as a non-progressive central nervous system disorder characterized by abnormal muscle tone in at least 1 extremity and abnormal control of movement or posture	CP (moderate to severe) assessed at age 18 to 22 months corrected age: PVL: OR 10.5 (7.2 – 15.2) IVH grade III-IV: Significantly increased risk but risk estimate not reported; Postnatal steroids: OR 2.02 (1.4-2.92) BPD: Significantly increased risk but risk estimate not reported; Sepsis: Insignificant association but risk estimate not reported Antenatal steroids: 0.66 (0.47-0.92)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			maternal education <12 years vs. >=12 years; private health insurance vs. public; conventional ventiolation vs. none; adjusted age at the time of assessment; centre; and the 4 interventions of interest: antenatal steroids (yes, no), high-frequency ventilation vs. none; days to regain birth weight, and postnatal steroids (yes, no).			
Adams-Chapman 2008 (USA)	Prospective cohort study	N= 6161 children born at between < 25wks and ≥ 33 weeks GA	IVH III/shunt IVH IV/shunt -adjusted for: study center, gestational age, birth weight, gender, race, caesarean section delivery, multiple birth, antenatal steroid exposure, postnatal steroid exposure,	CP	CP assessed at 18 to 22 months corrected age: IVH III/shunt versus IVH III/no shunt: OR 2.08 (1.63-2.66) IVH III/shunt versus no IVH/no shunt: OR 3.44 (2.76-4.29) IVH IV/shunt versus IVH IV/no shunt: OR 1.83 (1.47-2.28) IVH IV/shunt versus no IVH no shunt: OR 3.96 (3.19 – 4.92)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			surfactant use, respiratory distress syndrome, bronchopulmonary dysplacia (BPD), patent ductus arteriosus, periventricular leukomalacia (PVL), infection group, caregivers' education.			
Carlo 2011 (USA)	Prospective cohort study	N= 4924 children born at 22 to 25 weeks GA	Antenatal steroids: -adjusted for: Gender and race	CP: exact definition not reported	Moderate to severe CP assessed at age 18-22 months corrected age: Among children born at < 22-25wks GA: Antenatal steroids: OR 0.76 (0.59-0.98) Among children born at 22 weeks GA: Antenatal steroids: OR 0.88 (0.23-3.34) Among children born at 23 weeks GA: Antenatal steroids: OR 0.5 (0.3-0.85) Among children born at 24 weeks GA: Antenatal steroids: OR 0.71 (0.47-1.08) Among children born at 25 weeks GA:	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
					Antenatal steroids: OR 0.97 (0.62-1.5)	
Stoll 2004 (USA)	Prospective cohort study	N= 6314 pre-term children	Sepsis -adjusted for: study centre, gestational age, birth weight, sex, race/ethnicity, rupture of membranes >24 h, CS, multiple birth, antenatal antibiotics, antenatal steroids, postnatal steroids, surfactant use, respiratory distress syndrome, bronchopulmonary dysplasia, patent ductus arteriosus, intraventricular haemorrhage grade 3-4, periventricular leukomalacia, maternal age at time of delivery, caregiver's level of education	CP: defined as non-progressive disorder of movement and posture	CP assessed at age 18-22 months corrected age: Sepsis alone: OR 1.4 (1.1-1.8) Sepsis plus NEC: OR 1.7 (1.2-2.5) Meningitis with or without sepsis: OR 1.6 (1-2.5)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Vohr 2000 (USA)	Prospective study	N= 1151 preterm children born at 22-32 weeks GA	IVH/PVL grade III-IV; NEC	CP: non progressive central nervous system disorder characterized by abnormal muscle tone in at least 1 extremity and abnormal control of movement or posture. Moderate to severe CP included children who were non ambulatory or required an assistance device for ambulation	CP assessed at age 18-22 months corrected age: IVH/PVL grade III-IV: 3.05 (2.03-4.57) NEC: OR 2.01 (1.05-3.73)	Low
Shankaran 2004 (USA)	Prospective study	N= 246 children born at less or equal to 24 weeks GA	ICH grade III-IV; PVL; Any antenatal steroids BPD - Adjusted for: risk factors were adjusted for each other, plus surfactant administration, steroids for BPD, Medicaid, No high school degree, 2-parent household;	CP: Cerebral palsy was defined as a non-progressive central nervous system disorder characterized by abnormal muscle tone in at least 1 extremity and abnormal control of movement and posture.	CP assessed at age 18-22 months corrected age: ICH grade III-IVH: OR 1.9 (0.9-4.1) PVL; OR 4.4 (1.4-13.5) Any antenatal steroids: 1.1 (0.6-2.3) BPD: nonsignificant association was found	Low
Tommiska 2003 (Finland)	Prospective cohort study	N=208 children born at 27.3 months (mean) GA	Antenatal steroids Sepsis NEC Brain abnormalities	CP: defined as progressive motor impairment with spastic or dystonic muscle tone, brisk tendon reflexes, positive Babinski's sign	CP assessed at age 18 months: Antenatal steroids: OR 3.6 (1.3-10) Sepsis: nonsignificant association was found	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			-adjusted for: antenatal steroids, vaginal delivery, sepsis, NEC, brain abnormalities	and persistent primitive reflexes.	NEC with perforation: nonsignificant association was found IVH grade II-IV: nonsignificant association was found	
Van Marter 2011 (USA)	Prospective cohort study	N= 1047 children born at < 28wks' GA	BPD -adjusted for: It was not clearly reported	Cerebral palsy (CP), assessed through a neurological examination and an assessment for the Gross Motor Function Classification System (GMFCS) to assess the severity of the motor disability related to CP. CP classifications: quadriparesis diparesis hemiparesis	CP assessed at age 24 months corrected age: CP quadriparesis: BPD, only O2: OR 1.6 (0.8-3.2) BPD, with mechanical ventilation: OR 5.7 (2.5-13) CP diparesis: BPD, only O2: 2.1 (0.8-5) BPD, with mechanical ventilation: OR 4.2 (1.3-14) CP hemiparesis: BPD, only O2: OR 2.7 (0.7-11) BPD, with mechanical ventilation: OR 1.2 (0.1-13)	Moderate
Allred 2014 (USA)	Prospective cohort study	n=1085 Children born at < 28wks' GA	ROP -adjusted for: gestational age, birth weight z-score categories, hyperoxemia (a PaO2 in the highest quartile on 2 of the first 3 postnatal days), Score of Neonatal Acute Physiology-II (SNAP-II) in the	CP: topographic diagnosis of CP was based on an algorithm using the data of quadriparesis, diparesis, hemiparesis	CP assessed at age 24 months: CP quadriparesis : ROP stage 3+: OR 1.2 (0.7 -2) ROP plus disease: OR 1.2 (0.6 - 2.6) ROP zone 1: OR 0.9 (0.4 - 2.3) ROP threshold: OR 1.3 (0.3 -4.8) ROP pre-threshold: OR 0.9 (0.5 - 1.9) CP diparesis: ROP stage 3+: OR 1.2 (0.5 -2.7)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			highest quartile, culture-proven bacteremia in the first 28 days, mechanical or high frequency on 14 or more days, and growth velocity in the lowest quartile		ROP plus disease: OR 2.4 (0.99 - 5.9) ROP zone 1: OR 2.1 (0.8 -6) ROP threshold: OR 1.5 (0.3 -7.6) ROP pre-threshold: OR 2.2 (0.9 - 5.2) CP hemiparesis: ROP stage 3+: OR 1.1 (0.4 -3.1) ROP plus disease: OR 1.3 (0.3 - 4.9) ROP zone 1: OR 1 (0.2 -5.1) ROP threshold: NR NR NR ROP pre-threshold: OR 0.9 (0.2 - 3.3)	
Toome 2013 (Estonia)	Prospective cohort study	N= 187 children born at 22-31 weeks GA	Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV -adjusted for: antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4,	CP: was defined according to the guidelines of the Surveillance of Cerebral Palsy in Europe collaborative group	CP assessed at age 2 years: Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV: OR 43.2 (8.2-226.5)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family			
Wood 2005 (USA)	Prospective study	N= 283 children born between 20-25 weeks GA	Antenatal steroids ROP Postnatal steroids -Adjusted for: Risk factors were adjusted for each other although this was not clearly reported	Cerebral palsy was classified retrospectively, being defined as a non-progressive disorder of movement and posture.	CP assessed at age 30 months corrected age: Significantly abnormal ultrasound scan (defined as parenchymal pathology and/or ventriculomegaly): OR 4.95 (2.25 -10.85) Antenatal steroids: nonsignificant association Treatment for ROP: nonsignificant association Postnatal steroids for 1-14 days (vs none): OR 0.92 (0.3-2.82) Postnatal steroids for 15-28 days (vs none): OR1.06 (0.4 -2.84) Postnatal steroids for 29-42 days (vs none): OR 1.09 (0.35-3.4)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
					Postnatal steroids for 43-56 days (vs none): OR 0.68 (0.13 -3.4) Postnatal steroids for >=57 days (vs none): OR 4.77 (1.29 -17.56)	
Mikkola 2005 (Finland)	Prospective cohort study	N= 193 Children born at 27.3 (± 2.1) weeks' GA	Antenatal steroids -adjusted for: maternal smoking, high social class, preeclampsia, absence of antenatal steroids, multiple birth, gestational age, birth weight, gender, SGA, vaginal delivery, Apgar score <4 at 5 min, university hospital area, birth outside a tertiary hospital, IVH grade 3-4, perforated NEC, O2 dependency at 36 weeks, ROP grades 3-4	Cerebral palsy (CP), defined as a non-progressive motor disorder with abnormal muscle tone, persistent or exaggerated primitive reflexes, or a positive Babinski sign associated with delayed motor development.	CP assessed at age 5 years: Antenatal steroids: OR 3.4 (1.3-9)	Moderate
Victorian Infant Collaborative Study Group 2000 (Australia)	Prospective cohort study	N= 280 children born at < 28wks' GA	Postnatal steroids -adjusted for: ruptured membranes >24h, cystic PVL, and surgery during the	CP was assessed by a paediatrician	CP assessed at age 5 years: Postnatal steroids: OR 7.8 (2.9-21)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			primary hospitalization			
Foix-L'Helias 2008 (France)	Prospective cohort study	N= 2855 children born at 24- 32 weeks' GA	Antenatal steroids: -adjusted for: gestational age, social class, sex and pregnancy complications. A propensity score adjusted for general characteristics (maternal age, parity, tobacco consumption, region and level of neonatal intensive care), maternal complications and pregnancy etc.	CP: the definition of cerebral palsy was that established by the European Cerebral Palsy Network,	CP assessed at age 5 years: Among children born at 24-32wks GA: Antenatal steroids (any): OR 0.99 (0.65-1.52) Among children born at 24-27wks GA: Antenatal steroids (any): OR 1.69 (0.67-4.62) Among children born at 28-32wks GA: Antenatal steroids (any): OR 0.86 (0.54-1.38) Among children born at 24-32wks GA: Antenatal steroids (complete course): OR 0.83 (0.52-1.31) Among children born at 24-27wks GA: Antenatal steroids (complete course):: OR 1.22 (0.46-3.26) Among children born at 28-32wks GA: Antenatal steroids (complete course):: OR 0.71 (0.42-1.19)	Moderate
Andrews 2008 (US)	Prospective study	N= 375 children born at 23-31 weeks GA	IVH grade III-IV; NEC -adjusted for: gestational age and ethnicity	CP Cerebral palsy was defined as an abnormal muscle tone in at least one extremity and abnormal control of movement and posture.	CP assessed at age 6 years: IVH grade III-IV: OR 25.6 (3.8-172.2) NEC: OR 5.7 (0.9-34.1)	HIgh

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Hansen 2004 (Denmark)	Prospective study	N= 252 children born at 24.1-34.3 weeks GA	IVH grade III-IV; NCE -adjusted for: risk factors were adjusted for each other in the multivariate analysis, as well as CRIB-score (high), chronic lung disease, and mechanic ventilation	CP: Cerebral palsy was diagnosed in accordance with the criteria as defined in the Surveillance of cerebral palsy in Europe Visual disability:	CP assessed at age 5 years: IVH grade III-IV: OR 19.9 (6.1-64.8) NEC: OR 19.1 (3.3-111.3)	Moderate
Beaino 2010 (France)	Prospective study	N= 1812 children born at 22-32wks GA	IVH grade I IVH grade II IVH grade III or echodensities or ventricular dilatation Cystic PVL or intraparenchymal haemorrhage NEC BPD Postnatal steroids -adjusted for: “obstetric and neonatal factors" but it is not stated which factors these were.	CP: the definition of CP was that proposed by the Surveillance of Cerebral Palsy in Europe	CP assessed at age 5 years: IVH grade I: OR 1.76 (0.9 -3.45) IVH grade II: OR 2.56 (1.27 - 5.18) IVH grade III or echodensities or ventricular dilatation: OR 3.4 (2.07 -5.6) Cystic PVL or intraparenchymal haemorrhage: OR 28.41 (15.65 - 51.59) NEC: OR 1.51 (0.64 -3.55) BPD: 0.95 (0.53 -1.71) Postnatal steroids: OR 1.41 (0.82 -2.43)	Moderate
Hirvonen 2014 (Finland)	Prospective study	N- 6347 children born between < 32	Antenatal steroids Sepsis -adjusted for:	The definition of CP was that proposed by the Surveillance of Cerebral	CP assessed at age 7 years:	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
		and 36 weeks GA	maternal age, maternal smoking status, primiparous, previous C-section, maternal diabetes, multiple pregnancy, order of foetuses, assisted reproductive technology, cervical cerclage, chorionic villus sampling, PROM, preeclampsia, time of birth, antenatal steroid use, place of birth, mode of delivery, gender, gestational weight	Palsy in Europe (SCPE) collaborative group	Antenatal steroids among children born at < 32 weeks GA; OR: 0.8 (0.49-1.3) Sepsis among children born at < 32 weeks GA: OR 0.94 (0.62-1.43) Intracranial haemorrhage among children born at < 32 weeks GA: 3.05 (2.08-4.47) Antenatal steroids among children born at 32-33 weeks GA: OR 0.27 (0.09-0.8) Sepsis among children born at 32-33 weeks GA: OR 1.35 (0.6-3.05) Intracranial haemorrhage among children born at 32-33 weeks GA: OR 7.18 (3.6-14.3) Antenatal steroids among children born at 34-36 weeks GA: OR 1.01 (0.35-2.91) Sepsis among children born at 34-36 weeks GA: OR 1.5 (0.73-3.1) Intracranial haemorrhage among children born at 34-36 weeks GA: OR 12.8 (5.58-29.2)	
Intellectual disability						
Hintz 2005 (USA)	Retrospective cohort study	N= 2948 extremely low birth weight infants, mean GA not reported;	NEC -adjusted for: network centre, use of antenatal glucocorticoids, rupture of	Intellectual disability: defined as MDI < 70 assessed through the Bayley Scales of Infant Development-II (BSID-II)	MDI < 70 assessed at 18-22 months corrected age among children born extremely low birth weight: NEC surgical: OR 1.61 (1.05-2.5)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			membranes >24h, outborn status, estimated gestational age, gender, race, birth weight, small for gestational age, surfactant therapy, intraventricular haemmorrhage grade 3 or 4 or cystic periventricular leukomalacia, sepsis, postnatal steroid treatment, bronchopulmonary dysplasia, and highest level of education attained by the primary caregiver		NEC medical: OR 1.16 (0.74-1.81)	
O' Shea 2008 (USA)	Prospective cohort study	n=1017 children born at < 28 weeks GA	IVH Early PVL Cystic PVL Periventricular hemorrhagic infarction IVH Early PVL Cystic PVL Periventricular hemorrhagic infarction	MDI < 70 assessed through the Bayley Scales of Infant Development-II (BSID-II)	MDI < 70 assessed at age 24 months corrected age: IVH: OR 1.7 (1.2 -2.5) Early PVL: OR 1.3 (0.8 -2.1) Cystic PVL: OR 1.9 (0.98 -3.5) Periventricular hemorrhagic infarction: OR 2.2 (1.2 – 4)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			-adjusted for: risk factors were adjusted for each other in the multivariate analysis			
Payne 2013 (USA)	Prospective cohort study	N= 1472 children born at < 27 weeks' GA	Low grade PIVH Severe PIVH Antenatal steroids Sepsis Postnatal steroids -adjusted for: PIVH severity (3 levels), gestational age, sex, race/ethnicity, maternal education, chorioamnionitis, sepsis, antenatal steroid exposure, postnatal steroid exposure, high frequency ventilation and patent ductus arteriosus	Cognitive impairment defined as a score of <70 on the Bayley Scales of Infant Development 3rd edition (Bayley III).	Cognitive impairment assessed at 18-22 months corrected age: Low grade PIVH versus no PIVH: OR 0.94 (0.54-1.61) Severe PIVH no PIVH: OR 1.37 (0.79-2.37) Severe PIVH versus low grade PIVH: OR 1.46 (0.74-2.88) Antenatal steroids: OR 0.64 (0.39-1.13) Sepsis: OR 2.28 (1.49--3.48) Postnatal steroids: OR 2.28 (1.41-3.69)	Moderate
Shah 2012 (USA)	Prospective cohort study	N= 865 children born at 25.7-26.2 GA	NEC	Impaired mental development defined as a MDI score <70 assessed through Bayley III.	MDI assessed at age 18 to 22 months corrected age: NEC >=IIA: OR 2.04 (0.96 -4.34) NEC >=IIA: OR 2.64 (1.18 -5.91)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
					NEC >=IIA surgically managed: NS	
Vohr 2005 (USA)	Prospective cohort study	N= 3785 children born at 22 to 32 weeks' GA	PVL; IVH grade III-IV: Postnatal steroids: BPD: Sepsis: Antenatal steroids: -adjusted for: gestational age group; birth weight; gender; small for gestational age; multiple births; surfactant; grades 3 to 4 IVH; PVL; sepsis; oxygen requirement at 36 weeks; white vs. non-white race; outborn vs. inborn status ceasarean section vs. vaginal delivery; maternal education <12 years vs. >=12 years; private health insurance vs. public; conventional ventiolation vs.	MDI score < 70 assessed through Bayley II	MDI <70 (moderate to severe) assessed at age 18 to 22 months corrected age: PVL: only reported significant association was found IVH grade III-IV: only reported significant association was found Postnatal steroids: OR 1.29 (1.04-1.61) BPD: only reported significant association was found Sepsis: NS Antenatal steroids: NS	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			none; adjusted age at the time of assessment; centre; and the 4 interventions of interest: antenatal steroids (yes, no), high-frequency ventilation vs. none; days to regain birth weight, and postnatal steroids (yes, no).			
Adams-Chapman 2008 (USA)	Prospective cohort study	N= 6161 children born at between < 25wks and ≥ 33 weeks GA	IVH III/shunt IVH IV/shunt -adjusted for: study center, gestational age, birth weight, gender, race, caesarean section delivery, multiple birth, antenatal steroid exposure, postnatal steroid exposure, surfactant use, respiratory distress syndrome, bronchopulmonary dysplacia (BPD), patent ductus arteriosus,	Cognitive impairment assessed through Bayley IIR: MDI < 70	MDI assessed at 18 to 22 months corrected age: IVH III/shunt versus IVH III/no shunt: OR 1.19 (0.97-1.44) IVH III/shunt versus no IVH/no shunt: OR 1.41 (1.18-1.68) IVH IV/shunt versus IVH IV/no shunt: OR 1.48 (1.24-1.78) IVH IV/shunt versus no IVH no shunt: OR 1.72 (1.47-2.02)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			periventricular leukomalacia (PVL), infection group, caregivers' education.			
Allred 2014 (USA)	Prospective cohort study	n=1085 Children born at < 28wks' GA	ROP -adjusted for: gestational age, birth weight z-score categories, hyperoxemia (a PaO2 in the highest quartile on 2 of the first 3 postnatal days), Score of Neonatal Acute Physiology-II (SNAP-II) in the highest quartile, culture-proven bacteremia in the first 28 days, mechanical or high frequency on 14 or more days, and growth velocity in the lowest quertile	Cognitive impairment assessed through Bayley II, MDI < 55, or 56-69	MDI <55 assessed at age 24 months: ROP stage 3+: OR 1.9 (1.2-2.9) ROP plus disease: OR 1.9 (1.1-3.2) ROP zone 1: OR 1.5 (0.8-2.9) ROP threshold: OR 2.2 (0.8-6.2) ROP pretreshold: OR 1.7 (1-2.7) MDI 56-69 ROP stage 3+: OR 11.3 (0.8-2.1) ROP plus disease: OR 2.1 (1.1-4) ROP zone 1: OR 2.4 (1.2-4.7) ROP threshold: OR 3.6 (1.3-10) ROP pretreshold: OR 2.1 (1.2-3.8)	Moderate
Carlo 2011 (USA)	Prospective cohort study	N= 4924 children born at 22 to 25 weeks GA	Antenatal steroids: -adjusted for: Gender and race	Cognitive impairment: MDI < 70 by Bayley III; and Bayley III cognitive composite score <70	MDI < 70 assessed at age 18-22 months corrected age: Among children born at < 22-25wks GA: Antenatal steroids: OR 0.93 (0.78 -1.12)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
					Among children born at 22 weeks GA: Antenatal steroids: OR 2.16 (0.36 -13.1) Among children born at 23 weeks GA: Antenatal steroids: OR 1.27 (0.79-2.03) Among children born at 24 weeks GA: Antenatal steroids: OR 0.85 (0.62-1.16) Among children born at 25 weeks GA: Antenatal steroids: OR 0.91 (0.69-1.2) Baley III cognitive impairment < 70 assessed at age 18-22 months corrected age: Among children born at < 22-25wks GA: Antenatal steroids: OR 0.63 (0.34 -1.17) Among children born at 22 weeks GA: Antenatal steroids: OR 1.28 (0.06-27.5) Among children born at 23 weeks GA: Antenatal steroids: OR 0.31 (0.09 -0.998) Among children born at 24 weeks GA:	

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
					Antenatal steroids: OR 0.57 (0.17 -1.91) Among children born at 25 weeks GA: Antenatal steroids: OR 0.88 (0.34 -2.24)	
Stoll 2004 (USA)	Prospective cohort study	N= 6314 pre-term children	Sepsis -adjusted for: study center, gestational age, birth weight, sex, race/ethnicity, rupture of membranes >24 h, CS, multiple birth, antenatal antibiotics, antenatal steroids, postnatal steroids, surfactant use, respiratory distress syndrome, bronchopulmonary dysplasia, patent ductus arteriosus, intraventricular haemorrhage grade 3-4, periventricular leukomalacia, maternal age at time of delivery,	Mental developmental index (MDI) <70, assessed with Bayley Scales of Infant Development II (BSID-II)	MDI<70 assessed at age 18-22 months corrected age: Sepsis alone: OR 1.3 (1.1-1.6) Sepsis plus NEC: OR 1.6 (1.2-2.2) Meningitis with or without sepsis: OR 1.6 (1.1-2.3)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			caregiver's level of education			
Toome 2013 (Estonia)	Prospective cohort study	N= 187 children born at 22-31 weeks GA	Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV -adjusted for: antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education	Cognitive composite score assessed through the Bayley Scales of Infant and Toddler Development (-2SD below the mean)	Cognitive composite score < -2SD assessed at age 2 years: Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV: OR 9.8 (1.9-49.5) NEC grade II-III: OR 7.4 (1.5-37.2)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			and low income of the family			
Natarajan 2012 (USA)	Prospective study	N= 963 children born at 25.2-26.2 weeks GA	NEC Brain abnormalities BPD Antenatal steroids Sepsis -adjusted for: small for gestational age status, surgical NEC, severe IVH or cystic PVL, bloodstream infection, and antenatal steroids	Cognitive impairment: measured by Bayley Scales of Infant Development III, cognitive score < 70 was defined as cognitive impairment	Cognitive impairment assessed at 18 to 22 months corrected age: Surgical NEC: OR 3.35 (1.42 - 7.91) IVH or PVL: OR 3.97 (2.4 -6.55) BPD: OR 2.41 (1.4- 4.13) Antenatal steroids: NS Blood stream infection: NS	Moderate
Shankaran 2004 (USA)	Prospective study	N= 246 children born at less or equal to 24 weeks GA	ICH grade III-IV; PVL; Any antenatal steroids BPD - Adjusted for: risk factors were adjusted for each other, plus surfactant administration, steroids for BPD, Medicaid, No high school degree, 2-parent household;	MDI < 70 assessed through BSID II	MDI assessed at age 18-22 months corrected age: ICH grade III-IV: OR 1.8 (0.9-3.6) PVL: OR 3.4 (1 - 10.8- Any antenatal steroids: OR 0.9 (0.5 -1.7) BPD: NS	Low

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Kallen 2015 (Sweden)	Prospective study	N=456 children born at less than 27 weeks GA	Antenatal steroids -adjusted for gestational age and for birth weight standard deviation score	Intellectual disability: Mental developmental delay was defined as a cognitive or language Bayley III scale <2SD below the mean,	Mental developmental delay assessed at 2.5 yrs corrected age: Antenatal steroids: OR 0.7 (0.3-1.9)	Moderate
Vohr 2000 (USA)	Prospective cohort study	N= 1185 children born at 22 to 32 weeks' GA	PVL; IVH grade III-IV: BPD: Sepsis: Antenatal steroids: -adjusted for: gestational age group; birth weight; gender; small for gestational age; multiple births; surfactant; grades 3 to 4 IVH; PVL; sepsis; oxygen requirement at 36 weeks; white vs. non-white race; outborn vs. inborn status caesarean section vs. vaginal delivery; maternal education <12 years vs. >=12 years; private health insurance vs. public;	MDI < 70, Bayley II	MDI < 70 assessed at age 18 to 22 months corrected age: IVH/PVL grade III-IV: Significantly increased odds Postnatal steroids: Significantly increased odds, BPD: Significantly increased odds Antenatal steroids NS Early-onset sepsis NS Late-onset sepsis NS NEC:NS	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			conventional ventilation vs. none; adjusted age at the time of assessment; centre; and the 4 interventions of interest: antenatal steroids (yes, no), high-frequency ventilation vs. none; days to regain birth weight, and postnatal steroids (yes, no).			
Hoffman 2015 (USA)	Retrospective study	N= 1934 children born at < 27wks GA	Antenatal steroids -adjusted for: not clearly reported, only reported " infant and maternal characteristics that varied significantly between groups"	Cognitive impairment BSID – III cognitive composite score < 70	BSID cognitive composite score < 70 assessed at age 18-22 months corrected age: Antenatal steroids: OR 0.94 (0.57-1.52)	Moderate
Laughon 2009 (USA)	Retrospective study	n=children born at < 28wks GA	Sepsis NEC BPD -adjusted for: it was reported that risk factors were adjusted for each other in a temporal pattern	MDI < 55 assessed through Bayley Scales of Infant Development-2nd Edition (BSID-II),	Outcomes assessed at age 24 months MDI < 55: Late bacteraemia: OR 1.8 (1.3 - 2.5) NEC >=stage II: OR 2.1 (1.2 - 3.7) BPD without mechanical ventilation: OR 1.1 (0.8 -1.4)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
					BPD with mechanical ventilation: OR 1.2 (0.7 -2.3)	
Mikkola 2005 (Finland)	Prospective cohort study	N= 193 Children born at 27.3 (± 2.1) weeks' GA	Antenatal steroids NEC BPD -adjusted for: maternal smoking, high social class, preeclampsia, absence of antenatal steroids, multiple birth, gestational age, birth weight, gender, SGA, vaginal delivery, Apgar score <4 at 5 min, university hospital area, birth outside a tertiary hospital, IVH grade 3-4, perforated NEC, O2 dependency at 36 weeks, ROP grades 3-4	Cognitive impairment: defined as IQ score <70, assessed by the Wechsler Preschool and Primary Scale of Intelligence-revised (WPPSI-R)	Cognitive impairment assessed at age 5 years: Antenatal steroids: OR 3.93 (1.3-12.2) NEC perforated: OR 12.47 (2.4-64) BPD: 5.62 (1.8-17.8)	Moderate
Beaino 2011 (France)	Prospective population based cohort. (EPIPAGE)	n=2901 All preterm infants 22-32 weeks gestation. Follow-up at 5 years of age.	NEC BPD Cerebral lesions Postnatal steroids -adjusted for:	Cognitive deficiency: Kaufman Assessment Battery for Children (K-ABC): Severe when the MPC score was below 70 (-2SD below the norm).	Severe cognitive deficiency assessed at age 5 years: NEC No: Reference Yes: OR 0.84 (0.33-2.15) BPD	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			Neonatal cerebral lesions, gestational age of 28 weeks or less, gender, small for gestational age, Apgar score below 7 at one minute, NEC, BPD at 36 weeks, acute anaemia, late-onset anaemia, postnatal corticosteroid, parental socioeconomic status, number of siblings and breast feeding.		No: Reference Yes: OR 1.09 (0.62-1.90) Grade I IVH No: Reference Yes: OR 1.39 (0.74-2.60) Grade II IVH No: Reference Yes: OR 1.88 (0.95-3.72) Grade III IVH or echodensities or ventricular dilatation No: Reference Yes: OR 2.51 (1.53-4.11) Cystic PVL or IPH No: Reference Yes: OR 6.37 (2.46-16.54) Postnatal steroids: OR 1.14 (0.66-1.97)	
Foix-L'Helias 2008 (France)	Prospective cohort study	N= 2855 children born at 24- 32 weeks GA	Antenatal steroids: -adjusted for: gestational age, social class, sex and pregnancy complications. A propensity score adjusted for general characteristics (maternal age, parity, tobacco consumption, region and level of neonatal	Cognitive ability was assessed using the mental processing composite (MPC) of the Kaufman Assessment Battery for Children. MPC scores of less than 70 indicate cognitive impairment.	MPC < 70 assessed at age 5 years: Among children born at 24-32wks GA: Antenatal steroids (any): OR 0.82 (0.54-1) Among children born at 24-27wks GA: Antenatal steroids (any): OR 1.61 (0.55-1.24) Among children born at 28-32wks GA: Antenatal steroids (any): OR 0.76 (0.48-1.18)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			intensive care), maternal complications and pregnancy etc.		Among children born at 24-32wks GA: Antenatal steroids (complete course): OR 0.91 (0.58-1.42) Among children born at 24-27wks GA: Antenatal steroids (complete course):: OR 1.78 (0.59-5.38) Among children born at 28-32wks GA: Antenatal steroids (complete course):: OR 0.85 (0.52-1.38)	
Hansen 2004 (Denmark)	Prospective study	N= 252 children born at 24.1-34.3 weeks GA	IVH grade III-IV; NCE -adjusted for: risk factors were adjusted for each other in the multivariate analysis, as well as CRIB-score (high), chronic lung disease, and mechanic ventilation	Intellectual disability: Intellectual development was defined as IQ score below -2 standard deviations from the mean of a reference group, and classified children with intellectual disabilities.	Intellectual disability IQ < -2SD assessed at age 5 years: IVH grade III-IV: OR 6.2 (2.3-16.5) NEC: OR 4.1 (0.8-20.8)	Moderate
Andrews 2008 (US)	Prospective study	N= 375 children born at 23-31 weeks GA	PVL -adjusted for: gestational age and ethnicity	IQ < 70 assessed with WISC-IV	IQ < 70 on WISC assessed at age 6 years: PVL: 4.9 (0.9-26)	Moderate
Speech and Language disorders						
Payne 2013 (USA)	Prospective cohort study	N= 1472 children born at < 27 weeks' GA	Low grade PIVH Severe PIVH Antenatal steroids	Speech and Language disorders defined as a score of <70 on the Bayley III.	Speech and language disorders (<70 on Bayley < 70) assessed at 18-22 months corrected age:	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			Sepsis Postnatal steroids -adjusted for: PIVH severity (3 levels), gestational age, sex, race/ethnicity, maternal education, chorioamnionitis, sepsis, antenatal steroid exposure, postnatal steroid exposure, high frequency ventilation and patent ductus arteriosus		Low grade PIVH versus no PIVH: OR 1 (0.61-1.64) Severe PIVH no PIVH: OR 3.43 (2.24-5.27) Severe PIVH versus low grade PIVH: OR 3.44 (1.96-5.98) Antenatal steroids: OR 0.69 (0.42-1.14) Sepsis: OR 1.48 (1.03-2.11) Postnatal steroids: OR 1.44 (0.92-2.26)	
Toome 2013 (Estonia)	Prospective cohort study	N= 187 children born at 22-31 weeks GA	Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV -adjusted for: antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or	Language composite score < -2SD, the Bayley Scales of Infant and Toddler Development	Language composite score -2SD (Bayley) assessed at age 2 years: Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV: OR 19 (4.8-75.1)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family			
Hoffman 2015 (USA)	Retrospective study	N= 1934 children born at < 27wks GA	Antenatal steroids -adjusted for: not clearly reported, only reported " infant and maternal characteristics that varied significantly between groups"	BSID – III language composite < 70 score;	BSID III language composite score < 70 assessed at age 18-22 months corrected age: Antenatal steroids: OR 0.66 (0.46-0.96)	Moderate
Autism spectrum disorder (ASD)						
Kuzniewicz 2014 (USA)	Retrospective study	n=3807 children born at < 34 weeks GA	Sepsis ICH grade I-II ICH grade III – IV Cystic PVL NEC	Autism spectrum disorder: Kaiser Permanente (KP) Autism Registry. This contains the location, provider, provider speciality and date of any ASD	Autism spectrum disorder assessed at age 2 to 11 years: Sepsis: OR 1.6 (0.8 -3.4) ICH grade I-II: OR 1.9 (1.1 -3.4) ICH grade III-IV: OR 3.4 (1.4 -8.6) Cystic PVL: OR 1.7 (0.2 -12.4)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			-adjusted for gestational age, sex, maternal age, maternal education	diagnosis recorded in the KP outpatient databases		
Hwang 2013 (Taiwan)	Prospective cohort study	N= 1078 children born at < 37wks GA	BPD -adjusted for : it was reported that "potential confounding factors of the relationship between significant risk factors on autism prevalence in preterm children"	Infantile autism based on ICD-9-CM coded by their doctors	Infantile autism assessed at age 8 to 11 years: BPD: OR 1.5 (0.8-2.9)	Low
Specific learning difficulties						
Kiechl-Kohlendorfer 2013 (Austria)	Prospective cohort study	N=161 children born at < 32wks GA	ICH all grades BPD -adjusted for : Smoking in pregnancy SGA Sex Neonatal Intracerebral haemorrhage BDP- bronco pulmonary dysplasia (chronic lung disease [CLD] at 36 weeks)	Specific learning difficulties: delay in numerical skills was assessed individually with the TEDI-MATH which is a multi-componential dyscalculia test based on cognitive neuropsychological models of number processing and calculation	Delayed numerical skills assessed at age 5 years: ICH, all grades: OR 4.66 (1.56 -13.93) BPD: OR 4.35 (1.11 -17.01)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			Necrotizing enterocolitis – NEC (stage II or worse) Sepsis (Pneumothorax; Late bacteremia) ROP - Retinopathy of prematurity			
Mental and behavioural disorders						
Johnson 2010 (UK & Ireland)	Prospective cohort study	N=307 children born at < 26 weeks GA	NEC -adjusted for: fetal heart rate >100 beats per minute at 5 minutes, need for oxygen at 36 weeks, gestational age, male gender, prolonged rupture of membranes, maternal age, externalizing behaviour problems at 2.5 years, internalizing behaviour problems at 2.5 years, pervasive attentional problems (at 6 years), serious	Mental and behavioural disorder: the Development and Well Being Assessment (DAWBA), and summary sheets and clinical transcripts were then reviewed by two child and adolescent psychiatrists who assigned DSM-IV and ICD-10 consensus diagnoses.	Any psychiatric disorder assessed at age 11 years: NEC: OR 7.15 (1-51)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			functional disability (at 6 years)and pervasive conduct problems (at 6 years).			
Visual impairment						
Adams-Chapman 2008 (USA)	Prospective cohort study	N= 6161 children born at between < 25wks and ≥ 33 weeks GA	IVH III/shunt IVH IV/shunt -adjusted for: study centre, gestational age, birth weight, gender, race, caesarean section delivery, multiple birth, antenatal steroid exposure, postnatal steroid exposure, surfactant use, respiratory distress syndrome, bronchopulmonary dysplacia (BPD), patent ductus arteriosus, periventricular leukomalacia (PVL), infection group, caregivers' education.	Visual impairment, defined as the need for corrective lenses or blindness in 1 or both eyes.	Blindness assessed at 18 to 22 months corrected age: IVH III/shunt versus IVH III/no shunt: OR 1.26 (0.87-1.8/2) IVH III/shunt versus no IVH/no shunt: OR 1.65 (1.18 – 2.31) IVH IV/shunt versus IVH IV/no shunt: OR 1.72 (1.19-2.46) IVH IV/shunt versus no IVH no shunt: OR 2.39 (1.71 – 3.35)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Carlo 2011 (USA)	Prospective cohort study	N= 4924 children born at 22 to 25 weeks GA	Antenatal steroids: -adjusted for: gender and race	Visual impairment:: blindness (blind with no useful vision in either eye) • deafness (functional hearing impairment with aids on both ears)	Blindness assessed at age 18-22 months corrected age: Among children born at < 22-25wks GA: Antenatal steroids: OR 0.61(0.36 - 1.03) Among children born at 22 weeks GA: Antenatal steroids: Not reported Among children born at 23 weeks GA: Antenatal steroids: OR 0.31 (0.1-0.93) Among children born at 24 weeks GA: Antenatal steroids: OR 1.17 (0.48-2.83) Among children born at 25 weeks GA: Antenatal steroids: OR 0.46 (0.19-1.1)	Moderate
Stoll 2004 (USA)	Prospective cohort study	N= 6314 pre-term children	Sepsis -adjusted for: study centre, gestational age, birth weight, sex, race/ethnicity, rupture of membranes >24 h, CS, multiple birth, antenatal antibiotics, antenatal	Vision impairment, defined as blindness in one or both eyes or need for corrective lenses.	Blindness assessed at age 18-22 months corrected age: Sepsis alone: OR 1.7 (1.3-2.2) Sepsis plus NEC: OR 2 (1.3-3) Meningitis with or without sepsis: OR 2.2 (1.4-3.6)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			steroids, postnatal steroids, surfactant use, respiratory distress syndrome, bronchopulmonary dysplasia, patent ductus arteriosus, intraventricular haemorrhage grade 3-4, periventricular leukomalacia, maternal age at time of delivery, caregiver's level of education			
Mikkola 2005 (Finland)	Prospective cohort study	N= 193 Children born at 27.3 (± 2.1) weeks' GA	ROP -adjusted for: maternal smoking, high social class, preeclampsia, absence of antenatal steroids, multiple birth, gestational age, birth weight, gender, SGA, vaginal delivery, Apgar score <4 at 5 min, university hospital area,	Severe visual impairment, classified as bilateral or unilateral amaurosis (loss of sight without apparent lesion of the eye), or amblyopia ("lazy eye", uncorrectable decrease in vision in one or both eyes with no apparent structural abnormality seen to explain), or a combination.	Visual impairment assessed at age 5 years: ROP grade III-IV: OR 10.6 (3.2 – 31.5)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			birth outside a tertiary hospital, IVH grade 3-4, perforated NEC, O2 dependency at 36 weeks, ROP grades 3-4			
Hearing impairment						
Adams-Chapman 2008 (USA)	Prospective cohort study	N= 6161 children born at between < 25wks and ≥ 33 weeks GA	IVH III/shunt IVH IV/shunt -adjusted for: study centre, gestational age, birth weight, gender, race, caesarean section delivery, multiple birth, antenatal steroid exposure, postnatal steroid exposure, surfactant use, respiratory distress syndrome, bronchopulmonary dysplacia (BPD), patent ductus arteriosus, periventricular leukomalacia (PVL), infection group, caregivers' education.	Hearing impairment, defined by hearing aid use in 1 or both ears.	Deafness assessed at 18 to 22 months corrected age: IVH III/shunt versus IVH III/no shunt: OR 0.33 (0.09-1.3) IVH III/shunt versus no IVH/no shunt: OR 0.88 (0.23-3.35) IVH IV/shunt versus IVH IV/no shunt: OR 1.41 (0.56-3.59) IVH IV/shunt versus no IVH no shunt: OR 2.13 (0.96-4.76)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Carlo 2011 (USA)	Prospective cohort study	N= 4924 children born at 22 to 25 weeks GA	Antenatal steroids: -adjusted for: gender and race	Deafness (functional hearing impairment with aids on both ears)	Deafness assessed at age 18-22 months corrected age: Among children born at < 22-25wks GA: Antenatal steroids: OR 0.76 (0.5-1.16) Among children born at 22 weeks GA: Antenatal steroids: Not reported Among children born at 23 weeks GA: Antenatal steroids: OR 0.39 (0.17-0.93) Among children born at 24 weeks GA: Antenatal steroids: OR 0.93 (0.45-1.9) Among children born at 25 weeks GA: Antenatal steroids: OR 0.91 (0.46-1.81)	Moderate
Stoll 2004 (USA)	Prospective cohort study	N= 6314 pre-term children	Sepsis -adjusted for: study center, gestational age, birth weight, sex, race/ethnicity, rupture of membranes >24 h, CS, multiple birth, antenatal antibiotics, antenatal steroids,	Deafness: hearing impairment, defined as hearing aids in one or both ears.	Hearing impairment assessed at age 18-22 months corrected age: Sepsis alone: OR 1.8 (1-3.1) Sepsis plus NEC: OR 3.4 (1.6-7.3) Meningitis with or without sepsis: OR 0.8 (0.2-2.8)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			postnatal steroids, surfactant use, respiratory distress syndrome, bronchopulmonary dysplasia, patent ductus arteriosus, intraventricular haemorrhage grade 3-4, periventricular leukomalacia, maternal age at time of delivery, caregiver's level of education			
Composite outcomes						
Hintz 2005 (USA)	Retrospective cohort study	N= 2948 extremely low birth weight infants, mean GA not reported;	NEC -adjusted for: network centre, use of antenatal glucocorticoids, rupture of membranes >24h, outborn status, estimated gestational age, gender, race, birth weight, small for gestational age, surfactant therapy, intraventricular	Composite outcome: (neurodevelopmental impairment): Composite outcome was defined as one of the following: motor, MDI < 70 or PDI < 70, blindness, deafness.	Neurodevelopmental impairment assessed at 18-22 months corrected age among children born extremely low birth weight: NEC surgical: OR 1.78 (1.17-2.73) NEC medical: OR 1.06 (0.69-1.63)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			haemmorrhage grade 3 or 4 or cystic periventricular leukomalacia, sepsis, postnatal steroid treatment, bronchopulmonary dysplasia, and highest level of education attained by the primary caregiver			
Merhar 2012 (USA)	N= 166 children born at 26wk GA (mean)	Prospective cohort study	IVH grade III IVH grade IV Postnatal steroids Sepsis Bilateral IVH -adjusted for: gender, race, birth weight, presence of bronchopulmonary dysplasia, postnatal steroids, early or late culture positive sepsis, necrotising enterocolotis requiring surgery	Composite outcome: neurodevelopmental impairment was defined as one of the following: motor, MDI < 70 or PDI < 70, blindness, deafness	Neurodevelopmental impairment assessed at 18-22 months corrected age: IVH grade II (vs IVH grade I): OR 0.4 (0.06 -2.6) IVH grade III (vs IVH grade I): OR 1.6 (0.52 - 4.9) IVH grade IV (vs IVH grade I): OR 3.5 (1.2 -10.4) Postnatal steroids: OR 2.8 (1.2 - 6.3) Sepsis: OR 2.4 (1-5.3) Bilateral IVH (vs unilateral IVH): OR 2.1 (0.93 -4.6)	Moderate
Payne 2013 (USA)	Prospective cohort study	N= 1472 children born at < 27 weeks' GA	Low grade PIVH Severe PIVH Antenatal steroids	A composite measure is having any one of the following: moderate-severe CP, severe visual	Composite outcome (Neurodevelopmental impairment) assessed at 18-22 months corrected age:	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			Sepsis Postnatal steroids -adjusted for: PIVH severity (3 levels), gestational age, sex, race/ethnicity, maternal education, chorioamnionitis, sepsis, antenatal steroid exposure, postnatal steroid exposure, high frequency ventilation and patent ductus arteriosus	impairment, deafness, or cognitive score <70 (-2SD) on the Bayley III.	Low-grade PIVH (vs no PIVH): OR 0.82 (0.51 -1.31) Severe PIVH (vs no PIVH): OR 1.68 (1.06 -2.65) Severe PIVH (vs low-grade PIVH): OR 2.04 (1.15 -3.64) Antenatal steroids: OR 0.84 (0.51 -1.4) Sepsis: OR 1.99 (1.4 -2.83) Postnatal steroids: OR 1.62 (1.06 -2.48)	
Perrot 2003 (Canada)	Prospective study	N= 253 children born at 22-30 weeks GA	PVL -adjusted for: Hypernatremia; and surgery.	A composite measure is having any one of the following: moderate-severe CP, severe visual impairment, deafness, or cognitive score MDI <70 (-2SD) on the Bayley III.	Composite outcome (Neurodevelopmental impairment) assessed at age 22-30 months: Cystic PVL: OR 31.1 (8.8-110.3)	Low
Shah 2012 (USA)	Prospective cohort study	N= 865 children born at 25.7-26.2 GA	NEC -adjusted for: birth weight, race, gender, multiple births, antenatal steroids, surfactant, bronchopulmonary dysplasia,	"Any disability" defined as a composite variable including any one of the following conditions: MDI score <70 PDI score <70 Cerebral palsy (CP), Hearing impairment, and	Composite outcome (Neurodevelopmental impairment) assessed at age 18 to 22 months corrected age: NEC >=IIA: OR 2.59 (1.44 -4.66) NEC >=IIA surgically managed: NS NEC >=IIA medically managed: NS	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			sepsis, and any intraventricular hemorrhage	Visual impairment;		
Vohr 2005 (USA)	Prospective cohort study	N= 3785 children born at 22 to 32 weeks' GA	PVL; IVH grade III-IV: BPD: Sepsis: Antenatal steroids: -adjusted for: gestational age group; birth weight; gender; small for gestational age; multiple births; surfactant; grades 3 to 4 IVH; PVL; sepsis; oxygen requirement at 36 weeks; white vs. non-white race; outborn vs. inborn status caesarean section vs. vaginal delivery; maternal education <12 years vs. >=12 years; private health insurance vs. public; conventional ventilation vs. none; adjusted	Neurodevelopmental impairment (NDI), defined as the presence of any of the following: moderate to severe CP; hearing loss requiring bilateral amplification; bilateral blindness (not defined); MDI <70; PDI <70;	Neurodevelopmental impairment assessed at age 18-22 months corrected age: PVL: Significant, NR IVH grade III-IV: Significant, NR Postnatal steroids: Significant, NR BPD: Significant, NR Sepsis: NS Antenatal steroids: NS	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			age at the time of assessment; centre; and the 4 interventions of interest: antenatal steroids (yes, no), high-frequency ventilation vs. none; days to regain birth weight, and postnatal steroids (yes, no).			
Adams-Chapman 2008 (USA)	Prospective cohort study	N= 6161 children born at between < 25wks and ≥ 33 weeks GA	IVH III/shunt IVH IV/shunt -adjusted for: study center, gestational age, birth weight, gender, race, caesarean section delivery, multiple birth, antenatal steroid exposure, postnatal steroid exposure, surfactant use, respiratory distress syndrome, bronchopulmonary dysplacia (BPD), patent ductus arteriosus, periventricular	Neurodevelopmental impairment (NDI), a composite outcome defined as 1 or more of the following: MDI <70, PDI <70, CP, blind in both eyes, or hearing aids in both ears	Neurodevelopmental impairment assessed at 18 to 22 months corrected age: IVH III w/ shunt (vs IVH III no shunt): OR 1.29 (1.11 -1.48) IVH III w/ shunt (vs no IVH no shunt): OR 1.57 (1.38-1.78) IVH IV w/ shunt (vs IVH IV no shunt): OR 1.44 (1.27 -1.64) IVH IV w/ shunt (vs no IVH no shunt): OR 1.81 (1.62 -2.03)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			leukomalacia (PVL), infection group, caregivers' education.			
Carlo 2011 (USA)	Prospective cohort study	N= 4924 children born at 22 to 25 weeks GA	Antenatal steroids: -adjusted for: Gender and race	Neurodevelopmental impairment at 18-22 months defined as 1 or more of the following: a Bayley II Mental Developmental index (MDI) <70; a Bayley II Psychomotor Development index (PDI) <70;moderate-severe cerebral palsy (CP); deafness	Neurodevelopmental impairment assessed at age 18-22 months corrected age: Among children born at < 22-25wks GA: Antenatal steroids: OR 0.83 (0.7 - 0.99) Among children born at 22 weeks GA: Antenatal steroids: OR 1.14 (0.39 -3.28) Among children born at 23 weeks GA: Antenatal steroids: OR 1.11 (0.72 -1.71) Among children born at 24 weeks GA: Antenatal steroids: OR 0.8 (0.6 - 1.08) Among children born at 25 weeks GA: Antenatal steroids: OR 0.81 (0.62 -1.04)	Moderate
Goldstein 2013 (USA)	Multicentre retrospective cohort study	n=5456 Preterm infants born at 23-28 weeks. Follow-up at 18-22 months	NEC -adjusted for: Gestational age, Apgar score at 5 minutes, antenatal steroids, early	Neurodevelopmental impairment (NDI) was defined as at least one of: moderate/severe cerebral palsy with Gross Motor Function score 3-5, Mental	neurodevelopmental impairment assessed at 18-22 months corrected age: NEC: OR 6.89 (1.44-32.88)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			infection, postnatal steroids, NEC, late onset infection, cystic PVL, ventriculoperitoneal shunt insertion, maternal education, Medicaid status and BPD at 36 weeks.	Development Index or Psychomotor Development Index < 70 on the BSID-II at 18-22 months corrected age, blindness (no functional vision in both eyes) or deafness		
Leversen 2010 (Norway)	Prospective population based cohort.	n=376 preterm infants (22-27+6 weeks or birthweight 500-999g)	Sepsis BPD NEC IVH PVL ROP -adjusted for: Gestational age, gender, multiple pregnancy, chorioamnionitis, preeclampsia, antenatal steroids, PROM, Caesarean section, SGA, illness severity score (a score of the lowest and highest FiO2 requirements and the largest base	Neurosensory disabilities". This includes cerebral palsy, blindness (classified as legally blind) or complete deafness.	Neurosensory disability (CP/ blindness/ deafness) assessed at age 2 years: Antenatal steroids: OR 0.5 (0.2 - 1.6) Sepsis: OR 0.7 (0.2 -2.3) BPD: OR 0.9 (0.3 -2.9) NEC: OR 2 (0.3-11.9) Minor pathology in cranial ultrasound (periventricular haemorrhage grade I-II, eventually 1-2 small PVL): OR 2.5 (0.7 -9.7) Major pathology in cranial ultrasound (periventricular haemorrhage grade III-IV and/or multicystic PVL): OR 110.2 (23.4 - 518.5) ROP grade I-II: OR 3.5 (1.1 -11.6) ROP >II°: OR 5.8: (1 -32.5) Postnatal steroids <21 days: OR 0.9 (0.2 -3.7)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			deficit during the first 12 hours of life), septicaemia, BPD, patent ductus arteriosus, NEC, postnatal steroids, cranial ultrasound findings and ROP.		Postnatal steroids >=21 days: OR 5 (0.9 - 27.8)	
Toome 2013 (Estonia)	Prospective cohort study	N= 187 children born at 22-31 weeks GA	Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV -adjusted for: antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at	neurodevelopmental impairment includes any one (or more) of the following criteria: CP with GMFCS level 2,3,4 or 5; cognitive and/or language composite scores of ≤-2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.	Neurodevelopmental impairment assessed at age 2 years: Severe cerebral lesions, including IVH grade III-IV and/or PVL grade II-IV: OR 33.4 (8.6-129.9)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family			
Stoll 2004 (USA)	Prospective cohort study	N= 6314 pre-term children	Sepsis -adjusted for: study center, gestational age, birth weight, sex, race/ethnicity, rupture of membranes >24 h, CS, multiple birth, antenatal antibiotics, antenatal steroids, postnatal steroids, surfactant use, respiratory distress syndrome, bronchopulmonary dysplasia, patent ductus arteriosus, intraventricular haemorrhage grade 3-4, periventricular	Neurodevelopmental impairment (NDI, a composite outcome, defined as one or more of the following: MDI <70, PDI <70, CP, bilateral blindness or bilateral hearing impairment.	Neurodevelopmental impairment assessed at age 18-22 months corrected age: Sepsis alone: OR 1.5 (1.2-1.7) Sepsis plus NEC: OR 1.8 (1.4-2.5) Meningitis with or without sepsis: OR 1.6 (1.1-2.3)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			leukomalacia, maternal age at time of delivery, caregiver's level of education			
Shankaran 2004 (USA)	Prospective study	N= 246 children born at less or equal to 24 weeks GA	ICH grade III-IV; PVL; Any antenatal steroids BPD -Adjusted for: risk factors were adjusted for each other, plus surfactant administration, steroids for BPD, Medicaid, No high school degree, 2-parent household;	Composite outcome: Neurodevelopmental impairment: 1 or more of the following: motor, cognitive, visual, hearing)	Neurodevelopmental impairment assessed at age 18-22 months corrected age: ICH grade III-IV: OR 2.5 (1.2 -5.2) PVL: OR 2.4 (0.6 - 9.5) Any antenatal steroids: OR 1.4 (0.7 -2.6) BPD: OR 1.7 (0.9 -3.3)	Low
Walsh 2005 (UK)	Prospective cohort study	N= 3041 children born at 25.8 (mean) weeks GA	PVL IVH grade III-IV Postnatal steroids Antenatal steroids NEC -adjusted for: Risk factors were adjusted for each other in the multiple regression model	Composite outcome: (Neurodevelopmental impairment) the Bayley Scales of Infant Development - II, including the mental scale, psychomotor scale, and the behaviour rating scale, were administered by developmental specialist. 1 or more of the following were assessed: (motor, cognitive, visual, hearing)	Neurodevelopmental impairment assessed at age 18-22 months corrected age: PVL: OR 3.72 (2.52-5.5) IVH grade III-IV: OR 1.3 (1.06 -1.69) Postnatal steroids: OR 1.13 (0.91 -1.4) Antenatal steroids: OR 0.81 (0.65 -1) NEC: NS	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
Bolisetty 2014 (Australia)	Retrospective cohort study	N=1472 children born at 23-28 weeks GA	IVH grade I-II IVH grade III-IV Proven systemic infection NEC ROP grade III-IV -adjusted for: IVH, gestation (23-25 weeks versus 26-28 weeks), SGA, male gender, outborn, PVL, chronic lung disease, pregnancy induced hypertension, proven systemic infection, NEC and ROP grade 3-4	Neurosensory impairment: moderate or severe neurosensory impairment was defined as the presence of developmental delay (Griffiths Mental Developmental Scale General Quotient or Bayley Scales of Infant Development MDI between 2 and 3 SD below the mean), cerebral palsy (able to walk with the assistance of aids), deafness or bilateral blindness	Neurosensory impairment assessed at age 2-3 corrected years: IVH grade I-II: OR 1.61 (1.14 - 2.28) IVH grade III-IV: OR 3.81 (2.3-6.3) Proven systemic infection: OR 1.2 (.88-1.65) NEC: OR 1.09 (0.65-1.82) ROP grade III-IV: OR 2.13 (1.44 - 3.14)	Moderate
Kallen 2015 (Sweden)	Prospective study	N=456 children born at less than 27 weeks GA	Antenatal steroids -adjusted for gestational age and for birth weight standard deviation score	Neurosensory impairment: Bayley III scale (1 or more of the following impairments: motor, vision, hearing)	Neurosensory impairment assessed at 2.5 yrs corrected age: Antenatal steroids: OR 1.1 (0.3-4.8)	Moderate
Wong 2014 (Australia)	Retrospective study	N=1473	Antenatal steroids -adjusted for: Significant and clinically important baseline	Moderate/severe functional disability (Neurodevelopmental impairment), defined as one or more of the following:	Functional disability (Neurodevelopmental impairment) assessed at age 2-3 years: Antenatal steroids: 1.056 (0.785-1.42)	Moderate

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			population characteristics: maternal age, pregnancy-induced hypertension, gestational age, birth weight, gender, outborn status and assisted conception.	developmental delay (<2SD below the mean for adjusted age determined by the GMDS or BSID-II); cerebral palsy (unable to walk without aids); bilateral blindness (visual acuity <6/60 in better eye); bilateral deafness (requiring bilateral hearing aids or cochreal implants		
Victorian Infant Collaborative Study Group 2000 (Australia)	Prospective cohort study	N= 280 children born at < 28wks' GA	Postnatal steroids -adjusted for: ruptured membranes >24h, cystic PVL, and surgery during the primary hospitalization	Severe sensorineural impairment, composite outcome, defined as having 1 or more of the following: bilateral blindness. CP with the child unlikely ever to walk, IQ score <-3SD, IQ score assessed by Wechsler Preschool and Primary Scale of Intelligence - Revised (WPPSI-R) or other psychological test when WPPSI-R was unavailable (not specified).	sensorineural impairment assessed at age 5 years: Postnatal steroids: OR 3.2 (1.6-6.4)	Moderate
Herbat – Jonat 2014 (Germany)	Prospective cohort study	n=79 children born at 22-24 weeks GA	Intracerebral haemorrhage >II° ROP >II° NEC >IIB	Composite neurodevelopmental impairment including components of motor,	Composite outcome (Neurodevelopmental impairment assessed at age 7-10 yrs:	Low

Study	Data Source	Sample and population studied	Risk factor (s) and adjustment	Measures of outcomes	Prognostic outcomes	Study quality
			Chronic lung disease/BPD -adjusted for: all variables above	vision, cognitive, hearing assessed by	Intracerebral haemorrhage >II°: Not reported ROP >II°: OR 3.18 (1.09 - 9.31) NEC >IIB: NS Chronic lung disease/BPD: NS	

Table 18: Summary of studies on the associating between social, environmental and maternal factors and developmental disorders

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Cerebral palsy						
Beaino 2011	Population based prospective cohort study (EPIPAGE)	n=1812 preterm babies born at 24-32 weeks	Children were classified as having CP if they had involuntary movements (dyskinetic CP), loss of coordination (ataxic CP), or at least two of the following: abnormal posture or movement, increased tone or hyperreflexia (spastic CP).	Obstetric and neonatal factors (not specified further). From the text it is assumed that they are: cystic PVL, intraparenchymal haemorrhage, gestational age, gender, SGA, multiple pregnancy, PPROM or preterm labour, maternal hypertension, RDS, NEC, maternal-foetal infection, BPD at 36 weeks, acute anaemia	**At 5 years of age** Cerebral palsy **Multiple pregnancy** No: Reference Yes: OR 0.67 (0.43-1.03) **Maternal age** Not significant on univariate analysis	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
				and postnatal corticosteroid use.		
Hirvonen 2014	Retrospective cohort study using national registry data	n=53078 preterm infants	All inpatient and outpatient visits due to a CP diagnosis in public hospitals were registered. The diagnosis of CP in Finland is based on medical history, ultrasound and MRI data, and multidisciplinary evaluations in the paediatric neurology units of 20 secondary level central hospitals and 5 tertiary level university hospitals. The diagnosis is included in the database as soon as it has been established. A case of CP was recorded if the individual was detected in the Hospital Discharge Register and/or in the Reimbursement Register of the	Period of study (1991-1995, 1996-2001 or 2002-2008), maternal age, maternal smoking status, primiparous, previous C-section, maternal diabetes, multiple pregnancy, order of foetuses, assisted reproductive technology, cervical cerclage, chorionic villus sampling, PROM, preeclampsia, time of birth, antenatal steroid use, place of birth,	**Up to the age of 7 years** Cerebral palsy *Within very preterm infants, <32 weeks gestation* **Maternal age** < 40 years: Reference ≥ 40 years: OR 1.14 (0.69-1.89) **Multiple pregnancy** Singleton: Reference Twins: OR 0.94 (0.70-1.26) Higher order multiples: OR 1.24 (0.63-2.45) *Within moderately preterm infants, 32+0 to 33+6 weeks gestation* **Maternal age** < 40 years: Reference ≥ 40 years: OR 0.85 (0.33-2.17) **Multiple pregnancy** Singleton: Reference Twins: OR 0.83 (0.48-1.44) Higher order multiples: OR 0.88 (0.28-2.81) *Within late preterm infants, 34+0 to 36+6 weeks gestation*	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			Social Insurance Institution.	mode of delivery, gender, gestational weight, birth weight <1500g, Apgar score, umbilical artery pH, admission to neonatal unit, ventilator, resuscitation at birth, phototherapy, antibiotic therapy, RDS, sepsis, intracranial haemorrhage, convulsions and hyperbilirubinaemia.	**Maternal age** < 40 years: Reference ≥ 40 years: OR 1.40 (0.70-2.78) **Multiple pregnancy** Singleton: Reference Twins: OR 0.77 (0.47-1.27) Higher order multiples: OR 0.51 (0.07-3.92)	
Marret 2007	Population based prospective cohort study (EPIPAGE)	n=1461 preterm infants (30-34^{+6} weeks)	Cerebral palsy was defined as at least two of: abnormal posture or movement, increased tone and hyperreflexia. When the diagnosis of cerebral palsy was	Gestational age, multiple pregnancy, intrauterine growth restriction (IUGR), maternal hypertension,	**At 5 years of age** Cerebral palsy **Multiple pregnancy** No: Reference Yes: OR 1.6 (0.7-3.8)	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			in doubt, a panel of trained paediatricians met to discuss the case.	haemorrhage, preterm labour, preterm prolonged rupture of the membranes (PROM), antenatal corticosteroid exposure, gender and socioeconomic status.		
Miyazaki 2016	Retrospective cohort study using national registry data	n=2201 preterm infants born at <34 weeks of gestation	CP was defined as a non-progressive central nervous system disorder characterised by abnormal muscle tone in at least one extremity and abnormal control of movement and posture.	Maternal age, parity, maternal diabetes, premature rupture of membranes, preeclampsia, non-reassuring fetal status, mode of birth, administration of antenatal steroids, gestational age at birth, birth weight, SGA and sex.	**At 3 years of age (chronological age)** Cerebral palsy **Histological chorioamnionitis** No: Reference Yes: OR 0.91 (0.75-1.30)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Pappas 2014	Multicentre retrospective cohort study	n=2235 preterm infants born at <27 weeks' gestation	Cerebral palsy was defined as a non-progressive central nervous system disorder with abnormal muscle tone in at least one extremity and abnormal control of movement and posture that interfered with age-appropriate activities.	Adjusted by reduced models that contained covariates for centre, sex, antenatal steroids, SGA and hypertension.	**At 18-22 months' corrected age** Cerebral palsy **Histological chorioamnionitis** No: Reference Yes: OR 0.80 (0.42-1.53) **Histological chorioamnionitis plus clinical chorioamnionitis** No: Reference Yes: OR 1.39 (0.67-2.87)	High
Shankaran 2004	Multicentre prospective cohort study	n=246 preterm infants ≤24 weeks' gestation and ≤750g	Cerebral palsy was defined as a non-progressive central nervous system disorder characterized by abnormal muscle tone in at least 1 extremity and abnormal control of movement and posture.	ICH grade 3-4, PVL, any antenatal steroids, male gender, ethnicity, household income < 20K, BPD, surfactant administration, steriods for BPD, Medicaid, no high school degree and 2-parent household.	**At 18-22 months' corrected age;** Cerebral palsy **Household income < 20K:** OR 1 (0.4-2.4)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Tommiska 2003	Population based prospective cohort study	n=208 preterm infants <1000g	Cerebral palsy was defined as a non-progressive motor impairment with spastic or dystonic muscle tone, brisk tendon reflexes, positive Babinski's sign and persistent primitive reflexes.	Multiparity, pre-eclampsia, premature rupture of membranes, maternal infection, antenatal steroid treatment, hyperstimulation or in vitro fertilisation, maternal age below 20 or above 40, smoking, marital status, social class 1-4, birth in secondary level hospital, catchment area for the different hospitals, vaginal delivery, birth weight (100g groups), intrauterine growth restriction, gestational age, male gender, multiple birth, anomalies, respiratory distress syndrome, septicaemia,	**At 18 months' corrected age** Cerebral palsy **Multiple birth** Not a significant independent predictor on multivariate analysis **Maternal age** Not a significant independent predictor on multivariate analysis **Socioeconomic status** Not a significant independent predictor on multivariate analysis	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
				necrotising enterocolitis with perforation and intraventricular haemorrhage grades 2-4.		
Toome 2013	Population based prospective cohort study	n=187 preterm infants <32 weeks gestation	Cerebral palsy was defined according to the guidelines of the Surveillance of Cerebral Palsy in Europe collaborative group, and the Gross Motor Function Classification System (GMFCS) was used to quantify motor function in infants with CP.	Antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and	**At 2 years' corrected age** Cerebral palsy **Maternal age** Not a significant independent predictor on multivariate analysis **Low income of the family** Not a significant independent predictor on multivariate analysis **Multiple births** Not a significant independent predictor on multivariate analysis	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
				low income of the family).		
Wood 2005	Population based prospective cohort study (EPICure)	n=283 preterm babies <26 weeks	Cerebral palsy was classified retrospectively, being defined as a non-progressive disorder of movement and posture.	OR are stated to be adjusted. Factors adjusted for are not stated in the text.	**At 30 months correct age** Cerebral palsy **Chorioamnionitis** No: Reference Yes: OR 0.39 (0.16 to 0.96) (according to analysis of variables known at birth)	Moderate
Intellectual disability						
Beaino 2010	Population based prospective cohort study (EPIPAGE)	n=1503 preterm babies born at 24-32 weeks	Mental Processing Composite (MPC) of the Kaufmann Assessment Battery for Children (K-ABC) was used to assess intellectual disability. Scores of between 1 and 2 SD below the mean were identified as “mild cognitive deficiency”. Scores of <2SD below the mean were identified as “severe cognitive deficiency”	Medical factors (neonatal cerebral lesions, gestational age of 28 weeks or less, gender, small for gestational age, Apgar score below 7 at one minute, NEC, BPD at 36 weeks, acute anaemia, late-onset anaemia and postnatal corticosteroid), social factors (parental socioeconomic status, number of siblings) and breast feeding.	**At age 5 years** Mild cognitive deficiency High socioeconomic status: Reference High-intermediate socioeconomic status: OR 1.42 (0.88-2.28) Low-intermediate socioeconomic status: OR 2.19 (1.26-3.82) Low socioeconomic status: OR 3.43 (2.01-5.83) Severe cognitive deficiency High socioeconomic status: Reference High-intermediate socioeconomic status: OR 1.23 (0.65-2.32) Low-intermediate socioeconomic status: OR 2.89 (1.42-5.88) Low socioeconomic status:	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
					OR 2.60 (1.29-5.24)	
Hoffman 2015	Retrospective cohort study	Sample recruited - n=3790 infants born at <27 weeks (456 born to adolescent mothers + 3364 born to adult mothers)	The primary study outcomes were BSID-III composite cognitive and language scores.	Adjustment for infant and maternal characteristics that varied significantly between groups	**At 18-22 months** Intellectual disability (Cognitive Composite <70 and <85; Language Composite <70 and <85; and Motor Composite <70) **Adolescent mother<20 y old** Cognitive Composite <70 - (RR [95% CIs]) Referent group is not reported 1.42 (0.88–2.29) Motor Composite <70 - (RR [95% CIs]) Referent group is not reported1.01 (0.67–1.52)	Moderate
Källén 2015	Population based prospective cohort study (EXPRESS)	n=456 preterm infants <27 weeks	Mental developmental delay was defined as a cognitive or language Bayley III scale <2SD below the mean, or moderate or severe developmental delay according to chart review.	Gestational age	**At 2.5 years corrected age** Mental developmental delay **Chorioamnionitis/Prolonged and premature rupture of membranes** No: Reference Yes: OR 0.9 (0.5-1.7) **Multiple birth** No: Reference Yes: OR 1.5 (0.8-2.7)	Moderate
Marret 2007	Population based prospective cohort study (EPIPAGE)	n=1461 preterm infants (30-34^{+6} weeks)	The Kaufman Assessment Battery for Children (K-ABC) was used to identify cognitive ability, recorded as a mental processing composite	Gestational age, multiple pregnancy, intrauterine growth restriction (IUGR),	**At 5 years of age** Moderate/severe cognitive impairment **Multiple pregnancy** No: Reference Yes: OR 1.0 (0.6-1.7)	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			score (MPC). Scores on the MPC of <2SD below the mean were defined as moderate/severe cognitive impairment.	maternal hypertension, haemorrhage, preterm labour, preterm prolonged rupture of the membranes (PROM), antenatal corticosteroid exposure, gender and socioeconomic status.	**Socioeconomic status of the family** Professional: Reference Intermediate: OR 1.9 (0.7-5.4) Office worker or self-employed: OR 2.8 (1.0-7.6) Service worker or shop assistant: OR 4.5 (1.6-12.3) Manual worker or unemployed: OR 6.0 (2.3-15.6)	
Miyazaki 2016	Retrospective cohort study using national registry data	n=2201 preterm infants born at <34 weeks of gestation	Cognitive function was assessed using the Kyoto Scale of Psychological Development (KSPD) test by psychologists. When development quotient (DQ) was <70, the child was considered to have cognitive delay, according to the protocol of the Society for Follow-up Study of High-risk Infants.	Maternal age, parity, maternal diabetes, premature rupture of membranes, preeclampsia, non-reassuring fetal status, mode of birth, administration of antenatal steroids, gestational age at birth, birth weight, SGA and sex.	**At 3 years of age (chronological age)** DQ <70 **Histological chorioamnionitis** No: Reference Yes: OR 1.27 (0.90-1.79)	Low
Pappas 2014	Multicentre retrospective cohort study	n=2235 preterm infants born at <27 weeks' gestation	Infants underwent a comprehensive follow-up assessment at 18-22 months corrected	Adjusted by reduced models that contained covariates for	**At 18-22 months' corrected age** MDI <70	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			age. Psychometric testing was performed using the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley III). A score of less than 70 represents <2SD below the mean. Children who were so severely developmentally delayed that they could not be assessed were assigned scores (54 for severe cognitive delay and 46 for severe language delay).	centre, sex, antenatal steroids, SGA and hypertension.	**Histological chorioamnionitis** No: Reference Yes: OR 1.07 (0.62-1.85) **Histological chorioamnionitis plus clinical chorioamnionitis** No: Reference Yes: OR 2.00 (1.10-3.64)	
Shankaran 2004	Multicentre prospective cohort study	n=246 preterm infants ≤24 weeks' gestation and ≤750g	The Bayley Scales of Infant Development (BSID-II), including the Mental Developmental Index (MDI) was administered.	ICH grade 3-4, PVL, any antenatal steroids, male gender, ethnicity, household income < 20K, BPD, surfactant administration, steroids for BPD, Medicaid, no high school degree and 2-parent household.	**Assessment at 18-22 months' corrected age;** Cognitive impairment (MDI < 70) **Household income < 20K:** OR1.2 (0.5-2.5)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Singer 2001	Population based prospective cohort study	N=82 very low birth weight infants (41 mothers cocaine-positive + 41 mothers cocaine-negative)	The Bayley Scales of Infant Development that is described as widely used assessment toll of infant development. The Mental Development Index (MDI) is a standard score reflecting memory, learning and problem-solving abilities.		**At 3 years** Intellectual disability (MDI <70) When baseline differences were controlled, the effects of cocaine on intellectual disability remained significant	Low
Toome 2013	Population based prospective cohort study	n=187 preterm infants <32 weeks gestation	The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below the mean for cognitive and language composite scores.	Antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age,	**At 2 years' corrected age** Cognitive composite score <70 **Maternal age** Not a significant independent predictor on multivariate analysis **Low income of the family** Not a significant independent predictor on multivariate analysis **Multiple births** Not a significant independent predictor on multivariate analysis	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
				maternal higher education, single mother, paternal age, paternal higher education and low income of the family).		
Speech and/or language disorder						
Hoffman 2015	Retrospective cohort study	Sample recruited - n=3790 infants born at <27 weeks (456 born to adolescent mothers + 3364 born to adult mothers)	The primary study outcomes were BSID-III composite cognitive and language scores.	Adjustment for infant and maternal characteristics that varied significantly between groups	At 18-22 ,months Intellectual disability (Cognitive Composite <70 and <85; Language Composite <70 and <85; and Motor Composite <70) **Adolescent mother <20 years old** Language Composite <70 - (RR [95% CIs]) Referent group is not reported0.97 (0.64–1.47)	Moderate
Toome 2013	Population based prospective cohort study	n=187 preterm infants <32 weeks gestation	The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below	Antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4,	**At 2 years' corrected age** Language composite score <70 **Maternal age** Not a significant independent predictor on multivariate analysis **Low income of the family** Not a significant independent predictor on multivariate analysis **Multiple births**	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			the mean for cognitive and language composite scores.	BPD, ROP stage 3-5 with laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family).	Not a significant independent predictor on multivariate analysis	
Hearing impairment						
Miyazaki 2016	Retrospective cohort study using national registry data	n=2201 preterm infants born at <34 weeks of gestation	Severe hearing impairment including need for hearing aids was assessed at the participating centre.	Maternal age, parity, maternal diabetes, premature rupture of membranes, preeclampsia, non-reassuring fetal status, mode of birth, administration of antenatal steroids, gestational age at birth, birth weight, SGA and sex.	**At 3 years of age (chronological age)** Severe hearing impairment (including need for hearing aids) **Histological chorioamnionitis** No: Reference Yes: OR 1.28 (0.49-3.32)	Low

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
Visual impairment						
Miyazaki 2016	Retrospective cohort study using national registry data	n=2201 preterm infants born at <34 weeks of gestation	Visual impairment, defined as unilateral or bilateral blindness diagnosed by an ophthalmologist.	Maternal age, parity, maternal diabetes, premature rupture of membranes, preeclampsia, non-reassuring fetal status, mode of birth, administration of antenatal steroids, gestational age at birth, birth weight, SGA and sex.	**At 3 years of age (chronological age)** Visual impairment (unilateral or bilateral blindness) **Histological chorioamnionitis** No: Reference Yes: OR 1.08 (0.65-1.78)	Low
Composite outcomes						
Källén 2015	Population based prospective cohort study (EXPRESS)	n=456 preterm infants <27 weeks	Composite outcome of neurosensory impairment, defined as moderate/severe cerebral palsy or moderate/severe impairment regarding vision or hearing.	Gestational age	**At 2.5 years corrected age** Neurosensory impairment **Chorioamnionitis/Prolonged and premature rupture of membranes** No: Reference Yes: OR 0.8 (0.3-2.0) **Multiple birth** No: Reference Yes: OR 0.8 (0.3-2.1)	Moderate
Leversen 2010	Population based prospective cohort study	n=373 preterm infants (22-27^{+6} weeks)	Composite outcome of "major neurosensory	Gestational age, gender, multiple pregnancy,	**At 2 years of age** Major neurosensory disability	Moderate

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			disabilities". This includes cerebral palsy, blindness (classified as legally blind) or complete deafness.	chorioamnionitis, preeclampsia, antenatal steroids, PROM, Caesarean section, SGA, illness severity score (a score of the lowest and highest FiO2 requirements and the largest base deficit during the first 12 hours of life), septicaemia, BPD, patent ductus arteriosus, NEC, postnatal steroids, cranial ultrasound findings and retinopathy of prematurity.	**Multiple pregnancy** No: Reference Yes: OR 1.5 (0.4-5.8) **Chorioamnionitis** No: Reference Yes: OR 5.3 (1.4-20.4)	
Pappas 2014	Multicentre retrospective cohort study	n=2235 preterm infants born at <27 weeks' gestation	Infants underwent a comprehensive follow-up assessment at 18-22 months corrected age. Psychometric testing was perforemd using the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley	Adjusted for maternal age, multiple birth, parity, antenatal steroids, maternal hypertension, antepartum haemorrhage, sex, gestational	**At 18-22 months' corrected age** Neurodevelopmental impairment **Histological chorioamnionitis** No: Reference Yes: OR 0.89 (0.56-1.42)†	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			III). A score of less than 70 represents <2SD below the mean. Children who were so severely developmentally delayed that they could not be assessed were assigned scores (54 for severe cognitive delay and 46 for severe language delay). Cerebral palsy was defined as a nonprogressive central nervous system disorder with abnormal muscle tone in at least one extremity and abnormal control of movement and posture that interfered with age-appropriate activities. Disabling CP was classified as GMFCS ≥ level 2. Neurodevelopmental impairment was defined by one or more of disabling CP, Bayley scores <70, GMFCS level II or greater, blindness or permanent hearing	age, SGA status, insurance, race and centre.	**Histological chorioamnionitis plus clinical chorioamnionitis** No: Reference Yes: OR 1.51 (0.88-2.59)†	

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			loss that did not permit the child to understand or communicate despite amplification.			
Shankaran 2004	Multicentre prospective cohort study	n=246 preterm infants ≤24 weeks' gestation and ≤750g	Neurodevelopmental impairment (NDI) was defined as CP, MDI or PDI < 70, bilateral blindness, or hearing impaired with amplification.	ICH grade 3-4, PVL, any antenatal steroids, male gender, ethnicity, household income < 20K, BPD, surfactant administration, steriods for BPD, Medicaid, no high school degree and 2-parent household.	**At 18-22 months' corrected age;** Neurodevelopmental impairment **Household income < 20K:** OR 1.3 (0.6-2.8)	Low
Toome 2013	Population based prospective cohort study	n=187 preterm infants <32 weeks gestation	Cerebral palsy was defined according to the guidelines of the Surveillance of Cerebral Palsy in Europe collaborative group, and the Gross Motor Function Classification System (GMFCS) was used to quantify motor function in infants with CP. The Bayley Scales of Infant and Toddler	Antenatal steroids, multiple births, gestational age, birthweight, small for gestational age, male gender, surfactant, postnatal steroids, IVH grade 3-4 and/or PVL grade 2-4, BPD, ROP stage 3-5 with	**At 2 years' corrected age** Neurodevelopmental impairment **Maternal age** Not a significant independent predictor on multivariate analysis **Low income of the family** Not a significant independent predictor on multivariate analysis **Multiple births**	High

Study	Data Source	Sample and Population studied	Measures of Outcomes	Adjustment	Prognostic outcomes	Study Quality
			Development were used to generate composite scores for cognitive, language and motor skills, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below the mean for cognitive and language composite scores. A composite outcome measure of neurodevelopmental impairment was used. This includes any one (or more) of the following criteria: CP with GMFCS level 2,3,4 or 5; cognitive and/or language composite scores of ≤-2SD below the norm; hearing loss corrected with hearing aids or deafness; vision moderately reduced or blindness.	laser therapy, positive blood culture sepsis, NEC stage 2-3, weight<10th percentile at discharge, maternal age, maternal higher education, single mother, paternal age, paternal higher education and low income of the family).	Not a significant independent predictor on multivariate analysis	

1

4.1.2.3 Economic evidence

No health economic search was undertaken for this review question and consequently no evidence was found. This question focused on the risk of various developmental disorders rather than whether any strategy for the management of these disorders represents a cost-effective use of resources. Therefore, this question is not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review.

4.1.2.4 Evidence statements

Cerebral Palsy

CP in relation to gestational age

- Evidence from 4 studies showed an increase in the risk of cerebral palsy for preterm infants.
- Moderate quality evidence from 1 study (n=141321) showed a significant increase in the risk of cerebral palsy for children born preterm (30-33 weeks and 34-36 weeks) as compared to term children, during a follow-up period of up to 5.5 years.
- Moderate quality evidence from 1 study (n=6145357) also showed an increased risk of cerebral palsy for preterm children, regardless of gestation (32-36 weeks, 38-31 weeks and <28 weeks) as compared to those born at term.
- Low quality evidence from 1 study (n=1018302) also showed a significant increase in the risk of cerebral palsy (at the age of 7 years) for preterm infants of <32 weeks, 32 to 33^{+6} weeks, and 34 to 36^{+6} weeks as compared to term babies.
- Similarly, moderate quality evidence from 1 study (n=13843) showed a significant increase in the risk of cerebral palsy (at the age of 7 years) for preterm infants of 32-36 weeks compared to term babies.

CP in relation to biological factors

Moderate quality evidence from 5 studies (sample sizes ranging from 187 to 53078) showed mixed results on the association between being born SGA and CP.

Moderate quality evidence from 1 study (n=2971) showed a significant increase in the risk of moderate or severe cerebral palsy for children who were small for gestational age (SGA, versus not SGA) during a follow-up period of 18-22 months corrected age. Moderate quality evidence from 1 study (n=2846) showed that there was no increase in the risk of cerebral palsy in children born SGA (versus appropriate for gestational age) at 24-28 weeks or 29-32 weeks gestational age. Moderate quality evidence from 1 study (n=187) showed that there was no association between being born SGA (versus appropriate for gestational age) and CP among children born preterm at 2 years. Moderate quality evidence from 1 study (n=53078) showed that the risk of cerebral palsy in children born at <32 weeks of gestation who were SGA (versus appropriate for gestational age) was lowered. Among children born at 32-33 weeks, there was no association with SGA and CP, however, among children born at 34-36 weeks, there was an increased risk of CP among preterms born SGA.

Low to moderate quality evidence from 6 studies (sample sizes ranging from 187 to 53078) showed mixed results on the association between sex of the child born preterm and CP.

Moderate quality evidence from 1 study (n=208) showed that there was no significant risk of cerebral palsy in male infants (versus female) assessed at 18-22 months corrected age born at ≥22 weeks gestational age. Low quality evidence from 1 study (n=246) found no association between male sex and risk of CP among children born at <25 weeks of gestation and assessed at 18-22 months corrected age. Moderate quality evidence from 1 study (n=187) showed that there was no significant risk of cerebral palsy for male children (versus female) at follow-up of 2 years. Moderate quality evidence from 1 study (n=252) showed that there was no significant risk of cerebral palsy in males (versus females) born <28 weeks gestational age at follow-up of 5 years. Moderate quality evidence from 1 study (n=2457) showed that there was no increase in risk of cerebral palsy in male children born 30-34 weeks gestational age assessed at 5 years of age. Low quality evidence from 1 study (n=53078) showed that there was a significant increase in the risk of cerebral palsy in males (versus females) who were born at <32 weeks gestational age and assessed at 7 years of age. In the same study, no significant association was found between being male and CP among children born at 32-33 weeks of gestation.

High quality evidence from 1 study (n=375) showed that there was a lowered risk of CP among children of African American origin (versus not African American) among children born between 23 and 32 weeks gestational age followed up at 6 years of age.

CP in relation to neonatal factors

Moderate to high quality evidence from 10 studies (sample sizes ranging from 187 to 6161) largely showed increased risk in CP in children exposed to IVH grade III-IV, severe PIVH, PVL, IVH/shunt, IVH grade III-IV and/or grade II-IV, parenchymal pathology and/or ventriculomegaly, IVH grade III or echodensities or ventricular dilation, cystic PVL or intraparentchymal, intracranial haemorrhage compared with those unexposed to those risk factors. Children in these 11 studies were born at different gestational ages and assessed at age 18 months, 24 months, 18 to 22 months corrected age, 2 years, 30 months, 5 years, 6 years, and 7 years. Only 1 study (n=246) found no significant association between IVH grade III-IVH and CP when children were assessed at 18-22 months corrected age (moderate quality).

Moderate quality evidence from 5 studies (sample sizes ranging from 208 to 6347) showed mixed findings with regard to the association between sepsis and CP.

Two studies showed that preterm children exposed to sepsis were at an increased risk for CP in comparison with those unexposed when assessed at age 18-22 months corrected (moderate quality evidence). However, another 3 studies showed no significant association between the two when preterm children were assessed at age 18 to 22 months corrected, 18 months, and 7 years (moderate quality evidence).

Moderate quality evidence from 2 studies (n=1085; n=283) showed no significant association between ROP and the risk of CP when children were assessed at age 24 months and 30 months. The same non-significant association was found when ROP of different severities (such as ROP threshold, ROP pre-threshold) and the various forms of CP (for example CP quadriparesis, CP diparesis, and CP hemiparesis) were assessed in 1 of the studies (moderate quality evidence).

Low to high quality evidence from 6 studies (sample sizes ranging from 252 to 2948) showed mixed findings regarding the risk of CP in relation to NEC. Four studies found no significant association between NEC and CP when children born preterm were assessed at 18-22 months corrected age, age 18 months, and age 5 and 6 years. However, significantly increased risk in CP among those exposed to NEC compared with those unexposed was found in 2 studies when children were assessed at 18 to 22 months corrected age and 5 years, respectively.

Low to moderate quality evidence from 10 studies (sample sizes ranging from 193 to 6347) reported mixed findings regarding the association between antenatal steroids and CP. Seven studies found no significant association between those exposed to antenatal steroids and CP compared with those unexposed and when children were assessed at age 24 months, 18 to 22 months corrected age; 18 months; 30 months, 5 years, and 7 years. However, moderate to low quality evidence from three studies (sample size ranged from 193 to 1924) showed a significantly reduced risk in CP associated with antenatal steroids when children born at 27.3 (mean) weeks' GA were assessed at age 18 months and 5 years, respectively; and children born at 22-25 weeks' GA were assessed at age 18-22 months corrected.

Moderate quality evidence from 6 studies (sample sizes ranging from 280 to 6347) reported mixed findings. Three studies (n=280; n=672; n=3785) found a significantly increased risk in CP among those exposed to postnatal steroids compared with those unexposed when children were assessed at age 18 -22 months corrected, 24 months, and 5 years. However, nonsignificant association between postnatal steroids and CP was reported in another three studies (n=1472; n=1812; n=283) when children were assessed at age 18-22 months corrected age, 30 months corrected age, and 5 years.

Moderate quality evidence from 4 studies (sample sizes ranging from 246 to 3785) reported mixed findings on the risk of CP in relation to BPD at 36 weeks. No association was found in 4 studies when children born at 22-32 weeks' GA, <28 weeks GA were assessed at age 18-22 months corrected, 24 months corrected, and 5 years. However, in 1 study, when BPD with mechanical ventilation was assessed, no significant association was found between it and CP when children born at <28 weeks GA were assessed at age 24 months corrected.

CP in relation to social, environmental and maternal factors

High quality evidence from 1 study (n=2235) showed no significant impact of chorioamnionitis on the risk of cerebral palsy in a group of very preterm babies (<27 weeks' gestation) at 18-22 months of corrected age. Moderate quality evidence from 1 study (n=283) did not find an association between chorioamnionitis and CP among children born before 26 weeks of gestation and assessed at 30 months corrected age. Low quality evidence from 1 study (n=2202) showed no association between histological chorioamnionitis and cerebral palsy in children born before 34 weeks of gestation at 3 years of age (uncorrected).

Evidence from 3 studies (n=641) showed no impact of socioeconomic status on the risk of cerebral palsy (Shankaran 2004; Tommiska 2003; Toome 2013). The quality of evidence from these studies ranged from low to high.

High quality evidence from 1 study (n=208) showed that maternal age did not affect the risk of cerebral palsy in a group of extremely low birth weight infants assessed at 18 months corrected age. High quality evidence from another study (n=187) showed no significant effect of maternal age on the risk of cerebral palsy at 2 years (corrected age) among children born before 32 weeks of gestation. Low quality evidence from 1 study (n=53078) showed no association between maternal age and CP among children born preterm.

High quality evidence from 1 study (n=208) showed that multiple pregnancy did not significantly affect the risk of cerebral palsy in a group of extremely low birth weight infants assessed at 18 months corrected age. High quality evidence from another study (n=187) showed no significant effect of multiple pregnancy on the risk of cerebral palsy at 2 years (corrected age). Moderate quality evidence from 1 study (n=1461) reported no significant change in the risk of cerebral palsy for multiple pregnancies (as compared to singletons) born at 30-34 weeks. Further analysis of the same cohort included preterm infants from 24-32 weeks (n=812). This also showed no significant change in the risk of cerebral palsy for multiple pregnancies or with maternal age (moderate quality evidence). Low quality evidence from 1 study (n=53078) reported no association between multiple birth and cerebral palsy.

Developmental coordination disorder (DCD)

DCD in relation to gestational age

No evidence was found.

DCD in relation to biological factors

Low quality evidence from 1 study (n=560) showed that an increase in the risk of developmental coordination disorder in male children (versus female) born before 28 weeks of gestation and assessed at 8 to 9 years age.

DCD in relation to neonatal factors

No evidence was found.

DCD in relation to social, environmental and maternal factors

No evidence was found.

Intellectual disability

Intellectual disability in relation to gestational age

Low to high quality evidence from 7 studies (sample sizes ranging from 1157 to 141321) show that children born preterm are at an increased risk of intellectual disability.

Moderate quality evidence from 1 study (n=7500) also showed a significantly increased risk of developmental delay (mild and severe) in children born at 34-36 weeks' gestation as compared to term controls at the age of 2 years. Moderate quality evidence from 1 study (n=1157) also showed a significantly increased risk of mild cognitive impairment, and mild or moderate developmental delay in children born before 27 weeks' gestation as compared to term controls at the age of 2.5 years. High quality evidence from 1 study (n=1854) showed a significant increase in intellectual disability at age 5 years in preterm children born at 22-32 weeks, compared to term controls. Moderate quality evidence from 1 study (n=141321) showed a significantly increased risk of developmental delay in preterm children (30-33 weeks and 34-36 weeks) when compared to term children, up to the age of 5.5 years. Low quality evidence from 1 study (n=85535) showed a significant increase in the risk of intellectual disability in children born preterm (<37 weeks) as compared to term controls when parents were asked if a doctor had ever told that their preterm child (2 to 17 years old) has intellectual disability.

Moderate quality evidence from 1 study (n=1506) showed no significant increased risk of developmental delay (mild or severe) in early preterm children born at 23-24 weeks as compared to children born at 25-26 weeks and assessed at 2 years corrected.

Intellectual disability in relation to biological factors

Moderate quality evidence from 5 studies (sample sizes ranging from 187 to 2846) showed somewhat mixed results on the association between being born SGA and intellectual disability among children born preterm.

Moderate quality evidence from 1 study (n=963) found a significant increase in risk of cognitive impairment (MDI <70) in children born before 27 weeks of gestation who were SGA (versus appropriate for gestational age) at 18-22 months corrected age. Low quality evidence from 1 study (n=1151) did not find an association between SGA (versus appropriate for gestational age) and cognitive impairment in children born before 27 weeks of gestation and assessed at 18-22 months corrected age. High quality evidence from 1 study (n=187) showed that three was no significant increase in the risk of cognitive impairment in children born SGA born at a mean 28.8 weeks gestational age and assessed at 2 years (corrected age). Moderate quality evidence from 1 study (n=1503) found an increased risk of severe cognitive impairment in children born SGA (versus appropriate for gestational age) between 24 to 32 weeks gestational age and assessed at 5 years of age. Moderate quality evidence from 1 study (n=2846) showed that there was no increased risk of cognitive impairment at 5 years in children born SGA at 24-28 weeks gestational age, however, there was a significant increase in the risk of impairment at 29-32 weeks gestational age.

Low to moderate quality evidence from 8 studies (sample sizes ranging from 187 to 14147) showed somewhat mixed findings on the association between the sex of the preterm child and intellectual disability.

Moderate quality evidence from 1 study (n=963) showed that there was no association between male sex and cognitive impairment (MDI <70) in children born before 27 weeks of gestation and assessed at 18-22 months corrected age. High quality evidence from 1 study (n=246) showed that there was no increased risk of cognitive impairment (MDI<70) in male children born before 25 weeks of gestation (versus females) at 18-22 months corrected age. Moderate quality evidence from 1 study (n=14147) showed that there was a significant increase in risk of intellectual disability in male children (versus female) with birth weight of 401-1000 grams (mean gestational age 25.5 weeks) at 18-22 months corrected age. Low quality evidence from 1 study (n=1151) did not find an association between male sex (versus female) and cognitive impairment in children born before 27 weeks of gestation and assessed at 18-22 months corrected age. High quality evidence from 1 study (n=187) showed no significant increase in the risk of cognitive impairment in male children (versus female) born at a mean 28.8 weeks gestational age and assessed at 2 years (corrected age). Moderate quality evidence from 1 study (n=1506) showed that there was a significant increase in the risk of cognitive impairment in male children (versus female) born before 28 weeks of gestation and assessed at 2 years (corrected age). Moderate quality evidence from 1 study (n=1503) found no association between male sex (versus female) and mild or severe cognitive impairment in children born between 24 to 32 weeks gestational age and assessed at 5 years of age. Moderate quality evidence from 1 study (n=252) found no association between male sex (versus female) and cognitive impairment in children born before 28 weeks of gestation assessed at 5 years of age.

Low to moderate quality evidence from 4 studies (sample sizes ranging from 246 to 3790) showed mixed findings on the association between ethnicity and intellectual disability in children born preterm.

Low quality evidence from 1 study (n=246) showed that there was no increased risk of cognitive impairment (MDI<70) in children of black ethnicity (versus non-black) born before 25 weeks of gestation assessed at 18-22 months corrected age. Moderate quality evidence from 1 study (n=3790) showed no significant increase in the risk of cognitive impairment in children of non-white race (versus white) at 18-22 months corrected age. Low quality evidence from 1 study (n=1151) did not find an association between black ethnicity (versus non-black) and cognitive impairment in children born before 27 weeks of gestation and assessed at 18-22 months corrected age. However, moderate quality evidence from 1 study (n=1506) showed that there was a significant increase in the risk of cognitive impairment in children of non-white ethnicity (versus white) born before 28 weeks of gestation and assessed at 2 years (corrected age).

Intellectual disability in relation to neonatal factors

Low to moderate quality evidence from 11 studies (sample sizes ranging from 187 to 6161) largely showed an increased risk in intellectual disability defined in different ways across studies associated with PVL, IVH and infarct. Children in those studies were assessed at age 18 to 22 months corrected, 24 months corrected, 2 years, and 5 years. However, non-significant association was found in 2 studies when children were assessed at age 18-22 months corrected and 5 years.

Moderate quality evidence from 6 studies (sample sizes ranging from 1472 to 6314) reported mixed findings. Three studies found a significantly increased risk in intellectual disabilities associated with sepsis when children were assessed at age 18 to 22 months corrected age.

However, another three studies (sample size ranged from 963 to 3785) reported non-significant association between the two when children assessed also at age 18-22 months corrected.

Moderate quality evidence from 1 study (n=1085) showed mixed results when different degrees of ROP and intellectual disability of different levels were assessed among children aged 24 months. ROP stage 3 showed an increased risk associated with MDI <55 (Bayley). However, when MDI 56-69 was assessed as the outcome, the significantly increased risk associated with ROP was found for ROP zone 1, ROP threshold, and ROP pre-threshold.

Moderate quality evidence from 9 studies (sample sizes ranging from 193 to 6314) reported mixed findings regarding the association between NEC and intellectual disability defined in different methods. Six studies showed an increased risk in MDI < 70 associated with NEC (e.g., NEC surgery, NEC perforation) when children were assessed at age 18 to 22 months corrected, 2 years, 5 years. However, another 3 studies showed non-significant association between the two when children were assessed at age 18 to 22 months corrected, 5 years.

Low to moderate quality evidence from 10 studies (sample sizes ranging from 193 to 4924) showed largely non-significant association between antenatal steroids and intellectual disability measured in different ways when children were assessed at age 18-22 months corrected and 5 years. In 1 study (n=193), antenatal steroids were found to be associated with an IQ score <70 when children were assessed at age 5 years.

Moderate quality evidence from 4 studies (sample sizes ranging from 2901 to 3705) showed mixed results regarding the association between postnatal steroids and intellectual disability. Three studies found an increased risk in MDI < 70 associated with postnatal steroids when children were assessed at age 18 to 22 months corrected. However, 1 study (n=2901) found no significant association between it and severe cognitive deficiency assessed by Kaufman Assessment Battery for Children (K-ABC) scale when children at 5 years were assessed.

Low to moderate quality evidence from 7 studies (sample sizes ranging from 193 to 3785) reported mixed findings. Four studies found a significantly increased risk in intellectual disability associated with BPD at 36 weeks when children were assessed at age 18 to 22 months corrected. However, 3 studies found no significant associations between BPD with or without mechanical ventilation and intellectual disability when children were assessed at age 18 to 22 months corrected, and at age 24 months.

Intellectual disability in relation to social, environmental and maternal factors

High quality evidence from 1 study (n=2235) showed a significant increase in the risk of cognitive impairment at 2 years of age for preterm infants with chorioamnionitis that was diagnosed both clinically and histopathologically. Moderate quality evidence from 1 study (n=456) showed no significant effect of chorioamnionitis on cognitive function at 2.5 years among children born before 27 weeks of gestation. Low quality evidence from another study (n=2202) showed no association between histological chorioamnionitis and cognitive function in children born before 34 weeks of gestation at 3 years of age (uncorrected).

Low quality evidence from 1 study (n=246) showed no association between low socioeconomic status (household income <$20K) and cognitive impairment at 18-22 months corrected age among children born before 25 weeks of gestation. High quality evidence from another study (n=187) showed no significant effect of socioeconomic status on the risk of cognitive impairment at 2 years (corrected age) among children born before 32 weeks of gestation. Moderate quality evidence from 1 study (n=1503) found a significant increase in the risk of mild and severe intellectual disability for preterm infants (24-32 weeks) of families with lower socioeconomic status. Further analysis of the same study (n=1461) also showed a significant increase in moderate/severe cognitive deficiency for moderately preterm infants (30-34 weeks) born to families of lower socioeconomic status.

Moderate quality evidence from 1 study (n=3790) showed no association between maternal age and cognitive impairment at 18-22 months corrected age among children born before 27 weeks of gestation. High quality evidence from 1 study (n=187) showed no significant effect of maternal age on the risk of cognitive impairment at 2 years (corrected age) among children born before 32 weeks of gestation.

Low quality evidence from 1 study (n=82) found that maternal use of cocaine significantly increased the risk of intellectual disability among children born preterm at 3 years of age.

Moderate to high quality evidence from 2 studies (n=643) showed no significant effect of multiple birth on the risk of cognitive impairment at 2 and 2.5 years of age among children born before 27 weeks and before 32 weeks of gestation.

Specific learning impairment

Specific learning impairment in relation to gestational age

Moderate quality evidence from 1 study (n=372) showed a significant increase in the risk of reading impairment and mathematical impairment in early preterm children (<26 weeks) as compared to full term controls, at the age of 11 years.

Specific learning impairment in relation to biological factors

No evidence was found.

Specific learning impairment in relation to neonatal factors

Moderate quality evidence from 1 study (n=161) showed an increased risk in delayed numerical skills associated with ICH of all grades when children born preterm were assessed at age 5 years.

Moderate quality evidence from 1 study (n=161) showed an increased risk in delayed numerical skills associated with BPD at 36 weeks when children born preterm were assessed at age 5 years.

Specific learning impairment in relation to social, environmental and maternal factors

No evidence was found.

Speech or language impairment

Speech or language impairment in relation to gestational age

Low to moderate quality evidence from 3 studies (sample sizes ranging from 468 to 38802) showed mixed results.

Moderate quality evidence from 1 study (n=1157) showed an increase in the risk of mild or moderate language impairment in children born before 27 weeks of gestation as compared to term controls at 2.5 years of age.

Low quality evidence from 1 study (n=38802) showed an increased risk of developmental speech and/or language delay between the ages of 3 and 5 years in children born at 34 to 36 weeks' gestation compared to children born at term.

Low quality evidence from 1 study (n=468) showed no association between being born extremely preterm (<25 weeks) and serious impairment in language abilities at 6 years of age compared to those born at term.

Speech or language impairment in relation to biological factors

High quality evidence from 1 study (n=187) showed that there was a significant increase in the risk of language impairment in male children (compared to female) born at a mean gestationa age of 28.8 weeks at 2 years of age.

Moderate quality evidence from 1 study (n=3790) showed no association between being of non-white ethnic background and language impairment at 18-22 months' corrected age in children born preterm when compared to children born preterm of white ethnicity.

Speech or language impairment in relation to neonatal factors

Moderate quality evidence from 2 studies (n= 1472; n=187) showed an increased risk in speech and language disorders associated with severe PIVH and IVH grade III/IV or PVL grade II-IV when children born pre-term were assessed at age 18-22 months corrected age and 2 years.

Moderate quality evidence from 1 study (n=1472) found an increased risk in speech and language disorders associated with sepsis when children born pre-term were assessed at age 18-22 months corrected age.

Moderate quality evidence from 2 studies (n= 1472; n=1934) found no significant association between antenatal steroids and language disorders when children born pre-term were assessed at age 18-22 months corrected age.

Moderate quality evidence from 2 studies (n= 1472; n=1934) found no significant association between antenatal steroids and language disorders when children born pre-term were assessed at age 18-22 months corrected age.

Speech or language impairment in relation to social, environmental and maternal factors

High quality evidence from 1 study (n=187) showed no significant effect of socioeconomic status on the risk of language impairment at 2 years (corrected age) among children born before 32 weeks of gestation.

Moderate quality evidence from 1 study (n=3790) showed no significant effect of maternal age on the risk of language impairment at 18-22 months corrected age among children born before 27 weeks of gestation. High quality evidence from another study (n=187) showed no significant effect of maternal age on the risk of language impairment at 2 years (corrected age) among children born before 32 weeks of gestation.

High quality evidence from 1 study (n=187) showed no significant effect of multiple pregnancy on the risk of language impairment at 2 years (corrected age) among children born before 32 weeks of gestation.

Mental disorders

Mental disorders in relation to gestational age

Low to moderate quality evidence from 4 studies (sample sizes ranging from 193 to 85535) showed mixed results.

Low quality evidence from 1 study (n=193) showed an increased risk of any anxiety diagnosis at 3 to 6 years of age in children born at 34 to 36 weeks' gestation compared to children born at term. The same study found no association between being born preterm and conduct disorder (including oppositional defiant disorder) or major depressive disorder.

Low quality evidence from 1 study (n=85535) showed an increase in the risk of conduct disorder, anxiety and depression in children born preterm (<37 weeks) as compared to term controls. The outcomes were measured by asking parents of 2 to 17 year-old children born preterm if their doctor had ever told that their child has a particular disorder.

Moderate quality evidence from 1 study (n=371) showed no association between being born before 26 weeks' gestation and major depression, conduct disorder or oppositional defiant disorder at the age of 11 years.

Low quality evidence from 1 study (n=372) showed no association between being born before 28 weeks' gestation and anxiety or mood disorder at the age of 18 years.

Mental disorders in relation to biological factors

No evidence was found.

Mental disorders in relation to neonatal factors

Moderate quality evidence from 1 study (n=307) showed an increased risk in any psychiatric disorder associated with NEC when children born preterm were assessed at age 11 years.

Mental disorders in relation to social, environmental and maternal factors

No evidence was found.

Autism Spectrum Disorder (ASD)

ASD in relation to gestational age

Low to high quality evidence from 2 studies (n=85535; n=195021) showed children born preterm to be at an increased risk of autism spectrum disorder compared to term born children.

High quality evidence from 1 study (n=195021) showed a significant increase in the risk of autism spectrum disorder for preterm children (born at 34-36 weeks', 27-33 weeks' and 24-26 weeks' gestation) as compared to term children, at 2 to 11 years of age.

Low quality evidence from 1 study (n=85535) also showed a significant increase in the risk of autism spectrum disorder in children born preterm (<37 weeks) as compared to term controls when asked from parents if the doctor had ever told that their child born preterm aged 2 to 17 years had ASD.

ASD in relation to biological factors

High quality evidence from 2 studies (n=235198; n=21717) showed mixed findings on the association between being born SGA and ASD.

High quality evidence from 1 study (n=235198) showed that there was a significant increase in the risk of ASD diagnosis in children born preterm who were born small for gestational age compared to children born preterm appropriate for gestational age. High quality evidence from 1 study (n=21717) showed no association between being born SGA and autism among children born preterm (at 23-31 weeks', 32-33 weeks', and 34-36 weeks' gestation) at 11 years of age.

Low quality evidence from 2 studies (n=1078; n=85535) showed an increased risk of ASD in male preterm children compared to female preterm children.

Low quality evidence from 1 study (n=1078) showed that there was a significant increase in the risk of infantile autism among male children born preterm/extremely low birth weight at 8-11 years follow-up compared to their female peers. Low quality evidence from 1 study (n=85535) showed that there was a significant increase in the risk of autism spectrum disorder in males born preterm (compared to females) when asked from parents if the doctor had ever told that their child born preterm aged 2 to 17 years had ASD.

Low quality evidence from 1 study (n=95535) showed mixed results regarding association between ethnicity and ASD in children born preterm. No association was found in Hispanic or non-Hispanic mixed race children compared to non-Hispanic white children. A reduced risk of ASD was reported among non-Hispanic black children compared to non-Hispanic white children. The study measured ASD by asking parents of children born preterm if the doctor had ever told that their child born preterm aged 2 to 17 years had ASD.

ASD in relation to neonatal factors

Moderate quality evidence from 1 study (n=3807) showed an increased risk in autism associated with IVH grade III-IV when children born preterm were assessed at age 2 to 11 years, However no significant association between cystic PVL and autism was found in the same study.

Moderate quality evidence from 1 study (n=3807) showed no significant association between sepsis and autism when children born preterm were assessed at age 2 to 11 years.

Moderate quality evidence from 1 study (n=1078) showed no significant association between BPD at 36 weeks and autism when children born preterm were assessed at age 8 to 11 years.

ASD in relation to social, environmental and maternal factors

No evidence was found.

Attention Deficit Hyperactivity Disorder (ADHD)

ADHD in relation to gestational age

Low to moderate quality evidence from 5 studies (sample sizes ranging from 193 to 85535) showed somewhat mixed results.

Moderate quality evidence from 1 study (n=371) showed a significant increase in the risk of ADHD and ADHD inattentive subtype in children born before 26 weeks' gestation (<26 weeks) at the age of 11 years, as compared to term controls. No significant differences in the risk of ADHD combined type were identified. The difference in ADHD and ADHD inattentive subtype persisted after exclusion of children with neurosensory impairment, but not after additionally excluding those with cognitive impairment.

Low quality evidence from 1 study (n=372) showed a significant increase in the risk of any type of ADHD in early preterm/extremely low birth weight children (<28 weeks) as compared to normal birth weight controls, at the age of 18 years. The same study showed no increase in the risk of combined type of ADHD, inattentive or hyperactive/impulsive subtypes of ADHD.

Low quality evidence from 1 study (n=85535) also showed a significant increase in the risk of ADHD in children born preterm (<37 weeks) as compared to term controls when asked from parents if the doctor had ever told that their child born preterm aged 2 to 17 years had ADHD.

Low quality evidence from 2 studies (n=193; n=38802) showed no association between being born at 34-36 weeks' gestation and ADHD at 3 to 6 years of age.

ADHD in relation to biological factors

Low quality evidence from 1 study (n=85535) showed an increase in the risk of ADHD among male children born preterm (compared to female) when asked from parents if the doctor had ever told that their child born preterm aged 2 to 17 years had ADHD. The same study reported a reduced risk of ADHD, as reported by parents, among children born preterm of Hispanic and non-Hispanic black ethnicity compared to children born preterm of non-Hispanic white ethnicity.

ADHD in relation to neonatal factors

No evidence was found.

ADHD in relation to social, environmental and maternal factors

No evidence was found.

Vision impairment

Vision impairment in relation to gestational age

No evidence was found.

Vision impairment in relation to biological factors

Moderate quality evidence from 1 study (n=297) showed a significant increase in the risk of blindness (<20/200 vision bilaterally) among children born at 23-26 weeks' gestation who were born small for gestational age compared to children of the same gestation age who were born appropriate for gestational age.

Vision impairment in relation to neonatal factors

Moderate quality evidence from 1 study (n=6161) showed an increased risk in blindness associated with IVH grade III/shunt when children born preterm were assessed at age 18-22 months corrected.

Moderate quality evidence from 1 study (n=6161) showed an increased risk in blindness associated with sepsis, meningitis with our without sepsis when children born preterm were assessed at age 18-22 months corrected.

Moderate quality evidence from 1 study (n=193) showed an increased risk in blindness associated with ROP when children born preterm were assessed at age 5 years.

Moderate quality evidence from 1 study (n=6161) showed no significant association between antenatal steroids and blindness when children born preterm were assessed at age 18-22 months corrected.

Vision impairment in relation to social, environmental and maternal factors

Low quality evidence 1 study (n=2202) showed no association between histological chorioamnionitis and visual impairment in children born before 34 weeks of gestation at 3 years of age (uncorrected).

Hearing impairment

Hearing impairment in relation to gestational age

No evidence was found.

Hearing impairment in relation to biological factors

Moderate quality evidence from 1 study (n=2971) showed no association between being born small for gestational age and hearing loss among children born at 23 to 26 weeks' gestation.

Hearing impairment in relation to neonatal factors

Moderate quality evidence from 1 study (n=6161) showed no significant association between IVH grade III/shunt and deafness when children born preterm were assessed at age 18-22 months corrected.

Moderate quality evidence from 1 study (n=6314) showed an increased risk in deafness associated with sepsis when children born preterm were assessed at age 18-22 months corrected. However, the same study showed no significant association between meningitis with our without sepsis and deafness.

Moderate quality evidence from 1 study (n=4924) showed no significant association between antenatal steroids and deafness when children born preterm were assessed at age 18-22 months corrected.

Hearing impairment in relation to social, environmental and maternal factors

Low quality evidence 1 study (n=2202) showed no association between histological chorioamnionitis and severe hearing impairment in children born before 34 weeks of gestation at 3 years of age (uncorrected).

- Composite outcome

Composite outcome in relation to gestational age

High quality evidence from 1 study (n=1473) showed a significant increase in the risk of neurodevelopmental disorder (including 1 or more of the following: developmental delay, cerebral palsy, blindness or deafness) at 2 to 3 years corrected age in children born at 22-26 weeks' gestation when compared with born preterm at 27-28 weeks' gestation.

Composite outcome in relation to biological factors

Moderate to high quality evidence from 3 studies (sample sizes ranging from 187 to 1473) showed mixed results on the association between being born SGA and composite neurodevelopmental or neurosensory outcome in children born preterm.

High quality evidence from 1 study (n=1473) showed an increased risk of moderate to severe functional disability (1 or more of the following: developmental delay, cerebral palsy, bilateral blindness, or bilateral deafness) among SGA children (compared to children born appropriate to gestational age) born before 29 weeks' gestation and assessed at 2-3 years corrected age. Moderate quality evidence from 1 study (n=373) showed no association between being born SGA and major neurosensory disability (1 or more of the following: cerebral palsy, blindness, or complete deafness) at 2 years in children born at 22-27 weeks' gestation.

Low to high quality evidence from 4 studies (sample sizes ranging from 246 to 3041) showed mixed findings on the association between the sex of the child and composite neurodevelopmental or neurosensory outcome in children born preterm.

High quality evidence from 1 study (n=1473) showed an increased risk of moderate to severe functional disability (1 or more of the following: developmental delay, cerebral palsy, bilateral blindness, or bilateral deafness) among males (compared to females) born before 29 weeks' gestation and assessed at 2-3 years corrected age.

Moderate quality evidence from 1 study (n=3041) showed an increased risk of neurodevelopmental disability (1 or more of the following: mental developmental index score or physomotor developmental index score < 70, moderate or severe cerebral palsy, bilateral blindness, or deafness) among males (compared to females) born at a mean gestational age of 25.8 weeks and assessed at 18 to 22 months corrected age.

Moderate quality evidence from 1 study (n=373) showed no association between the sex of the child and major neurosensory disability (1 or more of the following: cerebral palsy, blindness, or complete deafness) at 2 years in children born at 22-27 weeks' gestation.

Low quality evidence from 1 study (n=246) showed no association between the sex of the child and neurodevelopmental impairment (1 or more of the following: cerebral palsy, mental developmental index score or psychomotor developmental index score < 70, bilateral blindness, or hearing impaired with amplification) at 18 to 22 months corrected age in children born before 25 weeks' gestation.

Low to moderate quality evidence from 2 studies (n=246; 2=3041) showed mixed findings on the association between ethnicity and composite neurodevelopmental outcome in children born preterm.

Moderate quality evidence from 1 study (n=3041) showed an increased risk of neurodevelopmental disability (1 or more of the following: mental developmental index score or physomotor developmental index score < 70, moderate or severe cerebral palsy, bilateral blindness, or deafness) among children of non-white ethnicity (compared to children of white ethnicity) born at a mean gestational age of 25.8 weeks and assessed at 18 to 22 months corrected age.

Low quality evidence from 1 study (n=246) showed no association between ethnicity and neurodevelopmental impairment (1 or more of the following: cerebral palsy, mental developmental index score or psychomotor developmental index score < 70, bilateral blindness, or hearing impaired with amplification) at 18 to 22 months corrected age in children born before 25 weeks' gestation.

Composite outcome in relation to neonatal factors

Moderate quality evidence from 11 studies (sample sizes ranging from 166 to 6161) showed largely increased risk in neurodevelopmental impairment or neurosensory impairment associated with IVH grade III, IVH grade IV, IVH grade III-IV, severe PIVH, cystic PVL, IVH III/shunt, severe cerebral lesions when children born preterm were assessed at age 18-22 months corrected, 22-30 months, 2 years, and 2-3 corrected year.

Moderate quality evidence from 6 studies (sample sizes ranging from 166 to 6314) reported mixed findings. Three studies showed an increased risk in neurodevelopmental/neurosensory impairment associated with sepsis when children were assessed at 18-22 months corrected age. However, 3 studies found no significant difference between those exposed to sepsis and those who were not when children were assessed at 18-22 months corrected age and 2 years.

Moderate quality evidence from 3 studies (sample sizes ranging from 79 to 1472) showed a borderline increased or increased risk in neurodevelopmental impairment and or neurosensory impairment associated with ROP when children born preterm were assessed at age 2 years, 2 to 3 corrected year, and 7 to 10 years.

Moderate quality evidence from 7 studies reported mixed findings regarding the relationship between NEC and composite outcomes either measured as neurodevelopmental impairment or neurosensory impairment. Five studies showed an increased risk in neurodevelopmental impairment or neurosensory impairment when children were assessed age 18 to 22 months corrected, 2 years, and 7 to 10 years, however 3 studies showed no significant associations when children were assessed at age 18-22 months corrected, 2 years, and 2-3 corrected years.

Low to moderate quality evidence from 8 studies (sample size ranging from 246 to 4924) showed no significant association between antenatal steroids and composite outcomes either measured as neurodevelopmental impairment or neurosensory impairment. This was the same when children were assessed at age 18-22 months corrected, 2 years, 2.5 corrected years, and 2-3 years,

Moderate quality evidence from 6 studies (sample sizes ranging from 166 to 3041) reported mixed findings regarding the relationship between postnatal steroids and neurodevelopmental impairment or neurosensory impairment. Four studies showed an increased risk in the composite outcomes associated with postnatal steroids when children were assessed at age 18-22 months corrected and 2 years. However 2 studies found no significant association between the two when children were assessed at age 18-22 months corrected and 2 years as well.

Low to moderate quality evidence from 4 studies (sample sizes ranging from 246 to 3785) reported mixed findings. Three studies found no significant association between BPD and neurodevelopment impairment or neurosensory impairment when children born preterm were assessed at age 18-22 months corrected, and 2 years. However, a significantly increased risk in neurodevelopmental impairment associated with BPD was found in 1 study when children born at 22-32 weeks' GA were assessed at age 18-22 months corrected.

Composite outcome in relation to social, environmental and maternal factors

High quality evidence from 1 study (n=2235) showed no impact of histological chorioamnionitis on the risk of a composite outcome measure of neurodevelopmental impairment (including CP, deafness, blindness and cognitive delay) at 18-22 months corrected age among children born before 27 weeks of gestation. This study also showed that infants with both clinical and histological chorioamnionitis also had no increase in the risk of neurodevelopmental impairment. Moderate quality evidence from 1 study (n=456) showed no significant effect of chorioamnionitis (including prolonged and premature rupture of membranes) on the risk of a neurosensory impairment (1 or more of the following: CP, moderate/severe visual impairment, or hearing impairment). However, moderate quality evidence from 1 study (n=373) showed a significant increase in the risk of major neurosensory disability (1 or more of the following: CP, blindness, or deafness) at 2 years of age in children born between 22-27 weeks of gestation with chorioamnionitis compared to those without.

Low quality evidence from 1 study (n=246) showed no significant association between low socioeconomic status (household income <$20K) and composite neurodevelopmental impairment outcome at 18-22 months corrected age among children born before 25 weeks of gestation. High quality evidence from 1 study (n=187) also showed no significant risk of neurodevelopmental impairment at 2 years (corrected age) among children born before 32 weeks of gestation from low income households (versus non-low income).

High quality evidence from 1 study (n=187) showed no effect of maternal age on the risk of neurodevelopmental impairment at 2 years (corrected age) among children born before 32 weeks of gestation.

Moderate quality evidence from 2 studies (n=829) showed no significant effect of multiple birth on the risk of neurosensory impairment (1 or more of the following: CP, moderate/severe visual, or hearing impairment) at 2 and 2.5 years corrected age among children born between 22-27 weeks of gestation. High quality evidence from 1 study (n=187) also showed no significant risk of neurodevelopmental impairment (1 or more of the following: intellectual disability, cerebral palsy, hearing impairment, or visual impairment) at 2 years (corrected age) for multiple pregnancies as compared to singletons among children born before 32 weeks of gestation.

4.1.2.5 Economic evidence

No health economic search was undertaken for this review question and consequently no evidence was found. This question focused on the risk of various developmental problems rather than whether any strategy for the management of these problems represents a cost-effective use of resources. Therefore, this question is not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review

4.1.2.6 Evidence statements

4.1.2.6.1 Prevalence of cerebral palsy

Cerebral palsy ≤ 28 completed weeks of gestation

Any cerebral palsy

Moderate to low quality evidence from four studies (sample size ranged from 141 to 373) showed that among children born at 22-27 weeks GA the prevalence of any CP varied from 7% (95% CI 4.6 to 10.10) to 11.3% (95%CI: 6.6 to 17.8) at 2 years (corrected age), 5 years and 8 years (corrected) (Leversen 2010; Leversen 2011; Roberts 2011; Anderson 2011).

Moderate quality evidence from four studies (sample size ranged from 75 to 244) showed that among children born at <27 weeks GA the prevalence of any CP varied from 14.7% (95%CI 7.6 to 24.7% to 24.7% (95%CI 15.6 to 35.8%) at age range 12 months CA to 9 years (Mikkola 2005; Stahlmann 2009; Sutton 1999; De Groote 2007).

Moderate to low quality evidence from four studies (sample size ranged from 275 to 331,154) showed that among children born at <28 weeks GA the prevalence of any CP varied from 6.7% (95%CI 5.1 to 8.6) to 16.6% (95%CI 12.5 to 21.3) (Larroque 2008; Ancel 2006; Glinianaia 2011; Anderson 2003).

Moderate quality evidence from one study (sample size 1718) showed that among children born at 24-27 weeks GA the prevalence of any CP was 14.7% (95%CI 10.6-19.5%) at 5 years age (Foix-Helias 2008).

Low quality evidence from one study (sample size 104) showed that among children born at 22-26 weeks GA the prevalence of any CP was 11.5% (95% CI 6.1-19.3%) at 18 months CA (Tommiska 2003)

Low quality evidence from one study (sample size 283) showed that among children born 22-25 weeks GA the prevalence of any CP was 17.7% (95% CI 13.4-22.6%) at a median age of 30 months (Wood 2000).

Moderate to very low quality evidence from three studies (sample size ranged from 19 to 189) showed that among children born at a mean GA range of 25.4 (±1) to 26.5 (±2) weeks the prevalence of any CP was 7.3% (95% CI 3.8-12.4%) to 37% (95%CI 16-62%) at age 2 years to 8 years (Hutchinson 2013; Doyle 2011; Rieger-Fackeldey 2010).

Low quality evidence from one study (sample size 219) showed that among children born at 23-27 weeks GA the prevalence of any CP was 11% (95%CI 7.2-15.9%) at 2 years age (Anon 1997).

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of CP was 19.0% (95%CI 12.9 to 26.5%) at 4 years age (Salakorpi 2001).

Mild cerebral palsy

Moderate to low quality evidence from two studies (sample size 77 to 456) showed that among children born at <27 weeks GA the prevalence of mild CP across the two studies (10.4% (95%CI 4.6 to 19.5) and 2.9% (95% CI 1.5 to 4.8)) at 2.5 years CA and 3 years age (De Groote 2007; Serenius 2013).

Moderate cerebral palsy

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks the prevalence was 7.1% (95%CI 4.2 to 11.1) at 6 years (Marlow 2005). The prevalence was varied in two studies of moderate to low quality in children (sample size 456 to 576) born at <27 weeks GA (2.6% (95%CI 1.5 to 4.3)) and 2.9% (95%CI 1.5 to 4.8)) (Moore 2012; Serenius 2013), whereas prevalence of CP was 13% (95% CI 6.4 to 22.6) in one study (at <27 weeks GA) (De Groote 2007).

Moderate to severe cerebral palsy

Moderate to low quality evidence from two studies (sample size 88 to 241) showed that among children born at <26 weeks GA the prevalence of CP (moderate/disabling or both ambulatory/non-ambulatory) was varied, with a prevalence of 6.8% (95%CI 2.5 to 14.3) at 11 years (Farooqi 2011) and 13.3% (95%CI 9.3 to 18.2) at 6 years (Marlow 2005). There was also variation of prevalence of moderate to severe CP in children born at <27 weeks GA at 2.5 years corrected age (4.2% (95%CI 2.5 to 6.4)) and at 3 years (7.8% (95%CI 5.8 to 10.3)) in two studies of moderate and low quality (Serenius 2013; Moore 2012).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-26 weeks GA the prevalence of moderate to severe CP (non-ambulatory or needing aids) was 11% (95%CI 9.8 to 12.4) at 18-22 months corrected age (Vohr 2005).

Severe cerebral palsy

Moderate to low quality evidence from two studies (sample size ranging from 77 to 456) showed that among children born at <27 weeks GA the prevalence of severe CP was 1.3% (95%CI 0.03 to 7%) at age 2.5 years CA to 3 years (Serenius 2013; De Groote 2007).

Moderate quality evidence from two studies (sample size ranging from 75 to 241) showed that among children born at <26 weeks and <27 weeks GA the prevalence of non-ambulatory CP was 6.2% (95%CI 3.5 to 7.4%) at 6 years age (Marlow 2005), and 10.7% (95%CI 4.7 to 19.9%) at 7-9 years age (Stahlmann 2009).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence for severe CP (GMFCS level 3-5) was 5.2% (95% CI 3.5-7.4%) at 3 years age (Moore 2012). Moderate quality evidence from one study (sample size 306) showed that among children born at 22-27 weeks GA the prevalence for severe CP (GMFCS level4-5) was 3.3% (95%CI 1.6-5.9%) at 5 years age (Leversen 2011).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe diplegia was 4.2 % (95%CI 2.2 to 7.3), severe hemiplegia was 0.4% (95%CI 0.01 to 2), and severe quadriplegia was 3.9% (95%CI 2 to 6.9) at 30 months corrected age (Wood 2000).

- Low quality evidence from one study (sample size 1718) showed that among children born at 24-27 weeks GA the prevalence for severe CP (unable to walk or only with aids) was 4.9% (95% CI 2.6 to 8.2%) at 5 years age (Foix-Helias 2008).
- Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence for severe motor impairment (GMFCS level 5, no self -mobility) was 1.9% (95%CI 1.1-3.1) at 10 years age (Joseph 2016b).

28-31 completed weeks of gestation

Any cerebral palsy

Moderate to low quality evidence from three studies (sample size varied from 1812 to 331,154) showed that among children born at 28-31 weeks the prevalence of any CP was varied, ranging from 5.9% (95%CI 4.9 to 7) to 9.5% (95%CI 7.8 to 11.4) across the three studies at 2-8 years (Larroque 2008; Ancel 2006; Glinianaia 2011).

Moderate to low quality evidence from two studies (sample ranged from 1455 to 1781) showed that among children born at 28-32 or 30-31 weeks, there was no difference in prevalence (7.7% (95%CI 5.8 to 9.9) and 7.9% (95%CI 6.6 to 9.3)) at 5 years (Marret 2007; Foix-Helias 2008). However, moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of CP was higher (11.6% (95%CI 10 to 13.3) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-32 weeks GA the prevalence of CP was 16% (95%CI 14.9 to 17.2) at 18-22 months corrected age (Vohr 2005). However, the prevalence of CP was lower (4.3% (95%CI 2.2 to 7.5)) in low quality evidence from one study (sample size 259) among children born at 23-32 weeks GA (Andrews 2008). There was minimal difference in prevalence in GA groups including 24-32 weeks (prevalence 8.9% (95%CI 7.6 to 10.3)) (Foix-Helias 2008), 25-32 weeks GA (prevalence 13.2 (95%CI 8.4 to 19.3)) (Burguet 1999), or <31(prevalence 16% (95%CI 14.9 to 17.2)) , <32 (prevalence 11% (95%CI 6.5 to 17), or <33 weeks GA (prevalence 8.8% (95%CI 7.5 to 10.2)) (Vincer 2014; Toome 2012; Larroque 2008).

Mild cerebral palsy

Low quality evidence from one study (sample size 801) showed that among children born at <31 weeks GA the prevalence of 6.7% (95%CI 5.1 to 8.7) for mild CP (GMFCS level1) at 12-42 months corrected age (Vincer 2014).

Moderate to severe cerebral palsy

Low quality evidence from one (sample size 801) showed that among children born at <31 weeks GA the prevalence of moderate to severe CP (GMFCS level 2-5) was 3.4% (95%CI 2.2-4.9%) at 12-42 months corrected age (Vincer 2014).

Low quality evidence from one study (sample size 155) showed that among children born at <32 weeks GA the prevalence of moderate to severe CP (GMFCS level 2-5) was 8.4% (95%CI 4.5-13.9%) at 2 years CA (Toome 2012).

Low quality evidence from one study (1455) showed that among children born at 30-31 weeks GA the prevalence of 5.7% (95%CI 4.1 to 7.7) for moderate to severe CP (bilateral spastic CP) at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence for moderate to severe CP (non-ambulatory or needing aids) was 7% (95%CI 5.8 to 8.4) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-32 weeks GA the prevalence of moderate to severe CP (non-ambulatory or needing aids) was 9.4% (95%CI 8.5-10.4%) at 18-22 months corrected age (Vohr 2005).

- Severe cerebral palsy

Moderate quality evidence from one study (sample size 1781) showed that among children born at 28-32 weeks GA the prevalence of severe CP (unable to walk or only with aids) was 2.4% (95%CI 1.7 to 3.4) at 5 years (Foix-Helias 2008). In the same study, the prevalence at 24-32 weeks was 2.8% (95%CI 2.1 to 3.7).

32-36 completed weeks of gestation

Any cerebral palsy

Low quality evidence from one study (sample size 1455) showed that among children born at 32-34 weeks GA the prevalence of any CP type was 3.4% (95%CI 2.3 to 5) at 5 years (Marret 2007).

Moderate to low quality evidence from three studies (sample size range from 741 to 331,154)) showed that among children born at 32-26 weeks GA the prevalence of any CP was similar (range from 0.8% (95%CI 0.7 to 0.9) to 1% (95%CI 0.8 to 1.1) across the studies at age up to 7 or 8 years (Odd 2013; Hirvonen 2014; Glinianaia 2011).

Moderate to severe cerebral palsy

Low quality evidence from one study (sample size 1455) showed that among children born at 32-34 weeks GA the prevalence of CP (bilateral spastic CP) was 2.2% (95% CI 1.3 to 3.5) at 5 years (Marret 2007).

Moderate quality from one study (sample size 53,078) showed that among children born at 32-36 weeks GA found the prevalence of CP (other types) was 0.35% (95%CI 0.3 to 0.4) at up to 7 years (Hirvonen 2014).

Low quality evidence from one study (sample size 331,154) showed that among children born at <37 weeks GA the prevalence of spastic-bilateral or unilateral CP was 1.3% (95%CI 1.1 to 1.5) and 0.4% (95%CI 0.3 to 0.5) respectively at up to 8 years (Glinianaia 2011).

Low quality evidence from one study (sample size 104) showed that among children born at 22.3-34.9 weeks GA/bw <1000g the prevalence of CP (ataxia/athetosis) was 1% (95%CI 0.1 to 3.4) at 18 months corrected age (Tommiska 2003).

4.1.2.6.2 *Small for gestational age*

Low quality evidence from one study (sample size 2357) showed that among children born at 24-28 weeks GA and small for gestational age, the prevalence of any CP was 18% (95%CI 5.2-40.3%). In the same study, the prevalence was 3.2% (95%CI 0.9-8%) at 5 years age (Guellec 2011).

4.1.2.6.3 *Hemiplegia*

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of hemiplegia was 1.8% (95%CI 0.6-4.1%) at median 30 months (Wood 2000). In the same study, the prevalence of severe hemiplegia was 0.4% (95%CI 0.01-2%).

Very low quality evidence from one study (sample size 167) showed that among children born at 25-32 weeks GA the prevalence of hemiplegia was 1.2% (95%CI 0.2-4.3%) at 2 years (corrected age) (Burguet 1999).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of hemiplegia was 3.9% (95%CI 0.8-11%) at 3 years age (De Groote 2007).

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of hemiplegia was 5.6% (95%CI 2.5 to 10.8%) at 4 years (Salakorpi 2001).

Low quality evidence from one study (sample size 1455) showed that among children born at gestational age ranging from 30 to 33 weeks the prevalence of hemiplegia ranged from 0.4% to 0.8% (95%CI range 0.01 – 4.1%) at 5 years age (Marret 2007).

Moderate quality evidence from one study (sample size 53,078) showed that among children born at <32 weeks GA the prevalence of hemiplegia was 1.3 % (95%CI 1-1.6%) at age up to 7 years (Hirvonen 2014). In the same study the prevalence of hemiplegia CP was 0.5% (95%CI 0.4-0.8%) at 32-33 weeks GA, 0.14% (95%CI 0.11-0.19%) at 34-36 weeks GA, and 0.2% (95%CI 0.16-0.25%) at 32-26 weeks GA (Hirvonen 2014).

4.1.2.6.4 *Diplegia*

Low quality evidence from one study (sample size 104) showed that among children born at 22.3 to 34.9 weeks GA the prevalence of diplegia was 7.2% (95%CI 4.1-11.6%) at 18 months corrected age (Tommiska 2003).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of diplegia was 9.5% (95%CI 6.4-13.6 %) at median 30 months (Wood 2000). In the same study, the prevalence of severe diplegia was 4.2% (95%CI 2.2-7.3%).

Very low quality evidence from one study (sample size 167) showed that among children born at 25-32 weeks GA the prevalence of spastic diplegia was 6% (95%CI 2.9-10.7%) at 2 years (corrected age) (Burguet 1999).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of diparesis was 11.7% (95%CI 5.5-21%) at 3 years age (De Groote 2007).

Low quality evidence from one study (sample size 155) showed that among children born at <32 weeks GA the prevalence of spastic diplegia was 4.5% (95%CI1.8-9.1%) at 2 years (corrected age) (Toome 2012).

Moderate quality evidence from one study (sample size 53,078) showed that among children born at <32 weeks GA the prevalence of diplegia was 3.4 % (95%CI 2.9-3.8%) at age up to 7 years (Hirvonen 2014). In the same study the prevalence of diplegia CP was 0.7% (95%CI 0.5-0.9%) at 32-33 weeks GA, 0.13% (95%CI 0.10-0.17%) at 34-36 weeks GA, and 0.2% (95%CI 0.17-0.26%) at 32-26 weeks GA (Hirvonen 2014).

4.1.2.6.5 *Triplegia*

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of triparesis was 2.6% (95%CI 0.3-9.1%) at 3 years age (De Groote 2007).

4.1.2.6.6 Diplegia or tetraplegia

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of bilateral spastic CP (diplegia or tetraplegia) was 10.6% (6.0 to 16.8%) at 4 years (Salakorpi 2001).

4.1.2.6.7 Tetraplegia

Low quality evidence from one study (sample size 104) showed that among children born at 22.3 to 34.9 weeks GA the prevalence of tetraplegia was 1.9% (95%CI 0.5-4.9%) at 18 months corrected age (Tommiska 2003).

Very low quality evidence from one study (sample size 167) showed that among children born at 25-32 weeks GA the prevalence of tetraplegia was 1.2% (95%CI 0.2-4.3%) at 2 years (corrected age) (Burguet 1999).

4.1.2.6.8 Quadriplegia

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of quadriplegia was 4.2% (95%CI 2.2-7.3 %) at median 30 months (Wood 2000). In the same study, the prevalence of severe quadriplegia was 3.9% (95%CI 2.0-6.9%).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of quadriplegia was 5.2% (95%CI 1.4-12.8%) at 3 years age (De Groote 2007).

Moderate quality evidence from one study (sample size 53,078) showed that among children born at <32 weeks GA the prevalence of quadriplegia was 0.6 % (95%CI 0.4-0.8%) at age up to 7 years (Hirvonen 2014). In the same study the prevalence of quadriplegia was 0.2% (95%CI 0.1-0.3%) at 32-33 weeks GA, 0.04% (95%CI 0.02-0.06%) at 34-36 weeks GA, and 0.06% (95%CI 0.04-0.08%) at 32-26 weeks GA (Hirvonen 2014).

4.1.2.6.9 Dystonic or athetoid type

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of dystonic or athetoid CP was 2.8% (95%CI 0.8 to 7.1%) at 4 years (Salakorpi 2001).

4.1.2.6.10 Prevalence of cerebral palsy by week of gestational age

Any cerebral palsy

Low quality evidence from one study (sample size 244) showed that among children born at 23 weeks GA the prevalence of any CP was 100% (95%CI 25 to 100%) at 12 months corrected age. However, the prevalence was 19.10% (95%CI 12 to 27.9%) for children who were born at 27 weeks GA (Sutton 1999).

Low quality evidence from one study (sample size 104) showed that among children born at 22-23 weeks GA the prevalence of any CP was 20% (95%CI 0.5 to 71.6%) compared to a prevalence of 10.6% (95%CI 3.6 to 23.10%) in children who were born at 26 weeks GA, assessed at the age of 18 months corrected age (Tommiska 2003).

Low quality evidence from one study (sample size 1954) showed that among children born at 24-25 weeks GA the prevalence of any CP was 19.4% (95%CI 10.4 to 31.4%) compared to a prevalence of 4.4% (95%CI 2.9 to 6.6%) in children who were born at 32 weeks GA, assessed at the age of 2 years (Ancel 2006).

Moderate quality evidence from one study (sample size 1812) showed that among children born at 24-25 weeks GA the prevalence of any CP was 18.3% (95%CI 9.5 to 30.4%)

compared to a prevalence of 4.1% (95%CI 2.6 to 6.2%) in children who were born at 32 weeks GA, assessed at the age of 5 years (Larroque 2008).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of any CP was 6.3% (95%CI 3.8 to 9.7%) compared to a prevalence of 3.7% (95%CI 1.2 to 8.4%) in children who were born at 34 weeks GA, assessed at the age of 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 6347) showed that among children born at <32 weeks GA the prevalence of any CP was 8.7% (95%CI 8.0 to 9.4%) compared to a prevalence of 0.56% (95%CI 0.49 to 0.64%) in children born at 34-36 weeks GA, assessed at up to the age of 7 years (Hirvonen 2014).

Moderate cerebral palsy

Low quality evidence from one study (sample size 576) showed that among children born at 24 weeks the prevalence of moderate CP was 4.1% (95%CI 1.1 to 10.1%) compared to a prevalence of 2% (95%CI 0.7 to 4.6%) in children who were born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks the prevalence of moderate CP was 12.5% (95%CI 2.7 to 32.4%) compared to a prevalence of 5.6% (95%CI 2.4 to 10.7%) in children who were born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate CP was 4.6% (95%CI 1.3 to 11.4%) compared to a prevalence of 2.0% (95%CI 0.4 to 5.7%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate to severe cerebral palsy

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks the prevalence of moderate to severe CP (GMFCS 2-5) was 10.5% (95%CI 2.9 to 24.8%) compared to a prevalence of 6.4% (95%CI 3.7 to 10.2%) in children who were born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe CP (bilateral spastic CP) was 4.2% (95%CI 2.2 to 7.2%) compared to a prevalence of 1.5% (95%CI 0.2 to 5.3%) in children who were born at 34 weeks GA, assessed at the age of 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks the prevalence of moderate to severe CP (ambulatory or non-ambulatory) was 16.7% (95%CI 4.7 to 37.4%) compared to a prevalence of 9.7% (95%CI 5.4 to 15.8%) in children who were born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate quality evidence from one study (sample size 6347) showed that among children born at <32 weeks GA the prevalence of moderate to severe CP (other types) was 3.5% (95%CI 3.0 to 4.0%) compared to a prevalence of 0.25% (95%CI 0.2 to 0.3%) in children born at 34-36 weeks GA, assessed at up to the age of 7 years (Hirvonen 2014).

Severe cerebral palsy

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks the prevalence of severe CP (GMFCS 3-5) was 10.5% (95%CI 2.9 to 24.8%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children who were born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of severe CP (GMFCS 4-5) was 9.2% (95%CI 4.1 to 17.3%) compared to a prevalence of 1.3% (95%CI 0.2 to 4.7%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate quality evidence from one study (sample size 1455) showed that among children born at ≤23 weeks the prevalence of severe CP (non-ambulatory) was 4.2% (95%CI 0.1 to 21.1%) compared to a prevalence of 4.2% (95%CI 1.5 to 8.9%) in children who were born at 25 weeks GA, assessed at 6 years age. The prevalence among children born at 24 weeks was higher (11% (95%CI 4.9 to 20.5%) (Marlow 2005).

Prevalence of cerebral palsy using per 1000 or 10,000 live births as denominator

Any cerebral palsy (<28 weeks GA)

Low quality evidence from one study (sample size 2858) showed that among children born at <28 weeks GA the rate of any CP was 112.7 per 1000 survivors (95%CI 50 to 210) (Drummond 2002).

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at <28 weeks GA the rate of any CP was 71.4 per 1000 livebirths (95%CI 42 to 112 per 1000 live births) at 4 to 8 years age (Himmelmann 2014).

Low quality evidence from one study (sample size 46) showed that among children born at <28 weeks GA the rate of any CP was 72.3 per 1000 live births (95%CI 39 to 120.3 per 1000 live births) at age 4-7 years (Nordmark 2001).

Moderate quality evidence from one study (sample size 975) showed that among children born at <28 weeks GA the rate of any CP in 1992-1994 was 131 per 1000 live births (95% CI 90-183/1000 live births) at age 2 years (confirmed at 3 years age) (Robertson 2007). In the same study, the rate of any CP decreased with the time points (years). From 1995-1997 and 1998-2000, the rate was 69 per 1000 live births (95%CI 41 to 108 per 1000 live births). From 2001-2003 the rate was 19 per 1000 live births (95%CI 6 to 44 per 1000 live births). Over the whole 11 years of the study, the rate was 70 per 1000 live births (95%CI 55 to 88 per 1000 live births) at 2 years age (Robertson 2007).

Severe cerebral palsy (<28 weeks GA)

Moderate quality evidence from one study (sample size 975) showed that among children born at <28 weeks GA the rate of non-ambulatory CP in 1992-1994 was 59 per 1000 live births (95% CI 32-99 per 1000 live births) at age 2 years (confirmed at 3 years age) (Robertson 2007). In the same study, the rate of any CP decreased with the time points in years. From 1995-1997 the rate was 16 per 1000 livebirths (95%CI 5-41 per 1000 livebirths) and from 1998-2000, the rate was 8 per 1000 live births (95%CI 1 to 29 per 1000 live births). From 2001-2003 the rate was 8 per 1000 live births (95%CI 1 to 27 per 1000 live births). Over the whole 11 years of the study, the rate was 22 per 1000 live births (95%CI 13 to 33 per 1000 live births) at 2 years age (Robertson 2007).

Any cerebral palsy (28-32 weeks GA)

Low quality evidence from one study (sample size 2858) showed that among children born at 28-32 weeks GA the rate of any CP was 56.3 per 1000 neonatal survivors (95%CI 33 to 90) (Drummond 2002).

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at 28-32 weeks GA the rate of any CP was 39.6 per 1000 livebirths (95%CI 25 to 59 per 1000 live births) at 4 to 8 years age (Himmelmann 2014).

Low quality evidence from one study (sample size 46) showed that among children born at 28-31 weeks GA the rate of any CP was 32.2 per 1000 live births (95%CI 18.1 to 52.2 per 1000 live births) at age 4-7 years (Nordmark 2001).

Any cerebral palsy (32-36 weeks GA)

Low quality evidence from one study (sample size 189) showed that among children (1991-1996 cohort in Norway) born at 33-36 weeks GA the rate of any CP was 13.8 per 1000 livebirths at earliest age of 4 years (Andersen 2011). In the same study the prevalence of any CP among children (1991-1998 cohort in Italy) was 8.8 per 1000 livebirths whereas in cohorts from Spain and Ireland the rate was 4 per 1000 livebirths (Andersen 2011).

Low quality evidence from one study (sample size 2858) showed that among children born at 32-36 weeks GA the rate of any CP was 9.6 per 1000 survivors (95%CI 6 to 14) (Drummond 2002).

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at 32-36 weeks GA the rate of any CP was 6.4 per 1000 livebirths (95%CI 4 to 9 per 1000 live births) at 4 to 8 years age (Himmelmann 2014). For children born at <37 weeks GA, the rate of any CP was 13 per 1000 live births (95%CI 10 to 16 per 1000 live births).

Low quality evidence from one study (sample size 46) showed that among children born at 32-36 weeks GA the rate of any CP was 4.6 per 1000 live births (95%CI 2.7 to 7.3 per 1000 live births) at age 4-7 years (Nordmark 2001).

4.1.2.6.11 *Diplegia or tetraplegia*

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at <37 weeks GA the rate of bilateral spastic CP was 7.5 per 1000 livebirths (95%CI 5 to 10 per 1000 live births) at 4 to 8 years age (Himmelmann 2014).

4.1.2.6.12 *Prevalence of intellectual disability*

Less than or equal to 28 completed weeks of gestation

Moderate intellectual disability

Moderate to low quality from 4 studies (sample size range from 165 to 576) showed that among children born at a range of 23 to 27 weeks GA the prevalence of intellectual disability (BSIDIII -2SD to -3SD) ranged from 6.4 (95%CI 4.6 to 8.8) to 24% (95%CI 20 to 29) (Doyle 2011; Moore 2012; Anon 1997; Serenius 2013). One further low quality study (sample size 77) used the Dutch version of BSIDII, which showed that the prevalence of intellectual disability was 10.4% (95%CI 4.6 to 19.5) (MDI 55-69) (De Groote 2007).

Moderate quality evidence from two studies (sample size range from 75 to 1508) showed that among children born at 24-27 weeks GA or <27 weeks GA the prevalence of intellectual disability (K-ABC 55-69) was 14.9% (95%CI 10.5 to 20.2) and 10.7% (95%CI 4.7 to 19.9) at 5 years and 7-9 years respectively (Foix-Helias 2008; Stahlmann 2009).

Moderate quality from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of intellectual disability (IQ -2 to -3SD on K-ABC, GMDS or NEPSY) was 19.9% (95%CI 15.1 to 25.5) at 6 years (Marlow 2005).

Moderate quality evidence from one study (sample size 306) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (Full scale IQ WPPSI-R 55-70) was 4.9% (95%CI 2.8 to 8) at 5 years (Leversen 2011). Low quality evidence from one study (sample size 141) showed that the prevalence (WISC-IV -2SD to -3SD) was 8.5%

(95%CI 4.4 to 14.1) in children born in the same gestational age range but assessed at 8 years (Roberts 2010).

Moderate to severe intellectual disability

Moderate to low quality evidence from 5 studies (sample size ranged from 19 to 1508) showed that among children born at GA range 24 to 28 weeks GA the prevalence of intellectual disability (MPC <70 or IQ <70 K-ABC) ranged from 17.6% (95%CI 12.8 to 23.2) to 41% (95%CI 18 to 67) at a range of 5-9 years (Beaino 2011; Foix-Helias 2008; Larroque 2008; Rieger-Fackeldey 2010; Stahlmann 2009).

Moderate to low quality evidence from 5 studies (sample size ranged from 77 to 3785) showed that among children born at a GA range 22-27 weeks GA the prevalence of intellectual disability (BSID <-2SD or MDI <70) ranged from 15.2%(95%CI 10.1 to 21.6) to 39% (95%CI 37 to 41) at 18-36 months (Doyle 2011; Moore 2012; Anon 1997; Vohr 2005; De Groote 2007).

Moderate to low quality evidence from two studies (sample size 203 to 1455) showed that among children born at 22-27 or <27 weeks GA the prevalence of intellectual disability (WPPSI-R IQ <70) was 11.8% (95%CI 6.2 to 19.7) and 5.6% (95%CI 3.3 to 8.8) respectively at 5 years (Mikkola 2005; Leversen 2011).

Low quality from one study (sample size 141) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (WISC-IV IQ <-2SD) was 14.6% (95%CI 9.3 to 21.4) at 8 years corrected age (Roberts 2010). Low quality evidence from one other study showed that the prevalence (using WISC-III <70) in 275 children born at <28 weeks GA was 5.1% (95%CI 2.8 to 8.4) (Anderson 2003).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks the prevalence of intellectual disability (IQ <-2SD [K-ABC, GMDS or NEPSY]) was 40.7% (95%CI 34.4 to 47.2) at 6 years (Marlow 2005).

Low quality evidence from one study (sample size 244) showed that among children born at <27 weeks GA the prevalence of intellectual disability (Griffiths <2SD) was 10.4% (95%CI 5.8 to 16.8) at 12 months corrected age (Sutton 1999).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence of intellectual disability (verbal, DAS II <=2SD) was 17% (95%CI 14.5 to 19.5) and 15% (95%CI 12.7 to 17.6) for non- verbal reasoning (DAS II <=2SD) at 10 years (Joseph 2016b).

Severe intellectual disability

Moderate quality evidence from two studies (sample size ranged from 75 to 1508) showed that among children born at <27 weeks or 24-27 weeks GA the prevalence of intellectual disability (IQ <55, K-ABC) was 14.7% (95% CI 7.6 to 24.7) and 2.7% (95%CI 1 to 5.8) at 5-9 years (Stahlmann 2009; Foix-Helias 2008).

Moderate to low quality evidence from 5 studies (sample size ranged from 77 to 576) showed that among children born at GA range 23 to 27 weeks the prevalence of intellectual disability (BSIDIII <-3SD or MDI <55) ranged from 3.6% (95%CI 1.4 to 7.8) to 18.2% (95%CI 10.3 to 28.6) across the studies (Moore 2012; Anon 1997; De Groote 2007; Serenius 2013; Doyle 2011).

Low quality evidence from one study (sample size 141) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (IQ <-3SD, WISC-IV) was 6.3% (95%CI 2.9 to 11.5) at 8 years corrected age (Roberts 2010).

Moderate quality evidence from one study (sample size 306) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (IQ <55, WPPSI-R) was 2.9% (95%CI 1.4 to 5.5) at 5 years (Leversen 2011).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of intellectual disability (IQ <-3SD, K-ABC, GMDS or NEPSY) was 20.8% (95%CI 15.8 to 26.4) at 6 years (Marlow 2005).

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of intellectual disability (IQ <71 WPPSI) was 4.2% (95%CI 1.6 to 9.0%) At 4 years (Salakorpi 2001).

28-31 completed weeks of gestation

Moderate intellectual disability

Moderate quality evidence from one study (sample size 1508) showed that among children born at 28-32 weeks GA the prevalence of intellectual disability (MPC 55-69) was 8.7% (95CI 7.2 to 10.4) at 5 years (Foix-Helias 2008). In the same study, the prevalence in children born at 24-32 weeks GA was 9.6% (95%CI 8.2 to 11.2).

Moderate to severe intellectual disability

Moderate to low quality evidence from 4 studies (sample size ranged from 1455 to 1812) showed that among children born at a gestational age range of 28-32 weeks the prevalence of intellectual disability (MPC <70, K-ABC) was similar across the studies (range 8.9% (95%CI 7.3 to 10.7) to 12.1% (95%CI 10 to 14.4)) at 5 years (Beaino 2011; Marret 2007; Foix-Helias 2008; Larroque 2008).

(*A number of studies reported intellectual disability in children born at <32 weeks GA. One study of moderate quality in 3785 children born at 22-32 weeks GA found that the prevalence for intellectual disability (MDI <70, BSIDII) was 33.8% (95%CI 32.3 to 35.4) at 18-22 months corrected age (Vohr 2005).*

Low quality evidence from two studies (sample size ranged from 203 to 259) showed that among children at 23-32 weeks or mean GA 27.3 (2.1) the prevalence of intellectual disability (IQ<70, WISC-IV or DAS, or IQ<70, WPPSI-R) was 15.8% (95%CI 11.6 to 20.9) and 9.4% (95%CI 5.7 to 14.2) respectively at 5 years (Andrews 2008; Mikkola 2005).

Moderate quality evidence from two studies (sample size ranged from 1508 to 1812) showed that among children born at 24-32 weeks and <33 weeks GA the prevalence was the same (11.9% (95%CI 10.3 to 13.7)) at 5 years (Foix-Helias 2008; Larroque 2008).

Moderate quality evidence from one study (sample size 402) showed that among children born at <32 weeks GA/<1500g the prevalence of intellectual disability (IQ<-2SD, revised Amsterdam Child Intelligence Test) was 6.2% (95%CI 4.1 to 9) at 5 years (de Kleine 2003).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of intellectual disability (MDI <70, BSIDII) was 25.9% (95%CI 23.7 to 28.2) at 18-22 months corrected age (Vohr 2005). Another study reported a prevalence of 17% (95%CI 11 to 24) at <32 weeks GA (Cognitive delay, <2SD BSID) (Toome 2012).

Low quality evidence from one study (sample size 347) showed that among children born at <33 weeks GA the prevalence of intellectual disability (DQ <70, Brunet-Lezine) was 2.3% (95%CI 1 to 4.5) at 2 years (corrected age) (Charkaluk 2010).

Severe intellectual disability

Moderate quality from one study (sample size 1508) showed that among children born at 28-32 weeks GA the prevalence of intellectual disability (MPC <55) was 2.3% (95%CI 1.5 to 3.2) at 5 years (Foix-Helias 2008). In the same study, the prevalence in children born at 24-32 weeks GA was 2.3% (95%CI 1.6 to 3.2).

32-36 completed weeks of gestation

Moderate to severe intellectual disability

Low quality evidence from one study (sample size 646) showed that among children born at 32-34 weeks GA the prevalence of intellectual disability (MPC<70) was 7.6% (95%CI 5.7 to 9.9) at 5 years (Marret 2007).

4.1.2.6.13 *Prevalence of intellectual disability by week of gestational age*

Moderate intellectual disability

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate intellectual disability (BSIDII -2 to -3 SD) was 13.2% (95%CI 4.4 to 28.1%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate intellectual disability (full scale IQ 55-70, WPPSI-R) was 6.9% (95%CI 2.6 to 14.4%) compared to a prevalence of 2.6% (95%CI 0.7 to 6.6%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of intellectual disability (IQ -2 to -3 SD, KABC GMDS or NEPSY) was 33.3% (95%CI 15.6 to 55.3%) compared to a prevalence of 18.8% (95%CI 12.7 to 26.1%) in children born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate to severe intellectual disability

Low quality evidence from one study (sample size 244) showed that among children born at 23 weeks GA the prevalence of moderate to severe intellectual disability (major developmental delay, Griffiths <2SD) was 100% (95%CI 25 to 100%) compared to a prevalence of 3.9% (95%CI 0.81 to 11%) in children born at 26 weeks GA, assessed at 12 months corrected age (Sutton 1999).

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate to severe intellectual disability (cognitive impairment BSIDIII ≤-2SD) was 31.6% (17.5 to 48.7%) compared to a prevalence of 12.0% (95%CI 8.2 to 16.6%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Low quality evidence from one study (sample size 1503) showed that among children born at 24-26 weeks GA the prevalence of moderate to severe intellectual disability (MPC<70, KABC) was 15.7% (95%CI 9.2 to 24.2) compared to a prevalence of 8.9% (95%CI 6.2 to 12.0%) in children born at 31-32 weeks GA, assessed at 5 years (Beaino 2011).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate to severe intellectual disability (full scale IQ <70, WPPSI-R) was 9.2% (95%CI 4.1 to 17.3%) compared to a prevalence of 2.6% (95%CI 0.7 to 6.6%) in children born at 26-27 weeks GA, assessed at 5 years (Leversen 2011).

Moderate quality evidence from one study (sample size 1534) showed that among children born at 24-25 weeks GA the prevalence of moderate to severe intellectual disability (MPC <70, KABC) was 12.5% (95%CI 4.7 to 25.3%) compared to a prevalence of 10.7% (95%CI 7.5 to 14.6%) in children born at 32 weeks GA. However, the prevalence was higher in children born at 26 weeks GA (prevalence 21.1% (95%CI 11.4 to 33.9%), 27 weeks (prevalence 18.6% (95%CI 12.1 to 26.9%), and 28 weeks GA (prevalence 20.7% (95%CI 14.5 to 28%) (Larroque 2008).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe intellectual disability (MPC <70, KABC) was 9.9% (95%CI 6.5 to 14.3%) compared to a prevalence of 5.3% (95%CI 2.0 to 11.2%) in children born at 34 weeks GA, assessed at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate to severe intellectual disability (IQ≤-=2SD, KABC GMDS or NEPSY) was 58.3% (95%CI 36.6 to 77.9%) compared to a prevalence of 35.4% (95%CI 27.6 to 43.8%) in children born at 25 weeks GA, assessed at 5 years (Marlow 2005).

Severe intellectual disability

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of severe intellectual disability (cognitive impairment, BSIDIII <-3SD) was 18.4% (95%CI 7.7 to 34.3%) compared to a prevalence of 7.6% (95%CI 4.6 to 11.6%) in children born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of severe intellectual disability (full scale IQ <55, WPPSI-R) was 4.6% (95%CI 1.3 to 11.4%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of severe intellectual disability (IQ <-3SD, KABC, GMDS or NEPSY) was 25.0% (95%CI 9.8 to 46.7%) compared to a prevalence of 16.7% (95%CI 11 to 23.8%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

4.1.2.6.14 Prevalence of speech and/or language disorder

≤ 28 completed weeks of gestation

Moderate and severe speech and/or language disorder

Moderate quality evidence from one study (sample size 456) showed that among children born at <27 weeks GA the prevalence of moderate language impairment (-2 to -3SD BSIDIII) was 9.4% (95%CI 6.7 to 12.7) (Serenius 2013).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of moderate communication impairment (-2SD to -3SD BSIDIII) was 5.4% (95%CI 3.7 to 7.6) at 3 years age (Moore 2012). In the same study, there was a prevalence of 11.6% (95%CI 9.1 to 14.5) in children with moderate to severe impairment (<=2SD BSIDIII).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe speech/communication impairment ranged from 1.10% to 5.3% depending on whether they could communicate by a systemised method or not at 30 months (median) (Wood 2000).

Low quality evidence from one study (sample size 241) showed that among children born at less tha or equal to 25+6 weeks GA the prevalence for total severe impairment (PLS <2SD) was 15.6% (95%CI 10.8 to 21.4) at a median age of 6 years (Wolke 2008). However, the prevalence of severe communication impairment and severe language impairment in children (sample size ranged from 456 to 576) born at <27 weeks was lower in two studies of moderate to low quality (6.30% (95%CI 4.4 to 8.6) and 6.60% (95%CI 4.4 to 9.5) respectively) at the age of 2.5 to 3 years age (Serenius 2013; Moore 2012).

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate communication impairment (-2 to -3 SD BSID III) was 10.5% (95%CI 2.3 to 24.8) compared to 4.4% (95%CI 2.2 to 7.7) at 26 weeks GA (at the age of 3 years). A similar trend was observed when severe communication impairment was assessed (<-3SD BSIDIII), with prevalence increasing with decreasing gestational age by week. At 22-23 weeks GA, the prevalence was 15.8% (95%CI 6 to 31.3) (Moore 2012) compared to the prevalence at 26 weeks GA, which was 4% (95%CI 1.9 to 7.2) (Moore 2012).

For moderate to severe impairment, there was a similar trend, prevalence in the 22-23 GA group was 26.5% (95%CI13.4 to 43.1) compared to 8.4% (95% CI 5.3 to 12.5) in the 26 weeks GA group (Moore 2012).

28-31 completed weeks of gestation

Low quality evidence from one study (sample size 155) showed that among children born at <32 weeks GA the prevalence of moderate language delay (<2SD BSIDIII) was 33% (95%CI 26 to 41) at 2 years (corrected age) (Toome 2012).

4.1.2.6.15 Prevalence of speech and language disorder by week of gestation

Moderate speech and language disorder

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate speech/language disability (communication impairment, BSIDII -2 to -3 SD) was 10.5% (95%CI 2.9 to 24.8%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate to severe speech and language disorder

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate to severe speech/language disability (communication impairment, BSIDII ≤-2 SD) was 26.3% (95%CI 13.4 to 43.1%) compared to a prevalence of 8.4% (95%CI 5.3 to 12.5%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Severe speech and language disorder

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of severe speech/language disability (communication impairment, BSIDII <-3 SD) was 15.8% (95%CI 6.0 to 31.3%) compared to a prevalence of 4.0% (95%CI 1.9 to 7.2%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

4.1.2.6.16 *Prevalence of attention deficit hyperactivity disorder*

Less than or equal to 28 completed weeks of gestation

Low quality evidence from two studies (sample size 205 to 219) showed that among children born at <26 weeks GA and x adolescents born at <28 weeks GA the prevalence of ADHD (including any type, DAWBA or ChIPs) was 11.5% (95%CI 7.3 to 17) at the age of 11 years and 14.6% (95%CI 10 to 20.2) at the age of 18 years respectively. In the same two studies, the prevalence of ADHD (combined) was 4.4% (95%CI 1.9 to 8.4) and 3.4% (95% CI 1.4 to 7) respectively at the ages of 11 years and at 18 years. Prevalence of ADHD (inattentive) in the two studies was 10.7% (95%CI 6.9 to 16) at the age of 11 years and 7.1% (95%CI 3.8 to 11.8) at the age of 18 years (Johnson 2010; Burnett 2014).

Low quality evidence from one study of (sample size 205) showed that among children born at <26 weeks GA the prevalence of ADHD (hyperactive/impulsive, ChIPs) was 0.5% (95%CI 0.01 to 2.7) at the age of 18 years (Burnett 2014).

4.1.2.6.17 *Prevalence of autism spectrum disorder*

Less than or equal to 28 completed weeks of gestation

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks GA the prevalence of ASD (any) was 8% (95%CI 4.6 to 12.6) at the age of 11 years. In the same study, the prevalence of autistic disorder was 6.5% (95%CI 3.5 to 10.8) and for atypical autism, the prevalence was 1.5% (95%CI 0.3 to 4.3) (Johnson 2010).

Moderate quality evidence from one study (sample size 857) showed that among children born at <28 weeks GA the prevalence of ASD (ADI-R and ADOS-2) was 9.2% (95%CI 7.4 to 11.4%) and 7.1% (95%CI 5.5 to 9.0) respectively at 10 years age (Joseph 2016a).

4.1.2.6.18 *Prevalence of specific learning difficulty*

Less than or equal to 28 completed weeks of gestation

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks GA the prevalence reading impairment (WIAT-II <-2SD) was 30.2% (95%CI 24.1 to 36.9) at the age of 11 years (Johnson 2011). However, in another study of low quality, 275 children who were born at <28 weeks GA had a lower prevalence of reading impairment (WRAT 3 <70) was lower (5.8% (95%CI 3.4 to 9.3)) when assessed at the age of 8 years (Anderson 2003). In the same two studies, there was a higher prevalence of arithmetic impairment (43.7% (95%CI 37 to 50.6)) in children born at <26 weeks GA compared with a prevalence of 6.6% (95%CI 4 to 10.2) in children born at <28 weeks GA (Johnson 2011; Anderson 2003)

Low quality evidence from one study (sample size 257) showed that among children born at <28 weeks GA the prevalence of spelling impairment was 2.5% (95%CI 1 to 5.2) assessed at the age of 8 years (Anderson 2003).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence of academic achievement (WIAT-III <=-2SD) was 14% (95%CI 11.7 to 16.5) for word reading, 16% (95%CI 13.7 to 18.6) for pseudoword decoding, 14% (95%CI 11.7-16.5) for spelling, and 17% (95%CI 14.5 to 19.6) for numeric operations when assessed at the age of 10 years (Joseph 2016b).

28-31 completed weeks of gestation

Low quality evidence from one study (sample size 135) showed that among children born at <33 weeks GA the prevalence of delayed numerical skills (TEDI-MATH <40) was 20% (95%CI 13.6 to 27.8) (at the age of 8 years (Kiechl-Kohlendorfer 2013).

4.1.2.6.19 Prevalence of developmental coordination disorder

Less than or equal to 28 completed weeks of gestation

Low quality evidence from one study (sample size 298) showed that among children born at 22-27 weeks GA the prevalence of DCD was higher in a cohort born in 1997 (16% (95%CI 10.1 to 23.3)) compared to a cohort born in 1991(sample size 298) (10% (95%CI 6.9 to 14.1)) (Roberts 2011).

28-31 completed weeks of gestation

Moderate to low quality evidence from two studies (sample size ranged from 280 to 402) showed that among children at <32 weeks GA the prevalence of DCD or motor delay was 22.3% (95%CI 18.3 to 26.7) at the age of 5 years and 30.7% (95%CI 25.4 to 36.5) at the age of 7-8 years. (de Kleine 2003; Foulder-Hughes 2003).

Moderate quality evidence from one study (sample size 168) showed that among children born between 24-31 weeks GA the prevalence of motor deficit was 17.9% (95%CI 12.4 to 24.5) at the age of 5 years (Agerholm 2011).

32-36 completed weeks of gestation

4.1.2.6.20 Prevalence of mental and behavioural disorders

Less than or equl to 28 completed weeks of gestation

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks GA the prevalence of emotional disorder (any) was highest among 11 year olds (9% (95%CI 5.4 to 13.6)), compared to conduct disorder (any), oppositional defiant disorder (5.5% (95%CI 2.9 to 9.4) and 5% (95%CI 2.5 to 8.8)), specific phobia (2.5% (95%CI 0.8 to 5.7)), or a number of disorders including specific phobia or social phobia, PTSD, generalised anxiety, disorder, childhood emotional disorder, and major depression (prevalence range from 0.5%(95%CI 0.01 to 2.8) to 2% (95%CI 0.5 to 5)) (DAWBA, Johnson 2011).

Low quality evidence from one study (sample size 205) showed that among children born at <28 weeks GA the prevalence of anxiety/mood disorder was highest (21% (95%CI 15.6 to 27.2)) in adolescents compared to mood disorder (16.1% (95%CI 11.4 to 22)), major depressive disorder (13.7% (95%CI 9.3 to 19.1)), anxiety disorder (BAI/CESD-R) (11.2% (95%CI 7.3 to 16.4)), co-morbid disorder (6.3% (95%CI 3.4 to 10.6)) and obsessive compulsive disorder (2% (95%CI 0.5 to 5)) (DSM-IV axis I, Burnett 2014).

4.1.2.6.21 Prevalence of visual impairment

≤ 28 completed weeks of gestation

Moderate quality evidence from one study (sample size 456) showed that among children born at <27 weeks GA the prevalence of visual impairment (any) was 3.7% (95%CI 2.2 to 5.9) at 2.5 years corrected age (Serenius 2013).

Moderate visual impairment

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of visual impairment (impaired but not blind) was 4.6% (95%CI 2.3 to 8) at 6 years age (Marlow 2005).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of visual impairment (functionally impaired vision) was 5.9% (95%CI 4.1 to 8;2) at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 456) showed that among children born at <27 weeks GA the prevalence of visual impairment (moderate impairment) was 2.9% (95% CI 1.5 to 4.8) at 2.5 years corrected age (Serenius 2013).

Moderate to severe visual impairment

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-26 weeks GA the prevalence of unilateral blindness was 2.7% (95%CI 2 to 3.4) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 242) showed that among children born at <28 weeks GA the prevalence of moderate to severe visual deficiency (<3/10, one or both eyes) was 7% (95%CI 4.1 to 11) at 5 years age (Larroque 2008).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of visual impairment (impaired or blind) was 7.1% (95%CI 4.2 to 11.1) at 6 years age (Marlow 2005).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of impaired vision (blind or functionally impaired) was 6.9% (95%CI 5 to 9.3) at 3 years (Moore 2012).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of visual impairment (little useful vision) was 9.1% (95%CI 3.7 to 17.8) at 3 years age (de Groote 2007).

Low quality evidence from one study (sample size 88) showed that among children born at <28 weeks the prevalence of severe visual impairment (uni- or bilateral blindness or visual acuity <20/200 without glasses in at least one eye) was 12.5% (95%CI 6.4 to 21.3) at 11 years (Farooqi 2011).

Moderate quality evidence from two studies (sample size 306) showed that among children born at either 22-27 weeks GA or 23-25 weeks the prevalence for severe visual impairment was 0.3% (95%CI 0.01 to 1.8) and 1.2% (95%CI 0.03 to 6.2) respectively at 5 years (Leversen 2011).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe visual impairment (blind or perceives light) was 2.5% (95%CI 1 to 5) at 30 months (median) (Wood 2000).

Moderate quality evidence from one study (sample size 411) showed that among children born at <27 weeks GA the prevalence of visual impairment (blind or able to only fixate and follow light binocularly) was 3.1% (95%CI 1.6 to 5.3) at 30 months corrected age (Holmstrom 2014).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of visual impairment (no useful vision) was 2.6% (95%CI 0.9 to 9.1) at 3 years age (De Groote 2007).

Low quality evidence from two studies (sample size ranged from 189 to 219) showed that among children born at 23-27 weeks GA and 22-27 weeks GA the prevalence for blindness (<6/60 in both eyes) was 2.3% (95%CI 0.8 to 5.3) and 1.6% (95%CI 0.3 to 4.6) at 2 years and 8 years (corrected) respectively (Anon 1997; Anderson 2011).

Moderate to low quality evidence from three separate studies (sample size ranged from 306 to 373) showed that among children born at 22-27 weeks GA and also 23-25 weeks GA the prevalence for blindness was varied, ranging from 5.8% (95%CI 1.9 to 12.9) in the lower GA group (Leversen 2011), and 1.6% (95%CI ranged from 0.5 to 3.8) in the two 22-27 GA groups (Leversen 2010; Leversen 2011).

Moderate to very low quality evidence from 8 studies (sample size ranged from 19 to 3785) showed that among children born at various gestational ages (ranging from <26 weeks to <28 weeks) the prevalence of blindness was varied, ranging from 0.9% (95%CI 0.24 to 2.3) to 11% (95%CI 1.3 to 33) (Vohr 2005; Roberts 2010; Marlow 2005; Moore 2012; Hutchinson 2013; Serenius 2013; Anderson 2003; Rieger-Fackeldey 2010).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence of severe visual impairment (functional blindness) was 0.8% (95%CI 0.3 to 1.7) at 10 years (Joseph 2016b).

28-31 completed weeks of gestation

Moderate to severe visual impairment

Low quality evidence from one study (sample size 1455) showed that among children born at 30-31 weeks GA the prevalence of visual impairment (visual acuity <3/10 in both eyes) was 1.5% (95%CI 0.7 to 2.8) at 5 years (Marret 2007)

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA found that the prevalence of visual impairment (unilateral blindness) was 1.3% (95%CI 0.8 to 2) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 971) showed that among children born at 28-31 weeks GA the prevalence of moderate to severe visual deficiency (<3/10 in one or both eyes) was 2.1% (95%CI 1.3 to 3.2) at 5 years age (Larroque 2008).

Studies reporting vision impairment at <32 weeks of gestation

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-32 weeks GA the prevalence of unilateral blindness was 2.1% (95%CI 1.7 to 2.6) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 1697) showed that among children born at <33 weeks GA the prevalence of moderate to severe visual deficiency (<3/10 in one or both eyes) was 2% (95%CI 1.4 to 2.8) at 5 years (Larroque 2008).

Low quality evidence from one study (sample size 93) showed that among children born at <32 weeks GA the prevalence of visual impairment (worst eye blind or able to fixate torch) was 2.2% (95%CI 0.3 to 7.6) at 2.5 years corrected age (Hreinsdottir 2013).

Low quality evidence from one study with (sample size 155) showed that among children born at <32 weeks GA found that the prevalence of visual impairment (moderately reduced/blindness) was 0.64% (95%CI 0.02 to 3.5) at 2 years (corrected age) (Toome 2012).

Severe visual impairment

Moderate quality evidence from on study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of visual impairment (bilateral blindness) was 0.7% (95%CI 0.3 to 1.2) at 18-22 months corrected age (Vohr 2005). In the same study, the prevalence of bilateral blindness in children born at 22-32 weeks GA was 1.2% (95%CI 0.9 to 1.6) (Vohr 2005).

Low quality evidence from one study (sample size 93) showed that among children born at <32 weeks GA the prevalence of visual impairment (best eye blind or only able to fixate a torch) was 1.1% (95%CI 0.03 to 5.9) at 2.5 years corrected age (Hreinsdottir 2013).

32-36 completed weeks of gestation

Moderate to severe visual impairment

Low quality evidence from on study (sample size 1455) showed that among children born at 32-24 weeks GA the prevalence of visual impairment (visual acuity <3/10 in both eyes) was 1.7% (95%CI 0.9 to 3) at 5 years age (Marret 2007).

4.1.2.6.22 *Prevalence of visual impairment by week of gestation*

Moderate visual impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate visual impairment (functionally impaired vision) was 15.8% (95%CI 6.0 to 31.3%) compared to a prevalence of 3.2% (95%CI 1.4 to 6.2%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate visual impairment (visually impaired, not blind) was 8.3% (95%CI 1.0 to 27.0%) compared to a prevalence of 2.8% (95%CI 0.8 to 7.0%) in children born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate quality evidence from one study (sample size 494) showed that among children born at 22-23 weeks GA the prevalence of visual impairment (any; best estimated visual acuity <20/40) was 23.8% (95%CI 12 to 40) compared to a prevalence of 13.4% (95%CI 6.9 to 22.7) at 24 weeks GA, prevalence of 7% (95%CI 3.4 to 12.6) at 25 weeks GA, and a prevalence of 5.1% (95%CI 2.1-1-.2) at 26 weeks GA (Hellgren 2016).

Moderate to severe visual impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate to severe visual impairment (functionally impaired vision) was 18.4% (95%CI 7.7 to 34.3%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Low quality evidence from on study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe visual impairment (visual acuity <3/10 in both eyes) was 0.7% (95%CI 0.1 to 2.6) compared to a prevalence of 0.8% (95%CI 0.02 to 4.1%) in children born at 34 weeks GA. The prevalence was higher at GA 31 weeks (2.2% (95%CI 0.8 to 4.3%), and 33 weeks GA (2.3% (95%CI 0.5 to 6.5%), assessed at 5 years age (Marret 2007).

Moderate quality evidence from one study (sample size 1817) showed that among children born at 24-25 weeks GA the prevalence of moderate to severe visual impairment (<3/10 one or both eyes) was 9.3% (95%CI 3.1 to 20.3%) compared to a prevalence of 1.9% (95%CI 0.9 to 3.5%) in children born at 32 weeks GA, assessed at 5 years age (Larroque 2008).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate to severe visual impairment (visually impaired, or blind) was 16.7% (95%CI 4.7 to 37.4%) compared to a prevalence of 3.5% (95%CI 1.1 to 7.9%) in children born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Severe visual impairment

Moderate quality evidence from one study (sample size 411) showed that among children born at 22-23 weeks GA the prevalence of severe visual impairment (blind or able to only fixate and follow light binocularly) was 4.8% (95%CI 0.6 to 16.2%) compared to a prevalence of 1.4% (95%CI 0.2 to 4.8%) in children born at 26 weeks GA, assessed at 30 months corrected age (Holmstrom 2014).

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of visual impairment (blindness) was 2.6% (95%CI 0.1 to 13.8%) compared to a prevalence of 1.2% (95%CI 0.3 to 3.5%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of severe visual impairment (blindness) was 8.3% (95%CI 1.0 to 27.0%) compared to a prevalence of 0.7% (95%CI 0.02 to 3.8%) in children born at 25 weeks GA assessed at 6 years age (Marlow 2005).

Prevalence of visual impairment using per 1000 or 10,000 live births as denominator

Moderate to severe visual impairment (<28 weeks GA)

Very low quality evidence from one study (sample size 172, 584 livebirths) showed that among children born at <28 weeks GA the prevalence of moderate to severe visual impairment (<=6/18 in better eye or worse) was 182.5 cases per 10,000 livebirths (95%CI 102.5 to 299.1) at 12 years (Bodeau-Livinec 2007).

Moderate to severe visual impairment (28-31 weeks GA)

Very low quality evidence from one study (sample size 172, 584 livebirths) showed that among children born at 29-32 weeks GA the prevalence of moderate to severe vision impairment (<=6/18 in better eye or worse) was 37.1 cases per 10,000 livebirths (95%CI 14.9 to 76.2)at 12 years age (Bodeau-Livinec 2007).

Moderate to severe visual impairment (32-36 weeks GA)

Very low quality evidence from one study (sample size 172, 584 livebirths) showed that among children born at 33-36 weeks GA the prevalence of moderate to severe vision impairment (<=6/18 in better eye or worse) was 27 cases per 10,000 livebirths (95%CI 17.3 to 40.1) at 12 years age (Bodeau-Livinec 2007).

4.1.2.6.23 Prevalence of hearing impairment

Less than or equal to 28 completed weeks of gestation

Moderate hearing impairment

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA, the prevalence of hearing loss (corrected with hearing aids) was 2.9% (95%CI 1.2 to 5.9) when assessed at 6 years age (Marlow 2005).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of hearing loss (improved by aids) was 5.2% (95%CI 3.5 to 7.4) when assessed at 3 years age (Moore 2012).

Low quality evidence from one study (sample size 77) showed that among children born at < 27 weeks GA the prevalence of hearing impairment (but useful hearing) was 3.9% (95%CI 0.8 to 11) (De Groote 2007).

Moderate to severe hearing impairment

Low quality evidence from one study (sample size 141) showed that among children born at 22-27 weeks GA the prevalence of hearing impairment was 2.1% (95%CI 0.4 to 6) at 8 years corrected age (Roberts 2010).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks the prevalence of moderate to severe hearing impairment was 5.8% (95%CI 3.2 to 9.6) at 6 years (Marlow 2005). In another study of low quality with 576 children born at <27 weeks GA the prevalence for severe hearing impairment was 5.4% (95%CI 3.7 to 7.6) at 3 years (Moore 2012).

Low quality evidence from one study (sample size 19) showed that among children born at mean 25.4 weeks GA the prevalence of hearing impairment (requiring hearing aid) was 11% (95%CI 1.3 to 33) at 5 years age (Rieger-Fackeldey 2010). Ten other studies (sample size ranged from 77 to 3785) of moderate to very low quality assessing hearing impairment or deafness (requiring hearing aids) in children born at a range of 22-28 weeks GA found that the prevalence was lower but varied, ranging from 0.7% (95%CI 0.14 to 2) to 5.7% (95%CI 1.9 to 12.8) (Farooqi 2011; Leversen 2011; Vohr 2005; Doyle 2011; Anderson 2011; De Groote 2007; Hutchinson 2013; Wood 2000; Serenius 2013; Anderson 2003).

Severe hearing impairment

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe hearing impairment (uncorrected without hearing aid) was 5.3% (95%CI 3.0 to 8.6) at 30 months (median) (Wood 2000).

Low quality evidence from one study (sample size 373) showed that among children born at 22-27 weeks GA the prevalence of deafness was 0.8% (95%CI 0.1 to 2.7) at 2 years (corrected age) (Leversen 2010). In another study (sample size 401) of low quality, the prevalence of deafness was 0.9% (95%CI 0.1 to 3.3) in children assessed at 2 years (Anon 1997). Prevalence of deafness was 0.2% (95%CI 0.01 to 1.2) in children (sample size 456) born at <27 weeks GA (moderate quality, Serenius 2013). At 5 years age, the prevalence of deafness was 1.0% (95%CI 0.2 to 2.8) in children (sample size 306) born at 22-27 weeks GA (moderate quality study, Leversen 2011).

Low quality evidence from one study (sample size 261) showed that among children born at <28 weeks GA the prevalence of severe hearing deficiency (>70 decibels in one or both ears or hearing aid) was 0.8% (95%CI 0.1 to 2.7) at 5 years age (Larroque 2008).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of profound sensorineural hearing loss (not improved by aids) was 0.2% (95%CI 0.1 to 1) at 3 years age (Moore 2012). In another moderate quality study (sample size 241) children born at <26 weeks GA found that the prevalence of profound sensorineural hearing loss was 2.9% (95%CI 1.2 to 5.9) at 6 years age (Marlow 2005).

28-31 completed weeks of gestation

Moderate to severe hearing impairment

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of permanent hearing loss (amplification in both ears) was 1.4% (95%CI 0.9 to 2.1) at 18-22 months corrected age (Vohr 2005).

Low quality evidence from one study (sample size 1455) showed that among children born at 30-31 weeks GA the prevalence for hearing loss >70 decibels was 0.30% (95%CI 0.04 to 1.1) at 5 years (Marret 2007).

Severe hearing impairment

Moderate quality evidence from one study (sample size 1020) showed that among children born at 28-31 weeks GA the prevalence for severe hearing deficiency (>70 decibels in one or both ears or hearing loss) was 0.5% (95%CI 0.2 to 1.1) at 5 years age (Larroque 2008).

4.1.2.6.24 Prevalence of hearing impairment by week of gestation

Moderate hearing impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate hearing impairment (hearing loss improved by aids) was 5.3% (95%CI 0.6 to 17.8%) compared to a prevalence of 5.2% (95%CI 2.8 to 8.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at 24 weeks GA the prevalence of moderate hearing impairment was 2.7 (95%CI 0.3 to 9.6%) compared to a prevalence of 3.5% (95%CI 1.1 to 7.9%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

Moderate to severe hearing impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate hearing impairment (hearing loss improved by aids) was 7.9% (95%CI 1.7 to 21%) compared to a prevalence of 5.2% (95%CI 2.8 to 8.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate to severe hearing impairment (hearing aid in both ears) was 2.3% (0.3 to 8.1%) compared to a prevalence of 1.3% (95%CI 0.2 to 4.7%) in children born at 26-27 weeks GA, assessed at 5 years (Leversen 2011).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe hearing impairment (hearing loss >70 decibels or aids in one or both ears) was 0.3% (95%CI 0.01 to 1.9%) compared to a prevalence of 1.5% (95%CI 0.2 to 5.3%) in children born at 34 weeks GA, assessed at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate to severe hearing impairment was 4.2% (95%CI 0.1 to 21.1%) compared to a prevalence of 4.9% (95%CI 2.0 to 9.8%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

4.1.2.6.25 Severe hearing impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of severe hearing impairment (profound sensorineural hearing loss not improved by aids) was 2.6% (95%CI 0.1 to 13.8%), assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 1817) showed that among children born at 24-25 weeks GA the prevalence of severe hearing impairment (>70 decibels in one or both ears or hearing aid) was 1.7% (95%CI 0.04 to 9.2%) compared to a prevalence of 0.2% (95%CI 0.01 to 1.1%) in children born at 32 weeks GA, assessed at 5 years (Larroque 2008).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of severe hearing impairment (profound sensorineural hearing loss) was 4.2% (95%CI 0.1 to 21.1%) compared to a prevalence of 1.4% (95%CI 0.1 to 4.9%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

4.1.3 Prevalence of problems

Review question:

What is the prevalence of developmental problems in babies, children and young people born preterm?

4.1.3.1 Description of clinical evidence

The aim of this review is to establish the prevalence and incidence of different developmental problems in relation to the different gestational ages in babies, children and young people born preterm. The developmental problems considered as outcomes are listed below:

- Sensory sensitivity (hypersensitivity and hyposensitivity) or sensory difficulties
- Functional problems (feeding, sleeping and toileting),
- Motor, developmental and language delay
- Problems specific to infancy (excessive crying, irritability, and poor-self regulation)
- Problems specific to childhood (behavioural, social and emotional problems, and special education needs)

Fifty-five studies were included in the review (Agerholm 2011; Anderson 2011; Anderson 2003; Anderson 2004; Arnaud 2007; Chan 2014; Charkaluk 2010; Chyi 2008; de Groote 2007; de Kleine 2003; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Downey 2015; Faebo Larsen 2013; Farooqi 2007; Foix-L'Helias 2008; Germa 2012; Guellec 2011; Guy 2015; Higa Diez 2016; Hornman 2016; Hutchinson 2013; Johnson 2010; Johnson 2016; Johnson 2015; Johnson 2015; Johnson 2011; Joseph 2016; Joseph 2016; Kan 2008; Kerstjens 2011; Larroque 2011; Mackay 2013; Mackay 2010; Mansson 2014; Moore 2012; Odd 2016; Odd 2013; Odd 2012; Peacock 2012; Plomgaard 2006; Potijk 2012; Potijk 2013; Quigley 2012; Rautava 2010; Raynes-Greenow 2012; Samara 2010; Samara 2008; Schendel 1997; Stahlman 2009; Stene-Larsen 2014; Stoelhorst 2003; Stoelhorst 2003; Wilson-Ching 2013;; Zhu 2012).

No evidence was found for the outcomes of functional problems (toileting), excessive crying, irritability, and poor self-regulation.

The sample size ranged from 77(de Groote 2007) to 403,106 (Raynes-Greenow 2012).

Seventeenstudies were from the UK or UK and Ireland (Chan 2014; Guy 2015; Johnson 2010; Johnson 2016; Johnson 2015; Johnson 2015; Johnson 2011; Mackay 2013; Mackay 2010; Moore 2012; Odd 2016; Odd 2013; Odd 2012; Peacock 2012; Quigley 2012; Samara 2010; Samara 2008;).

Eight studies were from France (Arnaud 2007; Charkaluk 2010; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Foix-Helias 2008; Germa 2012; Guellec 2011; Larroque 2011)

Seven studies were from the Netherlands (de Kleine 2003; Hornman 2016; Kerstjens 2011; Potijk 2012; Potijk 2013; Stoelhorst 2003; Stohelorst 2003).

Four studies were from Denmark (Agerholm 2011; Faebo Larsen 2013; Plomgaard 2006; Zhu 2012).

Two studies were from USA (Downey 2015; Schendel 1997)

One study each was from Australia (Wilson-Ching 2013), Belgium (de Groote 2007); Finland (Rautava 2010); Germany (Stahlman 2009); Japan (Higa Diez 2016); Norway (Stene-Larsen 2014); Sweden (Mansson 2014).

Majority of the publications used data from population-based (national, geographical or regional) prospective cohort studies (Anderson 2011; Anderson 2004; Arnaud 2007; Chan

2014; Charkaluk 2010; Chyi 2008; De Groote 2007; de Kleine 2003; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Downey 2015; Farooqi 2007; Foix-Helias 2008; Germa 2012; Guellec 2011; Guy 2015; Hutchinson 2013; Johnson 2010; Johnson 2015; Johnson 2015; Johnson 2011; Joseph 2016; Joseph 2016; Kerstjens 2011; Larroque 2011; Mansson 2014; Moore 2012; Odd 2016; Odd 2013; Odd 2012; Peacock 2012; Plomgaard 2006; Potijk 2012; Potijk 2013; Quigley 2012; Rautava 2010; Raynes-Greenow 2012; Samara 2010; Samara 2008; Schendel 1997; Stahlmann 2009; Stene-Larsen 2014; Wilson-Ching 2013;).

Four publications used data from regional birth cohort (Agerholm 2011; Anderson 2003; Kan 2008; Stoelhorst 2003; Stoelhorst 2003).

Two publications were from national birth cohorts (Faebo Larsen 2013; Zhu 2012).

Two publications were retrospective studies using national registry data (Mackay 2013; Mackay 2010).

Six studies reported on functional problems (Germa 2012; Johnson 2016; Potijk 2012; Raynes-Greenow 2012; Samara 2010; Stoelhorst 2003).

Eleven studies reported on motor problems (Agerholm 2011; Arnaud 2007; De Groote 2007; Faebo Larsen 2013; Kan 2008; Mansson 2014; Potijk 2013; Rautav a 2010; Schendel 1997; Stoelhorst 2003; Zhu 2012).

Seven studies reported on developmental delay (Agerholm 2011; Charkaluk 2010; Johnson 2015; Kerstjens 2011; Plomgaard 2006; Potijk 2013; Schendel 1997).

Six studies reported on language problems (Joseph 2016; Mansson 2014; Potijk 2013; Rautava 2010; Schendel 1997; Stene-Larsen 2014;).

Four studies reported on executive function (Anderson 2004; Anderson 2011; Joseph 2016; Rautava 2010).

Twenty-three studies reported on behavioural, social, and emotional problems (Anderson 2011; Anderson 2003; de Kleine 2003; Delobel-Ayoub 2009; Delobel-Ayoub 2006; Downey 2015; Farooqi 2007; Foix-Helias 2008; Guellec 2011; Guy 2015; Higa Diez 2016; Hornman 2016; Hutchinson 2013; Johnson 2010; Johnson 2015; Joseph 2016; Larroque 2011; Moore 2012; Potijk 2012; Rautava 2010; Samara 2010; Samara 2008; Stahlmann 2009; Stoelhorst 2003; Wilson-Ching 2013).

Fourteen studies reported on specialist educational needs (Chan 2014; Chyi 2008; Farooqi 2007; Guellec 2011; Johnson 2011; Larroque 2011; Mackay 2013; Mackay 2010; Odd 2016; Odd 2013; Odd 2012; Peacock 2012; Quigley 2012; Samara 2008).

Evidence from these are summarised in the summary of included studies table below (Table 19). See also the study selection flow chart in Appendix F:, study evidence tables in Appendix K: and exclusion list in Appendix G:.

The feasibility of combining study data using meta-analysis was assessed. Due to the following differences between studies, it was not considered appropriate to pool the results:

- the inclusion/exclusion criteria for participants
- ages of participants at the time of assessment
- outcome definitions and measurement tools
- consistency of results.

4.1.3.2 Summary of included studies

Table 19: Summary of included studies for prevalence of problems

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
Evidence on functional problems					
Germa 2012	Prospective population-based cohort	N=2901 born in 1997 N=247 born in 1998 n=2349 children born very preterm and followed n=1882 children followed because they attended the medical examination n=1711 children born followed who did not have head malformation and who underwent the medical examination at 5 years age were included	Palatal morphology was assessed by simple visual inspection as altered or not by the physicians, without any further indication. The assessment criteria for altered palatal morphology were left to the physicians' judgement.	At 5 years age Altered palatal morphology 22-33 weeks GA: 63/1711, 3.7% (95%CI 2.9-4.7)	Low
Johnson 2016	Prospective population-	N=628 late and	At 2 y corrected age, parents were asked to	At 2 years of corrected age Total eating difficulties	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
	based cohort study	moderately preterm (LMPT) children (32-36 weeks)	complete a questionnaire comprising measures to assess infants' eating behaviour, cognitive development, behaviour and emotional problems, and neurosensory impairment. A validated eating behaviour questionnaire (4) was used to assess the presence of eating difficulties in the 4 domains of refusal/picky eating (e.g., poor appetite, food refusal, selective eating), oral motor problems (e.g., problems biting, chewing, or swallowing; gagging; or choking on food), oral hypersensitivity (e.g., aversion to being touched around the mouth or having things put in the mouth), and eating behaviour problems (e.g., has tantrums or makes a mess during meals). For each of 17 items, parents were asked to state whether their child exhibited the problem behaviour never, occasionally, or often. Each item was scored 0, 1, or 2, respectively, from which a total eating difficulties score was computed (range: 0–34) and 4 subscale scores for refusal/picky eating (7 items;	32-36 weeks GA: 69/726, 9.5% (7.5-11.9%) Refusal picky eating 32-36 weeks GA: 48/744, 6.5% (4.8-8.5%) Oral motor problems 32-36 weeks GA: 41/749, 5.5% (4.0-7.4%) Oral hypersensitivity 32-36 weeks GA: 32/756, 4.2% (2.9-5.9%) Eating behaviour problems 32-36 weeks GA: 45/738, 6.1% (4.5-8.1%)	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			range: 0–14), oral motor problems (5 items; range: 0–10), oral hypersensitivity (2 items; range: 0–4), and eating behavior problems (3 items; range: 0–6); for all scales, higher scores indicate greater problems. >90th percentile of the term control group were used to identify children with clinically significant eating difficulties.		
Potijk 2012	Prospective cohort study	N=916 moderately preterm children	Behavioural and emotional problems were measures using the Dutch version of the Child Behaviour Checklist (CBCL) for ages 1.5-5. Problem scores were subdivided into three categories: normal range (<93rd percentile), subclinical or bordering range (93rd to 97th percentile), and clinical or elevated range (>97th percentile).	At 4 years age Sleep problems (CBCL, >97th perc) 32-35 weeks GA: 22/916, 2.4% (1.5-3.6%)	Moderate
Raynes-Greenow 2012	Population based linkage study	Sample recruited n=429305 Sample analysed after exclusions n=403106 (n=3115 children born	Data from births from 2000–2004 were obtained via the NSW Midwives Data Collection, a legislated population-based surveillance system that includes information on all babies born at ≥ 20 weeks gestation or weighing ≥ 400	Assessed at age 2.5 to 6 years Functional problems (sleep apnoea, ICD-10) <32 weeks GA:82/3115, 2.6% (95%CI 2.1-3.2) 32-36 weeks GA: 286/22,039, 1.3% (95%CI 1.2-1.5)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		at <32 weeks; n=22039 children born at 32-36 weeks; n=377952 children born at >36 weeks)	g. No further details reported. The primary outcome was sleep apnoea diagnosis in childhood, first diagnosed between 1 and 6 years of age. Children with sleep apnoea were identified from those hospital records with the ICD-10 code G47.3: sleep apnoea, central or obstructive.		
Samara 2010	National population based cohort study	n=308 children alive at 30 months age n=241 entered study n=223 completed eating questionnaire	When the child reached 6 years of age, parents completed a specially developed eating questionnaire. The scale included 19 items, which were grouped into four categories: refusal-faddy eating problems, oral motor problems, oral hypersensitivity problems and behavioural problems around meals. A total eating problems score was also constructed. Higher scores on each scale indicate more problems. To derive clinical categories, each scale was dichotomised into normal versus clinical (scores above the 90th centile or near according to the comparison group).	Assessed at 6 years age Total eating problems 25+6 weeks GA: 76/218, 34.9% (95%CI 29.0-41.6) ≤23 weeks GA: 9/22, 40.9% (95%CI 20.7-63.7) 24 weeks GA: 34/68, 50.0% (95%CI 37.6-62.4) 25 weeks GA: 33/128, 25.8% (95%CI 18.5-34.3) Oral motor problems 25+6 weeks GA: 72/215, 33.5% (95%CI 27.2-40.2) ≤23 weeks GA: 8/20, 40.0% (95%CI 19.1-64.0) 24 weeks GA: 27/66, 40.9% (95%CI 29.0-53.7) 25 weeks GA: 37/129, 28.7% (95%CI 21.1-37.3) Refusal faddy problems 25+6 weeks GA: 38/223, 17.0% (95%CI 12.4-22.6) ≤23 weeks GA: 3/22, 13.6% (95%CI 2.9-34.9) 24 weeks GA: 11/68, 16.2% (95%CI 8.4-27.1) 25 weeks GA: 24/133, 18.1% (95%CI 11.9-25.7) Hypersensitivity problems 25+6 weeks GA: 50/213, 23.5% (95%CI 18.0-30.0) ≤23 weeks GA: 4/22, 18.2% (95%CI 5.2-40.3) 24 weeks GA: 22/63, 34.9% (95%CI 23.3-48.0) 25 weeks GA: 24/128, 18.8% (95%CI 12.4-26.6)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
Stoelhorst 2003	Regional population-based prospective cohort study	N=158 children with completed CBCL questionnaires (N=266 children included in the cohort originally, N=235 survived)	The Child Behaviour Checklist (CBCL) for 2- to 3-y-old children was handed out to the parents during the 2-year check-up at the outpatient clinic and returned by mail. The CBCL had to be completed by one or both parents. For the total problem score, the internalizing and externalizing groups, scores above the 90th centile are defined as clinically abnormal, scores from the 85th through the 90th centile as borderline clinical.	At 2 years of corrected age Sleep problems (CBCL, 98th perc) <32 weeks GA: 5/158, 3.2% (1.0-7.2%)	Low
Evidence on motor delay					
Agerholm 2011	Regional birth cohort study	N=237 live born children with 24-31 weeks GA in the geographical area N=204 children survived N=175 children followed-up at 5 years of age (86% of the ones who survived)	Motor function was examined using the Movement Assessment Battery for Children (M-ABC), it measures three items in the area of manual dexterity, two items in the area of ball skills and three items in the area of balance. The items were scored from 0 to 5, where 0 was the optimum score. The test is standardised and the scores are presented in relation to the 5th and the 15th percentile in the reference group. A score above the 15th percentile show normal	At 5 years of age Motor function Uncertain motor function (M-ABC ≤15th percentile total score) 24-31 weeks GA: 31/168, 18.5% (12.9-25.2%) Combined cognitive and motor skills (Uncertain preschool skills, MAP, yellow) 24-31 weeks GA: 21/168, 12.5% (7.9-18.5%) Combined cognitive and motor skills (Deficit in preschool skills, MAP, red) 24-31 weeks GA: 12/168, 7.1% (3.8-12.1%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			motor skills. A score between the 5th and 15th percentile indicates need for observation for motor function deficit, and a score under 5th percentile indicates motor function deficit.		
Arnaud 2007	Prospective population-based cohort study	n=1662 children born before 33 weesk GA, examined at 5 years n=246, children born at 33 and 34 weeks GA, examined at 5 years	The short version of Touwen examination was used to assess at 5 years age, a 16 item assessment grouped into 4 subsets for posture and muscle tone, reflexes, coordination and balance, and motor and behaviour of the face and eyes. Each of the subsets was rated as optimal or nonoptimal. The children were then classified as healthy (MND-0), mild (MND-1) or moderate (MND-2) neuromotor dysfunctional signs. The test was designed to detect minor abnormalities.	Assessment at 5 years age Minor neuromotor dysfunction ((mild, MND-1, one or two items affected), Touwen assessment) ≤27 weeks GA: 93/178, 52.3% (95%CI 44.6-60.0) 28-30 weeks GA: 177/440, 40.2% (95%CI 35.6-45.0) 31 weeks GA: 107/263, 40.7% (95%CI 34.7-47.0) 32 weeks GA: 138/356, 38.8% (95%CI 33.7-44.0) 33-34 weeks GA: 60/195, 30.8% (95%CI 24.4-37.8) 28-31 weeks GA: 284/703, 40.4% (95%CI 36.8-44.1) 32-34 weeks GA: 198/551, 36.0% (95%CI 32.0-40.1) Minor neuromotor dysfunction ((moderate, MND-2, >2 items affected), Touwen assessment) ≤27 weeks GA: 9/178, 5.1% (95%CI 2.3-9.4) 28-30 weeks GA: 16/440, 3.6% (95%CI 2.1-5.8) 31 weeks GA: 6/263, 2.3% (95%CI 0.8-5.0) 32 weeks GA: 7/356, 2.0% (95%CI 0.8-4.0) 33-34 weeks GA: 1/195, 0.5% (95%CI 0.01-2.8) 28-31 weeks GA: 22/703, 3.1% (95%CI 2.0-4.7) 32-34 weeks GA: 8/551, 1.5% (95%CI 0.63-2.8) Postural/muscle tone regulation (consistent mild deviation in posture (≥2 items) and/or in muscle tone (≥1 item) ≤27 weeks GA: 36/178, 20.2% (95%CI 14.6-29.0) 28-30 weeks GA: 63/440, 14.3% (95%CI 11.2-18.0) 31 weeks GA: 14/263, 5.3% (95%CI 2.9-8.8) 32 weeks GA: 20/356, 5.6% (95%CI 3.5-8.5)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				33-34 weeks GA: 4.1% (95%CI 1.8-7.9) 28-31 weeks GA: 77/703, 11.0% (95%CI 8.7-13.5) 32-34 weeks GA: 28/551, 5.1% (95%CI 3.4-7.3) Reflex abnormalities (abnormal intensity and/or threshold or asymmetry in ≥1 item) ≤27 weeks GA: 26/178, 14.6% (95%CI 9.8-20.7)37.1 28-30 weeks GA: 41/440, 9.3% (95%CI 6.8-12.4) 31 weeks GA: 29/263, 11.0% (95%CI 7.5-15.5) 32 weeks GA: 29/356, 8.2% (95%CI 5.5-11.5) 33-34 weeks GA: 9/195 4.6% (95%CI 2.1-8.6) 28-31 weeks GA: 70/703, 10.0% (95%CI 7.8-12.4) 32-34 weeks GA: 38/551, 6.9% (95%CI 4.9-9.3) Coordination and balance (presence of age-inadequate performance on ≥2 tests) ≤27 weeks GA: 66/178, 37.1% (95%CI 30.0-44.6) 28-30 weeks GA: 121/440, 27.5% (95%CI 23.4-32.0) 31 weeks GA:74 /263, 28.1% (95%CI 22.8-34.0) 32 weeks GA: 90/356, 25.3% (95%CI 21.0-30.1) 33-34 weeks GA: 41/195, 21.0% (95%CI 15.5-27.4) 28-31 weeks GA: 195/703, 27.7% (95%CI 24.5-31.2) 32-34 weeks GA: 131/551, 23.8% (95%CI 20.3-27.6) Motor behaviour of face and eyes (≥1 abnormal item) ≤27 weeks GA: 28/178, 15.7% (95%CI 10.7-22.0) 28-30 weeks GA: 53/440, 12.1% (95%CI 9.2-15.5) 31 weeks GA: 36/263, 13.7% (95%CI 9.8-18.4) 32 weeks GA: 57/356, 16.0% (95%CI 12.4-20.2) 33-34 weeks GA: 20/195, 10.3% (95%CI 6.4-15.4) 28-31 weeks GA: 89/703, 12.7% (95%CI 10.3-15.4) 32-34 weeks GA: 77/551, 14.0% (95%CI 11.2-17.2)	
De Groote 2007	Population-based geographically	n=95 children that survived	The assessment at 3 years comprised of a detailed clinical examination and full	At 3 years Severe psychomotor developmental delay (PDI <55)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
	defined cohort study	to discharge from NICU n=77 children assessed at 3 years (n=3 died before follow-up, n=12 parents did not give consent, n=3 could not be reached), 81% follow-up rate (84% of the ones who were alive at follow-up).	developmental evaluation. The clinical evaluation included collecting the recent medical history and a global health and anthropometric assessment as well as standardised neurologic and sensory examination. The Dutch edition of the second version of the Bayley Scales of Infant Development (BSID-II-NL) was used to assess mental and psychomotor development. The BSID-II-NL is standardised on a mean score of 100 and a SD of 15 points. Moderate impairment is defined as a score of 55-69 and severe impairment as a score of <55.	<27 weeks GA: 21/77, 27.3% (17.7-38.6%) Moderate psychomotor developmental delay (PDI 55-69) <27 weeks GA: 16/77, 20.8% (12.4-31.5%) Moderate to severe psychomotor developmental delay (PDI <70) <27 weeks GA: 37/77, 48.1% (36.5-59.7%)	
Faebo Larsen 2013	Danish National Birth Cohort study	N=32097 children (including term and preterm children) included in analysis N=1234 moderately preterm (32-36 weeks) N=137 very preterm (23-31 weeks)	The outcome was based on the Developmental Coordination Disorder Questionnaire (DCDQ) '07 which is a parent questionnaire aimed at identifying children with motor problems. It enables classification of children into the categories 'indication possible or suspect for DCD' versus 'probably not DCD'. It captures three motor development areas: control	At 7 years of age Indication of, or suspect for DCD 23-31 weeks GA: 25/137, 18.3% (12.2-25.8%) 32-36 weeks GA: 79/1234, 6.4% (5.1-7.9%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			during movement, fine motor control/handwriting, and general coordination.		
Kan 2008	Regional cohort study	N=401 consecutive very preterm infants n=225 surviving to age 8 years n=210 assessed at age 8 years n=179 very preterm infants assessed in study	Assessment of motor function, using the Movement Assessment Battery for Children (Movement ABC), which yields a percentile score composed of cumulative scoring of manual dexterity, ball skills and balance tasks. Children with a percentile ranking <15 were considered to have poor motor performance	At age 8 years age Motor performance (MABC, <15th percentile) 23-27 weeks GA: 26/173, 15% (95%CI 10.1-21.2%)	Very low
Mansson 2014	Population based cohort study	N=707 n=461 eligible for follow-up n=399 children born at <27 weeks GA (after exclusions, surviving to age 2.5 years and had BSID III assessment)	Test scores were evaluated on the basis of the means and standard deviations of the controls. Function level was regarded as normal if the subtest scaled score was ≤+1 SD and ≥1 SD of the control mean. Mild delay was classed as ≤1SD to ≥2 SD, moderate delay was classed as <2SD to ≥3 SD, and severe delay was classed as <3SD.	At 2.5 years age Fine motor (BSID III mild -1SD to 2 SD) <27 weeks GA: 133/395, 33.7% (95%CI 29.0-39.0) Fine motor (BSID III moderate -2SD to 3SD) <27 weeks GA: 32/395, 8.1% (95%CI 5.6-11.2) Fine motor (BSID III moderate to severe -3SD) <27 weeks GA: 17/395, 4.3% (95%CI 2.5-6.8) Gross motor (BSID III mild -1 SD to 2SD) <27 weeks GA: 111/383, 29.0% (95%CI 24.5-33.8) Gross motor (BSID III moderate -2SD to 3SD) <27 weeks GA: 27/383, 7.0% (95%CI 4.7-10.1) Gross motor (BSID III moderate to severe -3SD) <27 weeks GA: 0/0	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
Potijk 2013	Prospective cohort study	N=926 moderately preterm children assessed at 4 years. (N=544 term born controls)	Developmental outcomes were measured using the Dutch version of the 48-month form of the Ages and Stages Questionnaire (ASQ) which is a validated, parent-completed developmental screening instrument. Five developmental domains: fine motor, gross motor, communication, problem-solving, and personal-social skills. For the total score and the domains scores cut-offs for normal and abnormal scores were set at 2 SD below the mean score of the Dutch reference group.	At 4 years of age Fine motor delay (ASQ, <-2SD) 32-35 weeks GA: 74/917, 8.1% (6.4-10.0%) Gross motor delay (ASQ, <-2SD) 32-35 weeks GA: 52/911, 5.7% (4.3-7.4%)	Moderate
Rautava 2010	Population based prospective cohort study	Original sample size: n=924 preterm/very low birth weight infants Included in follow-up: n=588 preterm/very low birth weight infants	Behavioural outcomes were assessed using the Five to Fifteen Questionnaire (FTF), which was completed by the parents. Questions on development and behaviour were rated by the parents as 0="does not describe", 1="describes to some extent" and 2="describes well" the individual child	Assessed at 5 years age Motor skills problems (FTF) <32 weeks GA: 49/588, 8.3% (95%CI 6.2-11.0)	Low
Schendel 1997	Regional prospective cohort study	n=367 very low birth weight children (<1500 g) with	The Denver II was used to screen for possible developmental delay by comparing the child's performance on various	At adjusted age 15 months (range 9-34 months) Fine motor-adaptive (Denver II) ≥1 cautions: VLBW/28.4 (3.0) weeks GA: 44/367, 12.0% (95%CI 9.0-15.8)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		Denver II assessment at follow-up n= 553 moderately low birth weight children (1500-2499 g) with Denver II assessment at follow-up	tasks with children of the same adjusted age. 9 outcomes indicating delay were based on four domains: Personal-social, language, fine motor-adaptive skills, and gross motor skills. The 9 outcomes reflected two types of delay: 1. A moderate delay (overall questionable performance + four domain specific outcomes for children who received one or more caution scores in a given domain); 2. Severe delay (abnormal overall test performance +the four domain specific outcomes for children who received one or more delay scores in a given domain The overall performance was based on total number of caution and/or delay scores across all domains and was categorised as: 1. questionable (two or more cautions and/or maximum of one delay score); 2. Abnormal (two or more delay scores).	MLBW/35.6 (2.8) weeks GA: 48/553, 8.7% (95%CI 6.5-11.3) ≥1 delays: VLBW/28.4 (3.0) weeks GA: 29/367, 7.9% (95%CI 5.4-11.1) MLBW/35.6 (2.8) weeks GA: 29/553, 5.2% (95%CI 3.5-7.5) Gross motor (Denver II) ≥1 cautions: VLBW/28.4 (3.0) weeks GA: 64/367, 17.4% (95%CI 13.7-21.7) MLBW/35.6 (2.8) weeks GA: 49/553, 9.0% (95%CI 6.6-11.6) ≥1 delays: VLBW/28.4 (3.0) weeks GA: 39/367, 10.6% (95%CI 7.7-14.2) MLBW/35.6 (2.8) weeks GA: 22/553, 4.0% (95%CI 2.5-6.0)	
Stoelhorst 2003	Regional population-based	N=163 with PDI data at 18 months CA,	Mental and psychomotor development were assessed using the Dutch version of	At 18 months of corrected age Severe psychomotor delay PDI (BSID-I, <-2SD) <32 weeks GA: 29/163, 17.8% (12.3-24.5%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
	prospective cohort study	N=144 with PDI data at 24 months CA (N=266 children included in the cohort originally, N=235 survived)	the Bayley Scales of Infant Development I (BSID-I). These scales have a population mean of 100 and a SD of 16. A PDI of >=84 was considered normal, PDI 68-84 (-2 to -1 SD) was considered moderate delay and <68 (<-2SD) was considered severe delay	Moderate psychomotor delay PDI (-2 to -1 SD) <32 weeks GA: 18/163, 11.0% (6.7-16.9%)) At 24 months of corrected age Severe psychomotor delay PDI (BSID-I, <-2SD) <32 weeks GA: 12/144, 8.3% (4.4-14.1%) Moderate psychomotor delay PDI (-2 to -1 SD) <32 weeks GA: 32/144, 22.2% (15.7-29.9%)	
Zhu 2012	National Birth Cohort	n=22, 898 children with data included in the analysis	The DCDQ, a 15-item parent questionnaire designed to screen for coordination disorders in children aged 5–15 years, including playing ball (throwing, catching, hitting), writing (fast, legibly, with proper effort) was used. Parents were asked to provide their responses on a five-point Likert scale when comparing the motor performance between their child and his/her peers. A high score suggests no DCD. In the study, DCD total score of 46 or below defined children having probable DCD.	At 7 year follow-up DCD ≤31 weeks GA: 14/99, 14.1% (95%CI 8.0-22.6) 32 weeks GA: 6/46, 13.0% (95%CI 5.0-26.3) 33 weeks GA: 7/77, 11.7$ (95%CI 3.7-17.8) 34 weeks GA: 14/125, 11.2% (95%CI 6.3-18.1) 35 weeks GA: 10/185, 5.4% (95%CI 2.6-9.7) 36 weeks GA: 18/411, 4.4% (95%CI 2.6-6.8) 32-36 weeks GA: 55/844, 6.5% (5.0-8.4%)	Low
Evidence on developmental delay					
Agerholm 2011	Regional birth cohort study	N=237 live born children with 24-31 weeks GA in	Preschool skills were assessed using the cognitive parts of the Miller Assessment for	At 5 years age Preschool skills Cognitive verbal skills (Uncertain preschool skills, MAP, yellow)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		the geographical area N=204 children survived N=175 children followed-up at 5 years of age (86% of the ones who survived) N=168 children included in analysis (7 children with CP could not be assessed)	Preschoolers (MAP) with four items in the cognitive verbal area, fice items in the cognitive non-verbal area and four items in the combined motor and cognitive area. MAP is standardised and the scores are presented in relation to two different percentiles within the three area and administered by colours according to the manual: green shows normal preschool skills, yellow indicates observation for deficit in preschool skills and red indicates deficit in preschool skills.	24-31 weeks GA: 23/168, 13.7% (8.9-19.8%) Cognitive verbal skills (Deficit in preschool skills, MAP, red) 24-31 weeks GA: 18/168, 10.7% (6.5-16.4%) Cognitive non-verbal skills (Uncertain preschool skills, MAP, yellow) 24-31 weeks GA: 11/168, 6.6% (3.3-11.4%) Cognitive non-verbal skills (Deficit in preschool skills, MAP, red) 24-31 weeks GA: 6/168, 3.6% (1.3-7.6%) Combined cognitive and motor skills (Uncertain preschool skills, MAP, yellow) 24-31 weeks GA: 21/168, 12.5% (7.9-18.5%) Combined cognitive and motor skills (Deficit in preschool skills, MAP, red) 24-31 weeks GA: 12/168, 7.1% (3.8-12.1%)	
Charkaluk 2010	Population based prospective cohort study	N=634 children born alive at GA <33 weeks. n=546 surviving children included at follow-up.	Developmental quotients were ascertained by the revised Brunet-Lezine scale, an early childhood psychomotor development scale covering four domains of development: gross motor function, fine motor function, language and sociability. Four separate DQs could be calculated for children aged 2-30 months, which can be combined to give a global DQ. (Global DQ cut off not reported in paper; DQ ≤70 is defined as moderate	At 2 years (corrected age) Global DQ/developmental delay <70 (severe) <33 weeks GA: 8/347, 2.3% (1.0-4.5%) Global DQ/developmental delay <85 (moderate) <33 weeks GA: 62/347, 17.9% (14.0-22.0%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			developmental delay; DQ <70 is defined as severe developmental delay) Children were considered to have an achievement discrepancy if the difference between the global DQ and at least one partial DQ was a value obtained by only 5% of the reference sample.		
Johnson 2015	Prospective cohort study (LAMBS)	n=1130 late/moderately preterm infants recruited n=638 late/moderately preterm infants included in analysis	At 2 years (corrected age), cognitive impairment was assessed using the Parent Report of Children's Abilities-Revised (PARCA-R). Scores for non-verbal cognition and expressive language were combined to give a total parent report composite. These scores are strongly correlated with scores on gold standard developmental tests. Moderate/severe cognitive impairment was identified as a score corresponding to with PRC scores < 2.5th percentile in the term reference group.	At 2 years of corrected age Cognitive impairment (PARCA-R , <2.5 percentile) 32-36 weeks GA: 40/638, 6.3% (4.5-8.4%)	Low
Kerstjens 2011	Population based prospective cohort study	Sample recruited: n=698 gestation < 32 weeks	The Dutch version of the age 48 month form of the Ages and Stages questionnaire was used to assess development. The ASQ covers five domains:	At 4 years Developmental delay (ASQ total score <-2 SD) <32 weeks GA: 76/512, 14.9% (11.9-18.2%) 32-35 weeks GA: 77/927, 8.3% (6.6-10.3%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		n=1145 gestation 32-35 weeks Sample analysed after exclusions: n=512 gestation < 32 weeks n=927 gestation 32-35 weeks	communication, fine motor function, gross motor function, personal-social functioning and problem solving. The total score was calculated by adding all the domain scores and dividing by five. The individual domain scores, and the total score were dichotomized at 2SD below the mean score of the Dutch reference group as normal/abnormal		
Plomgaard 2006	National cohort study	n=78 in group 1 (<26 weeks GA) invited to the study n=61 in group 1 returned questionnaire n=78 in group 2 (26-27 weeks GA) invited to the study n=57 in group 2 returned questionnaire	The Ages and Stages Questionnaire (ASQ) was used addressing the domains of communication, gross motor skills, fine motor skills, problem solving and personal-social skills. The questionnaire was appropriate for the child's age was completed by the parents at home partly from memory and partly after doing short exercises with their child. Severe developmental deficit was classed as <-3SD, moderate to severe was classed as <-2SD in both preterm groups.	At 12-60 months age Developmental delay (ASQ <-3SD) (after correction for parental education) <26 weeks GA: 8/58, 14% (95%CI 5-23) 26-27 weeks GA: 2/56, 4% (95%CI 0-8) Developmental delay (ASQ <-2SD) (after correction for parental education) <26 weeks GA:13/58, 22% (95%CI 12-33) 26-27 weeks GA: 7/56, 13% (95%CI 4-21) Developmental delay (ASQ <-3SD) (after exclusion of children with neurosensory deficit) <26 weeks GA: 3/51, 6% (95%CI 0-12) 26-27 weeks GA: 2/55, 4% (95%CI 0-9) Developmental delay (ASQ <-2SD) (after exclusion of children with neurosensory deficit) <26 weeks GA: 7/51, 14% (95%CI 0.5-23) 26-27 weeks GA: 7/55, 13% (95%CI 0-22)	Very low
Potijk 2013	Prospective cohort study	N=926 moderately preterm	Developmental outcomes were measured using the Dutch version of the 48-	At 4 years of age Developmental delay (ASQ total score <-2SD)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		children assessed at 4 years. (N=544 term born controls)	month form of the Ages and Stages Questionnaire (ASQ) which is a validated, parent-completed developmental screening instrument. Five developmental domains: fine motor, gross motor, communication, problem-solving, and personal-social skills. Each domain consists of 6 questions on developmental milestones. ASQ total score was computed by taking the mean of the 5 domain scores. For the total score and the domains scores cut-offs for normal and abnormal scores were set at 2 SD below the mean score of the Dutch reference group.	32-35 weeks GA: 74/891, 8.3% (6.6-10.3%)	
Schendel 1997	Regional prospective cohort study	n=367 very low birth weight children (<1500 g) with Denver II assessment at follow-up n= 553 moderately low birth weight children (1500-2499 g)	The Denver II was used to screen for possible developmental delay by comparing the child's performance on various tasks with children of the same adjusted age. 9 outcomes indicating delay were based on four domains: Personal-social, language, fine motor-adaptive skills, and gross motor skills. The 9 outcomes reflected two types of delay:	At adjusted age 15 months (range 9-34 months) Developmental delay (Overall performance, Denver II) Questionable (≥2 cautions and/or 1 delay score): VLBW/28.4 (3.0) weeks GA: 64/367, 17.4% (95%CI 13.7-21.7) MLBW/35.6 (2.8) weeks GA: 65/553, 11.8% (95%CI 9.2-14.7) Abnormal (≥2 delay scores): VLBW/28.4 (3.0) weeks GA: 40/367, 11.0% (95%CI 7.9-14.6) MLBW/35.6 (2.8) weeks GA: 32/553, 5.8% (95%CI 4.0-8.1)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		with Denver II assessment at follow-up	1. A moderate delay (overall questionable performance + four domain specific outcomes for children who received one or more caution scores in a given domain); 2. Severe delay (abnormal overall test performance +the four domain specific outcomes for children who received one or more delay scores in a given domain The overall performance was based on total number of caution and/or delay scores across all domains and was categorised as: 1. questionable (two or more cautions and/or maximum of one delay score); 2. Abnormal (two or more delay scores).		
Evidence on language delay					
Joseph 2016b	Prospective cohort study (ELGAN)	N=1506 infants n=1198 survived to age 10 years n=873 assessed at 10 years	Language ability: Expressive and receptive language skills were evaluated with the Oral and Written Language Scales, 30 which assess semantic, morphologic, syntactic, and pragmatic production and comprehension of elaborated sentences	At 10 years age Language (<28 weeks GA; <=-2SD) OWLS Listening Comprehension: 166/873, 19% (95%CI 16.5-21.8) OWLS Oral Expression: 166/873, 19% (95%CI 16.5-21.8)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
Mansson 2014	Population based cohort study	N=707 n=461 eligible for follow-up n=399 children born at <27 weeks GA (after exclusions, surviving to age 2.5 years and had BSID III assessment)	Bayley-III was used to assess five subtests: Cognition, Receptive Communication, Expressive Communication, Fine Motor, and Gross Motor. Test scores were evaluated on the basis of the means and standard deviations of the controls. Function level was regarded as normal if the subtest scaled score was ≤+1 SD and ≥1 SD of the control mean. Mild delay was classed as ≤1SD to ≥2 SD, moderate delay was classed as <2SD to ≥3 SD, and severe delay was classed as <3SD.	At 2.5 years age Receptive communication (BSID III mild -1SD to 2SD) <27 weeks GA:98/394, 24.9% (95%CI 20.7-30.0) Receptive communication (BSID III moderate -2SD to 3SD) <27 weeks GA: 36/394, 9.1% (95%CI 6.5-12.4) Receptive communication (BSID III moderate to severe -3SD) <27 weeks GA: 23/394, 5.8% (95%CI 3.7-8.6) Expressive communication (BSID III mild -1 SD to 2SD) <27 weeks GA: 123/393, 31.3% (95%CI 26.7-36.1) Expressive communication (BSID III moderate -2SD to 3SD) <27 weeks GA: 32/393, 8.1% (95%CI 5.6-11.3) Expressive communication (BSID III moderate to severe -3SD) <27 weeks GA: 25/393, 6.4% (95%CI 4.2-9.3)	Low
Potijk 2013	Prospective cohort study	N=926 moderately preterm children assessed at 4 years. (N=544 term born controls)	Developmental outcomes were measured using the Dutch version of the 48-month form of the Ages and Stages Questionnaire (ASQ) which is a validated, parent-completed developmental screening instrument. Five developmental domains: fine motor, gross motor, communication, problem-solving, and personal-social skills. Each domain consists of 6 questions on developmental milestones. ASQ total score was computed by taking the	At 4 years of age Communication delay (ASQ, <-2SD) 32-35 weeks GA: 86/906, 9.5% (7.7-11.6%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			mean of the 5 domain scores. For the total score and the domains scores cut-offs for normal and abnormal scores were set at 2 SD below the mean score of the Dutch reference group.		
Rautava 2010	Population based prospective cohort study	Original sample size: n=924 preterm/very low birth weight infants Included in follow-up: n=588 preterm/very low birth weight infants	Behavioural outcomes were assessed using the Five to Fifteen Questionnaire (FTF), which was completed by the parents. Questions on development and behaviour were rated by the parents as 0="does not describe", 1="describes to some extent" and 2="describes well" the individual child.	At 5 years age Language problems (FTF) <32 weeks GA: 27/588, 4.6% (95%CI 3.1-6.6)	Low
Schendel 1997	Regional prospective cohort study	n=367 very low birth weight children (<1500 g) with Denver II assessment at follow-up n= 553 moderately low birth weight children (1500-2499 g) with Denver II	The Denver II was used to screen for possible developmental delay by comparing the child's performance on various tasks with children of the same adjusted age. 9 outcomes indicating delay were based on four domains: Personal-social, language, fine motor-adaptive skills, and gross motor skills. The 9 outcomes reflected two types of delay: 1. A moderate delay (overall questionable performance +	At adjusted age 15 months (range 9-34 months) Language delay (Denver II) ≥1 cautions: VLBW/28.4 (3.0) weeks GA: 62/367, 17.0% (95%CI 13.2-21.1) MLBW/35.6 (2.8) weeks GA: 66/553, 11.9% (95%CI 9.4-14.9) ≥1 delays: VLBW/28.4 (3.0) weeks GA: 32/367, 8.7% (95%CI 6.0-12.1) MLBW/35.6 (2.8) weeks GA: 32/553, 5.8% (95%CI 4.0-8.1)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		assessment at follow-up	four domain specific outcomes for children who received one or more caution scores in a given domain); 2. Severe delay (abnormal overall test performance +the four domain specific outcomes for children who received one or more delay scores in a given domain The overall performance was based on total number of caution and/or delay scores across all domains and was categorised as: 1. questionable (two or more cautions and/or maximum of one delay score); 2. Abnormal (two or more delay scores).		
Stene-Larsen 2014	Prospective population-based pregnancy cohort study	questionnaires from gestational week 17 (n=101 624), child age 18 months (n=64 970) n=39,423 children (1673 born late preterm, 7109 born early preterm)	At 18 months, Child communication impairments were measured using selected items from the Ages and Stages Questionnaire (ASQ) which included receptive communication skills, and expressive communication skills. The selection of items for the MoBa study was performed a priori by specialists in clinical and developmental psychology. Mothers were asked to find	At age 18 months Communication impairment (ASQ) (≥ 2SD) 34-36 weeks GA: 122/1673, 7.3% (95%CI 6.1-8.6) At 36 months Communication impairment (ASQ ≥2SD) 34-36 weeks GA: 105/1673, 5.5% (95%CI 5.2-7.6)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			time to observe the child and rate the extent to which the child would typically show mastery of the skill in question, using the response categories "yes" (1), "very often" (2), "not yet" (3), and "I don't know" (missing). To identify those children at risk for clinically significant communication impairments, a cut-off at 2 SD above the cohort mean was set At 36 months, infants were assessed using 6 items from the ASQ measuring expressive (3 items) and receptive (3 items) communication skills. To identify the children at risk for clinically significant communication impairments, a cut-off of 2 SD above the cohort mean was set		
Evidence on executive function					
Anderson 2004	Geographically determined cohort study	N=275 final sample	Behaviour Rating Inventory of Executive Function (BRIEF) is a questionnaire that assesses behavioural manifestations of executive function. In this study the parent version was administered. Composite score (global executive composite) is derived from 8	At 8 years (corrected) Global executive composite (BRIEF, >=1.5SD above normative mean) <28 weeks GA/BW <1000 g: 32/245, 13.1% (9.1-17.9%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			clinical scales (inhibit, shift, emotional control, initiate, working memory, plan/organize, organization of materials and monitor) and 2 indices (metacognitive and behavioural regulation). Score >065 (>=1.5 SD above normative mean) is considered abnormal.		
Anderson 2011	Population-based cohort study	n=201 children survived to 8 years n=189 assessed at 8 years (94%)	Executive attention was categorised into 1) inhibitory control, which was assessed with the Opposite Worlds from the TEA-Ch, and the Inhibit scale from the parent form of the Behavioural Rating Inventory of Executive Function (BRIEF), 2) shifting attention, which was assessed with Creature Counting from the TEA-Ch, and the Shift scale from BRIEF, 3) divided attention, which was assessed with the Sky Search Dual Task from the TEA-Ch	At 8 years corrected age Executive attention 1) Inhibitory control: a) Opposite Worlds (<-1SD) 22-27 weeks GA/BW 1000 g: 10/167, 6.0% (2.9-10.7%)* b) BRIEF-Inhibit (T score >60) 22-27 weeks GA/BW 1000 g: 28/187 15.0% (10.2-20.9%)* 2) Shifting attention: a) Creature counting (<-1SD) 22-27 weeks GA/BW 1000 g: 46/170, 27.1% (20.5-34.4%)* b) BRIEF-Shift (T score >60) 22-27 weeks GA/BW 1000 g: 35/184, 19.0% (13.6-25.5%)* 3) Divided attention: Sky Search Dual Task (<1SD) 22-27 weeks GA/BW 1000 g: 62/168, 36.9% (29.6-44.7%)*	Low
Joseph 2016b	Population based cohort study (ELGAN)	N=1506 infants n=1198 survived to age 10 years n=873 assessed at 10 years	Executive function: Attention and executive functions were assessed with the DAS-II and the Developmental NEuroPSYchological Assessment-II (NEPSY-II).31 DAS-II Recall of Digits Backward and Recall of	At 10 years age Executive function (<28 weeks GA; <=-2SD): DAS-II Working Memory: 157/873, 18% (95%CI 15.5-20.7) NEPSY-II Auditory Attention: 201/873, 23% (95%CI 20.3-26.0) NEPSY-II Auditory Response Set: 175/873, 20% (95%CI 17.4-23) NEPSY-II Inhibition Inhibition: 297/873, 34% (95%CI 31-37)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			Sequential Order measured verbal working memory. The NEPSY-II Auditory Attention and Auditory Response Set evaluated auditory attention, set switching, and inhibition. NEPSY-III Inhibition Inhibition and Inhibition Switching assessed simple inhibition and inhibition in the context of set shifting, respectively. The NEPSY-II Animal Sorting measured concept generation and mental flexibility. Speed of processing: Speed of processing was assessed with NEPSY-II Inhibition Naming, a baseline measure of processing speed with no inhibitory component. Visual perception: NEPSY-II Arrows, which measures perception of line orientation, and Geometric Puzzles, a measure of mental rotation of complex visual spatial figures.	NEPSY-II Inhibition Switching: 236/979, 27% (95%CI 24.1-30.1) Processing speed (<28 weeks GA; <=-2SD): NEPSY-II Inhibition Naming: 270/873, 31% (95%CI 28-34) Visual perception (<28 weeks GA; <=-2SD): NEPSY-II Arrows: 227/873, 26% (95%CI 23-29) NEPSY-II Geometric Puzzles: 148/873, 17.0% (95%CI 14.5-19.6)	
Rautava 2010	Population based prospective cohort study	Original sample size: n=924 preterm/very low birth weight infants Included in follow-up:	Behavioural outcomes were assessed using the Five to Fifteen Questionnaire (FTF), which was completed by the parents. Questions on development and behaviour were rated by the parents as 0="does not describe",	At 5 years age Executive function problems (FTF) <32 weeks GA: 46/588, 7.8% (95%CI 5.8-10.3) Perception problems (FTF) <32 weeks GA: 23/588, 3.9% (95%CI 2.5-5.8) Memory problems (FTF) <32 weeks GA: 49/588, 8.3% (95%CI 6.2-11.0)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		n=588 preterm/very low birth weight infants	1="describes to some extent" and 2="describes well" the individual child.		
Evidence on behavioural, social, emotional, attention problems					
Anderson 2011	Population-based cohort study	n=201 children survived to 8 years n=189 assessed at 8 years (94%)	Selective attention was assessed with the Sky Search subtest from the Test of Everyday Attention for Children (TEA-Ch). Sustained attention was assessed with the Score Sub-test from the TEA-Ch. Attention encoding was assessed with the forward digit span from the Wechsler Intelligence Scale for Children (WISC-IV). Executive attention was categorised into 1) inhibitory control, which was assessed with the Opposite Worlds from the TEA-Ch, and the Inhibit scale from the parent form of the Behavioural Rating Inventory of Executive Function (BRIEF), 2) shifting attention, which was assessed with Creature Counting from the TEA-Ch, and the Sgift scale from BRIEF, 3) divided attention, which was assessed with the Sky Search Dual Task from the TEA-Ch.	At 8 years (corrected) Selective attention (TEA-Ch Sky Search, <-1SD) 22-27 weeks GA/BW 1000 g: 58/171, 33.9% (26.9-41.5%) Sustained attention (TEA-Ch Score, <-1SD) 22-27 weeks GA/BW 1000 g: 52/173, 30.1% (23.3-37.5%) Attention Encoding (TEA-Ch Forward digit span, <-1SD) 22-27 weeks GA/BW 1000 g: 71/178, 39.9% (32.6-47.5%) Executive attention 1) Inhibitory control: a) Opposite Worlds (<-1SD) 22-27 weeks GA/BW 1000 g: 10/167, 6.0% (2.9-10.7%)* b) BRIEF-Inhibit (T score >60) 22-27 weeks GA/BW 1000 g: 28/187 15.0% (10.2-20.9%)* 2) Shifting attention: a) Creature counting (<-1SD) 22-27 weeks GA/BW 1000 g: 46/170, 27.1% (20.5-34.4%)* b) BRIEF-Shift (T score >60) 22-27 weeks GA/BW 1000 g: 35/184, 19.0% (13.6-25.5%)* 3) Divided attention: Sky Search Dual Task (<1SD) 22-27 weeks GA/BW 1000 g: 62/168, 36.9% (29.6-44.7%)* ADHD symptoms CADS-P Inattentive symptoms (T score >60) 22-27 weeks GA/BW 1000 g: 18/56, 32.1% (20.3-46.0%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			Attention deficit hyperactivity disorder (ADHD) was assessed with the Conner's ADHD/DSM-IV Scales (CADS-P). The CADS-P consists of 26 items. For this study three scales were used: ADHD Index (items that best distinguish ADHD children from nonclinical children), DSM-IV Inattentive (items directly related to the DSM-IV symptoms of inattention), and DSM-IV Hyperactive-Impulsive (items directly related to DSM-IV symptoms of hyperactivity-impulsivity). Impairment was defined as scores more than 1 SD below the mean of the control group (term/normal birth weight peers) for the attention tasks and T scores >60 for the BRIEF and the CADS-P.	CADS-P Hyperactive-Impulsive symptoms (T score >60) 22-27 weeks GA/BW 1000 g: 23/55, 41.8% (28.7-55.9%) ADHD Index (CADS-P T score >60) 22-27 weeks GA/BW 1000 g: 24/55, 43.6% (30.3-57.7%)	
Anderson 2003 (Victorian Infant Collaborative Study group)	Prospective regional cohort study	N=568 consecutive live births of neonates with BW <1000g or <28 weeks GA. n=298 infants survived to 2,	The behaviour assessment system (BASC; parent and teacher rating scales) were used to assess children's adaptive and problem behaviours at home (parent) or at school (teacher). Both scales provide composite indexes for externalising	At 8 years age Behavioural problems- at risk (parent reported) <28 weeks GA: 41/275, 15% (95%CI 11.0-19.7) Behavioural problems-clinically significant (parent reported) <28 weeks GA: 19/275, 7% (95%CI 4.2-10.6)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		and 5 years assessment. n=275 children assessed at 8 years age.	problems, internalising problems, adaptive skills, and overall behavioural problems. For behavioural problems, T scores of 70 + are considered clinically significant, whereas T scores of 60-69 represent at risk range. For adaptive index, a T score of 30 or below is clinically significant, whereas a T score of 31-40 represents at risk range		
De Kleine 2003	Prospective cohort study	n=566 eligible children n=431 assessed at 5 years (76%) n=404 assessed for motor functioning (M-ABC) n=402 assessed for IQ (IQ test) n=407 assessed for behavioural problems (CBCL)	At 5 years, behavioural problems were assessed with the full Child Behaviour Checklist (CBCL) by trained child psychologists. Total scores up to and including 59 are considered normal, from 60 up to and including 63 intermediate and from 64 upwards "clinically important" disturbance of behaviour.	At 5 years Total behavioural problems (CBCL, score >=65) <32 weeks GA/bw <1500 g: 56/407, 56/407, 13.8% (10.6-17.5%)	Moderate
Delobel-Ayoub 2009	Population based prospective cohort study	n=2276 preterm infants born at 22-32 weeks	The French version of the Strengths and Difficulties Questionnaire (SDQ) was completed by one or both	At 5 years Total behavioural difficulties (SDQ, 10th percentile) 22-32 weeks GA: 240/1095, 21.9% (19.5-24.5%) Hyperactivity (SDQ, 10th perc)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		originally recruited n=1690 children's parent(s) completed questionnaire n=1102 preterm children included in analysis after exclusions	parents' (98%) or another caregiver (2%). Scores from the four symptom scales (hyperactivity/inattention, conduct, emotional and peer problems) are summed to provide a "total difficulties" score, with higher scores indicating poorer mental health. Cut-offs were defined based on the 10th percentile of the observed scores in the control group	22-32 weeks GA: 198/1096, 18.1% (15.8-20.5%) Conduct problem (SDQ, 10th perc) 22-32 weeks GA: 123/1097, 11.2% (9.4-13.2%) Emotional symptoms (SDQ, 10th perc) 22-32 weeks GA: 228/1096, 20.8% (18.4-23.3%) Peer problems (SDQ, 10th perc) 22-32 weeks GA: 220/1097, 20.1% (17.7-22.6%) Prosocial behaviour (SDQ, 10th perc) 22-32 weeks GA: 169/1095, 15.4% (13.3-17.7%)	
Delobel-Ayoub 2006	Population based prospective cohort study	N=2382 very preterm infants originally survived to discharge N=1880 children's parent(s) completed the questionnaire N=1228 very preterm singletons included in analysis after exclusions	The French version of the Strengths and Difficulties Questionnaire (SDQ) for 3- to 4-year-old children was completed by parents. Scores from the four symptom scales (hyperactivity/inattention, conduct, emotional and peer problems) are summed to provide a "total difficulties" score, with higher scores indicating poorer mental health. Cut-offs were defined based on the 10th percentile of the observed scores in the control group	At 3 years Total behavioural difficulties, (SDQ, 10th percentile) <33 weeks GA: 240/1202, 20.0% (17.7-22.3%) 24-28 weeks GA: 66/274, 24.1% (19.2-29.6%) 29-30 weeks GA: 57/338, 16.9% (13.0-21.3%) 31-32 weeks GA: 112/590, 19.0% (15.9-22.4%) Hyperactivity (SDQ, 10th perc) <33 weeks GA: 241/1205, 20.0% (17.8-22.4%) 24-28 weeks GA: 66/274, 24.1% (19.2-29.6%) 29-30 weeks GA: 58/339, 17.1% (13.3-21.6%) 31-32 weeks GA: 112/592, 18.9% (15.8-22.3%) Conduct problem (SDQ, 10th perc) <33 weeks GA: 193/1207, 16.0% (14.0-18.2%) 24-28 weeks GA: 44/274, 16.1% (11.9-21.0%) 29-30 weeks GA: 54/340, 15.9% (12.2-20.2%) 31-32 weeks GA: 89/593, 15.0% (12.2-18.1%) Emotional symptoms (SDQ, 10th perc) <33 weeks GA: 181/1207, 15.0% (13.0-17.1%) 24-28 weeks GA: 47/274, 17.2% (12.9-22.2%) 29-30 weeks GA: 48/340, 14.1% (10.6-18.3%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				31-32 weeks GA: 89/593, 15.0% (12.2-18.1%) Peer problems (SDQ, 10th perc) <33 weeks GA: 168/1203, 14.0% (12.1-16.1%) 24-28 weeks GA: 49/274, 17.9% (13.5-22.9%) 29-30 weeks GA: 44/339, 13.0% (9.6-17.0%) 31-32 weeks GA: 71/590, 12.0% (9.5-14.9%) Prosocial behaviour (SDQ, 10th perc) <33 weeks GA: 181/1205, 15.0% (13.1-17.2%) 24-28 weeks GA: 55/274, 20.1% (15.5-25.3%) 29-30 weeks GA: 54/339, 15.9% (12.2-20.3%) 31-32 weeks GA: 77/592, 13.0% (10.4-16.0%)	
Downey 2015	Population based cohort study (ELGAN)	N=826 children born preterm	At 24 months adjusted age, a parent/caregiver completed the CBCL for child behaviour problems. Five of the items on the CBCL are included in the attention problem scale (can't concentrate, can't sit still, clumsy, quickly shifts, wanders away). Scores between the 93rd and 97th percentile correspond to the borderline/subclinical range and are considered worthy of concern, and scores above the 97th percentile warrant definite concern. For this report, a child was considered to have an attention problem if his/her score was at or greater than the 93rd percentile.	At 24 months adjusted age Attention problems (assessed using CBCL =>93rd percentile) <28 weeks GA: 88/826, 10.7% (95%CI 8.6-13.0)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
Farooqi 2007	Nationally-representative population-based cohort study	Total sample: n=169 Extremely immature (EI) children born before 26 completed weeks of gestation (n= 83)	For assessment of the parents' and teachers' perceptions of the children's behaviour, the parents completed the Child Behaviour Checklist (CBCL) for ages 4 to 18 years and the teachers completed the analogous Teacher Report Form (TRF). For all TRF and CBCL problem subscales, scores above the 90th percentile for the control subjects of the same gender were classified as being in the abnormal range. The percentile distribution of the total CBCL problem scores for our control group was similar to that for a Swedish reference population. Children completed a self-report with a depression self-rating scale (DSRS).32 The DSRS is an 18-item self-report questionnaire composed of a psychiatric symptom checklist that measures anxiety and depression. The child is asked to rate his or her own situation during the past month, on a 3-point scale. Scores of 2, 1, and 0 refer to most of the time, sometimes, and never,	At 11 years Parents' report Total behavioural problems (CBCL, 90th perc) <26 weeks GA: 24/83, 28.9% (19.5-39.9%) Anxious/depressed (CBCL, 90th perc) <26 weeks GA: 22/83, 26.5% (17.4-37.4%) Withdrawn (CBCL, 90th perc) <26 weeks GA: 30/83, 36.1% (25.9-47.4%) Somatic complaints (CBCL, 90th perc) <26 weeks GA: 11/83, 13.3% (6.8-22.5%) Social problems (CBCL, 90th perc) <26 weeks GA: 21/83, 25.3% (16.4-36.0%)) Thought problems (CBCL, 90th perc) <26 weeks GA: 16/83, 19.3% (11.4-29.4%) Attention problems (CBCL, 90th perc) <26 weeks GA: 25/83, 30.1% (20.5-41.2%) Aggressive behaviour (CBCL, 90th perc) <26 weeks GA: 11/83, 13.3% (6.8-22.5%) Delinquent behaviour (CBCL, 90th perc) <26 weeks GA: 9/83, 10.8% (5.1-19.6%) Internalising (CBCL, 90th perc) <26 weeks GA: 27/83, 32.5% (22.7-43.7%) Externalising (CBCL, 90th perc) <26 weeks GA: 8/83, 9.6% (4.3-18.1%) Teachers' report Total behavioural problems (TRF, 90th perc) <26 weeks GA: 20/83, 24.1% (15.4-34.7%) Anxious/depressed (TRF, 90th perc) <26 weeks GA: 19/83, 22.9% (14.4-33.4%) Withdrawn (TRF, 90th perc)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			respectively. For the DSRS, scores above the 90th percentile for the control subjects of the same gender were classified as being in the abnormal range. School difficulties was defined as the child repeating a grade and/or using special educational resources (full-time or part-time). Attending special class or special school means attending a special school or training school for the physically disabled and severely mentally retarded or receiving full-time special education attached to the mainstream school.	<26 weeks GA: 19/83, 22.9% (14.4-33.4%) Somatic complaints (TRF, 90th perc) <26 weeks GA: 17/83, 20.5% (12.4-30.8%) Social problems (TRF, 90th perc) <26 weeks GA: 17/83, 20.5% (12.4-30.8%) Thought problems (TRF, 90th perc) <26 weeks GA: 25/83, 30.1% (20.5-41.2%) Attention problems (TRF, 90th perc) <26 weeks GA: 20/83, 24.1% (15.4-34.7%) Aggressive behaviour (TRF, 90th perc) <26 weeks GA: 17/83, 20.5% (12.4-30.8%) Delinquent behaviour (TRF, 90th perc) <26 weeks GA: 19/83, 22.9% (14.4-33.4%) Internalising (TRF, 90th perc) <26 weeks GA: 21/83, 25.3% (16.4-36.0%) Externalising (TRF, 90th perc) <26 weeks GA: 15/83, 18.1% (10.5-28.1%) Children's self-reported depression scale abnormal score (DSRS) <26 weeks GA: 10/83, 12.1% (5.9-21.0%)	
Foix-Helias 2008	Prospective population based cohort study	n=1645 children with data on behavioural difficulties (72% of the n=2300 survivors up to follow-up)	Total behavioural difficulties were assessed using the French version of the Strengths and Difficulties Questionnaire (SDQ) completed by parents. This questionnaire includes 25 items structured into five scales which assess hyperactivity-inattention, conduct problems, emotional symptoms, peer	At 5 years Total behavioural difficulties (SDQ, 10th percentile) 24-32 weeks GA: 348/1645, 21.2% (19.2-23.2%) 24-27 weeks GA: 52/234, 22.2% (17.1-28.1%) 28-32 weeks GA: 296/1411, 21.0% (18.9-23.2%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			problems and prosocial behaviour. Scores for the first four symptom scales are summed to provide an overall difficulties score with a range of 0-40. The cut-offs were defined such that about 10% of the children in contemporaneous reference group of children born at term (born between 39 and 40 weeks of GA) were considered at high risk of having a behavioural problem.		
Guellec 2011	Population based prospective cohort study	N=2855 live births at 24-32 weeks GA. n=2357 infants eligible for follow-up	Inattention-hyperactivity symptoms, assessed with the French version of the Strength and Difficulties Questionnaire completed by the parents. Total behavioural difficulties, including a sum score of scales on hyperactivity-inattention, conduct, emotional and peer problems, assessed with the French version of the Strength and Difficulties Questionnaire completed by the parents.	At 5 years age Inattention-hyperactivity symptoms SGA children (bw <10th percentile) 24-28 weeks GA: 4/21, 19% (5.5-42.0%) 29-32 weeks GA: 27/115, 23.5% (16.0-32.3%) MGA children (bw 10th-19th percentile) 24-28 weeks GA: 7/33, 21.2% (9.0-38.9%) 29-32 weeks GA: 19/121, 15.7% (9.7-23.4%) AGA (bw >=20th percentile) 24-28 weeks GA: 75/346, 21.7% (17.5-26.4%) 29-32 weeks GA: 156/1041, 15.0% (12.9-17.3%) Total behavioural difficulties SGA children (bw <10th percentile) 24-28 weeks GA: 7/21, 33.3% (14.6-57%) 29-32 weeks GA: 22/115, 19.1% (12.4-27.5%) MGA children (bw 10th-19th percentile)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				24-28 weeks GA: 9/33, 27.3% (13.3-45.5%) 29-32 weeks GA: 32/121, 26.5% (18.8-35.2%) AGA (bw >=20th percentile) 24-28 weeks GA: 82/346, 23.7% (19.3-28.5%) 29-32 weeks GA: 201/1037, 19.4% (17.0-21.9%)	
Guy 2015	Population-based prospective cohort	n=1130 late and moderately preterm infants recruited n=634 late and moderately preterm infants in the final sample	ASD/behaviour The M-CHAT 23 item parent questionnaire was used to identify early behaviours associated with ASD. Infants failing ≥2 of 6 critical items or ≥3 items overall screen positive for the risk of ASD. The interview took 5-15 minutes after which the MCHAT was re-scored and children with positive screens after follow-up were classified as true positives	At 2 years age ASD behaviour positive screen (MCHAT) 32-33 weeks GA: 8/86, 9.3% (95%CI 4.1-17.5) 34-36 weeks GA: 84/548, 15.3% (95%CI 12.4-18.6) 32-26 weeks GA: 92/634, 14.5% (95%CI 12.0-17.5)	Low
Higa Diez 2016	Prospective cohort design	n=34163 neonates born in Japan in 2001 of which n=356 born at <34 weeks n=1287 born at 34-36 weeks n=9885 born at 37-38 weeks n=22635 born at 39-41	Some questions of the standardised and validated version of the Child Behaviour Checklist 9CBCL) 4-18 for Japan was used. A total of 7 behavioural outcomes were used, three related to attention problems: 1) interrupting people, 2) inability for the child to wait his/her turn during play, and 3) failure to pay attention to the surrounding area when crossing a street; and four related to	At 8 years Attentional problems Interrupting people (CBCL) <34 weeks GA: 149/356, 41.9% (36.7-47.2%) 34-36 weeks GA: 519/1287, 40.3% (37.6-43.1%) 39-41 weeks GA (term): 8718/22635, 38.5% (37.9-39.2%) Inability to wait his/her turn <34 weeks GA: 45/356, 12.6% (9.4-16.6%) 34-36 weeks GA: 117/1287, 9.1% (7.6-10.8%) 39-41 weeks GA (term): 1359/22635, 6.0% (5.7-6.3%) Failure to pay attention crossing street <34 weeks GA: 81/356, 22.8% (18.5-27.5%) 34-36 weeks GA: 265/1287, 20.6% (18.4-22.9%) 39-41 weeks GA (term): 4306/22635, 19.0% (18.5-19.5%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		weeks (reference population	delinquent/aggressive behaviour: 1) lying, 2) destroying toys or books, 3) hurting other people, and 4) causing disturbances in public. Binary outcomes for each were used. Combined outcome for both attention and delinquent/aggressive behaviour was also used, defined as participants who present adverse for all attention or delinquent/aggressive behaviours.	Adverse outcomes for all attentional problems <34 weeks GA: 17/181, 9.4% (5.6-14.6%) 34-36 weeks GA: 38/683, 5.6% (4.0-7.6%) 39-41 weeks GA (term): 367/12119, 3.0% (2.7-3.4%) Delinquent/aggressive behaviours Lying <34 weeks GA: 100/356, 28.1% (23.5-33.1%) 34-36 weeks GA: 347/1287, 27.0% (24.6-29.5%) 39-41 weeks GA (term): 5621/22635, 24.8% (24.3-25.4%) Destroying toys/books <34 weeks GA: 54/356, 15.2% (11.6-19.3%) 34-36 weeks GA: 162/1287, 12.6% (10.8-14.5%) 39-41 weeks GA (term): 2088/22635, 9.2% (8.9-9.6%) Hurting other people <34 weeks GA: 51/356, 14.3% (10.9-18.4%) 34-36 weeks GA: 164/1287, 12.7% (11.0-14.7%) 39-41 weeks GA (term): 2381/22635, 10.5% (10.1-10.9%) Disturbance in public <34 weeks GA: 88/356, 24.7% (20.3-29.5%) 34-36 weeks GA: 327/1287, 25.4% (23.1-27.9%) 39-41 weeks GA (term): 4417/22635, 19.5% (19.0-20.0%) Adverse outcomes for all delinquent/aggressive behaviours <34 weeks GA: 11/194, 5.7% (2.9-9.9%) 34-36 weeks GA: 24/714, 3.4% (2.2-5.0%) 39-41 weeks GA (term): 273/13472, 2.0% (1.8-2.3%)	
Hornman 2016	Population-based cohort study	n=1054 preterm children (n=653 moderately preterm	Emotional and behavioural problems were assessed with the validated Dutch version of the Child Behaviour Checklist (CBCL), applicable for ages 1.5-5 years. The CBCL consists of	At 4 and 5 years of age Emerging total behavioural problems (CBCL >=84th percentile) (normal score at 4 years, abnormal score at 5 years) 25-35 weeks GA: 45/1054, 4.3% (3.1-5.7%) 25-31 weeks GA: 21/401, 5.2% (3.3-7.9%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		children [32-35 weeks] n=401 early preterm children [25-31 weeks]) n=389 term children as comparisons	99 problem items, each item can be rated by the parents as not true (0), somewhat/sometimes true (1), or very/often true (2). From these ratings, the total, internalising, and externalising problem scales were constructed. >=84th percentile of the scale was considered subclinical or clinical. The dichotomised CBCL outcomes at ages 4 and 5 years were combined, resulting in 4 categories: consistently normal (normal score at both 4 and 5 years), emerging problems (normal score at 4 years, abnormal score at 5 years), resolving problems (abnormal score at 4 years, normal score at 5 years), and persistent problems (abnormal score at both 4 and 5 years).	32-35 weeks GA: 24/653, 3.7% (2.4-5.4%) Resolving total behavioural problems (CBCL >=84th percentile) (abnormal score at 4 years, normal score at 5 years) 25-35 weeks GA: 79/1054, 7.5% (6.0-9.3%) 25-31 weeks GA: 22/401, 5.5% (3.5-8.2%) 32-35 weeks GA: 57/653, 8.7% (6.7-11.2%) Persistent total behavioural problems (CBCL >=84th percentile) (abnormal score at 4 and 5 years) 25-35 weeks GA: 76/1054, 7.2% (5.7-8.9%) 25-31 weeks GA: 33/401, 8.2% (5.7-11.4%) 32-35 weeks GA: 43/653, 6.6% (4.8-8.8%) Emerging internalising problems (CBCL >=84th percentile) (normal score at 4 years, abnormal score at 5 years) 25-35 weeks GA: 76/1054, 7.2% (5.7-8.9%) 25-31 weeks GA: 32/401, 8.0% (5.5-11.1%) 32-35 weeks GA: 44/653, 6.7% (4.9-8.9%) Resolving internalising problems (CBCL >=84th percentile) (abnormal score at 4 years, normal score at 5 years) 25-35 weeks GA: 78/1054, 7.4% (5.9-9.2%) 25-31 weeks GA: 29/401, 7.2% (4.9-10.2%) 32-35 weeks GA: 49/653, 7.5% (5.6-9.8%) Persistent internalising problems (CBCL >=84th percentile) (abnormal score at 4 and 5 years) 25-35 weeks GA: 113/1054, 10.7% (8.9-12.8%) 25-31 weeks GA: 47/401, 11.7% (8.7-15.3%) 32-35 weeks GA: 66/653, 10.1% (7.9-12.7%)	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				Emerging externalising problems (CBCL >=84th percentile) (normal score at 4 years, abnormal score at 5 years) 25-35 weeks GA: 56/1054, 5.3% (4.0-6.8%) 25-31 weeks GA: 21/401, 5.2% (3.3-7.9%) 32-35 weeks GA: 35/653, 5.4% (3.8-7.4%) Resolving externalising problems (CBCL >=84th percentile) (abnormal score at 4 years, normal score at 5 years) 25-35 weeks GA: 76/1054, 7.2% (5.7-8.9%) 25-31 weeks GA: 21/401, 5.2% (3.3-7.9%) 32-35 weeks GA: 55/653, 8.4% (6.4-10.8%) Persistent externalising problems (CBCL >=84th percentile) (abnormal score at 4 and 5 years) 25-35 weeks GA: 88/1054, 8.4% (6.8-10.2%) 25-31 weeks GA: 33/401, 8.2% (5.7-11.4%) 32-35 weeks GA: 55/653, 8.4% (6.4-10.8%)	
Hutchinson 2013 (Victorian Infant Collaborative study group)	Prospective cohort study	n=189 preterm/low birth weight cohort (94% eligible for follow-up; 12 children were not seen, but 10/12 were assessed at 2 years (corrected age)).	Behavioural outcomes were assessed by using Strengths and Difficulties Questionnaire (SDQ). This 25-item parent-rated questionnaire has 5 scales: emotional symptoms, conduct problems, hyperactivity/inattention, peers relationship problems and prosocial behaviour. Twenty of the items are combined to generate a "total difficulties" score. Normative data for children from the SDQ website was	At 8 years age Abnormal total behavioural difficulties score (SDQ, 90th percentile, SDQ norms as reference) <28 weeks GA/BW <1000 g: 34/189, 18.0% (12.8-24.2%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			used to determine those in the clinical range. Children with scores above 90th percentile were classified as being in the "abnormal" range, those between the 80th and 90th percentile were classified as "bordeline" and those below 80th percentile were classified as "normal".		
Johnson 2010	Population based cohort study (EPICURE)	N=307 surviving children at 11 years N=219 assessed at median age 10 years 11 months N=189 extremely preterm children (returned SCQ questionnaire)	Autism spectrum symptoms were assessed by using the Social Communication Questionnaire (SCQ) which was parent reported. Total scores were used to screen for symptoms (SCQ ≥15).	At 11 years age Autism spectrum problems (SCQ ≥15)	Low
Johnson 2015	A prospective geographical population-based study (LAMBS)	N=625 with completed BITSEA data (56% of originally recruited ones)	To assess behavioural outcome, parents completed the Brief infant Toddler Social Emotional Assessment (BITSEA).	At 2 years of corrected age Behaviour problems (BITSEA, >25th percentile) 32-36 weeks GA: 131/625, 21.0% (17.8-24.4%) 32-33 weeks GA: 17/84, 20.2% (12.3-30.4%) 34-36 weeks GA: 114/541, 21.1% (17.7-24.8%) Delayed social competence (BITSEA, <15th percentile) 32-36 weeks GA: 165/625, 26.4% (23.0-30.0%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				32-33 weeks GA: 23/84, 27.4% (18.2-38.2%) 34-36 weeks GA: 142/541, 26.3% (22.6-30.2%) Behaviour problem or delayed social competence (BITSEA) 32-36 weeks GA: 233/625, 37.3% (33.5-41.2%) 32-33 weeks GA: 34/84, 40.5% (29.9-51.8%) 34-36 weeks GA: 199/541, 36.8% (32.7-41.0%) Behaviour problem and delayed social competence (BITSEA) 32-36 weeks GA: 63/625, 10.1% (7.8-12.7%) 32-33 weeks GA: 6/84, 7.1% (2.7-14.9%) 34-36 weeks GA: 57/541, 10.5% (8.1-13.4%)	
Joseph 2016a	Population based cohort study (ELGAN)	N=1198 preterm infants surviving to 10 years n=966 children recruited for follow-up n=889 mothers of infants who agreed to participate	Participants were screened for ASD symptoms with the Social Communication Questionnaire (SCQ), the SCQ includes 39 ratings for children with simple sentence speech, and 33 ratings for those without simple sentence speech. To increase screener sensitivity, a score 11, recommended by the authors for individuals at higher-than-normal risk for ASD was used instead of the standard criterion of 15.	At 10 years ASD symptoms (assessed by SCQ): <27 weeks GA: 106/857, 12.4% (95% CI 10.2-14.8%)	High
Larroque 2011	Population based prospective cohort	Original sample: n=2901 very preterm	Parents filled in the French version of the Strengths and Difficulties Questionnaire (SDQ) to assess behavioural difficulties. It	At 8 years Total behavioural difficulties (SDQ, 10th perc) 24-32 weeks GA: 292/1387, 21.1% (18.9-23.3%) 24-28 weeks GA: 93/335, 27.8% (23.0-32.9%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		children (22-32 weeks) Included in follow-up: n=1439 preterm children	includes four scales that assess hyperactivity-inattention, conduct, emotional and peer problems, which are summed in a score of "total difficulties" and an additional scale assessing prosocial behaviour. Cut-offs were defined based on the 90th percentiles of the observed scores in the reference group (term children).	29-30 weeks GA: 65/378, 17.2% (13.5-21.4%) 31-32 weeks GA: 134/674, 19.9% (16.9-23.1%) Hyperactivity (SDQ, 10th perc) 24-32 weeks GA: 239/1387, 17.2% (15.3-19.3%) 24-28 weeks GA: 62/335, 18.5% (14.5-23.1%) 29-30 weeks GA: 57/378, 15.1% (11.6-19.1%) 31-32 weeks GA: 120/674, 17.8% (15.0-20.9%) Conduct problems (SDQ, 10th perc) 24-32 weeks GA: 131/1387, 9.4% (8.0-11.1%) 24-28 weeks GA: 30/335, 9.0% (6.1-12.5%) 29-30 weeks GA: 32/378, 8.5% (5.9-11.7%) 31-32 weeks GA: 69/674, 10.2% (8.1-12.8%) Emotional problems (SDQ, 10th perc) 24-32 weeks GA: 238/1387, 17.2% (15.2-19.3%) 24-28 weeks GA: 68/335, 20.3% (16.1-25.0%) 29-30 weeks GA: 54/378, 14.3% (10.9-18.2%) 31-32 weeks GA: 116/674, 17.2% (14.4-20.3%) Peer problems (SDQ, 10th perc) 24-32 weeks GA: 241/1387, 17.4% (15.4-19.5%) 24-28 weeks GA: 65/335, 19.4% (15.3-24.1%) 29-30 weeks GA: 72/378, 19.1% (15.2-23.4%) 31-32 weeks GA: 104/674, 15.4% (12.8-18.4%) Prosocial behaviour (SDQ, 10th perc) 24-32 weeks GA: 189/1387, 13.6% (11.9-15.6%) 24-28 weeks GA: 46/335, 13.7% (10.2-17.9%) 29-30 weeks GA: 36/378, 9.5% (6.8-12.9%)	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				31-32 weeks GA: 98/674, 14.5% (12.0-17.4%)	
Moore 2012	National population based cohort study	n=2035 EPT children born alive n=1031 survived to 2 years age n=559 completed questionnaires n=523 had completed MCHAT questionnaire	The 23-item MCHAT was used to assess children at age 16 to 30 months age to highlight behaviour that may indicate autistic traits and completed by the caregiver. If the child fails two or more of six critical items, or three or more items overall, he or she screens positive for autism and further investigation is warranted. The 'critical' items specifically address deficiencies in joint attention, pro-declarative pointing, and eye contact. These items have been found to predict the presence of autism	At age 2 years Positive screen for autistic traits (MCHAT) <27 weeks GA: 216/523, 41% (95%CI 37.0-45.7) 23 weeks GA: 17/31, 54.8% (95%CI 36.0-72.7) 24 weeks GA: 46/96, 47.9% (95%CI 37.6-58.4) 25wks GA: 67/168, 40.0% (95%CI 32.4-47.7) 26 weeks GA: 86/226, 38.1% (95%CI 31.7-44.7)	Low
Potijk 2012	Prospective cohort study	N=916 moderately preterm children assessed at 4 years.	Behavioural and emotional problems were measures using the Dutch version of the Child Behaviour Checklist (CBCL) for ages 1.5-5 For these scores, cut-offs for subclinical and clinical problems were set at 84th and 90th percentile, respectively, following the CBCL manual. Internalising problems consist of syndrome scales for emotionally reactive	At 4 years of age Total behavioural problems (CBCL, 90th perc) 32-35 weeks GA: 72/916, 7.9% (6.2-9.8%) Externalising problems (CBCL, 84th perc) 32-35 weeks GA: 87/916, 9.5%* (7.7-11.6%) Internalising problems (CBCL, 84th perc) 32-35 weeks GA: 89/916, 9.7% (7.9-11.8%) Emotionally reactive (CBCL, >97th perc) 32-35 weeks GA: 34/916, 3.7% (2.6-5.2%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			behaviour, anxious/depressed behaviour, somatic complaints and withdrawn behaviour. Externalising problems consist of syndrome scales for attention problems and aggressive behaviour.	Anxious/depressed (CBCL, >97th perc) 32-35 weeks GA: 11/916, 1.2% (0.6-2.1%) Somatic complaints (CBCL, >97th perc) 32-35 weeks GA: 54/916, 5.9% (4.5-7.6%) Withdrawn (CBCL, >97th perc) 32-35 weeks GA: 21/916, 2.3% (1.4-3.5%) Sleep problems (CBCL, >97th perc) 32-35 weeks GA: 22/916, 2.4% (1.5-3.6%) Attention problems (CBCL, >97th perc) 32-35 weeks GA: 38/916, 4.15% (3.0-5.7%) Aggressive behaviour (CBCL, >97th perc) 32-35 weeks GA: 31/916, 3.4% (2.3-4.8%)	
Rautava 2010	Population based prospective cohort study	Original sample size: n=924 preterm/very low birth weight infants Included in follow-up: n=588 preterm/very low birth weight infants	Behavioural outcomes were assessed using the Five to Fifteen Questionnaire (FTF), which was completed by the parents. Questions on development and behaviour were rated by the parents as 0="does not describe", 1="describes to some extent" and 2="describes well" the individual child.	At 5 years age Social skills problems (FTF) <32 weeks GA: 25/588, 4.3% (95%CI 2.7-6.2) Emotional and behavioural problems (FTF) <32 weeks GA: 20/588, 3.4% (95%CI 2.1-5.2)	Low
Samara 2010	National population based cohort study	n=308 children alive at 30 months age	When the child reached 6 years of age, parents completed a specially developed eating	At 6 years age Behavioural problems 25+6 weeks GA: 52/219, 23.7% (95%CI 18.3-30.0)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		n=241 entered study n=223 completed eating questionnaire	questionnaire. The scale included 19 items, which were grouped into four categories: refusal-faddy eating problems, oral motor problems, oral hypersensitivity problems and behavioural problems around meals. A total eating problems score was also constructed. Higher scores on each scale indicate more problems. To derive clinical categories, each scale was dichotomised into normal versus clinical (scores above the 90th centile or near according to the comparison group).	≤23 weeks GA: 8/22, 36.4% (95%CI 17.2-59.3) 24 weeks GA: 17/67, 25.4% (95%CI 15.5-37.5) 25 weeks GA: 27/130, 20.8 (95%CI 14.2-28.8)	
Samara 2008	A total-population prospective cohort study	N=224 children assessed at 6 years by parent-report N=215 children assessed at 6 years by teacher-report	Teachers and parents completed the respective versions of the Strengths and Difficulties Questionnaire (SDQ). The 25 SDQ items fall into 5 scales (with 5 items each), that is, emotional symptoms, conduct problems, hyperactivity, peer problems, and prosocial behaviour. For each scale except prosocial behaviour, higher scores indicate more problems. Additional items were adapted from the Conners Scales, the Child	At 6 years Parents' report Overall behavioural difficulties (SDQ, 90th perc) <26 weeks GA: 85/221, 38.5% (32.0-45.2%) Emotional problems (SDQ, 90th perc) <26 weeks GA: 60/222, 27.0% (21.3-33.4%) Conduct problems (SDQ, 90th perc) <26 weeks GA: 80/221, 36.2% (29.9-42.9%) Hyperactivity problems (SDQ, 90th perc) <26 weeks GA: 107/223, 48.0% (41.3-54.8%) Peer problems (SDQ, 90th perc)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			Behaviour Checklist, the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, and the International Classification of Diseases, 10th Revision, using the same Likert-scale format to assess components of attention-deficit/hyperactivity disorder (attention: teacher, 4 items; parents, 5 items; over-activity: 4 items each; impulsivity: teacher, 4 items; parents, 3 items). The total scores and subscale scores were dichotomized into normal/borderline versus clinical (score of 90th percentile, with respect to the control group). If the child scored at 90th percentile in both parent and teacher reports, then the behaviour was considered normal (no behaviour difficulty); mild difficulty refers to classification of the behaviour in the clinical range by either parent or teacher, whereas clinical pervasive behaviour refers to classification of the behaviour in the clinical range by both parent and teacher (severe behaviour difficulty).	<26 weeks GA: 80/222, 36.0% (29.7-42.7%) Prosocial behaviour (SDQ, 90th perc) <26 weeks GA: 40/219, 18.3% (13.4-24.0%) Additional scales Attention problems <26 weeks GA: 106/224, 47.3% (40.6-54.1%) Overactivity/impulsivity problems <26 weeks GA: 73/224, 32.6% (26.5-39.2%) School adaptation difficulties <26 weeks GA: 69.209, 33.0% (26.7-39.8%) Teachers' report Overall behavioural difficulties (SDQ, 90th perc) <26 weeks GA: 72/208, 34.6% (29.2-41.5%) Emotional problems (SDQ, 90th perc) <26 weeks GA: 63/211, 29.9% (23.8-36.5%) Conduct problems (SDQ, 90th perc) <26 weeks GA: 48/209, 23.0% (17.5-29.3%) Hyperactivity problems (SDQ, 90th perc) <26 weeks GA: 99/213, 46.5% (39.6-53.4%) Peer problems (SDQ, 90th perc) <26 weeks GA: 106/210, 50.5% (43.5-57.4%)	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				Prosocial behaviour (SDQ, 90th perc) <26 weeks GA: 43/209, 20.6% (15.3-26.7%) Additional scales Attention problems <26 weeks GA: 116/215, 54.0% (47.0-60.8%) Overactivity/impulsivity problems <26 weeks GA: 65/215, 30.2% (24.2-36.9%) School adaptation difficulties <26 weeks GA: 82/209, 39.2% (32.6-46.2%)	
Stahlmann 2009	A geographically defined cohort study	n=154 infants identified n=95 survived until discharge to home n=92 survived until follow-up at 7-9 years n=75 children were assessed at 7-9 years (81.5% of the surviving children)	Behavioural problems was assessed the Strengths and Difficulties Questionnaire (SDQ-Deu). The scoring was classified into normal, borderline and abnormal. Abnormal scores were based on the SDQ website's scoring instructions (according to the SDQinfo.com, in the total difficulties score, a score of 17-40 points is abnormal; for emotional symptoms, a score of 7-10 is abnormal; for hyperactivity-inattention, a score of 9-10 is abnormal; for conduct problems, a score of 6-10 is abnormal; for peer relationship problems, a score of 5-10 is abnormal; and for prosocial	At 7 to 9 years age Abnormal SDQ total difficulties (score 17-40) <27 weeks GA: 21/75, 28.0% (18.2-39.6%) Abnormal emotional symptoms (SDQ subscale score 7-10) <27 weeks GA: 20/75, 26.7% (17.1-38.1%) Abnormal hyperactivity-inattention score (SDQ subscale score 9-10) <27 weeks GA: 28/75, 37.3% (26.4-49.3%) Abnormal conduct problems score (SDQ subscale score 6-10) <27 weeks GA: 15/75, 20.0% (11.7-30.8%) Abnormal peer relationship score (SDQ subscale 5-10) <27 weeks GA: 15/75, 20.0% (11.7-30.8%) Abnormal prosocial behaviour score (SDQ subscale 0-5)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			behaviour, a score of 0-5 is abnormal. These are based on a population-based survey.)	<27 weeks GA: 7/75, 9.3% (3.8-18.3%)	
Stoelhorst 2003	Regional population-based prospective cohort study	N=158 children with completed CBCL questionnaires (N=266 children included in the cohort originally, N=235 survived)	The Child Behaviour Checklist (CBCL) for 2- to 3-y-old children was handed out to the parents during the 2-year check-up at the outpatient clinic and returned by mail. The CBCL had to be completed by one or both parents. In the six syndrome scales, scores above the 98th percentile are defined as clinically abnormal; scores from the 95th through the 98th percentile as borderline clinical. For the total problem score, the internalizing and externalizing groups, scores above the 90th centile are defined as clinically abnormal, scores from the 85th through the 90th centile as borderline clinical.	At 2 years of corrected age Total behavioural problems (CBCL, 90th perc) <32 weeks GA: 14/158, 8.9% (4.9-14.4%) Anxious/depressed (CBCL, 98th perc) <32 weeks GA: 1/158, 0.6% (0.02-3.5%) Withdrawn (CBCL, 98th perc) <32 weeks GA: 3/158, 1.9% (0.4-5.5%) Somatic problems (CBCL, 98th perc) <32 weeks GA: 3/158, 1.9% (0.4-5.5%) Aggressive behaviour (CBCL, 98th perc) <32 weeks GA: 3/158, 1.9% (0.4-5.5%) Destructive behaviour (CBCL, 98th perc) <32 weeks GA: 5/158, 3.2% (1.0-7.2%)	Low
Wilson-Ching 2013	Geographical cohort study	n=298 consecutive survivors	Attention problems (<-1.5 SD) Selective attention: The Telephone Search task of the Test of Everyday Attention was used. Participants were required to search simulated telephone	At 17 years age Attention problems Selective attention (<-1.5 SD) <28 weeks GA/ELBW: 71/199, 35.6% (95%CI 29-43) Sustained attention (<-1.5 SD) <28 weeks GA/ELBW, 16/174, 9.2% (95%CI 5.3-14.5)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			directory for pairs of shapes that looked the same. The number of targets detected (maximum-=20) and the time taken to complete the task were recorded. The Elevator with Distraction task, also from the Test of Everyday Attention, was used as a second measure with a maximum of 7 correct trials recorded. Sustained attention: The Test of Variables of Attention (TOVA) was used to measure how quickly the participants could see a target presented on the computer. Shifting attention: The Contingency Naming Test (CNT) was used to assess individuals by showing a page of coloured shapes embedded in a smaller shape. An efficiency score, which represents a ratio of the time taken to complete the task and the number of errors, was the variable of interest Divided attention: The Telephone Search while counting task on the Test of Everyday Attention was used. A divided attention	Shifting attention (<-1.5 SD) <28 weeks GA/ELBW, 86/209, 41.1% (95%CI 34.4-48.2) Divided attention (<-1.5 SD) <28 weeks GA/ELBW, 30/196, 15.3% (95%CI 10.6-21.1) Behavioural attention problems Inattentive (CADS parent report) (<-1.5 SD) <28 weeks GA/ELBW: 32/193, 16.6% (95%CI 11.6-22.6) Hyperactive (CADS parent report) (<-1.5 SD) <28 weeks GA/ELBW: 28/193, 14.5% (95%CI 9.9-20.1) ADHD DSM-IV (parent report) (<-1.5 SD) <28 weeks GA/ELBW: 34/193, 17.6% (95%CI 12.5-23.7) Shift (BRIEF parent report) (<-1.5 SD) <28 weeks GA/ELBW: 38/201, 19% (95%CI 13.7-25.0) Inhibit (BRIEF parent report) (<-1.5 SD) <28 weeks GA/ELBW: 35/201, 17.4% (95%CI 12.4-23.4) Inattentive (CADS self report) (<-1.5 SD) <28 weeks GA/ELBW: 17/192, 8.9% (95%CI 5.2-13.8) Hyperactive CADS (self report) (<-1.5 SD) <28 weeks GA/ELBW: 11/192, 5.7% (95%CI 3.0-10.0) ADHD DSM IV (self report) (<-1.5 SD) <28 weeks GA/ELBW:10/192, 5.2% (95%CI 2.5-9.4)	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			score was calculated by multiplying the proportion of correct targets found by the proportion of correct series of tones counted times 10, with a score of 10 signifying a perfect score Behavioural attention: The CADS-P consists of 26 items and the CADS-A of 30 items, and both provide 3 age standardized scales (inattentive behaviours, hyperactive behaviours, DSM-IV ADHD index) each with a mean of 50 and SD of 10 Behaviour rating inventory of executive function (BRIEF): Parent or self- reported behaviours related to executive functioning were assessed by evaluating specific behaviours relating to executive attention skills including ''shift'' and ''inhibit'' scales. Ability to flexibly move from a given activity or aspect of a problem to another as the situation demanded was evaluated. T scores were recorded for each of these scales (M=50; SD=10)	Shift (BRIEF self report) (<-1.5 SD) <28 weeks GA/ELBW: 10/180, 5.6% (95%CI 2.7-10.0) Inhibit (BRIEF self report) (<-1.5 SD) <28 weeks GA/ELBW: 17/180, 9.4% (95%CI 5.6-14.7)	
Evidence on special education needs					

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
Chan 2014 (MCS)	Prospective Cohort Study	Sample recruited - n=18818 Sample eligible for assessment - n=13543 Sample analysed after exclusions - n=6031 n=69 - Very preterm (<32 weeks) n=67 - Moderately preterm (32–33 weeks) n=360 - Late preterm (34–36 weeks) n=1258 - Early term (37–38 weeks)	School performance was investigated using the statutory Key Stage 1 (KS1) teacher assessments performed in the third school year in England. At KS1, children generally perform between level 1 (below expected level) to level 3 (considerably above the expected level), with adequate performance categorised as achieving level 2 or above	At 7 years age: Not achieving level 2 (expected) or above in reading, writing or mathematics (KS1) <32 weeks GA: 29/69, 42.0% (30.2-54.5%) 32-33 weeks GA: 18/67, 26.9% (16.8-39.1%) 34-36 weeks GA: 84/360, 23.3% (19.1-28.1%) Not achieving level 2 (expected) or above in reading (KS1) <32 weeks GA: 18/69, 26.1% (16.3-38.1%) 32-33 weeks GA: 13/67, 19.4% (10.8-30.9%) 34-36 weeks GA: 65/360, 18.1% (14.2-22.4%) Not achieving level 2 (expected) or above in writing (KS1) <32 weeks GA: 27/69, 39.1% (27.6-51.6%) 32-33 weeks GA: 16/67, 23.9% (14.3-35.9%) 34-36 weeks GA: 74/360, 20.6% (16.5-25.1%) Not achieving level 2 (expected) or above in speaking and listening (KS1) <32 weeks GA: 20/69, 29.0% (18.7-41.2%) 32-33 weeks GA: 11/67, 16.4% (8.5-27.5%) 34-36 weeks GA: 47/360, 13.1% (9.8-17.0%) Not achieving level 2 (expected) or above in mathematics (KS1) <32 weeks GA: - 32-33 weeks GA: - 34-36 weeks GA: 31/360, 8.6% (5.9-12.0%) No achieving level 2 (expected) or above in science (KS1) <32 weeks GA: 17/69, 24.6% (15.1-36.5%) 32-33 weeks GA: 11/67, 16.4% (8.5-27.5%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				34-36 weeks GA: 42/360, 11.7% (8.5-15.4%)	
Chyi 2008 (ECLS-K)	Population based cohort study (Early Childhood Longitudinal Study-Kindergarten Cohort)	N= 17,565 (ECLS-K cohort) n=988 preterms selected n=970 included in the analysis (after exclusions) Children born between 32 and 36 weeks GA and also children born between 32 and 33 weeks GA	Assessment included a battery of tests, including reading and math. Test items were adapted from the Peabody Individual Achievement Test-Revised, Peabody Picture Vocabulary Test-Revised, Primary Test of Cognitive Skills, the Test of Early Reading Ability, the Test of Early Mathematics Ability, and the Woodcock Johnson Tests of Achievement-Revised. Teacher academic ratings were also completed involving teacher evaluations of each student's reading and math ability	At various ages of assessment: Individualised education programme Kindergarten stage (3 years age?) 32-33 weeks GA: 19/146, 13.0% (95%CI 8.0-19.6) 34-36 weeks GA: 46/572, 8.0% (95%CI 6.0-10.6) 32-36 weeks GA: 65/718, 9.1% (95%CI 7.1-11.4) First grade (6-7 years age?) 32-33 weeks GA: 26/146, 17.8% (95%CI 12.0-25.0) 34-36 weeks GA: 61/579, 10.5% (95%CI 8.2-13.3) 32-36 weeks GA: 87/725, 12% (95%CI 9.7-14.6) Third grade (8-9 years age?) 32-33 weeks GA: 26/132, 19.7% (95%CI 13.3-27.5) 34-36 weeks GA: 64/528, 12.1% (95%CI 9.5-15.2) 32-36 weeks GA: 90/660, 13.6% (95%CI 11.1-16.5) Fifth grade (10-11 years age?) 32-33 weeks GA: 17/94, 18.1% (95%CI 10.9-27.4) 34-36 weeks GA: 49/402, 12.2% (95%CI 9.2-15.8) 32-36 weeks GA: 66/402, 16.4% (95%CI 12.9-20.4) Special education enrolment Kindergarten stage (3 years age?) 32-33 weeks GA: 16/199, 8.04% (95%CI 4.7-12.7) 34-36 weeks GA: 50/751, 6.7% (95%CI 5.0-8.7) 32-36 weeks GA: 66/956, 6.9% (95%CI 5.4-8.7) First grade (6-7 years age?) 32-33 weeks GA:23/193, 11.9% (95%CI 7.7-17.3)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				34-36 weeks GA: 46/734, 6.3% (95%CI 4.6-8.3) 32-36 weeks GA: 69/927, 7.4% (95%CI 5.8-9.3) Third grade (8-9 years age?) 32-33 weeks GA: 22/153, 14.4% (95%CI 9.2-21.0) 34-36 weeks GA: 57/623, 9.2% (95%CI 7.0-11.7) 32-36 weeks GA: 79/776, 10.0% (95%CI 8.0-12.3) Fifth grade (10-11 years age?) 32-33 weeks GA: 18/124, 14.5% (95%CI 8.8-22.0) 34-36 weeks GA: 52/506, 10.3% (95%CI 7.8-13.3) 32-36 weeks GA: 70/630, 11.1% (95%CI 8.8-13.8)	
Farooqi 2007 (Swedish national cohort)	Nationally-representative population-based cohort study	Total sample: n=169 n= 83 extremely immature (EI) children born before 26 completed weeks of gestation	School difficulties was defined as the child repeating a grade and/or using special educational resources (full-time or part-time). Attending special class or special school means attending a special school or training school for the physically disabled and severely mentally retarded or receiving full-time special education attached to the mainstream school.	At 11 years assessment: Special class or special school <26 weeks GA: 13/86, 15.1% (8.3-24.5%) Grade repetition <26 weeks GA: 13/83, 15.7% (8.6-25.3%) School difficulties (repeated year or special educational resources) <26 weeks GA: 51/86, 59.3% (48.2-69.8%)	Moderate
Guellec 2011(EPIPG AGE)	Population based prospective cohort study	N=2855 live births at 24-32 weeks GA. n=2357 infants eligible for follow-up	School difficulties were defined by special schooling (institution or special school, special class in mainstream school, mainstream class) or low grades. This was asked through a questionnaire sent	At 8 years assessment School difficulties SGA children (bw <10th percentile) 24-28 weeks GA: 6/17, 35.3% (14.2-61.7%) 29-32 weeks GA: 30/107, 28.0% (19.8-37.6%) MGA children (bw 10th-19th percentile)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			to the parents when the child was 8 years old.	24-28 weeks GA: 13/29, 44.8% (26.5-64.3%) 29-32 weeks GA: 24/104, 23.1% (15.4-32.4%)AGA chidlren (bw >=20th percentile) 24-28 weeks GA: 98/295, 33.2% (27.9-38.9%) 29-32 weeks GA: 163/887, 18.4% (15.9-21.1%)	
Johnson 2011 (EPICURE)	National population-based cohort study	n=219 children assessed at 11 years (data missing for some individuals in the outcomes of interest)	Teachers completed a questionnaire to elicit information detailing whether SEN provision was utilised by the child.	At 11 years assessment: Identified SEN <26 weeks GA: 134/215, 62.3% (55.5-68.8%) SEN provision <26 weeks GA: 132/215, 61.4% (54.5-67.9%) Children in main-steam schools only: Identified SEN <26 weeks GA: 105/186, 56.5% (49.0-63.7%) SEN provision <26 weeks GA: 103/186, 55.4% (47.9-62.7%)*	Low
Larroque 2011 (EPIPGAGE)	Population based prospective cohort	Original sample: n=2901 very preterm children (22-32 weeks) Included in follow-up: n=1439 preterm children	Schooling outcomes included whether the child attended an institution or special school, whether they were in a special class within mainstream schooling and whether they had repeated a school year. Support at school was defined according to whether the child was enrolled at a particular institution, special school or class, or a mainstream class with support at school (extra	At 8 years assessment: Schooling and special support: Institution or special school or special class 24-32 weeks GA: 75/1435, 5.2% (4.1-6.5%) 24-28 weeks GA: 32/340, 9.4% (6.5-13.0%) 29-30 weeks GA: 20/387, 5.2% (3.2-7.9%) 31-32 weeks GA: 23/708, 3.3% (2.1-4.8%) Support at school in mainstream class 24-32 weeks GA: 221/1435, 15.4% (13.6-17.4%) 24-28 weeks GA: 77/340, 22.7% (18.3-27.5%) 29-30 weeks GA: 40/387, 10.3% (7.5-13.8%) 31-32 weeks GA: 104/708, 14.7% (12.2-17.5%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			teacher in or outside of the class room, extra teaching hours at school, intervention of a psychologists or other person at school).	Special care since the age of 5 years (at least one of orthoptic, speech therapy, physical therapy, occupational therapy, psychologist/psychiatric therapy) 24-32 weeks GA: 794/1436, 55.3% (52.7-57.9%) 24-28 weeks GA: 223/341, 65.4% (60.1-70.4%) 29-30 weeks GA: 202/389, 51.9% (46.8-57.0%) 31-32 weeks GA: 369/706, 52.3% (48.5-56.0%) Special care since 5 years (see above) or support at school 24-32 weeks GA: 841/1438, 58.5% (55.9-61.1%) 24-28 weeks GA: 239/343, 69.7% (64.5-74.5%) 29-30 weeks GA: 208/388, 53.6% (48.5-58.7%) 31-32 weeks GA: 394/707, 55.7% (52.0-59.4%)	
Mackay 2013	Retrospective study using national registry data	Relevant sample included for this analysis n=237894 n=215935 full term (40-41 weeks) n=18035 preterm (33-36 weeks) n=3449 preterm (28-32 weeks) n=475 preterm (24-27 weeks)	Data on SEN were identified through the 2005 school census. SEN includes: language impairments; specific learning difficulties (such as dyslexia or dyscalculia); intellectual disabilities; other developmental disorders that impair learning (including autism, Asperger's syndrome and attention deficit hyperactivity	Assessed at 5-18 years Sensory SEN according to gestational age 24-27 weeks GA: 14/475, 3.0% (95%CI 1.6-4.9) 28-32 weeks GA: 17/3449, 0.49% (95% CI 0.29-0.79) 33-36 weeks GA: 40/18035, 0.2% (95%CI 0.16-0.3) Physical or motor SEN according to gestational age 24-27 weeks GA: 29/475, 6.1% (95%CI 4.1-8.7) 28-32 weeks GA: 98/3449, 2.8% (95%CI 2.3-3.5) 33-36 weeks GA: 84/18035, 0.47% (95%CI 0.37-0.58) Language SEN according to gestational age 24-27 weeks GA: 3/475, 0.63% (95%CI 0.13-1.83) 28-32 weeks GA: 13/3449, 0.38% (95%CI 0.2-0.6) 33-36 weeks GA: 42/18035, 0.2% (95%CI 0.2-0.3) Social, emotional or behavioural SEN according to gestational age	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				24-27 weeks GA: 6/475, 1.3% (95%CI 0.5-2.7) 28-32 weeks GA: 32/3449, 0.9% (95%CI 0.6-1.3) 33-36 weeks GA: 169/18035, 0.9% (95%CI 0.8-1.1) Specific learning difficulties SEN according to gestational age 24-27 weeks GA: 10/475, 2.1% (95%CI 1.0-3.8) 28-32 weeks GA: 49/3449, 1.4% (95%CI 1.1-1.9) 33-36 weeks GA: 235/18035, 1.3% (95%CI 1.1-1.5) Intellectual SEN according to gestational age 24-27 weeks GA: 67/475, 14.1% (95%CI 11.1-17.6) 28-32 weeks GA: 165/3449, 4.8% (95%CI 4.1-5.6) 33-36 weeks GA: 521/18035, 3.0% (95%CI 2.7-3.1) ASD SEN according to gestational age 24-27 weeks GA: 5/475, 1.1% (95%CI 0.3-2.4) 28-32 weeks GA: 34/3449, 1.0% (95%CI 0.7-1.4) 33-36 weeks GA: 75/18035, 0.4% (95%CI 0.3-0.5) Unspecified SEN according to gestational age 24-27 weeks GA: 6/475, 1.3% (95%CI 0.5-2.7) 28-32 weeks GA: 35/3449, 1.0% (95%CI 0.7-1.4) 33-36 weeks GA: 115/18035, 0.6% (95%CI 0.5-0.8)	
Mackay 2010	Retrospective study using national registry data	Relevant sample included for this analysis n=152757	Special educational need (SEN) was identified through the school census data. This includes information on children with learning disabilities (including dyslexia, dyspraxia, autism, Asperger's syndrome and	Assessed at age 5 to 18 years SEN according to gestational age 24-27 weeks GA: 140/475, 29.5% (95%CI 25.4-33.8) 28-32 weeks GA: 443/3449, 12.8% (95%CI 11.7-14.0) 33-36 weeks GA: 1281/18035, 7.1% (95%CI 6.7-7.5)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		n=18035 preterm (33-36 weeks) n=3449 preterm (28-32 weeks) n=475 preterm (24-27 weeks)	attention deficit hyperactivity disorder) as well as children with physical disabilities that impact on learning (including some children with hearing, motor and visual impairment).		
Odd 2016 (ALSPAC)	Regional prospective cohort study	N=775 preterm infants (<37 weeks)	Mandatory UK educational assessments done at 4 stages, the stages are Key Stage (KS) 1 at 5-7 years, KS2 at 7-11 years, KS3 at 11-14 years, and KS4 at 14-16 years. The test is done at the end of each stage. Governmental standards set the minimum standard expected at each stage of the first 3 stages and this was used as the cut-off for a low score. At the end of KS4 children take their school exams and an a-priori cut-off of 5 General Certificates of Secondary Education (GCSE) or equivalent at A* to C level was used to define a normal score at this age. At KS4, <5 passes at A* to C level was considered as poor/low attainment at KS4. Children identified as having special educational needs (SEN) in KS4 were identified	At 5-7 years Low score at KS1 <37 weeks GA: 210/662, 31.7% (28.2-35.4%) At 7-11 years Low score at KS2 <37 weeks GA: 239/675, 35.4% (31.8-39.2%) At 11-14 years Low score at KS3 <37 weeks GA: 251/631, 39.8% (35.9-43.7%) At 14-16 years Low score at KS4 <37 weeks GA: 276/701, 39.4% (35.7-43.1%) At 14-16 years SEN <37 weeks GA: 166/683, 24.3% (21.1-27.7%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			from the Pupil Level Annual School Census (PLASC).		
Odd 2013 (ALSPAC)	Regional prospective cohort study	n=722 preterm infants (<37 weeks)	At the age of 8 years, the child's teacher was sent a questionnaire, which asked the teacher to identify "has this child ever been recognised as having special educational needs?" (SEN)	Assessed at 8 years Low KS1 score <37 weeks GA: 227/722, 31.4% (95%CI 28.1-35.0) Special education needs <37 weeks GA: 256/722, 35.5% (95%CI 32.0-39.1)	Low
Odd 2012 (ALSPAC)	Regional prospective cohort study	N=741 moderate/late preterm children (32-36 weeks) in the cohort N=319 moderate/late preterm children with data on SEN (43%)	At the age of 8 years, the child's teacher was sent a standardized questionnaire which asked "Has this child ever been recognized as having special educational needs?"	At 8 years Special educational needs (reported by teacher) 32-36 weeks GA: 110/319, 34.5% (29.3-40.0%)	Very low
Peacock 2012 (ALSPAC)	Population-based longitudinal study	N=13,978 infants alive at 1 year n=596 born at 32-36 weeks included in analysis at 5 to 7 years age	Data on Key Stage 1 assessments were obtained from local education authorities. The results for the three assessment domains (reading, writing and mathematics) were dichotomized, with success defined as achieving at least level 2, the expected level of attainment. Overall KS1 score defined as having at	Assessed at 5 to 7 years age KS1 overall assessment among preterm group (below level 2 in reading, writing and mathematics) 32-36+6 weeks GA: 173/596, 29% (95%CI 25.4-33.0) KS1 reading assessment among preterm group (below level 2) 32-36+6 weeks GA: 132/596, 22.2% (95%CI 19.0-25.7 KS1 writing assessment among preterm group (below level 2) 32-36+6 weeks GA: 135/596, 22.7% (95%CI 19.4-26.2) KS1 mathematics assessment among preterm group (below level 2)	Very low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
			least level 2 in all three domains.	32-36+6 weeks GA: 108/596, 18.1% (95%CI 15.1-21.5)	
Quigley 2012 (MCS)	Population-based cohort	N=8728 total number of children in the study (all gestational ages) N=106 very preterm children (23-31 weeks) N=99 moderately preterm children (32-33 weeks) N=537 late preterm children (34-36 weeks)	The Foundation Stage Profile (FSP) records the child's achievement as measured by their teacher at the end of their first school year, 'foundation stage'. Children achieving a scale score of >=6 points are classified as "working securely with the Early Learning Goals" and are classified as having achieved a good level of development. Children who achieve a score of >=78 points across the 13 assessment scales (i.e. an average of 6 points per scale) and a score of >=6 in each of the three 'personal, social, and emotional development' scales and the four 'communication, language, and literacy' scales are classified as "reaching a good level of overall achievement".	At 5 years assessment Not good level of overall achievement in FSP 23-31 weeks GA: 56/84, 66.7% (55.5-76.6%) 32-33 weeks GA: 56/92, 60.9% (50.1-70.9%) 34-36 weeks GA: 276/471, 58.6% (54.0-63.1%) 39-41 weeks GA: 2853/5407, 52.8% (51.4-54.1%) 32-36 weeks GA: 332/563, 59.0 (54.8-63.1%) Not working securely in all three scales of personal, social and emotional development in FSP 23-31 weeks GA: 36/84, 42.9% (32.1-54.1%) 32-33 weeks GA: 30/92, 32.6% (23.2-43.2%) 34-36 weeks GA: 148/471, 31.4% (27.3-35.8%) 39-41 weeks GA: 1456/5407, 26.9% (25.8-28.1%) 32-36 weeks GA: 178/563, 31.6% (27.8-35.6%) Not working securely in all four scales of communication, language and literacy in FSP 23-31 weeks GA: 52/84, 61.9% (50.7-72.3%) 32-33 weeks GA: 53/92, 57.6% (46.9-67.9%) 34-36 weeks GA: 255/471, 54.1% (49.5-58.7%) 39-41 weeks GA: 2652/5407, 49.1% (47.7-50.4%) 32-36 weeks GA: 308/563, 54.7% (50.5-58.9%) Not working securely in all three scales of mathematical development in FSP 23-31 weeks GA: 46/84, 54.8% (43.5-65.7%) 32-33 weeks GA: 37/92, 40.2% (30.1-51.0%) 34-36 weeks GA: 174/471, 36.9% (32.6-41.5%) 39-41 weeks GA: 1745/5407, 32.3% (31.0-33.5%)	Moderate

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
				32-36 weeks GA: 211/563, 37.5% (33.5-41.6%) Not working securely in the "knowledge and understanding of the world" scale in FSP 23-31 weeks GA: 26/84, 31.0% (21.3-42.0%) 32-33 weeks GA: 23/92, 25.0% (16.6-35.1%) 34-36 weeks GA: 126/471, 26.8% (22.8-31.0%) 39-41 weeks GA: 1141/5407, 21.1% (20.0-22.2%) 32-36 weeks GA: 149/563, 26.5% (22.9-30.3%) Not working securely in the "physical development" scale in FSP 23-31 weeks GA: 18/84, 21.4% (13.2-31.7%) 32-33 weeks GA: 14/92, 15.2% (8.6-24.2%) 34-36 weeks GA: 67/471, 14.2% (11.2-17.7%) 39-41 weeks GA: 570/5407, 10.5% (9.7-11.4%) 32-36 weeks GA: 81/563, 14.4% (11.6-17.6%) Not working securely in the "creative development" in FSP 23-31 weeks GA: 32/84, 38.1% (27.7-49.3%) 32-33 weeks GA: 24/92, 26.1% (17.5-36.3%) 34-36 weeks GA: 117/471, 24.8% (21.0-29.0%) 39-41 weeks GA: 1077/5407, 19.9% (18.9-21.0%) 32-36 weeks GA: 141/563, 25.0% (21.5-28.8%)	
Samara 2008 (EPICURE)	A total-population prospective cohort study	N=224 children assessed at 6 years by parent-report N=215 children assessed at 6	Teachers and parents completed the respective versions of the Strengths and Difficulties Questionnaire (SDQ). The 25 SDQ items fall into 5 scales (with 5 items each), that is, emotional symptoms, conduct problems,	At 6 years Parent report School adaptation difficulties <26 weeks GA: 69.209, 33.0% (26.7-39.8%) Teacher report School adaptation difficulties <26 weeks GA: 82/209, 39.2% (32.6-46.2%)	Low

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (including GA at birth and age at assessment)	Study quality
		years by teacher-report	hyperactivity, peer problems, and prosocial behaviour. For each scale except prosocial behaviour, higher scores indicate more problems		

4.1.3.3 Economic evidence

No health economic search was undertaken for this review question and consequently no evidence was found. This question focused on the prevelance of various developmental problems rather than whether any strategy for the management of these problems represents a cost-effective use of resources. Therefore, this question is not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review.

4.1.3.4 Evidence statements

4.1.3.4.1 Prevalence of functional problems

Functional problems at <28 completed weeks of gestation

Feeding problems

Low quality evidence from one study (sample size 308) showed that among children born at 25^{+6} weeks of gestation, the prevalence of total eating problems was 34.9% (95%CI 29.0 to 41.6%) at 6 years age (Samara 2010). In the same study, prevalence for refusal faddy problems was 17% (95%CI 12.4 to 22.6%) and for oral motor problems, 33.5% (95%CI 27.2 to 40.2%).

Low quality evidence from one study (sample size 308) showed that among children born at or before 25 weeks+6 days of gestation, the prevalence of hypersensitivity problems (specific questionnaire) was 23.5% (95%CI 18.0 to 30.0%) at 6 years age (Samara 2010).

Functional problems at <32 completed weeks of gestation

Feeding problems

Low quality evidence from one study (sample size 1711) showed that among children born at <32 weeks of gestation, the prevalence of altered palatal morphology was 3.7% (95%CI 2.9 to 4.7%) at 5 years age (Germa 2012).

- Sleep problems

Low quality evidence from one study (sample size 158) showed that among children born at <32 weeks of gestation, the prevalence of sleeping problems (CBCL, 98th percentile) was 3.2% (95%CI 1.0 to 7.2%) at 2 years (corrected age) (Stoelhorst 2003).

Low quality evidence from one study (sample size 22039) showed that among children born at <32 weeks of gestation, the prevalence of sleep apnoea (ICD-10) was 2.6% (95%CI 2.1 to 3.2%) at 2.5 to 6 years age (Raynes-Greenow 2012).

Functional problems at 32 to 36 completed weeks of gestation

Feeding problems

Low quality evidence from one study (sample size 628) showed that among children born at 32-36 weeks of gestation the prevalence of total eating difficulties (parent questionnaire) was 9.5% (95%CI 7.5 to 11.9%) at 2 years (corrected age) (Johnson 2016). In the same study, prevalence for refusal or picky eating was 6.5% (95%CI 4.8 to 8.5%). Prevalence was also reported for oral motor problems (5.5% (95%CI 4.0-7.4%)), oral hypersensitivity (4.2% (95%CI 2.9 to 5.9%)), and eating behaviour problems (6.1% (95%CI 4.5 to 8.1%)) (Johnson 2016).

Sleep problems

Moderate quality evidence from one study (sample size 916) showed that among children born at 32-35 weeks of gestation the prevalence of sleeping problems (CBCL >97th percentile) was 2.4% (95%CI 1.5 to 3.6%) at 4 years age (Potijk 2012).

Low quality evidence from one study (sample size 22039) showed that among children born at 32-36 weeks of gestation the prevalence of sleep apnoea (ICD-10) was 1.3% (95%CI 1.2 to 1.5%) at 2.5 to 6 years age (Raynes-Greenow 2012).

Functional problems by gestational week

Feeding problems

Low quality evidence from one study (sample size 308) showed that among children born at 24 weeks of gestation the prevalence of total eating problems (parent reported) was 50% (95%CI 37.6 to 62.4%) at 6 years age, and the prevalence decreased at 25 weeks gestational age (25.8% (95%CI 18.5 to 34.3%) (Samara 2010). A similar trend was seen for oral motor problems at 24 weeks (40.9% (95%CI 29 to 53.7) and 25 weeks gestation (28.7% (95%CI 21.1 to 37.3%). The prevalence of refusal faddy problems was 13.6% (95%CI 2.9 to 34.9%) at ≤23 weeks, 16.2% (95%CI 8.4 to 27.1%) at 24 weeks, and 18.1% (95%CI 11.9 to 25.7%) at 25 weeks gestation (Samara 2010).

4.1.3.4.2 *Prevalence of motor problems*

Motor/coordination problems (DCD)

Moderate quality evidence from one study (sample size 237) showed that among children born at 24-31 weeks of gestation the prevalence of motor problems (MABC, <=15th percentile) was 36.3% (95%CI 29 to 44.1%) at 5 years age (Agerholm 2011).

Very low quality evidence from one study (sample size 401) showed that among children born at 23-27 weeks of gestation the prevalence of motor problems (MABC <15th percentile) was 15.0% (95%CI 10.1 to 21.2%) at 8 years age (Kan 2008).

Moderate quality evidence from one study (sample size 32097) showed that among children born at 32-36 weeks of gestation the prevalence of suspect or indicated DCD (DCDQ) was 6.4% (95%CI 5.1 to 7.9%) at 7 years age (Faebo Larsen 2013). In the same study the prevalence was higher among those children born at 23-31 weeks of gestation (18.3% (95%CI 12.2.to 25.8%).

Low quality evidence from one study (sample size 1662) showed that among children born at 32-34 weeks of gestation the prevalence of coordination and balance (presence of age-inadequate performance) was 23.8% (95%CI 20.3 to 27.6%) compared to the prevalence among those born at 28-31 weeks of gestation (27.7% (95%CI 24.5 to 31.2%)) (Arnaud 1997).

Motor problems at <28 completed weeks of gestation

Motor problems (PDI <55, 55-69, <70)

Moderate quality evidence from one study (sample size 95) showed that among children born at <27 weeks of gestation the prevalence of motor problems (PDI <55) was 27.3% (95%CI 17.7 to 38.6%) at 3 years age. In the same study, the prevalence of motor problems (PDI 55-69) was 20.8% (95%CI 12.4 to 31.5%) and 48.1% (95%CI 36.5 to 59.7%) (PDI <70) (De Groote 2008).

- Fine and gross motor problems

Low quality evidence from one study (sample size 707) showed that among children born at <27 weeks of gestation the prevalence of mild fine motor problems (Bayley -1SD to -2SD) was 33.7% (95%CI 29 to 39%) at 2.5 years age (Mansson 2014). In the same study, the prevalence of moderate motor problems (Bayley -2SD to -3SD) and moderate to severe motor problems was 8.1% (95%CI 5.6 to 11.2%) and 4.3% (95%CI 2.5 to 6.8%) respectively (Mansson 2014).

Low quality evidence from one study (sample size 707) showed that among children born at <27 weeks of gestation the prevalence of mild gross motor problems (Bayley -1SD to -2SD) was 29% (95%CI 24.5 to 33.8%) at 2.5 years age (Mansson 2014). In the same study, the prevalence for moderate gross motor problems (Bayley -2SD to -3SD) was 7% (95%CI 4.7 to 10.1%).

Motor problems at <32 weeks of gestation

Motor problems (PDI)

Low quality evidence from one study (sample size 158) showed that among children born at <32 weeks of gestation the prevalence of motor problems (mild to moderate; BSID -1 to -2SD, <-2SD) ranged from 11% (95%CI 6.7 to 16.9%) to 17.8% (95%CI 12.3 to 24.5%) at 18 months corrected age (Stoelhorst 2003b). At 24 months the prevalence (BSID -1 to -2SD), ranged from 22.2% (95%CI 15.7 to 29.9%) and 8.3% (95%CI 4.4 to 14.1%) (BSID <-2SD). In another study (sample size 924) the prevalence of motor skills problems (FTF) was 8.3% (95%CI 6.2 to 11%) among children born at <32 weeks of gestation, assessed at 5 years age (Rautava 2010).

Motor problems at 28-31 completed weeks of gestation

Low quality evidence from one study (sample size 1662) showed that among children born at 28-31 weeks of gestation the prevalence of minor neuromotor dysfunction (Touwen assessment, 1-2 items affected) was 40.4% (95%CI 36.8 to 44.1%) at 5 years age (Arnaud 2007). In the same study, the prevalence of moderate neuromotor dysfunction (Touwen, >2 items affected) was 3.1% (95%CI 2 to 4.7%). Prevalence of posture/muscle tone regulation and reflex abnormalities was 11% (95%CI 8.7 to 13.5%) and 10% (95%CI 7.8 to 12.4%) respectively. Prevalence of motor behaviour of face and eyes was 12.7% (95%CI 10.3 to 15.4%) (Arnaud 2007).

Motor problems at 32-36 completed weeks of gestation

Low quality evidence from one study (sample size 1662) showed that among children born at 32-36 weeks of gestation the prevalence of minor neuromotor dysfunction (Touwen assessment, 1-2 items affected) was 36%% (95%CI 32 to 40.1%) at 5 years age (Arnaud 2007). In the same study, the prevalence of moderate neuromotor dysfunction (Touwen, >2 items affected) was 1.5% (95%CI 0.6 to 2.8%). Prevalence of mild deviation of posture/muscle tone regulation and reflex abnormalities was 5.1% (95%CI 3.5 to 7.3%) and 6.9% (95%CI 4.9 to 9.3%) respectively. Prevalence of motor behaviour of face and eyes was 14% (95%CI 11.2 to 17.2%) (Arnaud 2007).

Fine and gross motor problems

Moderate quality evidence from one study (sample size 926) showed that among children born at 32-35 weeks of gestation the prevalence of fine motor problems (ASQ, <-2SD) was 8.1% (95%CI 6.4 to 10%) at 4 years age (Potijk 2013). In the same study, the prevalence of gross motor problems among this group of children was 5.7% (95%CI 4.3 to 7.4%).

Motor problems by week of gestational age

Motor problems

Low quality evidence from one study (sample size 1662) showed a trend of decreasing prevalence of minor motor dysfunction (Touwen, 1-2 items affected) with increasing gestational age, ranging from 52.3% (95%CI 44.6 to 60%) among those born at <28 weeks of gestation, to 30.8% (95%CI 24.4 to 37.8%) among those born at 33-34 weeks of gestation (Arnaud 2007). In the same study, there was a similar trend for the prevalence (although lower) of moderate motor dysfunction (Touwen, >2 items affected), which ranged from 5.1% (95%CI 2.3 to 9.4%) among those born at <28 weeks of gestation, to 0.5% (95%CI 0.01 to 2.8%) among those born at 33-34 weeks of gestation (Arnaud 2007). The prevalence of mild deviation of posture/muscle tone regulation was 20.2% (95%CI 14.6 to 29%) among those born at <28 weeks of gestation compared to those born at 33-34 weeks (4.1% (95%CI 1.8 to 7.9%)). The prevalence of reflex abnormalities among those born at <28 weeks gestation was 14.6% (95%CI 9.8 to 20.7) compared with those born at 33-34 weeks gestation (4.6% 95%CI 2.1 to 8.6%). The prevalence of motor behaviour (face and eyes) among those born at <28 weeks gestation was 15.7% (95%CI 10.7 to 22%) compared to those born at 33-34 weeks gestation (10.3% (95%CI 6.4 to 15.4%)) (Arnaud 2007).

- Fine and gross motor problems

Low quality evidence from one study (sample size 367) showed that among children born at mean gestational age of 28.4 (SD 3.0) the prevalence of fine motor problems (Denver II, 1 caution) was 12% (95%CI 9 to 15.8%) at 15 months (median) corrected age (Schendel 1997). Among those born at mean gestational age of 35.6 (SD 2.8) the prevalence of fine motor problems (Denver II, 1 caution) was 8.7% (95%CI 6.5 to 11.3%). For those with fine motor problems (Denver II, 1 delay) the prevalence among those born at 28.4 (SD 3.0) gestation was 7.9% (95% CI 5.4 to 11.1%) whereas the prevalence was 5.2% (95%CI 3.5 to 7.5%) among those born at 35.6 (SD 2.8) mean gestational age (Schendel 1997).

Low quality evidence from one study (sample size 367) showed that among children born at mean gestational age of 28.4 (SD 3.0) the prevalence of gross motor problems (Denver II, 1 caution) was 17.4% (95%CI 13.7 to 21.7%) at 15 months (median) corrected age (Schendel 1997). Among those born at mean gestational age of 35.6 (SD 2.8) the prevalence of gross motor problems was 9% (95%CI 6.6 to 11.6%). For those with gross motor problems (Denver II, 1 delay) the prevalence among those born at 28.4 (SD 3.0) gestation was 10.6% (95%CI 7.7 to 14.2%) whereas the prevalence was 4% (95%CI 2.5 to 6%) among those born at 35.6 (SD 2.8) mean gestational age (Schendel 1997).

Coordination/DCD problems

Low quality evidence from one study (sample size 1662) showed a trend of increasing prevalence of co-ordination and balance with decreasing gestational age ranging from 37.1% (95%CI 30 to 44.6%) among those born at <28 weeks of gestation, compared to those born at 33-34 weeks of gestation (21% (95%CI 15.5 to 27.4%) (Arnaud 2007).

Low quality evidence from one study (sample size 22898) showed a trend of increasing prevalence of probable DCD (DCDQ, =46) with decreasing gestational age, ranging from 14.1% (95%CI 8 to 22.6%) among those born at <32 weeks of gestation, compared to those born at 36 weeks of gestation (4.4% (95%CI 2.6 to 6.8%)) (Zhu 2012).

4.1.3.4.3 Prevalence of developmental delay

Developmental delay at <28 completed weeks of gestation

Very low quality evidence from one study (sample size 78) showed that among children born at <26 weeks of gestation the prevalence of developmental delay (ASQ, corrected for parental education, -2SD) was 22% (95%CI 12 to 33%) at 12-60 months age compared to those children born at 26-27 weeks of gestation (prevalence 13% (95%CI 4 to 21%)) (Plomgaard 2006).

Very low quality evidence from one study (sample size 78) showed that among children born at <26 weeks of gestation the prevalence of developmental delay (ASQ, corrected for parental education, -3SD) was 14% (95%CI 5 to 23%) at 12-60 months age compared to those children born at 26-27 weeks of gestation (prevalence 4% (95%CI 0 to 8%)) (Plomgaard 2006).

Very low quality evidence from one study (sample size 78) showed that among children born at <26 weeks of gestation the prevalence of developmental delay (ASQ, excluding children with neurosensory deficit, -2SD) was 14% (95%CI 0.5 to 23%) at 12-60 months age compared to those children born at 26-27 weeks of gestation (prevalence 13% (95%CI 0 to 22%)) (Plomgaard 2006).

Very low quality evidence from one study (sample size 78) showed that among children born at <26 weeks of gestation the prevalence of developmental delay (ASQ, excluding children with neurosensory deficit, -3SD) was 6% (95%CI 0 to 12%) at 12-60 months age compared to those children born at 26-27 weeks of gestation (prevalence 4% (95%CI 0 to 9%)) (Plomgaard 2006).

Developmental delay at 28-31 completed weeks of gestation

Low quality evidence from one study (sample size 367) showed that among children born at mean gestational age 28.4 weeks (SD3.0) the prevalence of overall developmental delay (Denver II, questionable ≥2 cautions and/or 1 delay) was 17.4% (95%CI 13.7 to 21.7%) at median 15 months corrected age (Schendel 1997). In the same study, the prevalence for overall developmental delay (Denver II, abnormal ≥2 delay scores) was 11% (95%CI 7.9 to 14.6%).

Developmental delay at <32 weeks of gestation

Low quality evidence from one study (sample size 698) showed that among children born at <32 weeks of gestation the prevalence of developmental delay (ASQ total score <-2SD) was 14.9% (95%CI 11.9 to 18.2%) at 4 years age (Kerstjens 2011).

Moderate quality evidence from one study (sample size 237) showed that among children born at 24-31 weeks of gestation the prevalence of uncertain cognitive verbal preschool skills (MAP) was 13.7% (95%CI 8.9 to 19.8%) at 4 years age (Agelholm 2011). In the same study, the prevalence of uncertain cognitive non-verbal preschool skills (MAP) was 6.6% (95%CI 3.3 to 11.4%), and the prevalence of uncertain combined cognitive and motor preschool skills (MAP) was 12.5% (95%CI 7.9 to 18.5%).

Moderate quality evidence from one study (sample size237) showed that among children born at 24-31 weeks of gestation the prevalence of deficit in cognitive verbal preschool skills (MAP) was 10.7% (95%CI 6.5 to 16.4%). The prevalence of deficit in cognitive non-verbal preschool skills (MAP) was 3.6% (95%CI 1.3 to 7.6%) whereas the prevalence of deficit in combined cognitive and motor preschool skills (MAP) was 7.1% (95%CI 3.8 to 12.1%) (Agelholm 2011).

Developmental delay at 32-36 completed weeks of gestation

Low quality evidence from one study (sample size 367) showed that among children born at a mean gestational age of 35.6 the prevalence of developmental delay (Questionable Denver II (≥2 cautions and/or 1 delay)) was 11.8% (95%CI 9.2 to 14.7%) at median 15 months corrected age (Schendel 1997). In the same study, the prevalence of developmental delay (Abnormal Denver II (≥2 delays)) was 5.8% (95%CI 4 to 8.1%).

Moderate quality evidence from one study (sample size 926) showed that among children born at 32-35 weeks of gestation the prevalence of problem-solving problems (ASQ, <-2SD) was 6.1% (95%CI 4.6 to 7.8%) at 4 years age (Potijk 2013).

Low quality evidence from one study (sample size 698) showed that among children born at 32-36 weeks of gestation the prevalence of developmental delay (ASQ total score <-2SD) was 8.3% (95%CI 6.6 to 10.3%) at 4 years age (Kerstjens 2011).

Low quality evidence from one study (sample size 634) showed that among children born at <33 weeks of gestation the prevalence of developmental delay (DQ <70, BLS) was 2.3% (95%CI 1 to 4.5%) whereas prevalence of developmental delay (DQ <85, BLS) was 17.9% (95%CI 14 to 22%) at 2 years (corrected age) (Charkaluk 2010).

Low quality evidence from one study (sample size 1130) showed that among children born at 32-36 weeks of gestation the prevalence of developmental delay (PARCA-R, <2.5th percentile) was 6.3% (95%CI 4.5 to 8.4%) at 2 years (corrected age) (Johnson 2015).

4.1.3.4.4 Prevalence of language delay

Language problems at <28 completed weeks of gestation

Receptive communication

Low quality evidence from one study (sample size 394) showed that among children born at <27 weeks of gestation the prevalence of receptive communication problems (Bayley, mild -1SD to -2SD) was 24.9% (95%CI 20.7 to 30.0%) at 2.5 years age. In the same study, the prevalence of moderate receptive communication problems (Bayley -2SD to -3SD) was 9.1% (95%CI 6.5 to 12.4%). The prevalence of moderate to severe (Bayley -3SD) receptive communication was 5.8% (95%CI 3.7 to 8.6%) (Mansson 2014).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks of gestation the prevalence of receptive communication problems (OWLS <=-2SD) was 19% (95% CI 16.5 to 21.8) when assessed at 10 years age (Joseph 2016b).

Expressive communication

Low quality evidence from one study (sample size 394) showed that among children born at <27 weeks of gestation the prevalence of expressive communication problems (Bayley, mild -1SD to -2SD) was 31.3% (95%CI 26.7 to 36.1%) at 2.5 years age (Mansson 2014). In the same study, prevalence of moderate expressive communication (Bayley moderate -2SD to -3SD) problems was 8.1% (95%CI5.6 to 11.3%), and for moderate to severe expressive communication problems (Bayley -3SD), the prevalence was 6.4% (95%CI 4.2 to 9.3%) (Mansson 2014).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks of gestation the prevalence of expressive communication problems (OWLS <=-2SD) was 19% (95% CI 16.5 to 21.8) when assessed at 10 years age (Joseph 2016b).

Language problems at 28-31 completed weeks of gestation

Low quality evidence from one study (sample size 367) showed that among children born at a mean gestational age of 28.4 (SD 3.0) the prevalence of language problems (Denver II ≥1 caution or ≥1 delay) was 17% (95%CI 13.2 to 21.1%) and 8.7% (95%CI 6.0 to 12.0%) respectively at a median 15 months corrected age (Schendel 1997).

Language problems at <32 weeks of gestation

Low quality evidence from one study (sample size 924) showed that among children born at <32 weeks of gestation the prevalence of language problems was 4.6% (95%CI 3.1 to 6.6%) at 5 years age (Rautava 2010).

Language problems at 32-36 completed weeks of gestation

Moderate quality evidence from one study (sample size 926) showed that among children born at 32-35 weeks of gestation the prevalence of communication problems (ASQ <-2SD) was 9.5% (95%CI 7.7 to 11.6%) at 4 years age (Potijk 2013).

Low quality evidence from one study (sample size 39423) showed that among children born at 34-36 weeks of gestation the prevalence of communication problems (ASQ 2SD) was 7.3% (95%CI 6.1 to 8.6%) at 18 months age, and 6.3% (95%CI 5.2 to 7.2%) at 36 months age (Stene-Larsen 2014).

Low quality evidence from one study (sample size 920) showed that among children born at a mean gestational age of 35.6 weeks (SD 2.8) the prevalence of language problems (Denver II ≥1 caution or ≥1 delay) was 11.9% (95%CI 9.4 to 14.9%) and 5.8% (95%CI 4.0 to 8.1%) respectively at median 15 months corrected age (Schendel 1997).

4.1.3.4.5 Prevalence of executive function problems

Executive function problems at <28 completed weeks of gestation

Moderate quality evidence from one study (sample size 275) showed that among children born at <28 weeks of gestation, the prevalence of executive function problems (BRIEF, >=1.5 SD above mean) was 13.1% (95%CI 9.1 to 17.9%) (Anderson 2004).

Low quality evidence from one study (sample size 201) showed that among children born at <28 weeks of gestation the prevalence of executive attention-inhibitory control (Opposite Worlds, <-1SD) was 6% (95%CI 2.9 to 10.7%) (Anderson 2011). In the same study executive attention-inhibitory control (BRIEF-Inhibit T score >60) was 15% (95%CI 10.2 to 20.9%). The prevalence of shifting attention (creature counting <-1SD) was 27.1% (95%CI 20.5 to 34.4%) whereas prevalence using BRIEF (T score >60) was 19% (95%CI 13.6 to 25.5%) (Anderson 2011).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks of gestation the prevalence of executive function regarding working memory (DAS-II <-2SD) was 18% (95%CI 15.5 to 20.7), auditory attention 23% (95%CI 20.3 to 26.0) (NEPSY-II <=-2SD), auditory response set 20% (95%CI 17.4 to 23), Inhibition 34% (95%CI 31-37) (NEPSY-II), inhibition switching 27% (95%CI 24.1 to 30.1) (NEPSY-II <=-2SD) (Joseph 2016b). In the same study, the prevalence of processing speed was 31% (95%CI 28-34) (NEPSY-II <=-2SD), and the prevalence of visual perception was 26% (95%CI 23-29) (NEPSY-II Arrows, <=-2SD) and 17% (95%CI 14.5 to 19.6) (NEPSY-II Geometric puzzles <=-2SD) (Joseph 2016b).

Executive function problems at <32 completed weeks of gestation

Low quality evidence from one study (sample size 924) showed that among children born at <32 weeks of gestation, the prevalence of executive function problems (FTF) was 7.8% (95%CI 5.8 to 10.3) (Rautava 2010). In the same study, the prevalence of memory problems was 8.3% (95%CI 6.2 to 11.0), and the prevalence of perception problems was 3.9% (95%CI 2.5 to 5.8)

4.1.3.4.6 Prevalence of behavioural, social and emotional problems

Behavioural, social and emotional problems at <28 completed weeks of gestation

Total behavioural problems

Low to moderate quality evidence from two separate studies (sample size 1645 to 2382) showed that among children born at 24-28 weeks and 24-27weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile) ranged from 24.1% (95%CI 19.2 to 29.6%) 22.2% (95%CI 17.1 to 28.1%) at 3 years age and 5 years age respectively (Delobel-Ayoub 2006; Foix-Helias 2008).

Low quality evidence from one study (sample size 224) showed that among children born at <26 weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile, parent reported) was 38.5% (95%CI 32.0 to 45.2%) and 34.6% (95%CI 29.2 to 41.5%) (teacher-reported) at 6 years age (Samara 2008).

Low quality evidence from one study (sample size 2901) showed that among children born at 24-28 weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile, parent reported) was 27.8% (95%CI 23.0 to 32.9%) at 8 years age (Larroque 2011). In another moderate quality study (sample size 189), among children born at <28 weeks of gestation, the prevalence of total behavioural difficulties was 18% (95%CI 12.8 to 24.2%) (Hutchinson 2013).

Low quality evidence from one study (sample size 568) showed that among children born at <28 weeks gestation the prevalence of total behavioural problems (at risk, BASC) was 15.0% (95%CI 11.0 to 19.7%) whereas in the same population those who had clinically significant behavioural problems (BASC) the prevalence was 7.0% (95%CI 4.2 to 10.6%) at 8 years corrected age (Anderson 2003).

Moderate quality evidence from one study (sample size 154) showed that among children born at <27 weeks of gestation the prevalence of total difficulties (SDQ score 17 to 40) was 28.0% (95%CI 18.2 to 39.6%) at 7-9 years age (Stahlman 2009).

Moderate quality evidence from one study (sample size 169) showed that among children born at <26 weeks of gestation the prevalence of total behavioural problems (CBCL, 90th percentile, parent reported) was 28.9% (95%CI 19.5 to 39.9%) and 24.1% (95%CI 15.4 to 34.7%) (teacher-reported CBCL, 90th percentile) at 11 years age (Farooqi 2007).

Low quality evidence from one study (sample size 2855) showed that among children born SGA at 24-28 weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile) was 33.3% (95%CI 14.6 to 57%) compared with those children born MGA (27.3% (95%CI 13.3 to 45.5%)) at 5 years age. For those children born AGA the prevalence was 23.7% (95%CI 19.3 to 28.5%) (Guellec 2011).

ADHD symptoms

Low quality evidence from one study (sample size 201) showed that among children born at <28 weeks of gestation the prevalence of ADHD symptoms (CADS-P, inattentive symptoms, T score >60) was 32.1% (95%CI 20.3 to 46%) at 8 years corrected age (Anderson 2011). In the same study, the prevalence of ADHD symptoms (hyperactivity-impulsive symptoms, T score >60) and ADHD index (CADS-P, T score >60) was 41.8% (95%CI 28.7 to 55.9%) and 43% (95%CI 30.3 to 57.7%) respectively (Anderson 2011).

Low quality evidence from one study (sample size 298) showed that among adolescents born at <28 weeks of gestation the prevalence of ADHD (DSM-IV, <-1.5 SD, parent reported) was 17.6% (95%CI 12.5 to 23.7%) at 17 years age whereas the prevalence of ADHD reported by adolescents themselves was 5.2% (95%CI 2.5 to 9.4%) at 17 years age (Wilson-Ching 2013).

ASD symptoms

Low quality evidence from one study (sample size 307) showed that among children born at <26 weeks of gestation the prevalence of positive ASD screen (SCQ ≥15, parent reported) was 15.8% (95%CI 10.9 to 22.0%) at 11 years age (Johnson 2010).

Moderate quality evidence from one study (sample size 1198) showed that among children born at <27 weeks of gestation the prevalence of positive ASD screen (SCQ ≥11, parent reported) was 12.4% (95% CI 10.2 to 14.8%) at 10 years age (Joseph 2016a).

Attention/hyperactivity symptoms

Low quality evidence from one study (sample size 2855) showed that among children born SGA at 24-28 weeks of gestation the prevalence of inattention/hyperactivity (SDQ 10th percentile) was 19% (95%CI 5.5 to 42%) compared with those children born MGA (21.2% (95%CI 9 to 38.9%)). For those children born AGA the prevalence was 21.7% (95%CI 17.5 to 26.4%)) (Guellec 2011).

Moderate quality evidence from one study (sample size 826) showed that among children born at <28 weeks of gestation the prevalence of attention problems (CBCL 93rd percentile) was 10.7% (95%CI 8.6 to 13.0%) at 24 months corrected age (Downey 2016).

Behavioural, social and emotional problems at 28-31 completed weeks of gestation

Total behavioural problems

Low quality evidence from one study (sample size 2382) showed that among children born at 29-32 weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile) was 18.2% (15.8 to 20.9%) at 3 years (Delobel-Ayoub 2006). At 5 years age (moderate quality evidence, sample size 1645), the prevalence of total behavioural problems (SDQ, 10th percentile) in children born at 28-32 weeks gestation was 21% (95%CI 18.9 to 23.2%) (Foix-Helias 2008).

Low quality evidence from one study (sample size 2901) showed that among children born at 29-32 weeks of gestation, the prevalence of total behavioural difficulties (SDQ, 10th percentile) was 18.9% (95%CI 16.6 to 21.4%) at 8 years age (Larroque 2011).

Behavioural, social and emotional problems at <32 weeks of gestation

Total behavioural problems

Low quality evidence from one study (sample size 235) showed that among children born at <32 weeks of gestation the prevalence of total behavioural problems (CBCL, >98th percentile) was 8.9% (95%CI 4.9 to 14.4%) at 2 years (corrected age) (Stoelhorst 2003a).

Low quality evidence from one study (sample size 2382) showed that among children born at <33 weeks of gestation the prevalence of total behavioural problems (SDQ, 10th percentile) was 20% (95%CI 17.7 to 22.3%) at 3 years age (Delobel-Ayoub 2006).

Low quality evidence from one study (sample size 1504) showed that among children born at 25-31 weeks of gestation the prevalence of emerging total behavioural problems (CBCL, >85th percentile) was 5.2% (95%CI 3.3 to 7.9%) at 4 and 5 years age (Hornman 2016). In the same study, the prevalence of resolving and persistent total behavioural problems (CBCL, 85th percentile) was 5.5% (95%CI 3.5 to 8.2%) and 8.2% (95%CI 5.7 to 11.4%) respectively (Hornman 2016).

Low to moderate quality evidence from two separate studies (sample size ranged from 566 to 924) showed that among children born at <32 weeks of gestation the prevalence of total behavioural problems (CBCL, >=55, parent reported) was 13.8% (95%CI 10.6 to 17.5%) and 3.4% (95%CI 2.1 to 5.2%) when measured on FTF for emotional and behavioural problems at 5 years age respectively (de Kleine 2003; Rautava 2010). In another study (sample size 1645) among children born at 24-32 weeks of gestation the prevalence of total behavioural problems (SDQ, 10th percentile) was 21.2% (95%CI 19.2 to 23.2%) at 5 years age (Foix-Helias 2008).

Low quality evidence from one study (sample size 2901) showed that among children born at 24-32 weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile) was 21.1% (95%CI 18.9 to 23.3%) at 8 years age (Larroque 2011).

Low quality evidence from one study (sample size 2855) showed that among children born SGA at 29-32 weeks of gestation the prevalence of total behavioural difficulties (SDQ 10th percentile) was 19.1% (95%CI 12.4 to 27.5%) compared to those children born MGA (26.5% (95%CI 18.8 to 35.2%)) at 5 years age. For those children born AGA the prevalence was 19.4% (95%CI 17 to 21.9%)) (Guellec 2011).

Attention/hyperactivity symptoms

Low quality evidence from one study (sample size 2855) showed that among children born SGA at 29-32 weeks of gestation the prevalence of inattention/hyperactivity was 23.5% (95%CI 16 to 32.3%) compared with those children born MGA (15.7% (95%CI 9.7 to 23.4%)) at 5 years age. For those children born AGA the prevalence was 15% (95%CI 12.9 to 17.3%)) (Guellec 2011).

Behavioural, social and emotional problems at 32-36 completed weeks of gestation

Total behavioural problems

Low quality evidence from one study (sample size 625) showed that among children born at 32-36 weeks of gestation the prevalence of behavioural problems (BITSEA, >25th percentile) was 21% (95%CI 17.8 to 24.4%) at 2 years (corrected age) (Johnson 2015). In the same study, the prevalence of delayed social incompetence (BITSEA <15th percentile) was 26.4% (95%CI 23 to 30%). For children who had behavioural problems or delayed social competence (BITSEA), or both, the prevalence was 37.3% (95%CI 33.5 to 41.2%) and 10.1% (95%CI 7.8 to 12.7%) respectively (Johnson 2015).

Moderate quality evidence from one study (sample size 916) showed that among children born at 32-35 weeks of gestation the prevalence of total behavioural problems (CBCL, 90th percentile) was 7.9% (95%CI 6.2 to 9.8%) at 4 years age (Potijk 2012).

Low quality evidence from one study (sample size 1504) showed that among children born at 32-35 weeks of gestation the prevalence of emerging total behavioural problems (CBCL, >85th percentile) was 3.7% (95%CI 2.4 to 5.4%) at 4 and 5 years age (Hornman 2016). In the same study, the prevalence of resolving or persistent total behavioural problems (CBCL, >85th percentile) was 8.7% (95%CI 6.7 to 11.2%) and 6.6% (95%CI 4.8 to 8.8%) respectively (Hornman 2016).

Behavioural, social and emotional problems by week or group of gestation

Total behavioural problems by gestational age (SDQ)

Low quality evidence from one study (sample 2382) showed that among children born at 24-28 weeks of gestation the prevalence of total behavioural difficulties (SDQ, 10th percentile) was 24.1% (95%CI 19.2 to 29.6%) at 3 years age (Delobel-Ayoub 2006). In the same study, the prevalence decreased to 16.9% (95%CI 13 to 21.3%) among children born at 29-30 weeks of gestation whereas there was an increase in prevalence of 19% (95%CI 15.9 to 22.4%) among children born at 31-32 weeks of gestation (Delobel-Ayoub 2006).

A similar pattern was observed in another low quality study (sample size 2901) showed that among children born at 24-28 weeks the prevalence of total behavioural difficulties (SDQ, 10th percentile) was 27.8% (95%CI 23 to 32.9%) at 8 years age (Larroque 2011). In the same study, the prevalence decreased to 17.2% (95%CI 13.5 to 21.4%) among children born at 29-30 weeks of gestation, whereas there was an increase in prevalence of 19.9% (95%CI 16.9 to 23.1%) among children born at 31-32 weeks of gestation (Larroque 2011).

ASD symptoms by week of gestation

Low quality evidence from one study (sample size 2035) showed that there was an increase in prevalence of positive autism screening (M-CHAT) with decreasing gestational age, ranging from 54.8% (95%CI 36 to 72.7%) at 23 weeks of gestation to 38.1% (95%CI 31.7 to 44.7%) at 26 weeks of gestation (assessed at 2 years age) (Moore 2012).

Low quality evidence from one study (Sample size 1130-2035) showed that among children born at <27 weeks of gestation the prevalence of autism (positive screen, M-CHAT) was 41% (95%CI 37 to 45.7%) at 2 years age (Moore 2012) compared to the prevalence of those children born at 32-33 or 34-36 weeks of gestation (9.3% (95%CI 4.1 to 17.5%) and 15.3% (95%CI 12.4 to 18.6%)) respectively (Guy 2015).

Total externalising behavioural problems by gestational group

Moderate quality evidence from one study (sample size 169) showed that among children born at <26 weeks of gestation the prevalence of externalising problems (CBCL, 90th percentile) was 9.6% (95%CI 4.3 to 18.1%) at 11 years age (Farooqi 2007). In the same study, the prevalence of externalising problems (TRF, 90th percentile) was 18.1% (95%CI 10.5 to 28.1%).

Moderate quality evidence from one study (sample size 916) showed that among children born at 32-35 weeks of gestation the prevalence of externalising problems (CBCL, 84th percentile) was 9.5% (95%CI 7.7 to 11.6%) at 4 years age (Potijk 2012).

Low quality evidence from one study (sample size 1054) showed that among children born at 25-31 weeks and 32-35 weeks of gestation the prevalence of emerging externalising problems (CBCL, >85th percentile) was 5.2% (95%CI 3.3 to 7.9%) and 5.4% (95%CI 3.8 to 7.4%) respectively at 4 and 5 years age (Hornman 2016). In the same study, the prevalence for resolving externalising problems at 25-31 and 32-35 weeks of gestation was 5.2% (95%CI 3.3 to 7.9%) and 8.4% (95%CI 6.4 to 10.3%) respectively. The prevalence of persistent externalising problems (CBCL, >85th percentile) at 25-31 and 32-25 weeks of gestation was 8.2% (95%CI 5.7 to 11.4%) and 8.4% (95%CI 6.4 to 10.8%) respectively at 4 and 5 years age (Hornman 2016).

Total internalising behavioural problems by gestational group

Moderate quality evidence from one study (sample size 169) showed that among children born at <26 weeks of gestation the prevalence of internalising problems (CBCL, 90th percentile) was 32.5% (95%CI 22.7 to 43.7%) at 11 years age (Farooqi 2007). In the same study, the prevalence of internalising problems (TRF, 90th percentile) was 25.3% (95%CI 16.4 to 36%).

Moderate quality evidence from one study (sample size 916) showed that among children born at 32-35 weeks of gestation the prevalence of internalising problems (CBCL, 84th percentile) was 9.7% (95%CI 7.9 to 11.8%) at 4 years age (Potijk 2012).

Low quality evidence from one study (sample size 1054) showed that among children born at 25-31 and 32-35 weeks of gestation the prevalence of emerging internalising problems (CBCL, >85th percentile) was 8% (95%CI 5.5 to 11.1%) and 6.7% (95%CI 4.9 to 8.9%) at 4 and 5 years age respectively (Hornman 2016). In the same study, the prevalence of resolving internalising problems (CBCL, >85th percentile) at 25-31 and 32-35 weeks of gestation was 7.2% (95%CI 4.9 to 10.2%) and 7.5% (95%CI 5.6 to 9.8%) respectively. The prevalence of persistent internalising problems at 25-31 and 32-35 weeks gestation was 11.7% (95%CI 8.7 to 15.3%) and 10.1% (95%CI 7.9 to 12.7%) respectively (Hornman 2016).

Attention/hyperactivity, conduct, and emotional problems

Attention/hyperactivity problems

Moderate quality evidence from one study (sample size 169 to 916) showed that among children born at <26 weeks of gestation the prevalence of attention problems (CBCL, 90th percentile) was 30.1% (95%CI 20.5 to 41.2%) and 24.1% (95%CI 15.4 to 34.7%) using the TRF (90th percentile) at 11 years age (Farooqi 2007). In another moderate quality study (sample size x) the prevalence of attention problems (CBCL, >97th percentile) was 4.15% (95%CI 3 to 5.7%) among children born at 32-35 weeks of gestation, assessed at 4 years age (Potijk 2012).

Low quality evidence from one study (sample size x) showed that among children born at <34 or 34-36 weeks of gestation the prevalence attention problems (failure to pay attention when crossing street (CBCL) was 22.8% (95%CI 18.5 to 27.5%) and 20.6% (95%CI 18.4 to 22.9%) respectively (Higa-Diez 2016). In the same study the prevalence of adverse outcomes for all attention problems (CBLC) was 9.4% (95%CI 5.6 to 14.6%) and 5.6% (95%CI 4 to 7.6%) among those born at <34 or 34-36 weeks of gestation, assessed at 8 years age.

Low quality evidence from three separate studies (sample range 201 to 224) showed a trend of higher prevalence of attention problems (using different tools) among children born at <26 or 28 weeks of gestation (range 30.1% (95%CI 23.3 to 37.5%) to 54% (95%CI 47 to 60.8%)) assessed at 6 and 8 years age respectively (Samara 2008; Anderson 2011).

Low quality evidence from one study (sample size) showed that among children born at <34 or 34-36 weeks of gestation the prevalence of interrupting people (CBCL) was 41.9% (95%CI 36.7 to 47.2%) and 40.3% (95%CI 37.6 to 43.1%) respectively (Higa-Diez 2016). In the same study, the prevalence of inability to wait turn (CBCL) was 12.6% (95%CI 9.4 to 16.6%) and 9,1% (95%CI 7.6 to 10.8%) respectively.

Adolescents (sample size 298) born at <28 weeks of gestation had a lower prevalence of hyperactive or inattentive (CADS <-1.5SD) problems, ranging from 14.5% (95%CI 9.9 to 20.1%) (Wilson-Ching 2013). In the same study, the prevalence of shifting attention (CNT, <-1.5SD) or divided attention (Telephone search wile counting/Test of Everyday Attention <-1.5SD) was 41.1% (95%CI 34.4 to 48.2%) and 15.3% (95%CI 10.6 to 21.1%) respectively (Wilson-Ching 2013).

Moderate to low quality evidence from four separate studies (sample size range 224 to 2901) showed a trend of high prevalence of hyperactivity problems (SDQ, >90th percentile) among those born at low gestational age of <26 weeks (48% (95%CI 41.3 to 54.8%)) (Samara 2008) compared to a lower prevalence among those born at higher gestational age of 24-32 weeks (17.2% (95%CI 15.3 to 19.3%) (Larroque 2011).

Low quality evidence from two separate studies (sample ranged from 2382 to 2901) showed that among children born at 24-28 weeks of gestation the prevalence of hyperactivity (SDQ, 10th percentile) was 24.1% (95%CI 19.2 to 29.6%) and 18.5% (95%CI 14.5 to 23.1%) at 3 years and 8 years age respectively (Delobel-Ayoub 2006; Larroque 2011). In the same two studies, the prevalence of hyperactivity ranged from 17.1% (95%CI 13.3 to 21.6%) to 15.1%(95%CI 11.6 to 19.1%) at 29-31 weeks of gestation, assessed at 3 and 8 years age respectively (Delobel-Ayoub 2006; Larroque 2011). The prevalence ranged from 18.5% (95%CI 14.5 to 23.1) to 17% (95%CI 15 to 20.9%) at 31-32 weeks of gestation, assessed at 3 and 8 years (Delobel-Ayoub 2006; Larroque 2011).

Conduct problems

Low to moderate quality evidence from four separate studies (sample range 224 to 2901) showed a general trend of decreasing prevalence of conduct problems (SDQ, 10th percentile) with increasing gestational age, ranging from 36.2% (95%CI 29.9 to 42.9%) (<26 weeks gestational age) (Samara 2008) to 9.4 % (95%CI 8.0 to 11.1%) (24-32 weeks gestational age) (Larroque 2011).

Low quality evidence from one study (Sample size 2382 to 2901) showed that the prevalence of conduct problems (SDQ, 10th percentile) decreased with increasing gestational age group from 16.1% (11.9 to 21%) at 24-28 weeks of gestation to 15% (95%CI 12.2 to 18.1%) at 31-32 weeks of gestation (assessed at 3 years age) (Delobel-Ayoub 2006). At 8 years, there was no clear trend of prevalence with gestational age group (Larroque 2011).

Emotional problems

Low quality evidence from two separate studies (sample size 2382 to 2901) showed that among children born at 24-28 weeks of gestation the prevalence of emotional symptoms (SDQ, 10th percentile) was 17.2% (95%CI 12.9 to 22.2%) and 20.3% (95%CI 16.1 to 25%) at 3 years and 8 years respectively (Delobel-Ayoub 2006; Larroque 2011). In the same two studies, the prevalence of emotional problems among children born at 29-30 weeks of gestation was 14.1% (95%CI 10.6 to 18.3%) and 14.3% (95%CI 10.9 to 18.2%) at 3 and 8 years age respectively. Prevalence of emotional problems among those born at 31-32 weeks of gestation was 15% (95%CI 12.2 to 18.1%) and 17.2% (95%CI 14.4 to 20.3%) at 3 and 8 years age (Delobel-Ayoub 2006; Larroque 2011).

Moderate quality evidence from one study (sample size 224 to 916) showed that among children born at 32-35 weeks of gestation the prevalence of emotionally reactive problems (CBCL, >97th percentile) was 3.7% (95%CI 2.6 to 5.2%) (Potijk 2012). In other studies, the prevalence of emotional problems was higher among those born at lower gestational age of <26 weeks of gestation (29.9% (95%CI 23.8 to 36.5%) (Samara 2008).

- Peer and prosocial problems

Low quality evidence from one study (sample size 224 to 2901) showed that among children born at <26 weeks of gestation the prevalence of peer problems (SDQ, >90th percentile) was 36% (95%CI 29.7 to 42.7%, parent reported) and 50% (95%CI 43.5 to 57.4%, teacher reported) respectively (Samara 2008). The prevalence of peer problems (SDQ, >90th percentile) was lower with varying gestational age groups, ranging from 17.4% (95%CI 15.4 to 19.5%) at 24-32 weeks of gestation (Larroque 2011) to 20% (95%CI 17.7 to 22.6%) in those born at 22-32 weeks of gestation (Delobel-Ayoub 2009).

Low quality evidence from two separate studies (Sample range from 2382 to 2901) showed a trend of decreasing prevalence of peer problems (SDQ, 10th percentile) with increasing gestational age, ranging from 17.9% (95%CI 13.5 to 22.9) among those born at 24-28 weeks of gestation to 12% (95%CI 9.5 to 14.9%) among those born at 31-32 weeks of gestation (Delobel-Ayoub 2006; Larroque 2011). A similar trend was observed in another low quality study (sample size 2382) with prevalence ranging from 19.4% (95%CI 15.3 to 24.1%) among those born at 24-28 weeks of gestation to 15.4% (95%CI 12.8 to 18.4%) among those born at 31-32 weeks of gestation (Larroque 2011).

Low quality evidence from one study (sample size 2382) showed a trend of decreasing prevalence of prosocial behaviour (SDQ, 10th percentile) with increasing gestational age, ranging from 20.1% (95%CI 15.5 to 25.3%) among those born at 24-28 weeks of gestation to 13% (95%CI 10.4 to 16%) among those born at 31-32 weeks of gestation (Delobel-Ayoub 2006), assessed at 3 years age.

4.1.3.4.7 *Prevalence of special education needs*

Special education needs at <28 completed weeks of gestation

Special education needs (overall and individual problems (sensory, physical/motor, language, intellectual, specific learning difficulties, ASD, social, emotional behavioural))

Low quality evidence from one study (sample size 152757) showed that among children born at 24-27 weeks of gestation the prevalence of SEN was 29.5% (95%CI 25.4 to 33.8%) at 5-18 years age (Mackay 2010).

Low quality evidence from one study (sample size 237894) showed that among children born at 24-27 weeks of gestation the prevalence of sensory SEN was 3% (95%CI 1.6 to 4.9%), physical or motor SEN was 6.1% (95%CI 4.1 to 8.7%), language SEN was 0.63% (95%CI 0.13 to 1.83%), intellectual SEN was 14.1% (95%CI 11.1 to 17.6%), specific learning difficulties SEN was 2.1% (95%CI 1.0 to 3.8%), ASD SEN was 1.1% (95%CI 0.3 to 2.4%), and social, emotional behavioural SEN was 1.3% (95%CI 0.5 to 2.7%) at 5-18 years (Mackay 2013).

School difficulties (low grade, repetition of grade, adaption difficulties)

Moderate quality evidence from one study (sample size 169) showed that among children born at <26 weeks of gestation, the prevalence of school difficulties (repetition of grade and/or use of SEN resources) was 59.3% (95%CI 48.2 to 69.8%) at 11 years age (Farooqi 2007). In the same study, the prevalence of grade repetition was 15.7% (95%CI 8.6 to 25.3%).

Low quality evidence from one study (sample size 2382) showed that among children born at <26 weeks of gestation the prevalence of school adaption difficulties (parent reported) was 33% (95%CI 36.7 to 39.8%) compared to a prevalence (teacher reported) of 39.2% (95%CI 32.6 to 46.2%) at 6 years age (Samara 2008).

Low quality evidence from one study (sample size 2855) showed that among children born at 24-28 weeks of gestation and were small for gestational age, the prevalence of school difficulties (special schooling or low grades, parent reported) was 35.3% (95%CI 14.2 to 61.7%) at 8 years age compared to those who were born MGA (prevalence 44.8% (95%CI 26.5 to 64.3%) (Guellec 2011).

Identified special education needs

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks of gestation the overall prevalence of identified SEN (teacher reported) was 62.3% (95%CI 55.5 to 68.8%) at 11 years age. In the same study, the prevalence of SEN identified in mainstream schools only (teacher reported) was 56.5% (95%CI 49 to 63.7%) (Johnson 2011).

Special school or special class

Moderate quality evidence from one study (sample size 169) showed that among children born at <26 weeks of gestation the prevalence of those in special class or special school was 15.1% (95%CI 8.3 to 24.5%) at 11 years age (Farooqi 2007).

Low quality evidence from one study (sample size 2901) showed that among children born at 24-28 weeks of gestation the prevalence of those in an institution or special school or special class (parent reported) was 9.4% (95%CI 6.5 to 13.0%) at 8 years age (Larroque 2011).

Special education needs provision/support at school

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks of gestation the prevalence of SEN provision (teacher reported) was 61.4% (95%CI 54.5 to 68.8%) at 11 years age (Johnson 2011). In the same study, among children who had SEN provision in mainstream school only (teacher reported) the prevalence was 55.4% (95%CI 47.9 to 62.7).

Low quality evidence from one study (sample size 2901) showed that among children born at 24-28 weeks of gestation the prevalence of support in mainstream school (parent reported) was 22.7% (95%CI 18.3 to 27.5%) at 8 years age (Larroque 2011). In the same study, the prevalence of children who had special care since 5 years age or support at school (parent reported) was 69.7% (95%CI 64.5 to 74.5%) at 8 years age. The prevalence of children who had special care since 5 years for more than one developmental problem (orthoptic, speech therapy, PT, OT or psychology) was 65.4% (95%CI 60.1 to 70.4%) at 8 years age (Larroque 2011).

Special education needs at 28-31 completed weeks of gestation

Special education needs (overall and individual problems (sensory, physical/motor, language, intellectual, specific learning difficulties, ASD, social, emotional and behavioural))

Low quality evidence from one study (sample size 152757) showed that among children born at 28-32 weeks of gestation the prevalence of SEN was 12.8% (95%CI 11.7 to 14%) at 5-18 years age (Mackay 2010).

Low quality evidence from one study (sample size 237894) showed that among children born at 28-32 weeks of gestation the prevalence of sensory SEN was 0.49% (95%CI 0.29 to 0.79%), physical or motor SEN was 2.8% (95%CI 2.3 to 3.5%), language SEN was 0.38% (95%CI 0.2 to 0.6%), intellectual SEN was 4.8% (95%CI 4.1 to 5.6%), specific learning difficulties SEN was 1.4% (95%CI 1.1 to 1.9%), ASD SEN was 1.0% (95%CI 0.7 to 1.4%), and social, emotional behavioural SEN was 0.9% (95%CI 0.6 to 1.3%) at 5-18 years (Mackay 2013).

School difficulties (special schooling or low grades)

Low quality evidence from one study (sample size 2855) showed that among children born at 29-32 weeks of gestation who were small for gestational age, the prevalence of school difficulties was 28% (95%CI 19.8 to 37.6%) at 8 years age compared to a prevalence of 23.1% (95%CI 15.4 to 32.4%) among children who were MGA (Guellec 2011).

Special school or special class

Low quality evidence from one study (sample size 2901) showed that among children born at 29-30 weeks of gestation the prevalence of those in an institution or special school or special class (parent reported) was 5.2% (95%CI 3.2 to 7.9%) at 8 years age (Larroque 2011).

Support at school

Low quality evidence from one study (sample size 2901) showed that among children born at 29-30 weeks of gestation the prevalence of support in mainstream school (parent reported) was 10.3% (95%CI 7.5 to 13.8%) at 8 years age (Larroque 2011). In the same study, the prevalence of children who had special care since 5 years age or support at school (parent reported) was 53.6% (95%CI 48.5 to 58.7%) at 8 years age.

Special education needs at <32 weeks gestation

Special school or special class

Low quality evidence from one study (sample size 2901) showed that among children born at 24-32 weeks of gestation the prevalence of those in an institution, special school or special class (parent reported) was 5.2% (95%CI 4.1 to 6.5%) at 8 years age (Larroque 2011). For those children born at 31-32 weeks of gestation, the prevalence of the same outcome was 3.3% (95%CI 2.1 to 4.8%) at 8 years age.

Support at school

Low quality evidence from one study (sample size 2901) showed that among children born at 24-32 weeks of gestation the prevalence of those supported at school in mainstream class (parent reported) was 15.4% (95%CI 13.6 to 17.4%) at 8 years age (Larroque 2011). For those children born at 31-32 weeks of gestation, the prevalence of the same outcome was 14.7% (95%CI 12.2 to 17.5%) at 8 years age.

In the same study among children born at 24-32 weeks of gestation, the prevalence among those who had special care since 5 years age or support at school (parent reported) was 58.5% (95%CI 55.9 to 61.1%) at 8 years age. Among children born at 31-32 weeks of gestation, the prevalence of the same outcome was 55.7% (95%CI 52 to 59.4%) at 8 years age (Larroque 2011).

Achievement (FSP or KS1)

Moderate quality evidence from one study (sample size 8728) showed that among children born at 23-31 weeks of gestation the prevalence of those not attaining a good overall level of achievement (teacher reported FSP) was 66.7% (95%CI 55.5 to 76.6%) at 5 years age (Quigley 2012). In the same study, among children who did not attain in all three scales of personal, social and emotional development (teacher reported FSP) the prevalence was 42.9% (95%CI 32.1 to 54.1%). Among children who did not attain in all 4 scales of communication, language and literacy, the prevalence was 61.9% (95%CI 43.5 to 65.7%). The prevalence was 54.8% (95%CI 43.5 to 65.7%) among children who did not attain in all 3 scales of mathematical development a 5 years age (Quigley 2012).

Moderate quality evidence from one study (sample size 18818) showed that among children born at <32 weeks of gestation not achieving level 2 or more in reading, writing or maths (teacher reported KS1) the prevalence was 42% (95%CI 30.2 to 54.5%) at 7 years age (Chan 2014). In the same study, the prevalence among children not achieving level 2 or more in speaking and listening was 29% (95%CI 18.7 to 41.2%) and for science, the prevalence was 24.6% (95%CI 15.1 to 36.5%) (Chan 2014).

Special education needs at 32-36 completed weeks of gestation

Special education needs (overall and individual problems (sensory, physical/motor, language, intellectual, specific learning difficulties, ASD, social, emotional and behavioural))

Low quality evidence from one study (sample size 152757) showed that among children born at 33-36 weeks of gestation the prevalence of SEN was 7.1% (95%CI 6.7 to 7.5%) at age 5-18 years age (Mackay 2010).

Very low quality evidence from one study (sample size 741) showed that among children born at 32-36 weeks of gestation the prevalence of SEN was 34.5% (95%CI 29.3 to 40%) at 8 years age (Odd 2012).

Individualised programme

Low quality evidence from one study (sample size 17565) showed that among children born at 32-36 weeks of gestation the prevalence of those who were enrolled on an individualised education programme (ECLS-K data) was 9.1% (95%CI 7.1 to 11.4%) at kindergarten stage (3 years age), 12% (95%CI 9.7 to 14.6%) at 1st grade (6-7 years), 13.6% (95%CI 11.1 to 16.5%) at 3rd grade (8-9 years) and 16.4% (95%CI 12.9 to 20.4%) at 5th grade (10-11 years) (Chyi 2008).

Low quality evidence from one study (sample size 17565) showed that among children born at 32-33 weeks of gestation the prevalence of those who were enrolled on an individualised education programme was 13% (95%CI 8 to 19%) at kindergarten stage (3 years age), 17.8% (95%CI 12 to 25%) at 1st grade (6-7 years), 19.7% (95%CI 13.3 to 27.5%) at 3rd grade (8-9 years) and 18.1% (95%CI 10.9 to 27.4%) at 5th grade (10-11 years) (Chyi 2008).

Low quality evidence from one study (sample size 17565) showed that among children born at 34-36 weeks of gestation the prevalence of those who were enrolled on an individualised education programme was 8% (95%CI 6 to 10.6%) at kindergarten stage (3 years age), 10.5% (95%CI 8.2 to 13.3%) at 1st grade (6-7 years), 12.1% (95%CI 9.5 to 15.2%) at 3rd grade (8-9 years) and 12.2% (95%CI 9.2 to 15.8%) at 5th grade (10-11 years) (Chyi 2008).

Special education enrolment

Low quality evidence from one study (sample size 17565) showed that among children born at 32-36 weeks of gestation the prevalence of those who were enrolled on a special education programme (ECLS-K data) was 6.9% (95%CI 5.4 to 8.7%) at kindergarten stage (3 years age), 7.4% (95%CI 5.8 to 9.3%) at 1st grade (6-7 years), 10% (95%CI 8 to 12.3%) at 3rd grade (8-9 years) and 11.1% (95%CI 8.8 to 13.8%) at 5th grade (10-11 years) (Chyi 2008).

Low quality evidence from one study (sample size 17565) showed that among children born at 32-33 weeks of gestation the prevalence of those who were enrolled on a special education programme (ECLS-K data) was 8% (95%CI 4.7 to 12.7%) at kindergarten stage (3 years age), 11.9% (95%CI 7.7 to 17.3%) at 1st grade (6-7 years), 14.4% (95%CI 9.2 to 21%) at 3rd grade (8-9 years) and 14.5% (95%CI 8.8 to 22%) at 5th grade (10-11 years) (Chyi 2008).

Achievement (FSP or KS1)

Moderate quality evidence from one study (sample size 8728) showed that among children born at 32-36 weeks of gestation the prevalence of those not with a good level of overall achievement (FSP, teacher reported) was 59% (95%CI 54.8 to 63.1%) at 5 years age (Quigley 2012). In the same study, the prevalence among children who did not achieve in all 3 scales of personal, social and emotional development (FSP, teacher reported) was 31.6%

(95%CI 27.8 to 35.6%). For those who did not achieve in all 4 scales of communication, language and literacy, the prevalence was 49.1% (95%CI 47.7 to 50.4%), and for mathematical development (not achieving in all 3 scales) the prevalence was 37.5% (95%CI 33.5 to 41.6%) at 5 years age (Quigley 2012).

Moderate quality evidence from one study (sample size 8728) showed that among children born at 32-33 weeks of gestation the prevalence of those not with a good level of overall achievement (FSP, teacher reported) was 60.9% (95%CI 50.1 to 70.9%) at 5 years age (Quigley 2012). In the same study, the prevalence among children who did not achieve in all 3 scales of personal, social and emotional development (FSP, teacher reported) was 32.6% (95%CI 23.2 to 43.2%). For those who did not achieve in all 4 scales of communication, language and literacy, the prevalence was 57.6% (95%CI 46.9 to 67.9%), and for mathematical development (not achieving in all 3 scales) the prevalence was 40.2% (95%CI 30.1 to 51%) at 5 years age (Quigley 2012).

Moderate quality evidence from one study (sample size 8728) showed that among children born at 34-36 weeks of gestation the prevalence of those not with a good level of overall achievement (FSP, teacher reported) was 58.6% (95%CI 54 to 63.1%) at 5 years age (Quigley 2012). In the same study, the prevalence among children who did not achieve in all 3 scales of personal, social and emotional development (FSP, teacher reported) was 31.4% (95%CI 27.3 to 35.8%). For those who did not achieve in all 4 scales of communication, language and literacy, the prevalence was 54.1% (95%CI 49.5 to 58.7%), and for mathematical development (not achieving in all 3 scales) the prevalence was 36.9% (95%CI 32.6 to 33.5%) at 5 years age (Quigley 2012).

Moderate quality evidence from one study (sample size 18818) showed that among children born at 32-33 weeks of gestation not achieving level 2 or more in reading, writing and mathematics (KS1, teacher reported), the prevalence was 26.9% (95%CI 16.8 to 39.1%) at 7 years age (Chan 2014). For those children not achieving level 2 or more in reading, writing, speaking/listening and science, the prevalence was 19.4% (95%CI 10.8 to 30.9%), 23.9% (14.3 to 35.9%), 16.4% (95%CI 8.5 to 27.5%) and 16.4% (95%CI 8.5 to 27.5%), respectively (Chan 2014).

Moderate quality evidence from one study (sample size 18818) showed that among children born at 34-36 weeks of gestation not achieving overall level 2 or more in reading, writing and mathematics was 23.3% (95%CI 19.1 to 28.1%) at 7 years age (Chan 2014). For those children not achieving level 2 in reading, writing, speaking and listening, mathematics, or science, the prevalence was 18.1% (95%CI 14.2 to 22.4%), 20.6% (95%CI 16.5 to 25.1%), 13.1% (95%CI 9.8 to 17%), 8.6% (95%CI 5.9 to 12%), and 11.7% (95%CI 8.5 to 15.4%), respectively (Chan 2014).

Very low quality evidence from one study (sample size 13978) showed that among children born at 32-36 weeks of gestation not achieving level 2 or more in reading, writing or maths (teacher reported KS1) the prevalence was 29% (95%CI 25.4 to 33%) at 5-7 years age (Peacock 2012). For those children not achieving level 2 in reading, writing and mathematics (individual items of KS1), the prevalence was 22.2% (95%CI 19 to 25.7%), 22.7% (95%CI 19.4 to 26.2%), and 18.1% (95%CI 15.1 to 21.5%) respectively (Peacock 2012).

Special education needs at <37 weeks of gestation

Overall special education needs

Low quality evidence from one study (sample size 722) showed that among children born at <37 weeks of gestation the prevalence of SEN was 35.5% (95%CI 32 to 39.1%) at 8 years age (Odd 2013).

Moderate quality evidence from one study (sample size 775) showed that among children born at <37 weeks of gestation the prevalence of SEN was 24.3% (95%CI 21.1 to 27.7%) at 14 to 16 years age (Odd 2016).

Low quality evidence from one study (sample size 722) showed that among children born at <37 weeks of gestation the prevalence of low achievement (KS1) was 31.4% (95%CI 28.1 to 35%) at 8 years age (Odd 2013).

4.1.4 Prevalence of developmental disorders

Review question:

What is the prevalence of developmental disorders in babies, children and young people born preterm?

4.1.4.1 Description of clinical evidence

The aim of this review is to establish the prevalence and incidence of different developmental disorders in relation to the different gestational ages in babies, children and young people born preterm. The developmental disorders considered as outcomes included cerebral palsy, intellectual disability, learning impairment, speech and language impairment, attention deficit hyperactivity disorder, autism spectrum disorder, DCD, mental and behavioural disorders, developmental co-ordination disorder and hearing and visual impairments.

Fifty-seven studies were included in the review (Agerholm 2011; Ancel 2006; Anderson 2003; Andersen 2011; Anderson 2011; Andrews 2008; Anonymous 1997; Beaino 2011; Bodeau-Livinec 2007; Burguet 1999; Burnett 2014; Charkaluk 2010; De Groote 2007; de Kleine 2003; Doyle 2011; Drummond 2002; Farooqi 2011; Foix-Helias 2008; Foulder-Hughes 2003; Glinianaia 2011; Guellec 2011; Hellgren 20116; Himmelmann 2014; Hirvonen 2014; Holmstrom 2014; Hreinsdottir 2013; Hutchinson 2013; Johnson 2009; Johnson 2010; Johnson 2011; Joseph 2016a; Joseph 2016b; Kiechl-Kohlendorfer 2013; Larroque 2008; Leversen 2010; Leversen 2011; Leversen 2012; Marlow 2005; Marret 2007; Mikkola 2005; Moore 2012; Nordmark 2001; Odd 2013; Rieger-Fackeldey 2010; Roberts 2010; Roberts 2011; Robertson 2007; Salakorpi 2001; Serenius 2013; Stahlmann 2009; Sutton 1999; Tommiska 2003; Toome 2012; Vincer 2014; Vohr 2005; Wolke 2008; Wood 2000).

The sample size ranged from 89 (Farooqi 2011) to 331,154 (Glinianaia 2011).

Twelve studies were from the UK or UK and Ireland (Bodeau-Livinec 2007; Drummond 2002; Foulder-Hughes 2003; Glinianaia 2011; Johnson 2009; Johnson 2010; Johnson 2011; Marlow 2005; Moore 2012; Odd 2013; Wolke 2008; Wood 2000). Six of the studies were part of the EPIcure study (Johnson 2009; Johnson 2010; Johnson 2011; Marlow 2005; Moore 2012; Wolke 2008; Wood 2000), one publication was from the ALSPAC study (Odd 2013), and another publication was from NECCPS study (Glinianaia 2011).

Nine studies were from Australia (Anderson 2003; Anderson 2011; Anonymous 1997; Burnett 2014; Doyle 2011; Hutchinson 2013; Roberts 2010; Roberts 2011; Sutton 1999). Four of the publications were from the Victorian Collaborative Study Group (Anderson 2003; Anonymous 1997; Burnett 2014; Roberts 2011).

Nine studies were from France (Ancel 2006; Andersen 2011; Beaino 2011; Burguet 1999; Charkaluk 2010; Foix-Helias 2008; Guellec 2011; Larroque 2008; Marret 2007). Seven of the publications were from the EPIPGAGE study (Ancel 2006; Beaino 2011; Charkaluk 2010; Foix-Helias 2008; Guellec 2011; Larroque 2008; Marret 2007).

Seven studies were from Sweden (Farooqi 2011; Hellgren 2016; Himmelman 2014; Holmstrom 2014; Hreinsdottir 2013; Nordmark 2001; Serenius 2013). Three of the publications were from the EXPRESS study (Hellgren 2016; Holmstrom 2014; Serenius 2013) and one publication was from the LOVIS study (Serenius 2013).

Four studies were from Finland (Hirvonen 2014; Mikkola 2005; Salakorpi 2001; Tommiska 2003), four publications were from USA (Andrews 2008; Joseph 2016a; Joseph 2016b; Vohr 2005). Two of the publications were from the ELGAN study (Jospeh 2016a; Joseph 2016b). Three studies were from Norway from the same author (Leversen 2010; Leversen 2011; Leversen 2012). Two studies were from Germany (Rieger-Fakeldey 2010; Stahlmann 2009), and another two publications were from Canada (Robertson 2007; Vincer 2014). There was one study each from Austria (Kiechl-Kohlendorfer 2013), Denmark (Agerholm 2011), Netherlands (de Kleine 2003), Belgium (de Groote 2007, EPIBEL study), and Estonia (Toome 2012).

Forty-five publications used data from population- based (national, geographical or regional prospective cohort studies (Agerholm 2011; Ancel 2006; Anderson 2011; anonymous 1997; Beaino 2011; Burguet 1999; Burnett 2014; Charkaluk 2010; de Groote 2007; de Kleine 2003; Doyle 2011; Farooqi 2011; Foix-Helias 2008; Foulder-Hughes 2003; Guellec 2011; Hellgren 2016; Hreinsdottir 2013; Hutchinson 2013; Johnson 2009; Johnson 2010; Johnson 2011; Joseph 2016a; Joseph 2016b; Kiechl-Kohlendorfer 2013; Larroque 2008; Leversen 2010; Leversen 2011; Leversen 2012; Marlow 2005; Mikkola 2005; Moore 2012; Nordmark 2001; Odd 2013; Rieger-Fackeldey 2010; Roberts 2011; Roberts 2010; Robertson 2007; Salakorpi 2001; Serenius 2013; Sutton 1999; Tommiska 2003; Toome 2012; Vincer 2014, Wolke 2008; Wood 2000).

Five publications used registry data (Anderson 2011; Bodeau-Livinec 2007; Drummond 2002; Himmelmann 2014; Hirvonen 2014)

One publication used data from a population based survey (Glinianaia 2011). One publication used data from a multicentre study (Vohr 2005).

Thirty- seven publications reported on CP (Ancel 2006; Andersen 2011; Anderson 2011; Andrews 2008; Anonymous (Victorian collaboration study) 1997; Burguet 1999; De Groote 2007; Doyle 2011; Drummond 2002; Farooqi 2011; Foix-Helias 2008; Glinianaia 2011; Guellec 2011; Himmelmann 2014; Hirvonen 2014; Hutchinson 2013; Larroque 2008; Leversen 2011; Marlow 2005; Marret 2007; Mikkola 2005; Moore 2012; Nordmark 2001; Odd 2013; Rieger-Fackeldey 2010; Roberts 2010; Robertson 2007; Salakorpi 2001; Serenius 2013; Stahlmann 2009; Sutton 1999; Tommiska 2003; Toome 2012; Vincer 2014; Vohr 2005; Wood 2000). Majority of studies reported assessment of CP by physical or neurological exam by trained physicians and paediatricians or psychologists (Ancel 2006; Anderson 2011; Anderson 2011; Andrews 2008; Burguet 1999; De Groote 2007; Farooqi 2011; Foix-Helias 2008; Glinianaia 2011; Guellec 2011; Himmelmann 2014; Larroque 2008; Marlow 2005; Marret 2007; Nordmark 2001; Robertson 2007; Salakorpi 2001; Sutton 1999; Vincer 2005; Wood 2000). Some of the studies used the European CP network for classification (Ancel 2006; Foix-Helias 2008; Larroque 2008; Marret 2007) or the Surveillance of CP in Europe classification (Anderson 2011; Glinianiaia 2011). Seven studies assessed CP using the Gross Motor Function Classification System (GMFCS) (Doyle 2011; Joseph 2016b; Leversen 2011; Moore 2012; Rieger-Fackeldey 2010; Stahlmann 2009; Toome 2012). One study used the Little Club definition for CP (Drummond 2002). One study used ICD-9 and ICD-10 codes for classification of CP (Hirvonen 2014), and one study used the Standard Recording of Central Motor Deficit for classification of CP (Odd 2013). Five studies reported results as total number of livebirths (Andersen 2011; Drummond 2002; Himmelmann 2014; Nordmark 2001; Robertson 2007).

Twenty-five publications reported intellectual disability (Anderson 2003; Andrews 2008; Anonymous (Victorian collaboration study) 1997; Beaino 2011; Charkaluk 2010; Doyle 2010; De Groote 2007; Foix-Helias 2008; de Kleine 2003; Joseph 2016b; Larroque 2008; Leversen 2011; Leversen 2012; Marlow 2005; Marret 2007; Mikkola 2005; Moore 2012; Rieger-Fackeldy 2010; Roberts 2010; Salakorpi 2001; Serenius 2013; Stahlmann 2009; Sutton 1999; Toome 2012; Vohr 2005). Three studies used the Wechsler Intelligence Scale for Children (WISC) version III (Anderson 2003), version IV (Roberts 2010) and version IV with Differential Ability Scale (DAS) (Andrews 2008; Joseph 2016b). Six studies used the Bayley Scale of Infant Development (BSID) version II or III (Anon (Victorian collaborative study) 2007; Doyle 2011; De Groote 2007; Moore 2012; Toome 2012; Vohr 2005). Seven studies used the Kaufmann Assessment Battery for Children (K-ABC)/Mental Processing Composite (MPC) score (Beaino 2011; Foix-Helias 2008; Larroque 2008; Marret 2007; Rieger-Fackeldey 2010; Serenius 2009; Stahlmann 2009)). One study used the K-ABC, NEPSY, and Griffiths Developmental Assessment (Marlow 2005). One study assessed major developmental delay using the Griffiths Developmental Assessment (Sutton 1999). Four studies used the Wechsler Preschool and Primary Scale of Intelligence revised (WPPSI-R) (Leversen 2011; Leversen 2012; Mikkola 2003; Salakorpi 2001). One study used the Brunte-Lezine scale (Charkaluk 2010) and another study used the revised Amsterdam Child Intelligence Test (de Kleine 2003).

Five publications reported on speech and/or language disorder (Moore 2012; Serenius 2013; Toome 2012; Wolke 2008; Wood 2000). One study assessed communication disability using the third edition of the Bayley Scales of Infant Development (BSID-III) (Moore 2012) and another study used BSID-II (Wood 2000). Two studies assessed language impairment by the BSID-III scale (Serenius 2013; Toome 2012). One study used the Pre-School Language Scale-3 (PLS-3) to assess language impairment (Wolke 2008).

Two publications reported on attention deficit hyperactivity disorder (Burnett 2014; Johnson 2010). One of the studies used the ADHD module of the Children's Interview for Psychiatric Syndromes (ChiPS) (Burnett 2014) whereas the other study used the Developmental and Well Being Assessment (DAWBA) to assess ADHD types (Johnson 2010).

Two publications reported on autism spectrum disorder (Johnson 2010; Joseph 2016a). One study assessed ASD by using the Developmental and Well Being Assessment (DAWBA) and the other study assessed ASD using the Autism Diagnostic Interview-Revised (ADI-R).

Four publications reported on specific learning difficulties (Anderson 2003; Johnson 2011; Joseph 2016b; Kiechl-Kohlendorfer 2013). One study assessed educational progress using the Wide Range Achievement Test (WRAT-3) and also the Comprehensive Scales of Student Abilities (CSSA) (Anderson 2003). One study assessed learning impairment using the Wechsler Individual Achievement Test-II (WIAT-II) (Johnson 2011). One study assessed academic achievement using the Wechsler Individual Achievement Test-III (WIAT-III) (Joseph 2016b). One study used TEDI-MAHT to assess delay in numerical skills (Kiechl-Kohlendorfer 2013).

Four publications reported on developmental coordination disorder (Agerholm 2011; de Kleine 2003; Foulder-Hughes 2003; Roberts 2011). All four studies assessed DCD or motor deficit with the Movement Assessment Battery for Children (M-ABC) tool.

Two publications reported on mental and behavioural disorders (Burnett 2014; Johnson 2010). One study used the Development and Well Being Assessment (DAWBA) tool to assess mental and behavioural disorders (Johnson 2010), whereas another study assessed anxiety, mood, and depressive or psychotic disorders using the DSM-IV Axis I disorders tool (Burnett 2014).

Twenty -four publications reported on vision impairment (Anderson 2003; Anderson 2011; Anonymous (Victorian collaborative study) 1997; Bodeau-Livinec 2007; De Groote 2007; Farooqi 2011; Hellgren 2016; Holmstrom 2014; Hreinsdottir 2013; Hutchinson 2013; Joseph 2016b; Larroque 2008; Leversen 2010; Leversen 2011; Marlow 2005; Marret 2007; Moore 2012; Rieger-Fackeldey 2010; Roberts 2010; Serenius 2013; Toome 2013; Tommiska 2003; Vohr 2005; Wood 2000). Severe vision impairment assessment was varied among studies. Three studies reported on vision impairment visual acuity in both eyes was assessed as worse than 6/60 (Anonymous (Victorian Collaborative Study) 1997) or visual acuity in the in the better eye of <6/60 (Bodeau-Livinec 2007; Roberts 2010). One study reported visual impairment as unilateral or bilateral blindness or visual acuity of <20/200 without glasses in at least one eye (Farooqi 2011; Rieger-Fackeldey 2010). One study assessed visual impairment with the Rossano test 12 and visual deficiency of <3/10 for both eyes (Larroque 2008; Marret 2007). Impaired vision was also defined as blindness in children who were not able to fixate and follow a light (Holsmstrom 2014; Hreinsodottir 2013) whereas other studies defined visual impairment as 'no useful vision' (De Groote 2007; Vohr 2005), 'legally blind' (Leversen 2010; Leversen 2011; Tommiska 2003), or 'blindness' (Marlow 2005; Moore 2012). One study reported results as total number of livebirths (Bodeau-Livinec 2007).

Nineteen publications reported on hearing impairment (Anderson 2003; Anderson 2011; Anonymous (Victorian collaborative study) 1997; De Groote 2007; Doyle 2011; Farooqi 2011; Hutchinson 2013; Larroque 2008; Leversen 2010; Marlow 2005; Marret 2007; Moore 2012; Rieger-Fackeldey 2010; Roberts 2010; Serenius 2013; Tommiska 2003; Toome 2012; Vohr 2005; Wood 2000). Hearing impairment assessment was varied among the studies. Two studies defined hearing impairment as hearing loss of more than 70 decibel (dB) for one or both ears (Larroque 2008; Marret 2007). Other studies defined hearing impairment as complete deafness (Leversen 2010), deafness or hearing loss (as a need of hearing aids or worse) (Anderson 2011; Doyle 2011; Farooqi 2011; Marlow 2005; Rieger-Fackeldey 2010; Roberts 2010; Tommiska 2003; Vohr 2005), 'no useful hearing or requiring hearing aids' (De Groote 2007), or profound sensorineural hearing loss not improved by aids (Moore 2012).

The feasibility of combining study data using meta-analysis was assessed. Due to the following differences between studies, it was not considered appropriate to pool the results:

- the inclusion/exclusion criteria for participants
- ages of participants at the time of assessment
- outcome definitions and measurement tools
- consistency of results.

.

4.1.4.2 Summary of included studies

Table 20: Summary of included studies

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Evidence on CP						
Ancel 2006 France	Prospective population-based cohort study (EPIPAGE).	n=1954 (83% of the eligible ones for the follow-up)	Each child was subjected to a detailed physical and neurologic examination assessing tone, reflexes, posture, and movements. A pre-coded standardised questionnaire, completed by each treating physician was designed to minimise the risk of ambiguous answers and trained paediatricians reviewed questionnaires for infants with abnormal neurologic examination results. The definition of CP proposed by the European Cerebral Palsy Network was used.	At 2 years (not reported if corrected or not) CP 24-25 weeks GA: 12/64, 19.4% (10.4-31.4%) 26 weeks GA: 18/82, 22.0% (13.6-32.5%) 27 weeks GA: 18/146, 12.3% (7.5-18.8%) 28 weeks GA: 21/191, 11.0% (6.9-16.3%) 29 weeks GA: 16/196, 8.2% (4.7-12.9%) 30 weeks GA: 26/315, 8.3% (5.5-11.9%) 31 weeks GA: 29/424, 6.8% (4.6-9.7%) 32 weeks GA: 24/538, 4.4% (2.9-6.6%) The following GA groups were calculated by the NGA technical team using the above data: <28 weeks GA: 48/290, 16.6% (12.5-21.3%) 28-31 weeks GA: 92/1126, 8.2% (6.6-9.9%)	Low	Children born 1997, assessed at 2 years.
Andersen 2011 France	Register-based study	n=903 children with CP born moderately preterm	Children with CP were identified and classified according to the definition and classification tree of the Surveillance of Cerebral Palsy in Europe (SCPE) database.	Age at assessment not reported but children were included in the register earliest at 4 years of age CP 1990-94 Grenoble, France 32-36 weeks GA: 8.2/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Cork, Ireland	Low	1980-1998 (but for this review only data between 1990-1998 is used).

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				32-36 weeks GA: 7.2/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Göteborg, Sweden 32-36 weeks GA: 6.1/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Copenhagen, Denmark 32-36 weeks GA: 7.2/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Rome, Italy 32-36 weeks GA: 13.0/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) 1995-1998 Grenoble, France 32-36 weeks GA: 5.6/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Cork, Ireland 32-36 weeks GA: 7.2/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Göteborg, Sweden 32-36 weeks GA: 6.6/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Copenhagen, Denmark 32-36 weeks GA: 6.1/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) Rome, Italy		

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				32-36 weeks GA: 8.6/1000 live births (number of cases and the number of live births not reported, not possible to calculate confidence intervals) 1991-1996 Tonsberg, Norway 32-36 weeks GA: 13.8/1000 live births (95% CI 7-25/1000 live births) (number of cases 10, the number of live births calculated to be 725) 1991-1998 Galway, Ireland 32-36 weeks GA: 4.0/1000 live births (95% CI 2-7/1000 live births) (number of cases 11, the number of live births calculated to be 2750) Madrid, Spain 32-36 weeks GA: 4.0/1000 live births (95% CI 2-7/1000 live births) (number of cases 14, the number of live births calculated to be 3500) 1992-1998 Bologna, Italy 32-36 weeks GA: 8.8/1000 live births (95% CI 5-15/1000 live births) (number of cases 15, the number of live births calculated to be 1705)		
Anderson 2011 Australia	Population-based cohort study	n=201 children survived to 8 years n=189 assessed at	8 years (corrected) by psychologists blind to perinatal details, predominantly in specialised follow-up clinics, although a few were tested at school or	At 8 years (corrected) CP 22-27 weeks GA/BW 1000g: 22/189, 11.6% (7.4-17.1%)	Low	Children born 1997, follow-up at 8 years of corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		8 years (94%)	home if they could not attend the clinics. CP, deafness and blindness were diagnosed by trained paediatricians who were blind to group membership (the study included a term-born control group).			
Andrews 2008 USA	Prospective cohort study	n=259 (around 70% of the 375 eligible and alive for the follow-up) with data on IQ n=257 with data on CP	CP was assessed with a complete physical and neurological examination including assessment of gross and fine motor function performed by certified nurse practitioner under the supervision of a developmental paediatrician. CP was defined as abnormal muscle tone in at least 1 extremity and abnormal control of movement and posture.	At 6 years CP 23-32 weeks GA: 11/257, 4.3% (2.2-7.5%)	Low	1996-1999
Anonymous (Victorian study) 1997 Australia	A geographically determined cohort study (Victoria, Australia)	n=401 live born children born at 23-27 weeks n=225 children survived to 2	A developmental paediatrician and a psychologist assessed the children at 2 years of age. They were blinded to the knowledge of prematurity. The paediatric assessment	At 2 years CP 23-27 weeks GA: 24/219, 11.0% (7.2-15.9%)	Low	Children born 1991-1992, follow-up at 2 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		years of age (56.1%) n=219 were assessed at 2 years (97.3% of the survivors)	included a neurological examination to determine outcomes such as cerebral palsy, and visual acuity. The criteria for cerebral palsy was not reported in this publication but in another publication: "Cerebral palsy was diagnosed in children with increased active tone, increased deep tendon reflexes, and, if affecting both lower limbs, positive Babinski reflexes." (Kitchen 1991 Changing two-year outcome of infants weighing 500 to 999 grams at birth: a hospital study. J Pediatr 118(6):938-43.)			
Burguet 1999 France	Prospective regional cohort study	Total number of live births in region=14,350 n=203 premature neonates were enrolled to the study	A physician examined the child at 2 years age, completed a questionnaire that was mailed to the inquirers. Abnormal infants were considered to have CP or sensorineural impairment when one or more of the following signs were observed: hemiplegia, diplegia, tetraplegia, dystonia, athetosis,	At 2 years age (corrected) CP 25-32 weeks GA: 22/167, 13.2% (8.4-19.3%) CP severe spastic tetraplegia with mental retardation 25-32 weeks GA: 8/167, 4.8% (2.1-9.2%) CP isolated spastic tetraplegia 25-32 weeks GA: 2/167, 1.2% (0.2-4.3%) CP spastic diplegia 25-32 weeks GA: 10/167, 6.0% (2.9-10.7%) CP hemiplegia 25-32 weeks GA: 2/167, 1.2% (0.2-4.3%)	Very low	Infants born from 1990 to 1992, assessed at 2 years age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=171 survived to 2 years age. n=167 surviving infants were evaluated at 2 years age	blindness, or neurosensory deafness			
De Groote 2007 Belgium	Population-based geographically defined cohort study (EPIBEL)	n=95 children that survived to discharge from NICU n=77 children assessed at 3 years (n=3 died before follow-up, n=12 parents did not give consent, n=3 could not be reached), 81% follow-up rate (84% of the ones who were alive at follow-up).	The assessment at 3 years comprised of a detailed clinical examination and full developmental evaluation. The clinical evaluation included collecting the recent medical history and a global health and anthropometric assessment as well as standardised neurologic and sensory examination. The classification of type and location of cerebral palsy was based on describing function, tone and reflexes in each limb. In addition, it comprised the results of the neurologic examination.	At 3 years CP total <27 weeks GA: 19/77, 24.7% (15.6-35.8%)* By type of CP: Spastic CP <27 weeks GA: 14/77, 18.2% (10.3-28.6%) Extrapyramidal dystonia CP <27 weeks GA: 3/77, 3.9% (0.8-11.0%) Hypotonic CP <27 weeks GA: 1/77, 1.3% (0.03-7.0%) Ataxia CP <27 weeks GA: 1/77, 1.3% (0.03-7.0%) By location of CP: CP hemiparesis <27 weeks GA: 3/77, 3.9% (0.8-11.0%) CP diparesis <27 weeks GA: 9/77, 11.7% (5.5-21.0%)	Low	Children born in 1999-2000, follow-up at 3 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				CP triparesis <27 weeks GA: 2/77, 2.6% (0.3-9.1%) CP quadriparesis <27 weeks GA: 4/77, 5.2% (1.4-12.8%) By severity of CP: Severe CP (regardless of type or location) <27 weeks GA: 1/77, 1.3% (0.03-7.0%) Moderate CP (regardless of type or location) <27 weeks GA: 10/77, 13.0% (6.4-22.6%) Mild CP (regardless of type or location) <27 weeks GA: 8/77, 10.4% (4.6-19.5%)		
Doyle 2011 Australia	A regional population-based cohort of extremely low birth weight infants in the state of Victoria, Australia	n=257 live births with bw 500-999 g (excl. cases with lethal anomalies) n=172 survived to 2 years n=165 assessed at 2 years (96%)	Survivors were assessed at 2 years by paediatricians and psychologists blinded to perinatal details. Criteria for diagnosis of CP included abnormal tone and loss of motor function, and its severity was assessed by the Gross Motor Function Classification System (GMFCS)	At 2 years (corrected age) CP BW 500-999 g (mean GA 25.7 [SD 2.3]): 12/165, 7.3% (3.8-12.4%)	Moderate	Children born 2005, follow-up at 2 years (corrected age).
Drummond 2002 UK	Epidemiological register data study	n=2858 singleton neonatal survivors in	The North of England Collaborative CP survey records all infants with CP born to mothers	Age at assessment not reported. Time period 1990-94 CP	Low	1970-1994 (only time period 1990-

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		1990-94 with <37 weeks of GA at birth	resident in Newcastle, North Tyneside and Northumberland at birth. The Little Club definition of CP is used (Mac Keith RC., MacKenzie ICK., Polani PE. (1959) The Little Club. Memorandum on terminology and classification of 'cerebral palsy'. Cereb Palsy Bull 1: 27–35.), updated by Bax (Bax MC. (1964) Terminology and classification of Cerebral Palsy. Dev Med Child Neurol 6: 295-7.). Spastic CP is classified as unilateral (hemiplegia and monoplegia) or bilateral (diplegia, quadriplegia and any other combination of bilateral spastic involvement) in line with the agreement of the European Collaboration	<37 weeks: 16.8/1000 neonatal survivors (95% CI 12-22) (number of cases 48, number for neonatal survivors 2858) <36 weeks: 24.5/1000 neonatal survivors (95% CI 18-33) (number of cases 42, number for neonatal survivors 1713) <35 weeks: 33.9/1000 neonatal survivors (95% CI 24-46) (number of cases 37, number for neonatal survivors 1093) <34 weeks: 50.5/1000 neonatal survivors (95% CI 36-69) (number of cases 37, number for neonatal survivors 732) <33 weeks: 61.8/1000 neonatal survivors (95% CI 42-87) (number of cases 31, number for neonatal survivors 502) <32 weeks: 67/1000 neonatal survivors (95% CI 44-99) (number of cases 24, number for neonatal survivors 355) 32-36 weeks: 9.6/1000 neonatal survivors (95% CI 6-14) (number of cases 24, number for neonatal survivors 2503) 28-31 weeks: 56.3/1000 neonatal survivors (95% CI 33-90) (number of cases 16, number for neonatal survivors 284) <28 weeks: 112.7/1000 neonatal survivors (95% CI 50-210) (number of cases 8, number for neonatal survivors 71)		94 used for the review).
Farooqi 2011 Sweden	Prospective national cohort study	n=89 children born at <26 weeks gestation and survived to follow-up (36% of all	Cerebral palsy (CP), classified as hemiplegia, diplegia, or quadriplegia. CP was categorized functionally as as mild (no evidence of clinically important functional difficulty related to gait or	At 11 years Moderate or disabling CP <26 weeks GA: 6/88, 6.8% (2.5-14.3%)	Low	Children born 1990-1992, follow-up at 11 years

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		247 children born at <26 weeks in Sweden of which the rest died) n=88 children with data (1 child was lost to follow-up, was followed-up but did not participate)	use of hands), moderate (independent walking but with an abnormal gait); or disabling (not walking, severe motor diability).			
Foix-Helias 2008 France	Prospective population based cohort study (EPIPAGE).	n=1781 children with data on CP (77% of n=2300 survivors up to follow-up) n=1508 children with data on cognition (66% of the n=2300 survivors up to follow-up)	Follow-up was at 5 years of age, and involved a medical and neuropsychological assessment. The assessment included a thorough physical examination and neurological assessment (tone, reflexes, posture and movements). Physicians recorded their findings on a standardized form. The definition of cerebral palsy was that established by the European Cerebral Palsy Network, which requires at least 2 of the following: abnormal posture or	At 5 years CP 24-32 weeks GA: 158/1781, 8.9% (7.6-10.3%) 24-27 weeks GA: 39/266, 14.7% (10.6-19.5%) 28-32 weeks GA: 119/1515, 7.9% (6.6-9.3%) Severe CP 24-32 weeks GA: 50/1781, 2.8% (2.1-3.7%) 24-27 weeks GA: 13/266, 4.9% (2.6-8.2%) 28-32 weeks GA: 37/1515, 2.4% (1.7-3.4%)	Moderate	Recruitment took place in 1997. Follow-up was at 5 years.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			movement, increased tone and hyperreflexia. Cerebral palsy was considered to be severe if infants were unable to walk, or only able to walk with assistance.			
Glinianaia 2011 UK	Prospective population-based survey (NECCPS)	n=331154 total study population (all live born neonatal survivors) n=18797 live born neonatal survivors born at <37 weeks of gestation (n=846 live born neonatal survivors born at <28 weeks of gestation n=2070 live born neonatal survivors born at 28-31 weeks of gestation n=15881 live born	CP is classified according to the agreement of the Surveillance of Cerebral Palsy in Europe: spastic CP (unilateral or bilateral), dyskinetic and ataxic. Data on CP was obtained from the North of England Collaborative Cerebral Palsy Survey (NECCPS) that prospectively records all infants with CP born to mothers resident in the region from 1991. Cases are notified to the survey by the District Convenors who are consultant community paediatricians. They coordinate services for such children and receive information about children needing services from other paediatricians, paediatric neurologists, physiotherapists, speech therapists, and the regional child	At age up to 8 years CP 1991-1995 <28 weeks GA: 28/463, 6.1% (4.1-8.6%) 28-31 weeks GA: 58/1111, 5.2% (4.0-6.7%) 32-36 weeks GS: 81/8276, 1.0% (0.8-1.2%) 1996-2000 <28 weeks GA: 29/383, 7.6% (5.1-10.7%) 28-31 weeks GA: 64/959, 6.7% (5.2-8.4%) 32-36 weeks GS: 70/7605, 0.9% (0.7-1.2%) 1991-2000 <28 weeks GA: 57/846, 6.7% (5.1-8.6%) 28-31 weeks GA: 122/2070, 5.9% (4.9-7.0%) 32-36 weeks GS: 151/15881, 1.0% (0.8-1.1%) CP non-spastic 1991-2000 <37 weeks GA: 13/18797, 0.07% (0.04-0.12%) CP spastic bilateral <37 weeks GA: 240/18797, 1.3% (1.1-1.5%) CP spastic unilateral	Low	Children born 1991-2000, follow-up up to 8 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		neonatal survivors born at 32-36 weeks of gestation)	development centre. The convenor completes the notification form. Further details are forwarded to the survey when the child reached 5 years of age to confirm the diagnosis and provide details of associated impairments. it is very unusual to for a case of CP to be diagnosed after age 6 years, however, the process of ascertainment by the convenor and the requirement to obtain parent consent means that sometimes children are added to the register up to age 8 years even though diagnosed a year or two earlier. Cases are notified from multiple sources, there is a regional network of interested clinicians and close links with the long standing prospective Perinatal Mortality Survey and Northern Congenital Abnormality Survey housed on the same premises. Every case of CP mentioned on a child death certificate and every case mentioned as a co-morbidity on a late	<37 weeks GA: 77/18797, 0.4% (0.3-0.5%)		

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			notification of a congenital abnormality is ascertained by the survey.			
Guellec 2011 France	Population based prospective cohort study (EPIPGAGE study)	n=2855 live births at 24-32 weeks GA. n=2357 infants eligible for follow-up.	Cerebral palsy (CP), defined according to the European CP Network definition, children were classified as having CP if they had abnormal posture or movement, increased tone or hyperreflexia (spastic CP), involuntary movements (dyskinetic CP), or loss of coordination (ataxic CP). Detaimedical and neurologic examination in which tone, reflexes, postures and movements were assessed. Trained paediatricians reviewed data for children with abnormal results on neurologic examination to validate the diagnosis of CP and assess the severity.	At 5 years age CP 24-28 weeks GA: CP: 22/542, 4.1% (2.6-6.1%) 24-28 weeks GA: SGA: 4/22, 18.1% (5.2-40.3%) 29-32 weeks GA: 125/1815, 6.9% (5.8-8.2%) 29-32 weeks GA: SGA: 4/125, 3.2% (0.9-8.0%)	Low	Children born 1997, assessed at 5 years.
Himmelmann 2014 Sweden	A population-based epidemiological study (using register data).	n=94466 live births in the region in 2003-2006, of which n=238 children born	CP was verified at 4 to 8 years of age by the local neuro-paediatrician. In doubtful cases, a second diagnostic assessment was performed by the	CP verified at 4 to 8 years of age CP <28 weeks GA: 71.4/1000 live births (95% CI 42-112/1000 live births) (number of cases 17, number of live births 238)	Moderate	Children born 2003-2006.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		at <28 weeks of gestation, n=581 children born at 28-31 weeks of gestation, n=4544 children born at 32-26 weeks of gestation	first author of the publication. The definition of CP was agreed at an international consensus meeting in Bethesda. The Swedish and internationally accepted classification of CP syndromes was applied, in parallel with the classification suggested by the Surveillance of Cerebral Palsy in Europe (SCPE) where hemiplegia corresponds to unilateral spastic CP and diplegia and tetraplegia are combined to create bilateral spastic CP.	28-31 weeks GA: 39.6/1000 live births (95% CI 25-59/1000 live births) (number of cases 23, number of live births 581) 32-36 weeks GA: 6.4/1000 live births (95% CI 4-9/1000 live births) (number of cases 29, number of live births 4544) <37 weeks GA: 13/1000 live births (95% CI 10-16/1000 live births) (number of cases 69, number of live births 5363) Bilateral spastic CP (diplegia and tetraplegia) <37 weeks GA: 7.5/1000 live births (95% CI 5-10/1000 live births) (number of cases 40, number of live births 5363)		
Hirvonen 2014 Finland	National register study	n=6347 children born at <32 weeks n=6799 children born at 32-33 weeks n=39932 children born at 34-36 weeks	A case with CP was recorded if the individual was detected in the Hospital Discharge Register (HDR) and/or in the Reimbursement Register of the Social Insurance Institution with ICD-10 codes G80 to G83 in 1996 to 2008 and ICD-9 codes 342 to 344 in 1991 to 1995. Subtypes of CP were defined by topographic involvement (hemiplegia, diplegia, guadriplegia and	Up to 7 years of age (Study period 1991-2008) CP (total) <32 weeks GA: 550/6347, 8.7% (8.0-9.4%) 32-33 weeks GA: 160/6799, 2.4% (2.0-2.7%) 34-36 weeks GA: 225/39932, 0.56% (0.49-0.64%) 32-36 weeks GA: 385/46731, 0.8% (0.7-0.9%) CP hemiplegia <32 weeks GA: 80/6347, 1.3% (1.0-1.6%) 32-33 weeks GA: 37/6799, 0.5% (0.4-0.8%) 34-36 weeks GA: 57/39932, 0.14% (0.11-0.19%) 32-36 weeks GA: 94/46731, 0.2% (0.16-0.25%)) CP diplegia <32 weeks GA: 213/6347, 3.4% (2.9-3.8%)	Moderate	Children born 1991-2008, followed up to 7 years or up to year 2009

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			other types) and sought from registers with corresponding ICD codes. All inpatient or outpatient visits due to a CP diagnosis in public hospitals were registered to the HDR. The diagnosis of CP in Finland is based on medical history, ultrasound and MRI data, and multidisciplinary evaluations in the paediatric neurology units of 20 secondary-level central hospitals and 5 tertiary-level university hospitals.	32-33 weeks GA: 48/6799, 0.7% (0.5-0.9%) 34-36 weeks GA: 52/39932, 0.13% (0.10-0.17%) 32-36 weeks GA: 100/46731, 0.2% (0.17-0.26%) CP quadriplegia <32 weeks GA: 37/6347, 0.6% (0.4-0.8%) 32-33 weeks GA: 11/6799, 0.2% (0.1-0.3%) 34-36 weeks GA: 16/39932, 0.04% (0.02-0.06%) 32-36 weeks GA: 27/46731, 0.06% (0.04-0.08%) CP other types <32 weeks GA: 220/6347, 3.5% (3.0-4.0%) 32-33 weeks GA: 64/6799, 0.9% (0.7-1.2%) 34-36 weeks GA: 100/39932, 0.25% (0.20-0.30%) 32-36 weeks GA: 164/46731, 0.35% (0.3-0.4%)		
Hutchinson 2013 Australia	Prospective cohort study (Victorian Infant Collaborative Study Group)	n=189 preterm/low birth weight cohort (94% eligible for follow-up	Definitions of measurement of CP, blindness or deafness were not reported in the study	At 8 years age CP EP/ELBW (GA 26.5±2.0): 24/189, 12.7% (8.3-18.3%)	Very low	Children born in 1997, assessed at 8 years age
Joseph 2016b USA	Prospective cohort study (ELGAN)	n=873 preterm children at 10 years follow-up	Severe gross motor function was defined as level 5 (GMFCS, no self-mobility)	At 10 years Severe motor impairment 22-27 weeks GA: 17/873, 1.9% (95%CI 1.1-3.1%)	Low	Children born in 2002-2004
Larroque 2008 France	A longitudinal cohort study (EPIPAGE)	n=1817 children born at 22-32 weeks were followed at 5 years of age	Cerebral palsy (CP): The European Cerebral Palsy Network definition of cerebral palsy was used. At 5 years of age, children were invited for a	At 5 years CP <33 weeks GA: 159/1812, 8.8% (7.5-10.2%) 24-25 weeks GA: 11/60, 18.3% (9.5-30.4%) 26 weeks GA: 13/72, 18.1% (10.0-28.9%)	Moderate	1997, follow-up at 5 years of age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		(77% of the population that survived) n=1812 children born at 22-32 weeks with data on CP outcome	check-up with a physician. A medical questionnaire was completed by the physician after the clinical assessment, which included a standardised neurological examination, and a questionnaire (regarding child's health, family situation) was completed by the parents. Questionnaires for children with abnormal findings from neurological examination were checked by a group of paediatricians to validate the diagnosis.	27 weeks GA: 16/136, 11.8% (6.9-18.4%) 28 weeks GA: 24/178, 13.5% (8.8-19.4%) 29 weeks GA: 23/189, 12.2% (7.9-17.7%) 30 weeks GA: 18/288, 6.3% (3.8-9.7%) 31 weeks GA: 33/379, 8.7% (6.1-12.0%) 32 weeks GA: 21/510, 4.1% (2.6-6.2%) <28 weeks GA: 40/268, 14.9% (10.9-19.8%) 28-31 week GA: 98/1034, 9.5% (7.8-11.4%)		
Leversen 2010	Prospective observational nationally representative cohort study	n=373 children born 22-27 weeks GA or with birthweight 500-999 g who survived	Limited information provided. At 2 years a paediatrician completed forms developed for the study on somatic health and neurological status. They were not blinded. Children who missed the planned follow-up, data were collected in retrospect from the medical records if a routine follow-up had been performed within 1 year of planned evaluation, and from an	CP 22-27 weeks GA or bw 500-999 g: 26/373, 7.0% (4.6-10.1%)	Low	Children born in 1999-2000, follow-up at 2 years' corrected age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			additional structures telephone interview. No definition or classification of CP provided.			
Leversen 2011 Norway	Prospective observational national cohort study	n=306 children assessed at 5 years (n=638 children born, of which n=376 survived to discharge, of which 3 died and n=373 were followed-up at 2 years, of which 1 died and 1 child with Down's syndrome were excluded and 65 were lost to follow-up)	CP (total, and classes 1-5) was assessed with the Gross Motor Function Classification System for Cerebral Palsy, which is a 5-level classification. Class 1 means that the child is freely ambulatory; class 2 means that the child is unable to run or jump; class 3 means that the child depends on devices for walking; and classes 4 and 5 means that the child has non-ambulatory CP.	At 5 years CP any class 22-27 weeks GA or bw 500-999 g: 29/306, 9.5% (6.4-13.3%) CP class 4-5 22-27 weeks GA or bw 500-999 g: 10/306, 3.3% (1.6-5.9%) 23-25 weeks GA: 8/87, 9.2% (4.1-17.3%) 26-27 weeks GA: 2/152, 1.3% (0.2-4.7%) >27 weeks GA (bw <1000 g): 0/67, 0% (0-5.4%) CP class 2-3 22-27 weeks GA or bw 500-999 g: 9/306, 2.9% (1.4-5.5%) 23-25 weeks GA: 4/87, 4.6% (1.3-11.4%) 26-27 weeks GA: 3/152, 2.0% (0.4-5.7%) >27 weeks GA (bw <1000 g): 1/67, 1.5% (0.04-8.0%)	Moderate	Children born 1999 and 2000, follow-up at 5 years.
Marlow 2005 UK and Ireland	Population-based national cohort study (EPICure)	n=241 (82% of the eligible ones, n=293)	The classification of CP was made retrospectively, at the completion of the study, according to the description of function for	At 6 years CP, non- ambulatory <26 weeks GA: 15/241, 6.2% (3.5-10.1%) <=23 weeks GA: 1/24, 4.2% (0.1-21.1%) 24 weeks GA: 8/73, 11.0% (4.9-20.5%)	Moderate	Children born 1995, follow-up at 6 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			each limb, by two assessors. Severe CP was defined as non-ambulant CP; moderate CP was defined as ambulant CP.	25 weeks GA: 6/144, 4.2% (1.5-8.9%) CP with disability, ambulatory <26 weeks GA: 17/241, 7.1% (4.2-11.1%) <=23 weeks GA: 3/24, 12.5% (2.7-32.4%) 24 weeks GA: 6/73, 8.2% (3.1-17.0%) 25 weeks GA: 8/144, 5.6% (2.4-10.7%) CP, non- ambulatory or ambulatory (calculated by the NGA technical team) <26 weeks GA: 32/241, 13.3% (9.3-18.2%) <=23 weeks GA: 4/24, 16.7% (4.7-37.4%) 24 weeks GA: 14/73, 19.2% (10.9-30.1%) 25 weeks GA: 14/144, 9.7% (5.4-15.8%)		
Marret 2007 France	Population based prospective cohort (EPIPAGE).	n=1455	The definition used for CP was developed by the European Cerebral Palsy Network, which requires at least two of the following: abnormal posture or movement, increased tone, and hyperreflexia. Three categories of CP were distinguished: bilateral spastic CP, hemiplegia, and other. When the diagnosis of CP was in doubt, a panel of trained paediatricians met to discuss the case.	At 5 years of age CP (any type) 30 weeks GA: 18/288, 6.3% (3.8-9.7%) 31 weeks GA: 33/379, 8.7% (6.1-12.0%) 32 weeks GA: 21/509, 4.1% (2.6-6.2%) 33 weeks GA: 5/135, 3.7% (1.2-8.4%) 34 weeks GA: 1/140, 0.7% (0.2-3.9%) Bilateral spastic CP 30 weeks GA: 12/288, 4.2% (2.2-7.2%) 31 weeks GA: 26/379, 6.9% (4.5-9.9%) 32 weeks GA: 14/509, 2.8% (1.5-4.6%) 33 weeks GA: 2/135, 1.5% (0.2-5.3%) 34 weeks GA: 1/140, 0.7% (0.2-3.9%)	Low	1997-2002. Cohort established in 1997. Follow-up at 5 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				CP hemiplegia 30 weeks GA: 1/288, 0.4% (0.01-1.9%) 31 weeks GA: 3/379, 0.8% (0.2-2.3%) 32 weeks GA: 4/509, 0.8% (0.2-2.0%) 33 weeks GA: 1/135, 0.7% (0.02-4.1%) 34 weeks GA: 0/140, 0% (0-2.6%)		
Mikkola 2005 Finland	National population-based prospective cohort study	n=203 children with birth weight <1000 g (of n=206 children who survived up to follow-up) n=102 children with <27 weeks GA	Cerebral palsy (CP), defined as a non-progressive motor disorder with abnormal muscle tone, persistent or exaggerated primitive reflexes, or a positive Babinski sign associated with delayed motor development. Data on CP was collected from hospital records and child welfare clinics.	At 5 years CP Children born with birth weight <1000 g (mean GA 27.3 (SD 2.1): 28/203, 13.8% (9.4-19.3%) <27 weeks GA: 19/102, 18.6% (11.6-27.6%)	Low	1996-1997, follow-up at 5 years of age.
Moore 2012 UK	Prospective national cohort study (EPICure 2, this publication also used data from the original EPICure when comparing children born in 2006 to children born in 1995).	n=576 children born 22-26 weeks' gestation, assessed at follow-up (n=38 born at 22-23 weeks; n=98 born at 24 weeks; n=189 born at 25 weeks; n=251 born at 26 weeks)	Motor disability: Cerebral palsy was identified by neurological examination using the Palisano method (a standardized methods of identifying CP). The functional motor outcomes for children with CP using the 5 levels defined in the Gross Motor Function Classification System (GMFCS) from 1 for minimal impairment to 5 for severe impairment with dependence on	At 3 years (generally, some assessments delayed) Severe motor disability (CP level 3-5 in GMFCS) 22-26 weeks GA: 30/576, 5.2% (3.5-7.4%) 22-23 weeks GA: 4/38, 10.5% (2.9-24.8%) 24 weeks GA: 5/98, 5.1% (1.7-11.5%) 25 weeks GA: 10/189, 5.3% (2.6-9.5%) 26 weeks GA: 11/251, 4.4% (2.2-7.7%) Moderate motor disability (CP level 2 in GMFCS) 22-26 weeks GA: 15/576, 2.6% (1.5-4.3%) 22-23 weeks GA: 0/38, 0% (0-9.3%) 24 weeks GA: 4/98, 4.1% (1.1-10.1%) 25 weeks GA: 6/189, 3.2% (1.2-6.8%) 26 weeks GA: 5/251, 2.0% (0.7-4.6%)	Low	Children born in 2006 (this publication also compared the children born in 2006 to children born in 1995).

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			carers for most daily activities. Severe motor disability comprises of any non-ambulant CP (GMFCS levels 3-5). Moderate motor disability comprises of ambulant CP (GMFCS level 2).	Moderate to severe motor disability (CP level 2-5 in GMFCS) 22-26 weeks GA: 45/576, 7.8% (5.8-10.3%) 22-23 weeks GA: 4/38, 10.5% (2.9-24.8%) 24 weeks GA: 9/98, 9.2% (4.3-16.7%) 25 weeks GA: 16/189, 8.5% (4.9-13.4%) 26 weeks GA: 16/251, 6.4% (3.7-10.2%)		
Nordmark 2001 Sweden	Population based study	n=145 children with CP (born in Sweden, all gestational ages) n=46 preterm children with CP (<37 weeks of gestation)	Children with CP were identified through medical files and diagnostic records from all paediatric departments and habilitation centres in the area. The CP status of children were classified according to the internationally widely accepted Swedish classification system and definitions. The classification was done by an experienced neuropaediatrician in agreement with the child's local doctor.	At 4-7 years old CP <28 weeks GA: 72.3/1000 live births (95% CI 39.0-120.3/1000 live births) (13 children with CP, the number of GA-specific total live births 180) 28-31 weeks GA: 32.2/1000 live births (95% CI 18.1-52.5/1000 live births) (15 children with CP, the number of GA-specific total live births 466) 32-36 weeks GA: 4.6/1000 live births (95% CI 2.7-7.3/1000 live births) (18 children with CP, the number of GA-specific total live births 3913)	Low	Children born 1990-1993.
Odd 2013 UK	Regional prospective cohort	n=741 moderate to late preterm children (Gestational age: 32-36 weeks (preterm))	CP was identified from hospital and community health service records and the diagnosis confirmed at age 4 years using the Standard Recording of Central Motor Deficit. No other details given.	At 7 years CP 32-36 weeks GA: 7/741, 0.9% (0.4-1.9%)	Low	April 1991 to December 1992

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Rieger-Fackeldey 2010 Germany	Prospective cohort study	n=107 initial cohort n=27 survived at 5 years follow-up	The Gross Motor Function Classification System (GMFCS) was used to assess mobility for CP, level 1 (normal) to level 5 (Lack of mobility).	At 5 years age CP ≥22 weeks GA/BW <501g; GMFCS level >1 (abnormal): 7/19, 37% (16-62%) ≥22 weeks GA/BW <501g; GMFCS level 2: 5/19, 26% (9-52%) ≥22 weeks GA/BW <501g; GMFCS level 3: 2/19, 11% (1.3-33%)	Low	Children born between 1998 and 2001, assessed at 5 years age
Roberts 2010 Australia	Regional cohort study	n=223 total live births n=151 consecutive live births at 22-27 weeks completed gestation n=144 survived to age 8 years	No information was provided how CP was diagnosed/assessed or how CP was defined but includes at least the following aspects: the child not walking, the child walking with considerable difficulty, with or without appliances, walking with minimal limitation.	At 8 years (corrected) CP 22-27 weeks GA: 16/141, 11.3% (6.6-17.8%)	Low	Children born in 1997, follow-up at 8 years of age (corrected)
Robertson 2007 Canada	A prospective population-based longitudinal outcome study	n=975 number of children who were live born between 1992-2003 n=506 number of children who survived to 2 years between 1992-2003	Throughout the 30 years of the whole study period, the diagnoses of CP was done by only 6 physicians in total, all which were reviewed by a single physician and all children with the diagnosis of CP have been seen by the same paediatric physiatrist (second author) and a consensus diagnosis of CP (spastic, dyskinetic, ataxic) and subtype	At 2 years of age (confirmed at 3 years of age) CP 1992-1994 22-27 weeks GA: 131/1000 live births (95% CI 90-183/1000 live births) (cases of CP 29, number of live births 221, number of survivors at 2 years who were assessed is not reported) 1995-1997 22-27 weeks GA: 69/1000 live births (95% CI 41-108/1000 live births) (cases of CP 17, number of live births 246, number of survivors at 2 years who were assessed is not reported) 1998-2000	Moderate	Children born 1974-2003 (only years 1992-2003 considered for the review), assessment of CP at 18-24 months corrected age (confirmation of

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		Number of children who were followed up at 2 years between 1992-2003 was not reported for these years. Over the whole study period 1974-2003, out of 881 survivors at 2 years, 23 were lost to follow-up.	(hemiplegic, diplegic, quadriplegic) made. Outcome of all children diagnosed with CP were confirmed after 3 years of age. The definition of CP was a disorder of movement and posture due to a defect or lesion of the immature brain. Children were grouped, using outcomes collected from those older than 3 years, as 1) ambulatory, i.e. capable of walking independently with or without ankle-foot orthoses, assistive mobility devices or both, or 2) non-ambulatory, i.e. requiring transportation or power mobility devices	22-27 weeks GA: 69/1000 live births (95% CI 41-108/1000 live births) (cases of CP 17, number of live births 246, number of survivors at 2 years who were assessed is not reported) 2001-2003 22-27 weeks GA: 19/1000 live births (95% CI 6-44/1000 live births) (cases of CP 5, number of live births 262, number of survivors at 2 years who were assessed is not reported) 1992-2003 22-27 weeks GA: 70/1000 live births (95% CI 55-88/1000 live births) (cases 68, number of live births 975, number of survivors or 2 years who were assessed is not reported) Non-ambulatory CP 1992-1994 22-27 weeks GA: 59/1000 live births (95% CI 32-99/1000 live births) (cases of CP 13, number of live births 221, number of survivors at 2 years who were assessed is not reported) 1995-1997 22-27 weeks GA: 16/1000 live births (95% CI 5-41/1000 live births) (cases of CP 4, number of live births 246, number of survivors at 2 years who were assessed is not reported) 1998-2000 22-27 weeks GA: 8/1000 live births (95% CI 1-29/1000 live births) (cases of CP 2, number of live births 246, number of survivors at 2 years who were assessed is not reported) 2001-2003 22-27 weeks GA: 8/1000 live births (95% CI 1-27/1000 live births) (cases of CP 2, number of live		diagnosis at 3 years or older).

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				births 262, number of survivors at 2 years who were assessed is not reported) 1992-2003 22-27 weeks GA: 22/1000 live births (95% CI 13-33/1000 live births) (cases 21, number of live births 975, number of survivors or 2 years who were assessed is not reported)		
Salakorpi 2001 Finland	Population-based cohort study.	n=228 extremely low birth weight infants born n=156 survived over the age of 12 months (corrected) (69%) n=142 followed up at 4 years (91% of ones who survived)	At 4 years age (+4 weeks) children were examined by a neurologist (with an assessment of motor skills, fine motor skills and drawing (handedness), eye movements, muscle tone, tendon reflexes and a positive Babinsky sign, persistent or exaggerated primitive reflexes, dyskinesia or ataxia were found.	At 4 years age CP Birthweight <1000g (mean GA 27 weeks): 27/142, 19.0% (12.9-26.5%) CP bilateral spastic (diplegia or tetraplegia) Birthweight <1000g (mean GA 27 weeks): 15/142, 10.6% (6.0-16.8%) CP hemiplegia Birthweight <1000g (mean GA 27 weeks): 8/142, 5.6% (2.5-10.8%) CP dystonic or athetoid type Birthweight <1000g (mean GA 27 weeks): 4/142, 2.8% (0.8-7.1%)	Moderate	Children born 1/1/1991-31/12/1994, assessed at 4 years age.
Serenius 2013 Sweden	Population-based prospective cohort study (EXPRESS group).	Sample recruited: n=707 live born preterm infants n=701 term controls Sample analysed after exclusions:	The definition of CP used was according to Bax et al. and characterised as hemiplegic, diplegic,tetraplegic, ataxic, or dyskinetic. Severity of CP was classified as mild in children who were able to walk without an aid, moderate in children able	At 2.5 years (corrected age) CP (formally assessed or assessed by chart review) <27 weeks GA: mild CP: 13/456, 2.9% (1.5-4.8%) <27 weeks GA: moderate CP: 13/456, 2.9% (1.5-4.8%) <27 weeks GA: severe CP: 6/456, 1.3% (0.48-2.8%) <27 weeks GA: moderate/severe CP: 19/456, 4.2% (2.5-6.4%) <27 weeks GA: any CP: 32/456, 7% (4.9-9.8%)	Moderate	Children born between 2004 and 2007, assessed at 2.5 years corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=456 preterm infants	to walk with an aid, and severe in children who were unable to walk even with an aid			
Stahlmann 2009 Germany	A geographically defined cohort	n=154 infants identified n=95 survived until discharge to home n=92 survived until follow-up at 7-9 years n=75 children were assessed at 7-9 years (81.5% of the surviving children)	All neurosensory examinations were conducted by the first author who was unaware of the neonatal course of the child and the outcome of the follow-up at 3-5 years. CP was assessed through Gross Motor Function Classification System (GMFCS). Non-ambulant CP was considered severe dysfunction (GMFCS III-V) and CP with low functional impairment (GMFCS I-II)	At 7-9 years CP <27 weeks GA: 11/75, 14.7% (7.6-24.7%) Non-ambulatory CP (GMFCS 3-5) <27 weeks GA: 8/75, 10.7% (4.7-19.9%)	Moderat e	Children born 1997-1999, follow-up at 7-9 years of age
Sutton 1999 Australia	Prospective population-based cohort study	n=1170 (including live and still births in 1992-1993). n=614 live births. n=434 admitted to tertiary NICU (180 died in	The neurological outcome at 12 months was expressed as normal, provisional diagnosis of cerebral palsy, or motor delay greater than expected with or without equivocal neurological signs.	At 12 months corrected age CP All <27 weeks GA: 22/139, 15.8% (10.2-23.0) 23 weeks GA: 1/1, 100% (25-100%) 24 weeks GA: 4/25, 16% (4.5-36.0%) 25 weeks GA: 7/36, 19,4% (8.2-36.0%) 26 weeks GA: 10/77, 13.0% (6.4-22.6%) 27 weeks GA: 20/105, 19.1% (12.0-27.9%)	Low	Infants born between 1992-1993, assessed at 12 months corrected age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		the labour ward at 12 months: n=244 infants had a neurological examination n=239 infants had a formal Griffiths development assessment n=255 data available for at least one follow-up				
Tommiska 2003 Finland	Prospective cohort study	n=208 extremely low birth weight infants (born with bw <1000 g) of which n=104 children were born at 22-26 weeks GA	CP was defined as a non-progressive motor impairment with spastic or dystonic muscle tone, brisk tendon reflexes, positive Babinski's sign, and persistent primitive reflexes. Four categories were used: diplegia, hemiplegia, tetraplegia, and ataxia or athetosis syndrome.	At 18 months corrected age CP 22-23 weeks GA: 1/5, 20.0% (0.5-71.6%) 24 weeks GA: 2/18, 11.1% (1.4-34.7%) 25 weeks GA: 4/34, 11.8% (3.3-27.5%) 26 weeks GA: 5/47, 10.6% (3.6-23.1%) 22-26 weeks GA: 12/104, 11.5% (6.1-19.3%) The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 23/208, 11.1% (7.1-16.1%) CP diplegia The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 15/208, 7.2% (4.1-11.6%)	Low	Recruitment from 1st January 1996 to 31st December 1997, follow-up at 18 months of corrected age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				CP tetraplegia The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 4/208, 1.9% (0.5-4.9%) CP hemiplegia The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 2/208, 1.0% (0.1-3.4%) CP ataxia/athetosis The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 2/208, 1.0% (0.1-3.4%)		
Toome 2012 Estonia	Population based national cohort study	n=187 very low gestational age infants (83% eligible for follow-up 155/187)	Families were invited for a physical assessment by a paediatrician, neurological examination by a child neurologist and an assessment of development by a child psychologist. Cerebral palsy was defined according to the guidelines of the Surveillance of Cerebral Palsy in Europe collaborative group, and the Gross Motor Function Classification System (GMFCS) was used to quantify motor function in infants with CP.	At 2 years age (corrected) CP <32 weeks GA: 17/155, 11% (6.5-17%) 22-25 weeks GA: 3/17, 18% (3.8-43.3%) 26-31 weeks GA: 2/17, 12% (1.5-36%) GMFCS level 2-5 <32 weeks GA: 13/17, 76.4% (50-93%) Spastic displegia <32 weeks GA: 7/17, 41% (18-67%)	Low	Children born 2007, assessed at 2 years (corrected age)
Vincer 2014 Canada	Population-based cohort study	n=1014 the whole cohort born in 1988-2007	A neurological examination between 12 and 42 months' corrected age was used to presence or absence of	CP Children born 1993-1997 <31 weeks GA: 23/288, 8.0% (5.1-11.7%) Children born 1998-2002	Low	1988-2007 (data from 1993 onwards

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			CP and to define the gross motor functional classification. CP was defined as a disorder of control of movement or posture secondary to a non-progressive brain lesion	<31 weeks GA: 42/251, 16.7% (12.3-21.9%) Children born 2003-2007 <31 weeks GA: 16/262, 6.1% (3.5-9.7%) Children born 1993-2007 <31 weeks GA: 81/801, 10.1% (8.1-12.4%) Level 1 CP (mild) Children born between 1993-1997 <31 weeks GA: 12/288, 4.2% (2.2-7.2%) Children born between 1998-2002 <31 weeks GA: 31/251, 12.4% (8.6-17.1%) Children born between 2003-2007 <31 weeks GA: 11/262, 4.2% (2.1-7.4%) Children born between 1993-2007 <31 weeks GA: 54/801, 6.7% (5.1-8.7%) Level 2-5 CP (moderate to severe) Children born between 1993-1997 <31 weeks GA: 11/288, 3.8% (1.9-6.7%) Children born between 1998-2002 <31 weeks GA: 11/251, 4.4% (2.2-7.7%) Children born between 2003-2007 <31 weeks GA: 5/262, 1.9% (0.6-4.4%) Children born between 1993-2007 <31 weeks GA: 27/801, 3.4% (2.2-4.9%)		used for this review)
Vohr 2005 USA	A multicentre cohort study	n=3785 infants included in analysis	CP was defined as non-progressive central nervous system disorder characterised by	At 18-22 months corrected age Disorders: CP	Moderat e	1993-1998, follow-up at 18 to 22 months of

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		(51% of the original sample, 79.5% of the ones who survived up to discharge or 120 days)	abnormal muscle tone in at least 1 extremity and abnormal control of movement or posture. Moderate to severe CP included children who were non- ambulatory or required an assistive device for ambulation.	Years 1993-94 22-26 weeks GA: 134/665, 20.1% (17.2-23.4%) 27-32 weeks GA: 55/444, 12.4%, (9.5-15.8%) Years 1995-96 22-26 weeks GA: 134/716, 18.7% (15.9-21.8%) 27-32 weeks GA: 60/538, 11.2% (8.6-14.1%) Years 1997-98 22-26 weeks GA: 165/910, 18.1% (15.7-20.8%) 27-32 weeks GA: 58/512, 11.3% (8.7-14.4%) All epochs, 1993-98 22-26 weeks GA: 433/2291, 18.9% (17.3-20.6%) 27-32 weeks GA: 173/1494, 11.6% (10.0-13.3%) 22-32 weeks GA: 606/3785, 16.0% (14.9-17.2%) Moderate to severe CP Years 1993-94 22-26 weeks GA: 80/665, 12.1% (10.0-14.8%) 27-32 weeks GA: 35/444, 7.8% (5.6-10.8%) Years 1995-96 22-26 weeks GA: 77/716, 10.8% (8.6-13.3%) 27-32 weeks GA: 38/538, 7.1% (5.1-9.6%) Years 1997-98 22-26 weeks GA: 95/910, 10.4% (8.5-12.6%) 27-32 weeks GA: 32/512, 6.3% (4.3-8.7%) All epochs, 1993-1998 22-26 weeks GA: 252/2291, 11.0% (9.8-12.4%) 27-32 weeks GA: 105/1494, 7.0% (5.8-8.4%) 22-32 weeks GA: 357/3785, 9.4% (8.5-10.4%)		corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Wood 2000 UK and Ireland	Population based prospective cohort study	n=4004 infants identified n=1185 survived at birth (843/1185 were admitted to NICU; 342/1185 died in the delivery room) n=283 assessed at follow-up	Cerebral palsy was classified retrospectively according to the description of function for each limb in children with abnormal results or neurological examination (diplegia, hemiplegia, quadriplegia, other non-spastic types (hypotonia, dyskinesia)).	At median age 30 months. CP (children with neuromotor disability) 22-25 weeks GA: 50/283, 17.7% (13.4-22.6%) Diplegia CP 22-25 weeks GA: 27/283, 9.5% (6.4-13.6%) Severe diplegia CP 22-25 weeks GA: 12/283, 4.2 (2.2-7.3%) Hemiplegia CP 22-25 weeks GA: 5/283, 1.8% (0.6-4.1%) Severe hemiplegia CP 22-25 weeks GA: 1/283, 0.4% (0.01-2.0%) Quadriplegia CP 22-25 weeks GA: 12/283, 4.2 (2.2-7.3%) Severe quadriplegia CP 22-25 weeks GA: 11/283, 3.9% (2.0-6.9%)	Low	Infants born 1995, assessed at median age 30 months
Evidence on intellectual disability						
Anderson 2003 Australia	Prospective regional cohort study (Victorian Infant Collaborative Study Group)	n=568 consecutive live births of neonates with BW <1000g or <28 weeks GA. n=298 infants survived to 2, and 5 years assessment. n=275 children	Cognitive ability was assessed using the Wechsler Intelligence Scale for Children (WISC-III). Full scale IQ was a measure of general intellectual ability. Major intellectual impairment was classified as an IQ below 70 (<-2SDs).	At 8 years Major intellectual impairment (WISC-III IQ<70, n=275) <28 weeks GA or ELBW: Full scale IQ: 14/275, 5.1% (2.8-8.4%)	Low	Infants born 1991-1992, assessed at 8 years age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		assessed at 8 years age.				
Andrews 2008 USA	Prospective cohort study	n=259 (around 70% of the 375 eligible and alive for the follow-up) with data on IQ	Each child was given a battery of tests assessing a wide range of psychometric measures (requiring approximately 3 hours to complete) including the Wechsler Intelligence Scale for Children-IV (WISC-IV) or the Differential Ability Scales (DAS, for children who were not yet six-years-old or were unable to complete the WISC-IV) used to assess IQ. The IQ score <70 on the WISC-IV or DAS was considered a cognitive impairment.	At 6 years IQ <70 (WISC-IV or DAS) 23-32 weeks GA: 41/259, 15.8% (11.6-20.9%)	Low	
Anonymous 1997 Australia	A geographically determined cohort study (Victoria, Australia)	n=401 live born children born at 23-27 weeks n=225 children survived to 2 years of age (56.1%) n=219 were assessed at 2 years (97.3% of the survivors)	The psychological assessment included the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development, or alternative psychological tests if the children were assessed by a psychologist where the Bayley Scales were not available. The test scores were expressed as standardised normal developmental quotients	At 2 years MDI <-3 SD 23-27 weeks GA: 12/219, 5.5% (2.9-9.4%) MDI -2 to -3SD 23-27 weeks GA: 28/219, 12.8% (8.7-18.0%) MDI <=-2SD 23-27 weeks GA: 40/219, 18.3% (13.4-24.0%)	Low	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			using the mean and standard deviation for the MDI obtained from the normal birthweight controls. The children were considered to have severe mental developmental impairment if the score was below <-3 SD and moderate impairment if the score was between -2 and -3 SD.			
Beaino 2011 France	Population based prospective cohort (EPIPAGE)	n=1503	Cognitive deficiency was classified as moderate to severe when the MPC score was below 70 (-2SD below the norm).	At 5 years Moderate to severe cognitive impairment (MPC<70) 24-26 weeks GA: 16/102, 15.7% (9.2-24.2%) 27-28 weeks GA: 50/263, 19.0% (14.5-24.3%) 29-30 weeks GA: 36/409, 8.8% (6.2-12.0%) 31-32 weeks GA: 65/729, 8.9% (7.0-11.2%)	Low	1997-2002. Cohort established in 1997. Follow-up at 5 years of age.
Charkaluk 2010 France	Population based prospective cohort study (EPIPAGE).	n=634 children born alive at GA <33 weeks. n=546 surviving children included at follow-up.	Developmental quotients were ascertained by the revised Brunet-Lezine scale, an early childhood psychomotor development scale covering four domains of development: gross motor function, fine motor function, language and sociability; DQ <70 is defined as severe developmental delay	At 2 years (corrected age) Global DQ/developmental delay <70 (severe) <33 weeks GA: 8/347, 2.3% (1.0-4.5%)	Low	Children born in 1997, assessed at 2 years (corrected age).
Doyle 2011 Australia	A population-based cohort study (in the	n=257 live births with bw 500-999	Development delay was assessed with the Bayley Scales of Infants and	At 2 years (corrected age)	Moderat e	Children born 2005, follow-up at

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
	State of Victoria).	g (excl. cases with lethal anomalies) n=172 survived to 2 years n=165 assessed at 2 years (96%)	Toddler Development (Bayley-III) and Cognitive Scale and Language Composite Scale. The scores for ELBW infants were compared with the term controls rather than the test norms. Moderate developmental delay was defined as a score on either scale from -3SD to -2SD. Severe developmental delay was defined as a score <-3SD.	Moderate developmental delay (Bayley-III), -3SD to -2SD BW 500-999 g (mean GA 25.7 [SD 2.3]): 19/165, 11.5% (7.1-17.4%) Severe developmental delay (Bayley-III), <-3SD BW 500-999 g (mean GA 25.7 [SD 2.3]): 6/165, 3.6% (1.4-7.8%) Moderate to severe developmental delay (<=2SD) BW 500-999 g (mean GA 25.7 [SD 2.3]): 25/165, 15.2% (10.1-21.6%)		2 years (corrected age).
De Groote 2007 Belgium	Population-based geographically defined cohort study (EPIBEL)	n=95 children that survived to discharge from NICU n=77 children assessed at 3 years (n=3 died before follow-up, n=12 parents did not give consent, n=3 could not be reached), 81% follow-up rate (84% of the ones who were	The Dutch edition of the second version of the Bayley Scales of Infant Development (BSID-II-NL) was used to assess mental and psychomotor development. The BSID-II-NL is standardised on a mean score of 100 and a SD of 15 points. Moderate impairment is defined as a score of 55-69 and severe impairment as a score of <55.	At 3 years Severe mental developmental delay (MDI <55) <27 weeks GA: 14/77, 18.2% (10.3-28.6%) Moderate mental developmental delay (MDI 55-69) <27 weeks GA: 8/77, 10.4% (4.6-19.5%) Moderate to severe mental developmental delay (MDI <70)* <27 weeks GA: 22/77, 28.6% (18.9-40.0%)	Low	Children born in 1999-2000, follow-up at 3 years of age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		alive at follow-up).				
Foix-Helias 2008 France	Prospective population based cohort study (EPIPAGE).	n=1781 children with data on CP (77% of n=2300 survivors up to follow-up) n=1508 children with data on cognition (66% of the n=2300 survivors up to follow-up)	Cognitive ability was assessed using the mental processing composite (MPC) of the Kaufman Assessment Battery for Children (K-ABC). This score is standardised to a mean (±SD) of 100 (±15) based on a reference population of French children born in the late 1990s. MPC scores of less than 70 indicate cognitive impairment.	At 5 years Moderate cognitive impairment (MPC 55-69) 24-32 weeks GA: 145/1508, 9.6% (8.2-11.2%) 24-27 weeks GA: 33/222, 14.9% (10.5-20.2%) 28-32 weeks GA: 112/1286, 8.7% (7.2-10.4%) Severe cognitive impairment (MPC <55) 24-32 weeks GA: 35/1508, 2.3% (1.6-3.2%) 24-27 weeks GA: 6/222, 2.7% (1.0-5.8%) 28-32 weeks GA: 29/1286, 2.3% (1.5-3.2%) Cognitive impairment (MPC <70) 24-32 weeks GA: 180/1508, 11.9% (10.3-13.7%) 24-27 weeks GA: 39/222, 17.6% (12.8-23.2%) 28-32 weeks GA: 141/1286, 11.0% (9.3-12.8%)	Moderate	Recruitment took place in 1997. Follow-up was at 5 years
de Kleine 2003 Netherlands	A prospective cohort study	n=566 eligible children n=431 assessed at 5 years (76%) n=404 assessed for motor functioning (M-ABC) n=402 assessed for IQ (IQ test)	At 5 years, cognitive delay was assessed with revised Amsterdam child intelligence test (IQ test) by trained child psychologists. The revised Amsterdam child intelligence test has been normalised for Dutch children between 4-7 years. Children with a score between -2 and -1 SD were considered at risk and those below -2 SD were abnormal.	At 5 years Cognitive delay (IQ <-2SD) <32 weeks GA/bw <1500 g: 25/402, 6.2% (4.1-9.0%)	Moderate	Children 1992-1995, assessed at 5 years.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=407 assessed for behavioural problems (CBCL)				
Joseph 2016b USA	Prospective cohort study (ELGAN)	n=873 preterm children at 10 years follow-up	Cognitive ability (IQ): School- Age Differential Ability Scales–II (DAS-II) 28 Verbal and Nonverbal Reasoning scales.	At 10 years General cognitive ability (<=-2SD): DAS-II Verbal: 22-27 weeks GA: 148/873, 17.0% (95%CI 14.5-19.6) DAS-II Nonverbal Reasoning: 22-27 weeks GA: 131/873, 15% (95%CI 12.7-17.6)	Low	Children born 2002-2004
Larroque 2008 France	A longitudinal cohort study (EPIPAGE).	n=1817 children born at 22-32 weeks were followed at 5 years of age (77% of the population that survived) n=1534 children born at 22-32 weeks with data on MPC score outcome	Cognitive function: At 5 years of age, children were invited for a check-up with a psychologist especially trained in use of the Kaufman assessment battery for children (K-ABC). The K-ABC13 was used to assess cognitive function. The mental processing composite (MPC) scale,13 which is considered to be equivalent to intelligence quotient (IQ), is a global measure of cognitive ability in two dimensions: a sequential processing scale and a simultaneous processing scale. The achievement scale assesses knowledge of facts, language ideas,	At 5 years Cognitive impairment (MPC <70) <33 weeks GA: 182/1534, 11.9% (10.3-13.6%) 24-25 weeks GA: 6/48, 12.5% (4.7-25.3%) 26 weeks GA: 12/57, 21.1% (11.4-33.9%) 27 weeks GA: 22/118, 18.6% (12.1-26.9%) 28 weeks GA: 31/150, 20.7% (14.5-28.0%) 29 weeks GA: 17/167, 10.2% (6.0-15.8%) 30 weeks GA: 25/252, 9.9% (6.5-14.3%) 31 weeks GA: 34/319, 10.7% (7.5-14.6%) 32 weeks GA: 35/423, 8.3% (5.8-11.3%) <28 weeks GA: 40/223, 17.9% (13.1-23.6%) 28-31 week GA: 107/888, 12.1% (10.0-14.4%)	Moderate	1997, follow-up at 5 years of age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			and skills related to school. Each scale is standardised to a mean of 100 (SD 15). MPC score <70 considered a cognitive impairment.			
Leversen 2011 Norway	Prospective observational national cohort study	n=306 children assessed at 5 years (n=638 children born, of which n=376 survived to discharge, of which 3 died and n=373 were followed-up at 2 years, of which 1 died and 1 child with Down's syndrome were excluded and 65 were lost to follow-up)	Cognitive abilities (verbal IQ, performance IQ, and full-scale IQ) were assessed with the Wechsler Preschool and Primary Scale of Intelligence - Revised (WPPSI-R). Reference means for the IQ scores are 100.	At 5 years Full-scale IQ <55 22-27 weeks GA or bw 500-999 g: 2/306, 0.7% (0.08-2.3%) 23-25 weeks GA: 2/87, 2.3% (0.3-8.1%) 26-27 weeks GA: 0/152, 0% (0-2.4%) >27 weeks GA (bw <1000 g): 0/67, 0% (0-5.4%) Full-scale IQ 55-70 22-27 weeks GA or bw 500-999 g: 15/306, 4.9% (2.8-8.0%) 23-25 weeks GA: 6/87, 6.9% (2.6-14.4%) 26-27 weeks GA: 4/152, 2.6% (0.7-6.6%) >27 weeks GA (bw <1000 g): 5/67, 7.5% (2.5-16.6%) Full-scale IQ <70 22-27 weeks GA or bw 500-999 g: 17/306, 5.6% (3.3-8.8%) 23-25 weeks GA: 8/87, 9.2% (4.1-17.3%) 26-27 weeks GA: 4/152, 2.6% (0.7-6.6%) >27 weeks GA (bw <1000 g): 5/67, 7.5% (2.5-16.6%)	Moderate	Children born 1999 and 2000, follow-up at 5 years
Leversen 2012 Norway	Prospective observational national cohort study	n=232 assessed for mental delay at both 2 and 5 years	Mental delay: At 2 years of corrected age, a qualified paediatrician assessed the child's mental function by addressing four specific	At 2 years of age (corrected) Mental delay (paediatrician's assessment on 4 specific issues) <28 weeks GA/bw <1000 g: 41/232, 17.7% (13.0-23.2%)	Low	Children born 1999-2000, follow-up at 2 and 5 years.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=260 assessed for motor delay at both 2 and 5 years	issues and was classified as delayed if they did not respond appropriately when asked to perform tasks such as fetching objects, did not understand and speak words, co-operate and concentrate and generally respond as expected for age. At 5 years of age (chronological), a psychologist assessed cognitive abilities with the Welchsler Preschool and Primary Scale of Intelligence -Revised (WPPSI-R). On the WPPSI-R, verbal IQ, performance IQ and full-scale IQ were calculated from the subscales. Reference means (SD) for the IQ scores are 100. IQ <85 was considered a delay.	Problems Motor delay (paediatrician's assessment on 8 milestone abilities) <28 weeks GA/bw <1000 g: 36/260, 13.9% (9.9-18.7%) At 5 years of age (chronological) Disorders Mental delay (WPPSI-R, IQ <85) <28 weeks GA/bw <1000 g: 63/232, 27.2% (21.5-33.4%)		
Marlow 2005 UK and Ireland	Population-based national cohort study (EPICure)	n=241 (82% of the eligible ones, n=293)	Cognitive impairment: when cognitive assessment was appropriate, it was made with the use of the Kaufman Assessment Battery for Children (K-ABC). If the child's disability precluded the	At 6 years Severe cognitive impairment (IQ <-3SD) <26 weeks GA: 50/241, 20.8% (15.8-26.4%) <=23 weeks GA: 6/24, 25.0% (9.8-46.7%) 24 weeks GA: 20/73, 27.4% (17.6-39.1%) 25 weeks GA: 24/144, 16.7% (11.0-23.8%) Moderate cognitive impairment (IQ -2 to -3SD)	Moderate	Children born 1995, follow-up at 6 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			use of the K-ABC, either the Griffiths Scales of Mental Development (n=35) or the neuropsychological instrument knows as NEPSY (n=6) were used. The results for these children were substituted for the missing values in the Mental Processing Composite of K-ABC to produce an overall cognitive score. The cognitive performance (IQ) was classified as severely impaired if the score was <-3 SD of the mean and moderate if the score of -2 to -3 SD.	<26 weeks GA: 48/241, 19.9% (15.1-25.5%) <=23 weeks GA: 8/24, 33.3% (15.6-55.3%) 24 weeks GA: 13/73, 17.8% (9.8-28.5%) 25 weeks GA: 27/144, 18.8% (12.7-26.1%) Moderate to severe cognitive impairment (IQ <=-2SD) <26 weeks GA: 98/241, 40.7% (34.4-47.2%) <=23 weeks GA: 14/24, 58.3% (36.6-77.9%) 24 weeks GA: 33/73, 45.2% (33.5-57.3%) 25 weeks GA: 51/144, 35.4% (27.6-43.8%)		
Marret 2007 France	Population based prospective cohort (EPIPAGE).	n=1455	Children were invited for a check -up at 5 years, and assessed by trained psychologists blinded to their perinatal data. The assessment used the Kaufman Assessment Battery for Children (K-ABC) test. Overall cognitive ability was evaluated by the Mental Processing Composite (MPC) score. Cognitive deficiency was classified as moderate to severe when the MPC score was	At 5 years Cognitive impairment (MPC <70) 30 weeks GA: 25/252, 9.9% (6.5-14.3%) 31 weeks GA: 34/319, 10.7% (7.5-14.6%) 32 weeks GA: 34/423, 8.0% (5.6-11.1%) 33 weeks GA: 9/110, 8.2% (3.8-15.0%) 34 weeks GA: 6/113, 5.3% (2.0-11.2%)	Low	1997-2002. Cohort established in 1997. Follow-up at 5 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			below 70 (-2SD below the norm).			
Mikkola 2005 Finland	National population-based prospective cohort study	n=203 children with birth weight <1000 g (of n=206 children who survived up to follow-up) n=102 children with <27 weeks GA	Cognitive impairment, defined as IQ score <70, assessed by the Wechsler Preschool and Primary Scale of Intelligence-revised (WPPSI-R).	At 5 years Cognitive impairment (IQ <70) Children born with birth weight <1000 g (mean GA 27.3 (SD 2.1): 19/203, 9.4% (5.7-14.2%) <27 weeks GA: 12/102, 11.8% (6.2-19.7%)	Low	1996-1997, follow-up at 5 years of age.
Moore 2012 UK		n=576 children born 22-26 weeks' gestation, assessed at follow-up (n=38 born at 22-23 weeks; n=98 born at 24 weeks; n=189 born at 25 weeks; n=251 born at 26 weeks)	Cognitive disability and communication disability: Cognitive and communication disability were assessed with the third edition of the Bayley Scales of Infant Development (BSID-III) cognitive and language scales by trained assessors. A subgroup of the cohort (=208) was evaluated using a combination of the cognitive and language scales of the BSID-III and the mental developmental index (MDI) from the second edition (BSID-II). As assessments were sometimes delayed,	At 3 years (generally, some assessments delayed) Severe cognitive disability (Bayley or WPPSI, <-3SD) 22-26 weeks GA: 57/576, 9.9% (7.6-12.6%) 22-23 weeks GA: 7/38, 18.4% (7.7-34.3%) 24 weeks GA: 11/98, 11.2% (5.7-19.2%) 25 weeks GA: 20/189, 10.6% (6.6-15.9%) 26 weeks GA: 19/251, 7.6% (4.6-11.6%) Moderate cognitive disability (Bayley or WPPSI, -2 to -3SD) 22-26 weeks GA: 37/576, 6.4% (4.6-8.8%) 22-23 weeks GA: 5/38, 13.2% (4.4-28.1%) 24 weeks GA: 6/98, 6.1% (2.3-12.9%) 25 weeks GA: 15/189, 7.9% (4.5-12.8%) 26 weeks GA: 11/251, 4.4% (2.2-7.7%) Moderate to severe cognitive disability (Bayley or WPPSI, <=-2SD) 22-26 weeks GA: 94/576, 16.3% (13.4-19.6%)	Low	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			children older than 42 months were evaluated using the Wechsler preschool and primary scales of intelligence (WPPSI), the assessors were trained and validated to administer the scales. Severe cognitive disability was defined as developmental score of <-3SD of the mean. Moderate cognitive disability was defined as developmental score of -2 to -3 SD of the mean.	22-23 weeks GA: 12/38, 31.6% (17.5-48.7%) 24 weeks GA: 17/98, 17.4% (10.4-26.3%) 25 weeks GA: 35/189, 18.5% (13.3-24.8%) 26 weeks GA: 30/251, 12.0% (8.2-16.6%)		
Rieger-Fackeldey 2010 Germany	Prospective cohort study	n=107 initial cohort n=27 survived at 5 years follow-up n=19 eligible for follow-up (8/27 were not able to be evaluated due to refusal of consent by parents (n=3), or family had moved away, failed	Cognitive function was assessed by a child psychologist with the Kaufmann Assessment Battery for Children (K-ABC), which comprises the mental processing composite (global measure of cognitive ability/IQ). IQ <85 (mild impairment); IQ <70 (severe impairment).	At 5 years age Cognitive development (Mental Processing Composite, IQ ≥22 weeks GA/BW <501g; IQ<85: 10/17, 59% (33-82%) ≥22 weeks GA/BW <501g; IQ<70: 7/17, 41% (18-67%)	Low	Children born between 1998 and 2001, assessed at 5 years age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		appointment, or moved to another follow-up care (n=5))				
Roberts 2010 Australia	A regional cohort study	n=223 total live births n=151 consecutive live births at 22-27 weeks completed gestation n=144 survived to age 8 years	Intelligence was assessed using the Welchsler Intelligence Scale for Children, 4th edition (WISC-IV) Severe intellectual disability was defined as IQ <-3SD; moderate intellectual disability was defined as IQ -3SD to <-2SD.	At 8 years (corrected) Severe intellectual impairment (IQ <-3SD) 22-27 weeks GA: 9/144, 6.3% (2.9-11.5%) Moderate intellectual impairment (IQ-3SD to <-2SD) 22-27 weeks GA: 12/144, 8.5% (4.4-14.1%) Intellectual impairment (IQ <-2SD) 22-27 weeks GA: 21/144, 14.6% (9.3-21.4%)	Low	Children born in 1997, follow-up at 8 years of age (corrected)
Serenius 2013 Sweden	Population-based prospective cohort study (EXPRESS group).	Sample recruited: n=707 live born preterm infants n=701 term controls Sample analysed after exclusions: n=456 preterm infants n=701 full term controls	At 2.5 years of corrected age, certified psychologists assessed cognitive impairment with the Bayley Scales of Infant and Toddler Development Mild: a score of between 1 and 2 SD below the norm Moderate: a score of between 2 and 3 SD below the norm Severe: a score of less than 3 SD below the norm	At 2.5 years corrected age Cognitive impairment <27 weeks GA: mild (scores 83-94): 258/399, 64.7% (60.0-70.0%) <27 weeks GA: moderate (scores 72-82): 96/399, 24% (20.0-29.0%) <27 weeks GA: severe (scores <72): 25/399, 6.3% (4.1-9.1%)	Moderate	Children born between 2004 and 2007, assessed at 2.5 years corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Stahlmann 2009 Germany	A geographically defined cohort study.	n=154 infants identified n=95 survived until discharge to home n=92 survived until follow-up at 7-9 years n=75 children were assessed at 7-9 years (81.5% of the surviving children)	Cognitive status was assessed with the Kaufman Assessment Battery for Children (K-ABC) German version. The Scale Mental Processing provides information about fundamental mental processes and represents the cognitive abilities, reported as intelligent quotient (IQ). Using the original test standardisation norms standard deviation (SD) was 15. We classified an IQ <55 severely impaired and IQ 55-69 as moderately impaired. In cases where the child had been recently tested (within the last year) with the K-ABC or another equivalent instrument (n=7), e.g. the Hamburg Wechsler Intelligence Test for Children (HAWIK), the Snijders-Oomen Nonverbal Intelligence Test (SON-R) or the Culture Fair Intelligence Tests (CFT) we used the reported results.	At 7-9 years Severe cognitive impairment (IQ <55) <27 weeks GA: 11/75, 14.7% (7.6-24.7%) Moderate cognitive impairment (IQ 55-69) <27 weeks GA: 8/75, 10.7% (4.7-19.9%) Moderate to severe cognitive impairment (IQ <70) <27 weeks GA: 19/75, 25.3% (16.0-36.7%)	Moderate	Children born 1997-1999, follow-up at 7-9 years of age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Sutton 1999 Australia	Prospective population-based cohort study.	n=1170 (including live and still births in 1992-1993). n=614 live births. n=434 admitted to tertiary NICU (180 died in the labour ward).	Babies were assessed by a developmental paediatrician with or without a clinical psychologist, and in some cases a developmentally trained physiotherapist, with a full physical examination and Griffiths developmental assessment. Major developmental disability was defined as a general quotient of ≥ 2 SD below the mean on the Griffiths scale.	At 12 months corrected age Major developmental delay (formal Griffiths assessment) All <27 weeks GA: 14/135, 10.4% (5.8-16.8%) 23 weeks GA: 1/1, 100% (25-100%) 24 weeks GA: 4/23, 17.4% (5.0-39%) 25 weeks GA: 6/34, 17.7% (6.8-34.5%) 26 weeks GA: 3/77, 3.9% (0.81-11%) 27 weeks GA: 12/104, 11.5% (6.1-19.3%)	Low	Infants born between 1992-1993, assessed at 12 months corrected age
Toome 2012 Estonia	Population based cohort.	n=187 very low gestational age infants (83% eligible for follow-up 155/187)	The Bayley Scales of Infant and Toddler Development were used to generate composite scores for cognitive, language and motor skills, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below the mean for cognitive composite scores	At 2 years (corrected age) Cognitive delay <32 weeks GA: 26/155, 17% (11-24%)	Low	Children born 2007, assessed at 2 years (corrected age)
Vohr 2005 USA	A multicentre cohort study	n=3785 infants included in analysis (51% of the	At 18-22 months corrected age, families were invited to participate in a comprehensive assessment that	At 18-22 months corrected age Bayley MDI <70 Years 1993-94 22-26 weeks GA: 278/665, 41.8% (38.0-45.7%)	Moderat e	1993-1998, follow-up at 18 to 22 months of

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		original sample, 79.5% of the ones who survived up to discharge or 120 days)	consisted of a battery of developmental, neurologic, and behavioural assessment, a medical and social history and parent interviews. Bayley Scales of Infant Development II (BSID-II) was administered by a certified examiner who was trained to reliability and previous formal training in test administration. The Mental Developmental Index (MDI) was derived, a score of <70 was considered abnormal.	27-32 weeks GA: 133/444, 29.9% (25.7-34.5%) Years 1995-96 22-26 weeks GA: 276/716, 38.5% (35.0-42.2%) 27-32 weeks GA: 137/538, 25.5% (21.8-29.4%) Years 1997-98 22-26 weeks GA: 339/910, 37.2% (34.1-40.5%) 27-32 weeks GA: 117/512, 22.8% (19.3-26.7%) All epochs, 1993-1998 22-26 weeks GA: 893/2291, 39.0% (37.0-41.0%) 27-32 weeks GA: 387/1494, 25.9% (23.7-28.2%) 22-32 weeks GA: 1280/3785, 33.8% (32.3-35.4%)		corrected age
Evidence on speech and/or language disorder						
Moore 2012 UK	Prospective national cohort study (EPICure 2, this publication also used data from the original EPICure when comparing children born in 2006 to children born in 1995).	n=576 children born 22-26 weeks' gestation, assessed at follow-up (n=38 born at 22-23 weeks; n=98 born at 24 weeks; n=189 born at 25 weeks; n=251 born at 26 weeks)	Communication disability were assessed with the third edition of the Bayley Scales of Infant Development (BSID-III) cognitive and language scales by trained assessors. A subgroup of the cohort (=208) was evaluated using a combination of the cognitive and language scales of the BSID-III and the mental developmental index (MDI) from the second	Severe communication disability (Bayley or WPPSI, <-3SD) 22-26 weeks GA: 36/576, 6.3% (4.4-8.6%) 22-23 weeks GA: 6/38, 15.8% (6.0-31.3%) 24 weeks GA: 7/98, 7.1% (2.9-14.2%) 25 weeks GA: 13/189, 6.9% (3.7-11.5%) 26 weeks GA: 10/251, 4.0% (1.9-7.2%) Moderate communication disability (Bayley or WPPSI, -2 to -3SD) 22-26 weeks GA: 31/576, 5.4% (3.7-7.6%) 22-23 weeks GA: 4/38, 10.5% (2.9-24.8%) 24 weeks GA: 5/98, 5.1% (1.7-11.5%) 25 weeks GA: 11/189, 5.8% (2.9-10.2%)	Low	Children born in 2006 (this publication also compared the children born in 2006 to children born in 1995).

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			edition (BSID-II). As assessments were sometimes delayed, children older than 42 months were evaluated using the Wechsler preschool and primary scales of intelligence (WPPSI), the assessors were trained and validated to administer the scales. Severe cognitive disability was defined as developmental score of <-3SD of the mean. Moderate cognitive disability was defined as developmental score of -2 to -3 SD of the mean	26 weeks GA: 11/251, 4.4% (2.2-7.7%) Moderate to severe communication disability (Bayley or WPPSI, <=-2SD) 22-26 weeks GA: 67/576, 11.6% (9.1-14.5%) 22-23 weeks GA: 10/38, 26.3% (13.4-43.1%) 24 weeks GA: 12/98, 12.2% (6.5-20.4%) 25 weeks GA: 24/189, 12.7% (8.3-18.3%) 26 weeks GA: 21/251, 8.4% (5.3-12.5%)		
Serenius 2013 Sweden	Population-based prospective cohort study (EXPRESS group).	Sample recruited: n=707 live born preterm infants n=701 term controls Sample analysed after exclusions: n=456 preterm infants	At 2.5 years of corrected age, certified psychologists assessed language development with the Bayley Scales of Infant and Toddler Development. Language development was considered normal if the composite score on the respective Bayley-III scale was within 1 SD of the norm, mildly impaired if the score was between 1 and 2SD below the norm, moderately	At 2.5 years (corrected) Language impairment (assessed by Bayley III) <27 weeks GA: moderate (scores 72-84): 37/393, 9.4% (6.7-12.7%) <27 weeks GA: severe (score <72): 26/393, 6.6% (4.4-9.5%)	Moderate	Children born between 2004 and 2007, assessed at 2.5 years corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=701 full term controls	impaired if the score was between 2 and 3 SD below the norm, and severely impaired if the score was < 3SD below the norm. Mental developmental delay was also included as an outcome and classified as follows: Mild: a score of between 1 and 2 SD below the norm on either the cognitive or the language composite score. Moderate: a score of between 2 and 3 SD below the norm on either the cognitive or language composite score. Severe: a score of less than 3 SD below the norm on either the cognitive of language composite score.			
Toome 2012 Estonia	Population based national cohort study	n=187 very low gestational age infants (83% eligible for follow-up 155/187) n=153 full term controls	The Bayley Scales of Infant and Toddler Development were used to generate composite scores for language, with a mean (SD) score of 100 (±15). Results are presented according to the number of participants with scores <2SD below the mean.	At 2 years (corrected age) Language delay <32 weeks GA: 51/155, 33% (26-41%)	Low	Children born 2007, assessed at 2 years (corrected age).
Wolke 2008	Prospective national cohort	n=241 children for	Repetitive and expressive language was	At median age 6 years and 4 months	Low	Children born 1995,

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
UK	study (EPICURE study group).	whom parents consented to the study	assessed using the Preschool Language Scale-3 (PLS-3).	Language abilities (PLS-3 score), serious impairment (<2SD) ≤25 weeks and 6 days GA: total PLS-3: 31/199, 15.6% (10.8-21.4%)		assessed at median age 6 years and 4 months
Wood 2000 UK and Ireland	Population based prospective cohort study.	N=4004 infants identified n=1185 survived at birth (843/1185 were admitted to NICU; 342/1185 died in the delivery room) n=283 assessed at follow-up	All children had clinical examination including detailed medical history obtained from semi-structured interview with family, and a neurologic assessment, classification of degree and type of disability, and functional classification of hearing and visual ability. Development was assessed using the Bayley Scales of Infant Development II (BSID II) for mental and psychomotor development (MDI or PDI; score <55 considered as severe impairment, 55-69 considered as moderate impairment, 70-84 considered as mild impairment). If the child was unable to complete the BSID II assessment, the paediatrician estimated the child's development level as severely or	At median age 30 months Speech/communication (severe disability, n=283) 22-25 weeks GA: communicating by systemised method only: 3/283, 1.1% (0.2-3.1%) 22-25 weeks GA: not communicating by speech or other method: 15/283, 5.3% (3.0-8.6%)	Low	Infants born 1995, assessed at median age 30 months.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			moderately impaired (equivalent to Bayley score <55 or 55-69) or as not impaired.			
Evidence on attention deficit hyperactivity disorder						
Burnett 2014 Australia	Prospective regional cohort study.	n=215 early preterm/extremely low birth weight infants n=157 normal birth weight (>2499 g) controls n=372 in total	Standardized face-to-face clinical interview and questionnaires were used to assess the mental health status in late adolescence: ADHD, any type (All ADHD types assessed with the ADHD module of the Children's Interview for Psychiatric Syndromes (ChIPS)) ADHD, combined type ADHD, inattentive type ADHD, hyperactive/impulsive type	At 18 years age Any ADHD diagnosis <28 weeks GA/<1000g: 30/205, 14.6% (10.0-20.2%) ADHD combined type < 28 weeks GA/<1000g: 7/205, 3.4% (1.4-7.0%) ADHD inattentive type < 28 weeks GA/<1000g: 22/205, 10.7% (6.9-16.0%) ADHD hyperactive/impulsive type < 28 weeks GA/<1000g: 1/205, 0.5% (0.01-2.7%)	Low	Adolescents born between 1991 and 1992, assessed at 18 years age.
Johnson 2010 UK and Ireland	Population-based cohort study	n=219 children born at <26 weeks of GA were followed up at 11 years	The Development And Well Being Assessment (DAWBA), a structured psychiatric evaluation regarding children's development and behaviour was administered to parents via telephone interview (92%) or online (8%) from which information required for assigning ICD-10 and DSM-IV-TR diagnoses of childhood psychiatric disorders was	At 11 years Any DSM-IV clinical diagnosis <26 weeks GA: 51/219, 23.3% (17.9-29.5%) Any ADHD <26 weeks GA: 21/183, 11.5% (7.3-17.0%) ADHD inattentive subtype <26 weeks GA: 13/183, 7.1% (3.8-11.8%) ADHD combined type <26 weeks GA: 8/183, 4.4% (1.9-8.4%)	Low	Children born 1995, follow-up at 11 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			obtained. Supplemental information was provided by teachers who completed a corresponding questionnaire-based version of the DAWBA.			
Evidence on autism spectrum disorder						
Johnson 2010 UK and Ireland	Population-based cohort study	n=219 children born at <26 weeks of GA were followed up at 11 years	The Development And Well Being Assessment (DAWBA), a structured psychiatric evaluation regarding children's development and behaviour was administered to parents via telephone interview (92%) or online (8%) from which information required for assigning ICD-10 and DSM-IV-TR diagnoses of childhood psychiatric disorders was obtained. Supplemental information was provided by teachers who completed a corresponding questionnaire-based version of the DAWBA. Multi-informant data were collated by study assessors (paediatricians and psychologist), and	At 11 years Any ASD <26 weeks GA: 16/201, 8.0% (4.6-12.6%) Autistic disorder <26 weeks GA: 13/201, 6.5% (3.5-10.8%) Atypical autism <26 weeks GA: 3/201, 1.5% (0.3-4.3%)	Low	Children born 1995, follow-up at 11 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			potential cases were identified using computer-generated scoring algorithms (www.dawba.com). Summary sheets and clinical transcripts (with any reference to birth status removed) were then reviewed by two child and adolescent psychiatrists who had no prior knowledge of the children or their birth status and were therefore blind to group allocation, and who assigned DSM-IV and ICD			
Joseph 2016a USA	Prospective cohort study (ELGAN)	n=1198 preterm infants surviving to 10 years n=966 children recruited for follow-up at 10 years n=889 mothers of infants who agreed to participate	Autism Diagnostic Interview–Revised (ADI-R), an in-depth parent interview that assesses symptoms in the core domains of communication, social, and repetitive behaviour, and classifies autism based on 30–36 ratings, depending on the child's language level. Children who met criteria for autism or ASD on the ADI-R were assessed with the Autism	At 10 years ASD (assessed by ADI-R): <28 weeks GA: 79/857, 9.2% (95% CI 7.4-11.4%) ASD (assessed by ADOS-2 criteria): <28 weeks GA: 61/857, 7.1% (95%CI 5.5-9.0) 23-24 weeks GA: 26/173, 15% (95%CI 10-21.2) 25-26 weeks GA: 25/386, 6.5% (95%CI 4.2-9.4) 27 weeks GA: 10/298, 3.4% (95%CI 1.6-6.1)	Moderat e	

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			Diagnostic Observation Schedule, a semi-structured, observation protocol in which the examiner interacts with the child to assess social-communicative and repetitive behaviour symptoms.			
Specific learning difficulties						
Anderson 2003 Australia	Prospective regional cohort study (Victorian Infant Collaborative Study Group)	n=568 consecutive live births of neonates with BW <1000g or <28 weeks GA. n=298 infants survived to 2, and 5 years assessment. n=275 children assessed at 8 years age.	Educational progress was assessed using the Wide Range Achievement Test (WRAT3: reading, spelling, arithmetic) and the Comprehensive Scales of Student Abilities (CSSA, teacher assessed for verbal thinking, speech, reading, writing, handwriting, maths, general facts, basic motor generalisations, social behaviour). For WRAT3 major impairment represented a score <70.	At 8 years Educational progress (WRAT3 score <70, n=275) <28 weeks GA or ELBW: major reading impairment: 16/275, 5.8% (3.4-9.3%) <28 weeks GA or ELBW: major spelling impairment: 7/275, 2.54% (1.0-5.2%) <28 weeks GA or ELBW: major arithmetic impairment: 18/275, 6.6% (4.0-10.2%)	Low	Infants born 1991-1992, assessed at 8 years age.
Johnson 2011 UK and Ireland	National population-based cohort study (EPICure)	n=219 children assessed at 11 years (data missing for some	At 11 years, children were assessed at school by a paediatrician and psychologist blind to group allocation. Examiners received training in administration	At 11 years Learning impairment in reading (WIAT-II reading composite score <-2SD) <26 weeks GA: 64/212, 30.2% (24.1-36.9%) Learning impairment in mathematics (WIAT-II mathematics composite score <-2SD)	Low	Children born between March and December 1995, follow-

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		individuals in the outcomes of interest) (of n=307 survivors at 11 years, 71%)	of standardised tests and achieved a high criterion for inter-rater reliability (>95% agreement across test items) prior to commencing study assessments. Academic attainment was assessed using the Wechsler Individual Achievement Test-II (WIAT-II) from which standardised scores (mean=100, SD=15) were obtained for Word Reading, Reading Comprehension, Pseudo-word Decoding, Numerical Operations, Mathematical Reasoning, and the composite scales of Reading and Mathematics. For children in whom severe cognitive deficit precluded testing (n=18), a score 1-point below the basal score for the Reading and Mathematics composite scales was substituted. Learning impairment was classified as score <2SD below the mean of the comparison group of	<26 weeks GA: 94/215, 43.7% (37.0-50.6%)		up at 11 years of age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			term-born classmates on each scale.			
Joseph 2016b USA	Prospective cohort study (ELGAN)	N=1506 infants n=1198 survived to age 10 years n=873 assessed at 10 years	Academic achievement: The Wechsler Individual Achievement Test–III (WIATIII) 32 Word Reading, Pseudoword Decoding, and Spelling subtests were used to assess proficiency in word recognition, decoding, and spelling, respectively. WIAT-III Numeric Operations was used to assess math related computational skills.	At 10 years age Academic achievement (<28 weeks GA; <=-2SD) WIAT-III Word Reading: 122/873, 14% (95%CI 11.7-16.5) WIAT-III Pseudoword Decoding: 140/873, 16% (95%CI 13.7-18.6) WIAT-III Spelling: 122/873, 14% (95%CI 11.7-16.5) WIAT-III Numeric Operations 148/873, 17.0% (95%CI 14.5-19.6)	Low	
Kiechl-Kohlendorfer 2013 Austria	Prospective population-based cohort study.	N=303 (children live birth with gestational age <32 weeks) n=223 n=161 (children whose parents consented to take part in the study). n=153 assessed at 5 years age.	Delay in numerical skills was assessed individually with the TEDI-MATH which is a multi-componential dyscalculia test based on cognitive neuropsychological models of number processing and calculation [11]. The TEDI-MATH consists of several subtests designed for the assessment of pre-schoolers: In the counting principles subtest, children's mastery of the verbal counting sequence	At 5 years At 5 years age Specific learning difficulty (delayed numerical skills) (n=135) <32 weeks GA: 27/135, 20% (13.6-27.8%)	Low	Children born between 2003 and 2006, assessed at 5 years age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=135 assessed for numerical skills.	and its flexibility is tested (e.g. counting in steps of two, and counting backwards). Delay in numerical skills was defined as a Sum T-score <40.			
Evidence on developmental coordination disorder						
Agerholm 2011	Regional birth cohort study	N=237 live born children with 24-31 weeks GA in the geographical area N=204 children survived N=175 children followed-up at 5 years of age (86% of the ones who survived) N=168 children included in analysis (7 children with CP could not be assessed)	Motor function was examined using the Movement Assessment Battery for Children (M-ABC), it measures three items in the area of manual dexterity, two items in the area of ball skills and three items in the area of balance. The items were scored from 0 to 5, where 0 was the optimum score. A score under 5th percentile indicates motor function deficit.	At 5 years age Motor deficit (M-ABC <5th percentile total score) (disorder) 24-31 weeks GA: 30/168, 17.9% (12.4-24.5%)	Moderate	Children born 1996-2000, assessed at 5 years.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
de Kleine 2003	Prospective cohort study	n=566 eligible children n=431 assessed at 5 years (76%) n=404 assessed for motor functioning (M-ABC) n=402 assessed for IQ (IQ test) n=407 assessed for behavioural problems (CBCL)	At 5 years, motor function delay was assessed with the Movement ABC. Total scores above 17.0 (5th centile) were considered abnormal.	At 5 years Motor function delay (M-ABC <5th centile) <32 weeks GA/bw <1500 g: 90/404, 22.3% (18.3-26.7%)	Moderat e	Children 1992-1995, assessed at 5 years.
Foulder-Hughes 2003 UK	Geographicall y determined cohort study	n=280 children born at <32 weeks	DCD: Fine and motor gross skills were assessed using age band 2 of the Movement Assessment Battery for Children (MABC). The test comprises eight items, two in each of four subsections: manual dexterity, ball skills, static balance, and dynamic balance. The scoring system for each item ranges from 0 (no impairment) to 5 (severe impairment). The scores	At 7-8 years DCD <32 weeks GA: 86/280, 30.7% (25.4-36.5%)	Low	Geographic ally determined cohort study

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			for each item are added and converted to centiles. A score <=5th centile was taken to indicate motor difficulties consistent with DCD.			
Roberts 2011 Australia	Prospective cohort study (The Victorian Infant Collaborative Study Group)	EP/ELBW (22-27) (1997 cohort) n=201 survivors to 8 years age out of 283 consecutive live births. EP/ELBW (1991-1992) cohort n=298 survivors to 8 years age out of 533 consecutive live births.	DCD was defined as motor impairment in the absence of CP or an intellectual impairment. Motor impairment was determined by using the Movement Assessment Battery for Children carried out by a paediatrician. Moderate motor impairment was defined as a total score that was less than the 5th centile.	At 8 years age Moderate DCD (1997 cohort) 22-27 weeks GA: 21/132, 16% (10.1-23.3%) Moderate DCD (1991-1992 cohort) 22-27 weeks GA: 30/298, 10% (6.9% to 14.1%)	Low	Children born 1997 assessed at 8 years age. Children born 1991-1992 assessed at 8 years age.
Mental and behavioural disorder						
Johnson 2010 UK and Ireland	Population-based cohort study	n=219 children born at <26 weeks of GA were followed up at 11 years	The Development And Well Being Assessment (DAWBA), a structured psychiatric evaluation regarding children's development and behaviour was administered to parents via telephone interview (92%) or online (8%) from which information	Any emotional disorder <26 weeks GA: 18/201, 9.0% (5.4-13.8%) Separation anxiety <26 weeks GA: 5/201, 2.5% (0.8-5.7%) Specific phobia <26 weeks GA: 3/200, 1.5% (0.3-4.3%) Social phobia	Low	Children born 1995, follow-up at 11 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			required for assigning ICD-10 and DSM-IV-TR diagnoses of childhood psychiatric disorders was obtained. Supplemental information was provided by teachers who completed a corresponding questionnaire-based version of the DAWBA. Multi-informant data were collated by study assessors (paediatricians and psychologist), and potential cases were identified using computer-generated scoring algorithms (www.dawba.com). Summary sheets and clinical transcripts (with any reference to birth status removed) were then reviewed by two child and adolescent psychiatrists who had no prior knowledge of the children or their birth status and were therefore blind to group allocation, and who assigned DSM-IV and ICD-10 consensus diagnoses.	<26 weeks GA: 1/200, 0.5% (0.01-2.8%) Posttraumatic stress disorder <26 weeks GA: 1/200, 0.5% (0.01-2.8%) Generalized anxiety disorder <26 weeks GA: 4/201, 2.0% (0.5-5.0%) Childhood emotional disorder NOS <26 weeks GA: 1/200, 0.5% (0.01-2.8%) Major depression <26 weeks GA: 3/200, 1.5% (0.3-4.3%) Any conduct disorder <26 weeks GA: 12/219, 5.5% (2.9-9.4%) Oppositional defiant disorder <26 weeks GA: 11/219, 5.0% (2.5-8.8%) Conduct disorder <26 weeks GA: 1/219, 0.5% (0.01-2.5%)		

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Burnett 2014 Australia	Prospective regional cohort study.	n=215 early preterm/extremely low birth weight infants n=157 normal birth weight (>2499 g) controls n=372 in total	Any anxiety or mood disorder (All DSM-IV Axis I disorders (mood, anxiety, substance use, psychotic, eating and adjustment disorders) assessed with the Structured Clinical Interview for DSM-IV Disorders, Axis 1 Non-Patient version (SCIP-I/NP), administered by 5 interviewers blinded to group. Experienced consultant psychiatrists, also blinded by group, were consulted extensively and consensus diagnoses were reached for all participants. These assessments were supplemented by questionnaires examining recent anxiety and depression symptoms: the Beck Anxiety Inventory (BAI) and the Centre for Epidemiologic Studies Depression Scale -Revised (CESD-R).) Any mood disorder Any anxiety disorder Co-morbid anxiety and mood disorder.	At 18 years Any SCID-I/NP diagnosis (n=205) < 28 weeks GA/<1000g: 47/205, 23.0% (17.4-29.3%) Any anxiety or mood disorder (n=205) < 28 weeks GA/<1000g: 43/205, 21.0% (15.6-27.2%) Any mood disorder (n=205) < 28 weeks GA/<1000g: 33/205, 16.1% (11.4-22.0%) Major depressive disorder (n=205) < 28 weeks GA/<1000g: 28/205, 13.7% (9.3-19.1%) Any anxiety disorder (n=205) < 28 weeks GA/<1000g: 23/205, 11.2% (7.3-16.4%) Obsessive-compulsive disorder (n=205) < 28 weeks GA/<1000g: 4/205, 2.0% (0.5-5.0%) Co-morbid anxiety and mood disorder (n=205) < 28 weeks GA/<1000g: 13/205, 6.3% (3.4-10.6%) Psychotic disorders (n=205) < 28 weeks GA/<1000g: 0/0	Low	Adolescents born between 1991 and 1992, assessed at 18 years age.
Evidence on vision impairment						

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Anderson 2003	Prospective regional cohort study (Victorian Infant Collaborative Study Group)	N=568 consecutive live births of neonates with BW <1000g or <28 weeks GA. n=298 infants survived to 2, and 5 years assessment. n=275 children assessed at 8 years age.	No outcome measurement was reported.	At 8 years age Blindness 3/275, 1.1% (0.2-3.2%)	Low	Infants born 1991-1992, assessed at 8 years age.
Anderson 2011 Australia	Population-based cohort study	n=201 children survived to 8 years n=189 assessed at 8 years (94%)	Blindness was diagnosed by trained paediatricians who were blind to group membership (the study included a term-born control group).	At 8 years age (corrected) Blindness 22-27 weeks GA/BW 1000 g: 3/189, 1.6% (0.3-4.6%)	Low	Children born 1997, follow-up at 8 years of corrected age
Anonymous 1997	A geographically determined cohort study (Victoria, Australia)	n=401 live born children born at 23-27 weeks n=225 children survived to 2 years of age (56.1%)	Children were considered blind if visual acuity in both eyes was assessed as worse than 6/60.	At 2 years Blind 23-27 weeks GA: 5/219, 2.3% (0.8-5.3%)	Low	Children born 1991-1992, follow-up at 2 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=219 were assessed at 2 years (97.3% of the survivors)				
Bodeau-Livinec 2007 UK	Population based register study.	n=172 584 live births in 1994-1998.	Vision impairment was defined as visual acuity in the better eye of 6/18 or less with glasses or aids if worn (moderate impairment). Severe visual impairment or blindness was defined as visual acuity in the better eye of <6/60 or no useful vision	At 12 years Vision impairment (including moderate and severe impairment***) <28 weeks GA: 182.5 (102.5 to 299.1) 29-32 weeks GA: 37.1 (14.9 to 76.2) 33-36 weeks GA: 27.0 (17.3 to 40.1) **the data above refers to the number of cases per 10,000 livebirths.	Very low	Children born 1994-1998.
De Groote 2007 Belgium	Population-based geographically defined cohort study (EPIBEL	n=95 children that survived to discharge from NICU n=77 children assessed at 3 years (n=3 died before follow-up, n=12 parents did not give consent, n=3 could not be reached), 81% follow-up rate (84% of the ones	Vision impairment was classified as "impaired, but some useful vision", "impaired, and little useful vision", and "no useful vision".	At 3 years Vision impairment and little useful vision <27 weeks GA: 7/77, 9.1% (3.7-17.8%) Vision impairment, no useful vision <27 weeks GA: 2/77, 2.6% (0.3-9.1%)	Low	Children born in 1999-2000, follow-up at 3 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		who were alive at follow-up).				
Farooqi 2011 Sweden	Prospective national cohort study	n=89 children born at <26 weeks gestation and survived to follow-up (36% of all 247 children born at <26 weeks in Sweden of which the rest died) n=88 children with data (1 child was lost to follow-up, was followed-up but did not participate)	Severe visual impairment, including unilateral or bilateral blindness or visual acuity <20/200 without glasses in at least one eye.	At 11 years Severe visual impairment <26 weeks GA: 11/88, 12.5% (6.4-21.3%)	Low	Children born 1990-1992, follow-up at 11 years
Hellgren 2016 Sweden	National cohort study (EXPRESS)	N=494 EPT (22-26 weeks of gestation) infants alive at 1 year n=486 EPT infants surviving at	Monocular and binocular distance linear visual acuity with habitual correction was assessed at 3 m. The best measurable VA was 20/10. For VA, at least 4 of 5 optotypes had to be correctly identified. Based on results of	At 6.5 years Any visual impairment (best estimated visual acuity <20/40 at age 6 years and up in younger ages, adjusted for age) 22-23 weeks GA: 10/42, 23.8% (95%CI 12-40) 24 weeks GA: 11/82, 13.4% (95%CI 6.9-22.7) 25 weeks GA: 10/142, 7% (95%CI 3.4-12.6) 26 weeks GA: 7/138, 5.1% (95%CI 2.1-10.2)	Moderate	GA ascertainment was not reported in the study Children were born between 2004 and 2007

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		6.5 years age n=434 EPT infants included in the study	monocular VA, a better eye an a worse eye were identified in children with unequal VA, and the right eye was chosen as the better eye in the remaining children. Visual impairment was defined according to the WHO criteria: blindness was best VA <20/400, severe visual impairment was <20/60, moderate visual impairment was defined as <20/40 VA.	Visual impairment according to WHO criteria (Best-estimated visual acuity below 20/60 at age 6 years and up in younger ages adjusted for age) 22-23 weeks GA: 7/42, 16.7% (95%CI 7.0-31.4) 24 weeks GA: 6/82, 7.3% (95%CI 2.7-15.3) 25 weeks GA:5/142, 3.5% (95%CI 1.2-8.0) 26 weeks GA: 3/138, 2.2% (95%CI 0.4-6.2)		
Holmstrom 2014 Sweden	Prospective national cohort study (the Extremely Preterm Infants in Sweden Study EXPRESS)	n=491 eligible children (<27 weeks GA) n=411 (83.7% of the eligible sample) were assessed at 30 months' corrected age	Ophthalmologic examination was scheduled at 30 months (+-3 months) corrected age. Visual impairment: defined as blind or able to only fixate and follow a light binocularly. Three different test with gradually decreasing difficulty were used: 1) ability to identity single optotypes 0.4 Lea Hyvarinen test at 3 m distance, 2) ability to fixate and follow a toy of 5 cm at 30 cm, and 3) ability to fixate and follow a light/torch at 30 cm. Children or eyes that	At 30 months' corrected age Visual impairment (blind or able to only fixate and follow a light binocularly) <27 weeks GA: 12/390, 3.1% (1.6-5.3%) 22-23 weeks GA: 2/42, 4.8% (0.6-16.2%) 24 weeks GA: 4/70, 5.7% (1.6-14.0%) 25 weeks GA: 4/131, 3.1% (0.8-7.6%) 26 weeks GA: 2/147, 1.4% (0.2-4.8%)	Moderat e	Children born between April 1, 2004 and March 31, 2007, follow-up at 30 months' corrected age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
			were not able to identify an optotype at 3 m or a toy at 30 cm were considered to have impaired vision. Children or eyes that were not able to fixate and follow a light were considered to be blind.			
Hreinsdottir 2013 Sweden	Population based prospective study (Longitudinal Multidisciplinary Study of Visuo motor Capacity in Very Preterm Infants (LOVIS study))	n=98 (90% eligible for follow-up) (eleven children were lost to follow-up as n=6 refused to take part in the study, and n=5 had moved from the area) n=25 control group (recruited from the department of psychology and consisted of healthy normally developed term-born children (GA	At 2.5 years CA, children were examined by paediatric ophthalmologists and orthoptists and testing of spatial function was carried out by the same orthoptist. Best corrected visual acuity was assessed using the Lea single optotypes test at 3 metre distance. Ability to fixate and follow a small toy at 30 cm was investigated, as well as ability to fixate and follow a torch at 30 cm. Impaired vision was defined as blind or only able to fixate a torch.	At 2.5 years (corrected age) Impaired vision (blind or only able to fixate a torch) Best eye <32 weeks GA: 1/93, 1.1% (0.03-5.9%) Worst eye <32 weeks GA: 2/93, 2.2% (0.3-7.6%)	Low	Children born from 1 January 2005 to 31 December 2007, assessed at 2.5 years CA.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		38-42) in Uppsala county).				
Hutchinson 2013	Prospective cohort study (Victorian Infant Collaborative Study Group)	n=189 preterm/low birth weight cohort (94% eligible for follow-up; 12 children were not seen, but 10/12 were assessed at 2 years (corrected age)).	Assessment of blindness was not reported.	At 8 years age Blindness (n=189) EP/ELBW (mean GA26.5 (±2)): 3/189, 1.6% (0.3-5.0%)	Very low	Children born in 1997, assessed at 8 years age.
Joseph 2016b USA	Prospective cohort study (ELGAN)	N=1506 infants (<28 weeks of gestation) n=873 assessed at age 10 years	Severe visual impairment was defined as uncorrected functional blindness in both eyes	At 10 years Functional blindness: 22-27 weeks GA: 7/873, 0.8% (95%CI 0.3-1.7)	Low	Gestational age ascertainment was not reported Children born between 2002 and 2004
Larroque 2008 France	A longitudinal cohort study (EPIPAGE).	n=1817 children born at 22-32 weeks were followed at 5 years of age (77% of the population	Moderate and severe visual deficiency: Vision was assessed, without correction, with the Rossano test12 and visual deficiency classified as severe (<3/10 for both eyes), and moderate (<3/10 for	At 5 years Moderate to severe visual deficiency <33 weeks GA: 34/1697, 2.0% (1.4-2.8%) 24-25 weeks GA: 5/54, 9.3% (3.1-20.3%) 26 weeks GA: 6/60, 10.0% (3.8-20.5%) 27 weeks GA: 6/128, 4.7% (1.7-9.9%) 28 weeks GA: 4/165, 2.4% (0.7-6.1%)	Moderate	1997, follow-up at 5 years of age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		that survived)	one eye). Children born very preterm who did not take the Rossano test were classified according to information obtained from the medical questionnaire, interviews with parents, and medical sources	29 weeks GA: 6/178, 3.4% (1.3-7.2%) 30 weeks GA: 2/280, 0.7% (0.09-2.6%) 31 weeks GA: 8/348, 2.3% (1.0-4.5%) 32 weeks GA: 9/484, 1.9% (0.9-3.5%) <28 weeks GA: 17/242, 7.0% (4.1-11.0%) 28-31 week GA: 20/971, 2.1% (1.3-3.2%)		
Leversen 2010 Norway	Prospective observational nationally representative cohort study	n=373 children born 22-27 weeks GA or with birthweight 500-999 g who survived	Limited information provided. At 2 years a paediatrician completed forms developed for the study on somatic health and neurological status. They were not blinded. Children who missed the planned follow-up, data were collected in retrospect from the medical records if a routine follow-up had been performed within 1 year of planned evaluation, and from an additional structures telephone interview. Blindness meaning that the child was classified as legally blind.	At 2 years (corrected age) Blindness 22-27 weeks GA or bw 500-999 g: 6/373, 1.6% (0.6-3.5%)	Low	Children born in 1999-2000, follow-up at 2 years' corrected age
Leversen 2011 Norway	Prospective observational national cohort study	All infants born at 22-27 weeks of gestation or with birth	Vision impairment: registered from the clinical examination or previous examinations. All children in Norway	At 5 years Blindness 22-27 weeks GA or bw 500-999 g: 5/306, 1.6% (0.5-3.8%)	Moderate	Children born 1999 and 2000, follow-up at 5 years

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		weight between 500 and 999 g born in Norway in 1999 and 2000.	have a vision screen at the age of 4 years at the public health care clinics, using methods and standards according to national guidelines. Any significant deviation results in a referral to an ophthalmologist. Minor visual deficits were squints, myopia, hypermetropia, astigmatism, or other visual deficits requiring glasses. Severe visual impairment was not defined but the most severe visual impairment was classified as legal blindness.	23-25 weeks GA: 5/87, 5.8% (1.9-12.9%) 26-27 weeks GA: 0/152, 0% (0-2.4%) >27 weeks GA (bw <1000 g): 0/67, 0% (0-5.4%) Severe visual impairment 22-27 weeks GA or bw 500-999 g: 1/306, 0.3% (0.01-1.8%) 23-25 weeks GA: 1/87, 1.2% (0.03-6.2%) 26-27 weeks GA: 0/152, 0% (0-2.4%) >27 weeks GA (bw <1000 g): 0/67, 0% (0-5.4%)		
Marlow 2005 UK and Ireland	Population-based national cohort study (EPICure)	n=241 (82% of the eligible ones, n=293) (also n=160 term controls)	Vision impairment: Severe vision impairment was defined as blindness, moderate vision impairment was defined as impaired vision but ability to see.	At 6 years Blind <26 weeks GA: 6/241, 2.5% (0.9-5.3%) <=23 weeks GA: 2/24, 8.3% (1.0-27.0%) 24 weeks GA: 3/73, 4.1% (0.9-11.5%) 25 weeks GA: 1/144, 0.7% (0.02-3.8%) Moderate vision impairment (not blind) <26 weeks GA: 11/241, 4.6% (2.3-8.0%) <=23 weeks GA: 2/24, 8.3% (1.0-27.0%) 24 weeks GA: 5/73, 6.9% (2.3-15.3%) 25 weeks GA: 4/144, 2.8% (0.8-7.0%) Visually impaired or blind <26 weeks GA: 17/241, 7.1 (4.2-11.1%)	Moderat e	Children born 1995, follow-up at 6 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				<=23 weeks GA: 4/24, 16.7% (4.7-37.4%) 24 weeks GA: 8/73, 11.0% (4.9-20.5%) 25 weeks GA: 5/144, 3.5% (1.1-7.9%)		
Marret 2007 France	Population based prospective cohort (EPIPAGE).	n=1455	Visual impairment was defined as visual acuity less than 3/10 in one or both eyes.	At 5 years of age Visual deficiency 30 weeks GA: 2/280, 0.7% (0.1-2.6%) 31 weeks GA: 7/335, 2.2% (0.8-4.3%) 32 weeks GA: 9/484, 1.9% (0.9-3.5%) 33 weeks GA: 3/132, 2.3% (0.5-6.5%) 34 weeks GA: 1/134, 0.8% (0.02-4.1%) 30-31 weeks GA: 9/615, 1.5% (0.7-2.8%) 32-34 weeks GA: 13/750, 1.7% (0.9-3.0%)	Low	1997-2002. Cohort established in 1997. Follow-up at 5 years of age.
Moore 2012 UK	Prospective national cohort study (EPICure 2, this publication also used data from the original EPICure when comparing children born in 2006 to children born in 1995).	n=576 children born 22-26 weeks' gestation, assessed at follow-up (n=38 born at 22-23 weeks; n=98 born at 24 weeks; n=189 born at 25 weeks; n=251 born at 26 weeks)	Vision disability: Severe vision disability defined as blindness. Moderate vision disability defined as functionally impaired vision. The publication reports that a standard set of definitions was used to record visual functions.	Severe vision disability (blind) 22-26 weeks GA: 6/576, 1.0% (0.4-2.3%) 22-23 weeks GA: 1/38, 2.6% (0.1-13.8%) 24 weeks GA: 1/98, 1% (0.03-5.6%) 25 weeks GA: 1/189, 0.5% (0.01-2.9%) 26 weeks GA: 3/251, 1.2% (0.3-3.5%) Moderate vision disability (functionally impaired vision) 22-26 weeks GA: 34/576, 5.9% (4.1-8.2%) 22-23 weeks GA: 6/38, 15.8% (6.0-31.3%) 24 weeks GA: 8/98, 8.2% (3.6-15.5%) 25 weeks GA: 12/189, 6.4% (3.3-10.8%) 26 weeks GA: 8/251, 3.2% (1.4-6.2%) Moderate to severe vision disability 22-26 weeks GA: 40/576, 6.9% (5.0-9.3%) 22-23 weeks GA: 7/38, 18.4% (7.7-34.3%)	Low	Children born in 2006 (this publication also compared the children born in 2006 to children born in 1995).

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				24 weeks GA: 9/98, 9.2% (4.3-16.7%) 25 weeks GA: 13/189, 6.9% (3.7-11.5%) 26 weeks GA: 11/251, 4.4% (2.2-7.7%)		
Rieger-Fackeldey 2010 Germany	Prospective cohort study.	n=107 initial cohort n=27 survived at 5 years follow-up n=19 eligible for follow-up (8/27 were not able to be evaluated due to refusal of consent by parents (n=3), or family had moved away, failed appointment, or moved to another follow-up care (n=5))	Visual acuity after best possible correction for ametropia by refractive lenses of <20/200 was defined as blindness.	At 5 years age Visual impairment (blindness) ≥22 weeks GA/BW <501g: 2/19, 11% (1.3-33%)	Low	Children born between 1998 and 2001, assessed at 5 years age
Roberts 2010 Australia	A regional cohort study	n=223 total live births n=151 consecutive live births at 22-27 weeks completed gestation	Blindness was defined as visual acuity <6/60 in the better eye).	At 8 years (corrected) Blindness 22-27 weeks GA: 3/144, 2.1% (0.4-6.0%)	Low	Children born in 1997, follow-up at 8 years of age (corrected).

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=144 survived to age 8 years				
Serenius 2013 Sweden	Population-based prospective cohort study (EXPRESS group).	Sample recruited: n=707 live born preterm infants n=701 term controls Sample analysed after exclusions: n=456 preterm infants n=701 full term controls	Children unable to fixate and follow a light with either eye were considered bilaterally blind. Children registered at low vision centres without blindness were recorded as having moderate visual impairment.	At 2.5 years (corrected) Vision impairment <27 weeks GA: moderate: 13/456, 2.9% (1.5-4.8%) <27 weeks GA: blindness: 4/456, 0.9% (0.24-2.3%) <27 weeks GA: any vision impairment: 17/456, 3.7% (2.2-5.9%)	Moderat e	Children born between 2004 and 2007, assessed at 2.5 years corrected age.
Tommiska 2003 Finland	Prospective cohort study	n=208 extremely low birth weight infants (born with bw <1000 g) of which n=104 children were born at 22-26 weeks GA	A national neurological follow-up program included an ophthalmologic assessment at 12-18 months (corrected), and examinations by a neurologist, physiotherapist and speech therapist at the corrected age of 18 months. Bilateral blindness ("legally blind") and unilateral blindness (has lost vision in one eye).	At 12-18 months corrected age Bilateral blindness** The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 1/197, 0.5% (0.01-2.8%) Unilateral blindness** The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 2/197, 1.0% (0.1-3.6%) **Data available for 197 children	Low	Recruitment from 1st January 1996 to 31st December 1997, follow-up at 18 months of corrected age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Toome 2013	Population based national cohort study (follow-up study)	n=187 very low gestational age infants (83% eligible for follow-up 155/187) n=153 full term controls	Vision impairment defined as moderately reduced or blind	Vision impairment <32 weeks GA: 1/155, 0.64% (0.02-3.5%)	Low	Children born 2007, assessed at 2 years (corrected age).
Vohr 2005 USA	A multicentre cohort study	n=3785 infants included in analysis (51% of the original sample, 79.5% of the ones who survived up to discharge or 120 days)	Detailed interim medical history was obtained, blindness is defined as blind with no functional vision.	At 18-22 months corrected age Unilateral blindness Years 1993-94 22-26 weeks GA: 28/665, 4.2% (2.8-6.0%) 27-32 weeks GA: 9/444, 2.1% (0.9-3.8%) Years 1995-96 22-26 weeks GA: 18/716, 2.5% (1.5-3.9%) 27-32 weeks GA: 6/538, 1.1% (0.4-2.4%) Years 1997-98 22-26 weeks GA: 15/910, 1.6% (0.9-2.7%) 27-32 weeks GA: 4/512, 0.8% (0.2-2.0%) All epochs, 1993-1998 22-26 weeks GA: 61/2291, 2.7% (2.0-3.4%) 27-32 weeks GA: 19/1494, 1.3% (0.8-2.0%) 22-32 weeks GA: 80/3785, 2.1% (1.7-2.6%) Bilateral blindness Years 1993-94 22-26 weeks GA: 15/665, 2.3% (1.3-3.7%) 27-32 weeks GA: 6/444, 1.4% (0.5-2.9%) Years 1995-96 22-26 weeks GA: 11/716, 1.5% (0.8-2.7%)	Moderat e	1993-1998, follow-up at 18 to 22 months of corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
				27-32 weeks GA: 2/538, 0.4% (0.05-1.3%) Years 1997-98 22-26 weeks GA: 9/910, 1.0% (0.5-1.9%) 27-32 weeks GA: 2/512, 0.4% (0.05-1.4%) All epochs, 1993-1998 22-26 weeks GA: 35/2291, 1.5% (1.1-2.1%) 27-32 weeks GA: 10/1494, 0.7% (0.3-1.2%) 22-32 weeks GA: 45/3785, 1.2% (0.9-1.6%)		
Wood 2000 UK and Ireland	Population based prospective cohort study	N=4004 infants identified n=1185 survived at birth (843/1185 were admitted to NICU; 342/1185 died in the delivery room) n=283 assessed at follow-up	All children had clinical examination including detailed medical history obtained from semi-structured interview with family, and a neurologic assessment, classification of degree and type of disability, and functional classification of hearing and visual ability.	At median age 30 months. Vision impairment (severe disability, n=283) 22-25 weeks GA: blind or perceives light: 7/283, 2.5% (1-5%)	Low	Infants born 1995, assessed at median age 30 months
Evidence on hearing impairment						
Anderson 2003 Australia	Prospective regional cohort study (Victorian Infant Collaborative Study Group)	N=568 consecutive live births of neonates with BW <1000g or	Deafness was defined as needing hearing aids or worse.	At 8 years age Hearing impairment (requiring hearing aids) 4/275, 1.5% (0.4-3.7%)	Low	Infants born 1991-1992, assessed at 8 years age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		<28 weeks GA. n=298 infants survived to 2, and 5 years assessment. n=275 children assessed at 8 years age				
Anderson 2011 Australia	Population-based cohort study	n=201 children survived to 8 years n=189 assessed at 8 years (94%)	The children were assessed at 8 years (corrected) by psychologists blind to perinatal details, predominantly in specialised follow-up clinics, although a few were tested at school or home if they could not attend the clinics. Deafness was defined as needing hearing aids or worse.	At 8 years (corrected) Deafness 22-27 weeks GA/BW 1000 g: 4/189, 2.1% (0.6-5.3%)	Low	Children born 1997, follow-up at 8 years of corrected age.
Anonymous 1997 Australia	A geographically determined cohort study (Victoria, Australia)	n=401 liveborn children born at 23-27 weeks n=225 children survived to 2	Children were usually screened for major hearing loss earlier at 7-8 months of corrected age by distraction testing with calibrated noise makers. Those who had not been screened, or those with suspected deafness at 2 years of age were	At 2 years Deaf 23-27 weeks GA: 2/219, 0.9% (0.1-3.3%)	Low	Children born 1991-1992, follow-up at 2 years of age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		years of age (56.1%) n=219 were assessed at 2 years (97.3% of the survivors)	referred again for audiological assessment			
De Groote 2007 Belgium	Population-based geographically defined cohort study (EPIBEL)	n=95 children that survived to discharge from NICU n=77 children assessed at 3 years (n=3 died before follow-up, n=12 parents did not give consent, n=3 could not be reached), 81% follow-up rate (84% of the ones who were alive at follow-up).	Hearing impairment was classified as "no useful hearing", "impairment but useful hearing", and "hearing aids".	At 3 years Hearing impairment but useful hearing <27 weeks GA: 3/77, 3.9% (0.8-11.0%) Hearing impairment, no useful hearing <27 weeks GA: 0/77, 0% (0-4.7%) Hearing impairment, use of hearing aids <27 weeks GA: 4/77, 5.2% (1.4-12.8%)	Low	Children born in 1999-2000, follow-up at 3 years of age
Doyle 2011 Australia	A population-based cohort study (in the State of Victoria).	n=257 live births with bw 500-999 g (excl. cases with	Deafness was defines as requiring hearing aids or more advanced requirements.	At 2 years (corrected age) Deafness BW 500-999 g (mean GA 25.7 [SD 2.3]): 4/165, 2.4% (0.7-6.1%)	Moderate	Children born 2005, follow-up at 2 years (corrected age)

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		lethal anomalies) n=172 survived to 2 years n=165 assessed at 2 years (96%)				
Farooqi 2011 Sweden	Prospective national cohort study	n=89 children born at <26 weeks gestation and survived to follow-up (36% of all 247 children born at <26 weeks in Sweden of which the rest died) n=88 children with data (1 child was lost to follow-up, was followed-up but did not participate)	Moderate, severe or profound hearing loss in both ears resulting in amplification.	At 11 years Moderate, severe or profound hearing loss in both ears requiring amplification <26 weeks GA: 5/88, 5.7% (1.9-12.8%)	Low	Children born 1990-1992, follow-up at 11 years
Hutchinson 2013 Australia	Prospective cohort study (Victorian Infant	n=189 preterm/low birth weight cohort (94%	Definitions of measurement of	At 8 years age Hearing impairment (requiring hearing aids, n=189) EP/ELBW: 4/189, 2.1% (0.6-5.3%)	Very low	Children born in 1997,

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
	Collaborative Study Group)	eligible for follow-up; 12 children were not seen, but 10/12 were assessed at 2 years (corrected age)).	deafness was not reported in the study.	26.5 (+/-2)		assessed at 8 years age.
Larroque 2008 France	Longitudinal cohort study (EPIPAGE).	n=1817 children born at 22-32 weeks were followed at 5 years of age (77% of the population that survived) n=1812 children born at 22-32 weeks with data on CP outcome n=1534 children born at 22-32 weeks with data on MPC score outcome	Severe auditory deficiency: Severe auditory deficit was defined as a hearing loss of more than 70 decibel (dB) for one or both ears, or the use of a hearing aid (reported in the medical questionnaire).	At 5 years Severe hearing deficiency <33 weeks GA: 8/1784, 0.45% (0.2-0.9%) 24-25 weeks GA: 1/58, 1.7% (0.04-9.2%) 26 weeks GA: 1/71, 1.4% (0.04-7.6%) 27 weeks GA: 0/132, 0% (0-2.8%) 28 weeks GA: 2/174, 1.2% (0.1-4.1%) 29 weeks GA: 1/185, 0.5% (0.01-3.0%) 30 weeks GA: 1/285, 0.4% (0.01-1.9%) 31 weeks GA: 1/376, 0.3% (0.01-1.5%) 32 weeks GA: 1/503, 0.2% (0.01-1.1%) <28 weeks GA: 2/261, 0.8% (0.1-2.7%%) 28-31 week GA: 5/1020, 0.5% (0.2-1.1%)	Moderat e	1997, follow-up at 5 years of age
Leversen 2010	Prospective observational nationally	n=373 children born 22-27 weeks	Complete deafness, not further defined.	At 2 years (corrected age) Deafness	Low	Children born in 1999-2000,

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
Norway	representative cohort study	GA or with birthweight 500-999 g who survived		22-27 weeks GA or bw 500-999 g: 3/373, 0.8% (0.2-2.3%)		follow-up at 2 years' corrected age
Marlow 2005 UK and Ireland	Population-based national cohort study (EPICure)	n=241 (82% of the eligible ones, n=293) (also n=160 term controls)	Hearing impairment: Severe hearing impairment was defined as profound sensorineural hearing loss, moderate hearing loss was defined as sensorineural hearing loss corrected with hearing aids.	At 6 years Moderate hearing impairment (use of hearing aids) <26 weeks GA: 7/241, 2.9% (1.2-5.9%) <=23 weeks GA: 0/24, 0% (0-14.3%) 24 weeks GA: 2/73, 2.7% (0.3-9.6%) 25 weeks GA: 5/144, 3.5% (1.1-7.9%) Moderate to severe hearing impairment <26 weeks GA: 14/241, 5.8% (3.2-9.6%) <=23 weeks GA: 1/24, 4.2% (0.1-21.1%) 24 weeks GA: 6/73, 8.2% (3.1-17.0%) 25 weeks GA: 7/144, 4.9% (2.0-9.8%)	Moderat e	Children born 1995, follow-up at 6 years of age.
Marret 2007 France	1997-2002. Cohort established in 1997. Follow-up at 5 years of age.	n=1455	Hearing impairment was defined as loss of more than 70 decibels or use of hearing aid in one or both ears.	At 5 years of age Hearing deficiency 30 weeks GA: 1/285, 0.3% (0.01-1.9%) 31 weeks GA: 1/376, 0.3% (0.01-1.5%) 32 weeks GA: 1/503, 0.2% (0.01-1.1%)10.3 33 weeks GA: 0/130, 0% 34 weeks GA: 2/135, 1.5% (0.2-5.3%) 30-31 weeks GA: 2/661, 0.3% (0.04-1.1%) 32-34 weeks GA: 3/768, 0.4% (0.1-1.1%)	Low	1997-2002. Cohort established in 1997. Follow-up at 5 years of age.
Moore 2012 UK	Prospective national cohort study (EPICure 2, this publication	n=576 children born 22-26 weeks' gestation, assessed at follow-up	Hearing disability: Severe hearing disability defined as profound sensorineural hearing loss not improved by aids. Moderate hearing	At 3 years (generally, some assessments delayed) Severe hearing disability (profound hearing loss not improved with aids) 22-26 weeks GA: 1/576, 0.2% (0-1.0%) 22-23 weeks GA: 1/38, 2.6% (0.1-13.8%)	Low	Children born in 2006 (this publication also compared

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
	also used data from the original EPICure when comparing children born in 2006 to children born in 1995).	(n=38 born at 22-23 weeks; n=98 born at 24 weeks; n=189 born at 25 weeks; n=251 born at 26 weeks)	disability defined as hearing loss improved by aids. The publication reports that a standard set of definitions was used to record auditory functions.	24 weeks GA: 0/98, 0% (0-3.7%) 25 weeks GA: 0/189, 0% (0-1.9%) 26 weeks GA: 0/251, 0% (0-1.5%) Moderate hearing disability (hearing loss improved with aids) 22-26 weeks GA: 30/576, 5.2% (3.5-7.4%) 22-23 weeks GA: 2/38, 5.3% (0.6-17.8%) 24 weeks GA: 5/98, 5.1% (1.7-11.5%) 25 weeks GA: 10/189, 5.3% (2.6-9.5%) 26 weeks GA: 13/251, 5.2% (2.8-8.7%) Moderate to severe hearing disability 22-26 weeks GA: 31/576, 5.4% (3.7-7.6%) 22-23 weeks GA: 3/38, 7.9% (1.7-21.4%) 24 weeks GA: 5/98, 5.1% (1.7-11.5%) 25 weeks GA: 10/189, 5.3% (2.6-9.5%) 26 weeks GA: 13/251, 5.2% (2.8-8.7%)		the children born in 2006 to children born in 1995).
Rieger-Fackeldey 2010 Germany	Prospective cohort study	n=107 initial cohort n=27 survived at 5 years follow-up n=19 eligible for follow-up (8/27 were not able to be evaluated due to refusal of consent by parents	Severe hearing disability was defined when a hearing aid for one or both ears was necessary.	At 5 years age Hearing impairment (requiring hearing aid) ≥22 weeks GA/BW <501g: 2/19, 11% (1.3-33%)	Low	Children born between 1998 and 2001, assessed at 5 years age

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		(n=3), or family had moved away, failed appointment, or moved to another follow-up care (n=5))				
Roberts 2010 Australia	A regional cohort study	n=223 total live births n=151 consecutive live births at 22-27 weeks completed gestation n=144 survived to age 8 years	Severe hearing impairment was defined as requiring hearing aids or worse). No details about how it was assessed.	At 8 years (corrected) Hearing impairment 22-27 weeks GA: 3/144, 2.1% (0.4-6.0%)	Low	Children born in 1997, follow-up at 8 years of age (corrected).
Serenius 2013 Sweden	Population-based prospective cohort study (EXPRESS group).	Sample recruited: n=707 liveborn preterm infants n=701 term controls Sample analysed after exclusions: n=456 preterm infants	Moderate auditory impairment was defined as hearing loss corrected with an aid and severe hearing impairment was defined as hearing loss that could not be corrected with hearing aids (deafness).	At 2.5 years corrected age Hearing impairment <27 weeks GA: impaired hearing, corrected with hearing aid: 3/456, 0.7% (0.14-2.0%) <27 weeks GA: deaf: 1/456, 0.2% (0.01-1.2%) <27 weeks GA: any hearing impairment: 4/456, 0.9% (0.24-2.2%)	Moderat e	Children born between 2004 and 2007, assessed at 2.5 years corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		n=701 full term controls				
Tommiska 2003 Finland	Prospective cohort study	n=208 extremely low birth weight infants (born with bw <1000 g) of which n=104 children were born at 22-26 weeks GA	Hearing impairment defined as necessitating hearing rehabilitation or the use of a hearing aid.	At 18 months corrected age Hearing impairment* The whole cohort of children born <1000 g (mean GA 27.3 with range 22.3-34.9): 6/195, 3.1% (1.1-6.6%) *Data available for 195 children.	Low	Recruitment from 1st January 1996 to 31st December 1997, follow-up at 18 months of corrected age
Toome 2012 Estonia	Population based national cohort study (follow-up study)	n=187 very low gestational age infants (83% eligible for follow-up 155/187) n=153 full term controls	Hearing impairment was defined as hearing aids or deafness;	At 2 years (corrected age) Hearing impairment <32 weeks GA: 2/155, 1% (0.16-4.6%)	Low	Children born 2007, assessed at 2 years (corrected age).
Vohr 2005 USA	A multicentre cohort study	n=3785 infants included in analysis (51% of the original sample, 79.5% of the ones who survived up	Permanent hearing loss is defined as a hearing loss requiring amplification in both ears	At 18-22 months corrected age Permanent hearing loss Years 1993-94 22-26 weeks GA: 23/665, 3.4% (2.2-5.1%) 27-32 weeks GA: 8/444, 1.7% (0.8-3.5%) Years 1995-96 22-26 weeks GA: 16/716, 2.3% (1.3-3.6%) 27-32 weeks GA: 4/538, 0.8% (0.2-1.9%) Years 1997-98 22-26 weeks GA: 16/910, 1.8% (1.0-2.8%)	Moderat e	1993-1998, follow-up at 18 to 22 months of corrected age.

Study	Data source	Sample and population studied	Measurement of outcome	Outcome (Prevalence n/N and % (with 95% CI) (incl. GA at birth and age at assessment)	Study quality	Comments
		to discharge or 120 days)		27-32 weeks GA: 9/512, 1.8% (0.8-3.3%) All epochs, 1993-1998 22-26 weeks GA: 55/2291, 2.4% (1.8-3.1%) 27-32 weeks GA: 21/1494, 1.4% (0.9-2.1%) 22-32 weeks GA: 76/3785, 2.0% (1.6-2.5%)		
Wood 2000 UK and Ireland		N=4004 infants identified n=1185 survived at birth (843/1185 were admitted to NICU; 342/1185 died in the delivery room) n=283 assessed at follow-up	All children had clinical examination including detailed medical history obtained from semi-structured interview with family, and a neurologic assessment, classification of degree and type of disability, and functional classification of hearing ability.	At median age 30 months. Hearing impairment (severe disability, n=283) 22-25 weeks GA: impaired, corrected with hearing aid:3/283, 1.1% (0.2-3.1%) 22-25 weeks GA: impaired, uncorrected even with hearing aid: 5/283, 1.8% (0.58-4.1%)	Low	Infants born 1995, assessed at median age 30 months.

4.1.4.3 Economic evidence

No health economic search was undertaken for this review question and consequently no evidence was found. This question focused on the prevalence of various developmental problems rather than whether any strategy for the management of these problems represents a cost-effective use of resources. Therefore, this question is not primarily about competing alternatives which have different opportunity costs and therefore was not considered suitable for a health economic review

4.1.4.4 Evidence statements

4.1.4.4.1 Prevalence of cerebral palsy

Cerebral palsy ≤ 28 completed weeks of gestation

Any cerebral palsy

Moderate to low quality evidence from four studies (sample size ranged from 141 to 373) showed that among children born at 22-27 weeks GA the prevalence of any CP varied from 7% (95% CI 4.6 to 10.10) to 11.3% (95%CI: 6.6 to 17.8) at 2 years (corrected age), 5 years and 8 years (corrected) (Leversen 2010; Leversen 2011; Roberts 2011; Anderson 2011).

Moderate quality evidence from four studies (sample size ranged from 75 to 244) showed that among children born at <27 weeks GA the prevalence of any CP varied from 14.7% (95%CI 7.6 to 24.7% to 24.7% (95%CI 15.6 to 35.8%) at age range 12 months CA to 9 years (Mikkola 2005; Stahlmann 2009; Sutton 1999; De Groote 2007).

Moderate to low quality evidence from four studies (sample size ranged from 275 to 331,154) showed that among children born at <28 weeks GA the prevalence of any CP varied from 6.7% (95%CI 5.1 to 8.6) to 16.6% (95%CI 12.5 to 21.3) (Larroque 2008; Ancel 2006; Glinianaia 2011; Anderson 2003).

Moderate quality evidence from one study (sample size 1718) showed that among children born at 24-27 weeks GA the prevalence of any CP was 14.7% (95%CI 10.6-19.5%) at 5 years age (Foix-Helias 2008).

Low quality evidence from one study (sample size 104) showed that among children born at 22-26 weeks GA the prevalence of any CP was 11.5% (95% CI 6.1-19.3%) at 18 months CA (Tommiska 2003)

Low quality evidence from one study (sample size 283) showed that among children born 22-25 weeks GA the prevalence of any CP was 17.7% (95% CI 13.4-22.6%) at a median age of 30 months (Wood 2000).

Moderate to very low quality evidence from three studies (sample size ranged from 19 to 189) showed that among children born at a mean GA range of 25.4 (±1) to 26.5 (±2) weeks the prevalence of any CP was 7.3% (95% CI 3.8-12.4%) to 37% (95%CI 16-62%) at age 2 years to 8 years (Hutchinson 2013; Doyle 2011; Rieger-Fackeldey 2010).

Low quality evidence from one study (sample size 219) showed that among children born at 23-27 weeks GA the prevalence of any CP was 11% (95%CI 7.2-15.9%) at 2 years age (Anon 1997).

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of CP was 19.0% (95%CI 12.9 to 26.5%) at 4 years age (Salakorpi 2001).

Mild cerebral palsy

Moderate to low quality evidence from two studies (sample size 77 to 456) showed that among children born at <27 weeks GA the prevalence of mild CP across the two studies (10.4% (95%CI 4.6 to 19.5) and 2.9% (95% CI 1.5 to 4.8)) at 2.5 years CA and 3 years age (De Groote 2007; Serenius 2013).

Moderate cerebral palsy

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks the prevalence was 7.1% (95%CI 4.2 to 11.1) at 6 years (Marlow 2005). The prevalence was varied in two studies of moderate to low quality in children (sample size 456 to 576) born at <27 weeks GA (2.6% (95%CI 1.5 to 4.3)) and 2.9% (95%CI 1.5 to 4.8)) (Moore 2012; Serenius 2013), whereas prevalence of CP was 13% (95% CI 6.4 to 22.6) in one study (at <27 weeks GA) (De Groote 2007).

Moderate to severe cerebral palsy

Moderate to low quality evidence from two studies (sample size 88 to 241) showed that among children born at <26 weeks GA the prevalence of CP (moderate/disabling or both ambulatory/non-ambulatory) was varied, with a prevalence of 6.8% (95%CI 2.5 to 14.3) at 11 years (Farooqi 2011) and 13.3% (95%CI 9.3 to 18.2) at 6 years (Marlow 2005). There was also variation of prevalence of moderate to severe CP in children born at <27 weeks GA at 2.5 years corrected age (4.2% (95%CI 2.5 to 6.4)) and at 3 years (7.8% (95%CI 5.8 to 10.3)) in two studies of moderate and low quality (Serenius 2013; Moore 2012).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-26 weeks GA the prevalence of moderate to severe CP (non-ambulatory or needing aids) was 11% (95%CI 9.8 to 12.4) at 18-22 months corrected age (Vohr 2005).

Severe cerebral palsy

Moderate to low quality evidence from two studies (sample size ranging from 77 to 456) showed that among children born at <27 weeks GA the prevalence of severe CP was 1.3% (95%CI 0.03 to 7%) at age 2.5 years CA to 3 years (Serenius 2013; De Groote 2007).

Moderate quality evidence from two studies (sample size ranging from 75 to 241) showed that among children born at <26 weeks and <27 weeks GA the prevalence of non-ambulatory CP was 6.2% (95%CI 3.5 to 7.4%) at 6 years age (Marlow 2005), and 10.7% (95%CI 4.7 to 19.9%) at 7-9 years age (Stahlmann 2009).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence for severe CP (GMFCS level 3-5) was 5.2% (95% CI 3.5-7.4%) at 3 years age (Moore 2012). Moderate quality evidence from one study (sample size 306) showed that among children born at 22-27 weeks GA the prevalence for severe CP (GMFCS level4-5) was 3.3% (95%CI 1.6-5.9%) at 5 years age (Leversen 2011).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe diplegia was 4.2 % (95%CI 2.2 to 7.3), severe hemiplegia was 0.4% (95%CI 0.01 to 2), and severe quadriplegia was 3.9% (95%CI 2 to 6.9) at 30 months corrected age (Wood 2000).

- Low quality evidence from one study (sample size 1718) showed that among children born at 24-27 weeks GA the prevalence for severe CP (unable to walk or only with aids) was 4.9% (95% CI 2.6 to 8.2%) at 5 years age (Foix-Helias 2008).
- Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence for severe motor impairment (GMFCS level 5, no self -mobility) was 1.9% (95%CI 1.1-3.1) at 10 years age (Joseph 2016b).

28-31 completed weeks of gestation

Any cerebral palsy

Moderate to low quality evidence from three studies (sample size varied from 1812 to 331,154) showed that among children born at 28-31 weeks the prevalence of any CP was varied, ranging from 5.9% (95%CI 4.9 to 7) to 9.5% (95%CI 7.8 to 11.4) across the three studies at 2-8 years (Larroque 2008; Ancel 2006; Glinianaia 2011).

Moderate to low quality evidence from two studies (sample ranged from 1455 to 1781) showed that among children born at 28-32 or 30-31 weeks, there was no difference in prevalence (7.7% (95%CI 5.8 to 9.9) and 7.9% (95%CI 6.6 to 9.3)) at 5 years (Marret 2007; Foix-Helias 2008). However, moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of CP was higher (11.6% (95%CI 10 to 13.3) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-32 weeks GA the prevalence of CP was 16% (95%CI 14.9 to 17.2) at 18-22 months corrected age (Vohr 2005). However, the prevalence of CP was lower (4.3% (95%CI 2.2 to 7.5)) in low quality evidence from one study (sample size 259) among children born at 23-32 weeks GA (Andrews 2008). There was minimal difference in prevalence in GA groups including 24-32 weeks (prevalence 8.9% (95%CI 7.6 to 10.3)) (Foix-Helias 2008), 25-32 weeks GA (prevalence 13.2 (95%CI 8.4 to 19.3)) (Burguet 1999), or <31(prevalence 16% (95%CI 14.9 to 17.2)) , <32 (prevalence 11% (95%CI 6.5 to 17), or <33 weeks GA (prevalence 8.8% (95%CI 7.5 to 10.2)) (Vincer 2014; Toome 2012; Larroque 2008).

Mild cerebral palsy

Low quality evidence from one study (sample size 801) showed that among children born at <31 weeks GA the prevalence of 6.7% (95%CI 5.1 to 8.7) for mild CP (GMFCS level1) at 12-42 months corrected age (Vincer 2014).

Moderate to severe cerebral palsy

Low quality evidence from one (sample size 801) showed that among children born at <31 weeks GA the prevalence of moderate to severe CP (GMFCS level 2-5) was 3.4% (95%CI 2.2-4.9%) at 12-42 months corrected age (Vincer 2014).

Low quality evidence from one study (sample size 155) showed that among children born at <32 weeks GA the prevalence of moderate to severe CP (GMFCS level 2-5) was 8.4% (95%CI 4.5-13.9%) at 2 years CA (Toome 2012).

Low quality evidence from one study (1455) showed that among children born at 30-31 weeks GA the prevalence of 5.7% (95%CI 4.1 to 7.7) for moderate to severe CP (bilateral spastic CP) at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence for moderate to severe CP (non-ambulatory or needing aids) was 7% (95%CI 5.8 to 8.4) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-32 weeks GA the prevalence of moderate to severe CP (non-ambulatory or needing aids) was 9.4% (95%CI 8.5-10.4%) at 18-22 months corrected age (Vohr 2005).

Severe cerebral palsy

Moderate quality evidence from one study (sample size 1781) showed that among children born at 28-32 weeks GA the prevalence of severe CP (unable to walk or only with aids) was

2.4% (95%CI 1.7 to 3.4) at 5 years (Foix-Helias 2008). In the same study, the prevalence at 24-32 weeks was 2.8% (95%CI 2.1 to 3.7).

32-36 completed weeks of gestation

Any cerebral palsy

Low quality evidence from one study (sample size 1455) showed that among children born at 32-34 weeks GA the prevalence of any CP type was 3.4% (95%CI 2.3 to 5) at 5 years (Marret 2007).

Moderate to low quality evidence from three studies (sample size range from 741 to 331,154)) showed that among children born at 32-26 weeks GA the prevalence of any CP was similar (range from 0.8% (95%CI 0.7 to 0.9) to 1% (95%CI 0.8 to 1.1) across the studies at age up to 7 or 8 years (Odd 2013; Hirvonen 2014; Glinianaia 2011).

Moderate to severe cerebral palsy

Low quality evidence from one study (sample size 1455) showed that among children born at 32-34 weeks GA the prevalence of CP (bilateral spastic CP) was 2.2% (95% CI 1.3 to 3.5) at 5 years (Marret 2007).

Moderate quality from one study (sample size 53,078) showed that among children born at 32-36 weeks GA found the prevalence of CP (other types) was 0.35% (95%CI 0.3 to 0.4) at up to 7 years (Hirvonen 2014).

Low quality evidence from one study (sample size 331,154) showed that among children born at <37 weeks GA the prevalence of spastic-bilateral or unilateral CP was 1.3% (95%CI 1.1 to 1.5) and 0.4% (95%CI 0.3 to 0.5) respectively at up to 8 years (Glinianaia 2011).

Low quality evidence from one study (sample size 104) showed that among children born at 22.3-34.9 weeks GA/bw <1000g the prevalence of CP (ataxia/athetosis) was 1% (95%CI 0.1 to 3.4) at 18 months corrected age (Tommiska 2003).

4.1.4.4.2 Small for gestational age

Low quality evidence from one study (sample size 2357) showed that among children born at 24-28 weeks GA and small for gestational age, the prevalence of any CP was 18% (95%CI 5.2-40.3%). In the same study, the prevalence was 3.2% (95%CI 0.9-8%) at 5 years age (Guellec 2011).

4.1.4.4.3 Hemiplegia

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of hemiplegia was 1.8% (95%CI 0.6-4.1%) at median 30 months (Wood 2000). In the same study, the prevalence of severe hemiplegia was 0.4% (95%CI 0.01-2%).

Very low quality evidence from one study (sample size 167) showed that among children born at 25-32 weeks GA the prevalence of hemiplegia was 1.2% (95%CI 0.2-4.3%) at 2 years (corrected age) (Burguet 1999).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of hemiplegia was 3.9% (95%CI 0.8-11%) at 3 years age (De Groote 2007).

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of hemiplegia was 5.6% (95%CI 2.5 to 10.8%) at 4 years (Salakorpi 2001).

Low quality evidence from one study (sample size 1455) showed that among children born at gestational age ranging from 30 to 33 weeks the prevalence of hemiplegia ranged from 0.4% to 0.8% (95%CI range 0.01 – 4.1%) at 5 years age (Marret 2007).

Moderate quality evidence from one study (sample size 53,078) showed that among children born at <32 weeks GA the prevalence of hemiplegia was 1.3 % (95%CI 1-1.6%) at age up to 7 years (Hirvonen 2014). In the same study the prevalence of hemiplegia CP was 0.5% (95%CI 0.4-0.8%) at 32-33 weeks GA, 0.14% (95%CI 0.11-0.19%) at 34-36 weeks GA, and 0.2% (95%CI 0.16-0.25%) at 32-26 weeks GA (Hirvonen 2014).

4.1.4.4.4 *Diplegia*

Low quality evidence from one study (sample size 104) showed that among children born at 22.3 to 34.9 weeks GA the prevalence of diplegia was 7.2% (95%CI 4.1-11.6%) at 18 months corrected age (Tommiska 2003).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of diplegia was 9.5% (95%CI 6.4-13.6 %) at median 30 months (Wood 2000). In the same study, the prevalence of severe diplegia was 4.2% (95%CI 2.2-7.3%).

Very low quality evidence from one study (sample size 167) showed that among children born at 25-32 weeks GA the prevalence of spastic diplegia was 6% (95%CI 2.9-10.7%) at 2 years (corrected age) (Burguet 1999).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of diparesis was 11.7% (95%CI 5.5-21%) at 3 years age (De Groote 2007).

Low quality evidence from one study (sample size 155) showed that among children born at <32 weeks GA the prevalence of spastic diplegia was 4.5% (95%CI1.8-9.1%) at 2 years (corrected age) (Toome 2012).

Moderate quality evidence from one study (sample size 53,078) showed that among children born at <32 weeks GA the prevalence of diplegia was 3.4 % (95%CI 2.9-3.8%) at age up to 7 years (Hirvonen 2014). In the same study the prevalence of diplegia CP was 0.7% (95%CI 0.5-0.9%) at 32-33 weeks GA, 0.13% (95%CI 0.10-0.17%) at 34-36 weeks GA, and 0.2% (95%CI 0.17-0.26%) at 32-26 weeks GA (Hirvonen 2014).

4.1.4.4.5 *Triplegia*

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of triparesis was 2.6% (95%CI 0.3-9.1%) at 3 years age (De Groote 2007).

4.1.4.4.6 *Diplegia or tetraplegia*

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of bilateral spastic CP (diplegia or tetraplegia) was 10.6% (6.0 to 16.8%) at 4 years (Salakorpi 2001).

4.1.4.4.7 *Tetraplegia*

Low quality evidence from one study (sample size 104) showed that among children born at 22.3 to 34.9 weeks GA the prevalence of tetraplegia was 1.9% (95%CI 0.5-4.9%) at 18 months corrected age (Tommiska 2003).

Very low quality evidence from one study (sample size 167) showed that among children born at 25-32 weeks GA the prevalence of tetraplegia was 1.2% (95%CI 0.2-4.3%) at 2 years (corrected age) (Burguet 1999).

4.1.4.4.8 Quadriplegia

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of quadriplegia was 4.2% (95%CI 2.2-7.3 %) at median 30 months (Wood 2000). In the same study, the prevalence of severe quadriplegia was 3.9% (95%CI 2.0-6.9%).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of quadriplegia was 5.2% (95%CI 1.4-12.8%) at 3 years age (De Groote 2007).

Moderate quality evidence from one study (sample size 53,078) showed that among children born at <32 weeks GA the prevalence of quadriplegia was 0.6 % (95%CI 0.4-0.8%) at age up to 7 years (Hirvonen 2014). In the same study the prevalence of quadriplegia was 0.2% (95%CI 0.1-0.3%) at 32-33 weeks GA, 0.04% (95%CI 0.02-0.06%) at 34-36 weeks GA, and 0.06% (95%CI 0.04-0.08%) at 32-26 weeks GA (Hirvonen 2014).

4.1.4.4.9 Dystonic or athetoid type

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of dystonic or athetoid CP was 2.8% (95%CI 0.8 to 7.1%) at 4 years (Salakorpi 2001).

4.1.4.4.10 Prevalence of cerebral palsy by week of gestational age

Any cerebral palsy

Low quality evidence from one study (sample size 244) showed that among children born at 23 weeks GA the prevalence of any CP was 100% (95%CI 25 to 100%) at 12 months corrected age. However, the prevalence was 19.10% (95%CI 12 to 27.9%) for children who were born at 27 weeks GA (Sutton 1999).

Low quality evidence from one study (sample size 104) showed that among children born at 22-23 weeks GA the prevalence of any CP was 20% (95%CI 0.5 to 71.6%) compared to a prevalence of 10.6% (95%CI 3.6 to 23.10%) in children who were born at 26 weeks GA, assessed at the age of 18 months corrected age (Tommiska 2003).

Low quality evidence from one study (sample size 1954) showed that among children born at 24-25 weeks GA the prevalence of any CP was 19.4% (95%CI 10.4 to 31.4%) compared to a prevalence of 4.4% (95%CI 2.9 to 6.6%) in children who were born at 32 weeks GA, assessed at the age of 2 years (Ancel 2006).

Moderate quality evidence from one study (sample size 1812) showed that among children born at 24-25 weeks GA the prevalence of any CP was 18.3% (95%CI 9.5 to 30.4%) compared to a prevalence of 4.1% (95%CI 2.6 to 6.2%) in children who were born at 32 weeks GA, assessed at the age of 5 years (Larroque 2008).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of any CP was 6.3% (95%CI 3.8 to 9.7%) compared to a prevalence of 3.7% (95%CI 1.2 to 8.4%) in children who were born at 34 weeks GA, assessed at the age of 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 6347) showed that among children born at <32 weeks GA the prevalence of any CP was 8.7% (95%CI 8.0 to 9.4%) compared to a prevalence of 0.56% (95%CI 0.49 to 0.64%) in children born at 34-36 weeks GA, assessed at up to the age of 7 years (Hirvonen 2014).

Moderate cerebral palsy

Low quality evidence from one study (sample size 576) showed that among children born at 24 weeks the prevalence of moderate CP was 4.1% (95%CI 1.1 to 10.1%) compared to a prevalence of 2% (95%CI 0.7 to 4.6%) in children who were born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks the prevalence of moderate CP was 12.5% (95%CI 2.7 to 32.4%) compared to a prevalence of 5.6% (95%CI 2.4 to 10.7%) in children who were born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate CP was 4.6% (95%CI 1.3 to 11.4%) compared to a prevalence of 2.0% (95%CI 0.4 to 5.7%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate to severe cerebral palsy

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks the prevalence of moderate to severe CP (GMFCS 2-5) was 10.5% (95%CI 2.9 to 24.8%) compared to a prevalence of 6.4% (95%CI 3.7 to 10.2%) in children who were born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe CP (bilateral spastic CP) was 4.2% (95%CI 2.2 to 7.2%) compared to a prevalence of 1.5% (95%CI 0.2 to 5.3%) in children who were born at 34 weeks GA, assessed at the age of 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks the prevalence of moderate to severe CP (ambulatory or non-ambulatory) was 16.7% (95%CI 4.7 to 37.4%) compared to a prevalence of 9.7% (95%CI 5.4 to 15.8%) in children who were born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate quality evidence from one study (sample size 6347) showed that among children born at <32 weeks GA the prevalence of moderate to severe CP (other types) was 3.5% (95%CI 3.0 to 4.0%) compared to a prevalence of 0.25% (95%CI 0.2 to 0.3%) in children born at 34-36 weeks GA, assessed at up to the age of 7 years (Hirvonen 2014).

Severe cerebral palsy

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks the prevalence of severe CP (GMFCS 3-5) was 10.5% (95%CI 2.9 to 24.8%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children who were born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of severe CP (GMFCS 4-5) was 9.2% (95%CI 4.1 to 17.3%) compared to a prevalence of 1.3% (95%CI 0.2 to 4.7%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate quality evidence from one study (sample size 1455) showed that among children born at ≤23 weeks the prevalence of severe CP (non-ambulatory) was 4.2% (95%CI 0.1 to 21.1%) compared to a prevalence of 4.2% (95%CI 1.5 to 8.9%) in children who were born at 25 weeks GA, assessed at 6 years age. The prevalence among children born at 24 weeks was higher (11% (95%CI 4.9 to 20.5%) (Marlow 2005).

Prevalence of cerebral palsy using per 1000 or 10,000 live births as denominator

Any cerebral palsy (<28 weeks GA)

Low quality evidence from one study (sample size 2858) showed that among children born at <28 weeks GA the rate of any CP was 112.7 per 1000 survivors (95%CI 50 to 210) (Drummond 2002).

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at <28 weeks GA the rate of any CP was 71.4 per 1000 livebirths (95%CI 42 to 112 per 1000 live births) at 4 to 8 years age (Himmelmann 2014).

Low quality evidence from one study (sample size 46) showed that among children born at <28 weeks GA the rate of any CP was 72.3 per 1000 live births (95%CI 39 to 120.3 per 1000 live births) at age 4-7 years (Nordmark 2001).

Moderate quality evidence from one study (sample size 975) showed that among children born at <28 weeks GA the rate of any CP in 1992-1994 was 131 per 1000 live births (95% CI 90-183/1000 live births) at age 2 years (confirmed at 3 years age) (Robertson 2007). In the same study, the rate of any CP decreased with the time points (years). From 1995-1997 and 1998-2000, the rate was 69 per 1000 live births (95%CI 41 to 108 per 1000 live births). From 2001-2003 the rate was 19 per 1000 live births (95%CI 6 to 44 per 1000 live births). Over the whole 11 years of the study, the rate was 70 per 1000 live births (95%CI 55 to 88 per 1000 live births) at 2 years age (Robertson 2007).

Severe cerebral palsy (<28 weeks GA)

Moderate quality evidence from one study (sample size 975) showed that among children born at <28 weeks GA the rate of non-ambulatory CP in 1992-1994 was 59 per 1000 live births (95% CI 32-99 per 1000 live births) at age 2 years (confirmed at 3 years age) (Robertson 2007). In the same study, the rate of any CP decreased with the time points in years. From 1995-1997 the rate was 16 per 1000 livebirths (95%CI 5-41 per 1000 livebirths) and from 1998-2000, the rate was 8 per 1000 live births (95%CI 1 to 29 per 1000 live births). From 2001-2003 the rate was 8 per 1000 live births (95%CI 1 to 27 per 1000 live births). Over the whole 11 years of the study, the rate was 22 per 1000 live births (95%CI 13 to 33 per 1000 live births) at 2 years age (Robertson 2007).

Any cerebral palsy (28-32 weeks GA)

Low quality evidence from one study (sample size 2858) showed that among children born at 28-32 weeks GA the rate of any CP was 56.3 per 1000 neonatal survivors (95%CI 33 to 90) (Drummond 2002).

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at 28-32 weeks GA the rate of any CP was 39.6 per 1000 livebirths (95%CI 25 to 59 per 1000 live births) at 4 to 8 years age (Himmelmann 2014).

Low quality evidence from one study (sample size 46) showed that among children born at 28-31 weeks GA the rate of any CP was 32.2 per 1000 live births (95%CI 18.1 to 52.2 per 1000 live births) at age 4-7 years (Nordmark 2001).

Any cerebral palsy (32-36 weeks GA)

Low quality evidence from one study (sample size 189) showed that among children (1991-1996 cohort in Norway) born at 33-36 weeks GA the rate of any CP was 13.8 per 1000 livebirths at earliest age of 4 years (Andersen 2011). In the same study the prevalence of any CP among children (1991-1998 cohort in Italy) was 8.8 per 1000 livebirths whereas in cohorts from Spain and Ireland the rate was 4 per 1000 livebirths (Andersen 2011).

Low quality evidence from one study (sample size 2858) showed that among children born at 32-36 weeks GA the rate of any CP was 9.6 per 1000 survivors (95%CI 6 to 14) (Drummond 2002).

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at 32-36 weeks GA the rate of any CP was 6.4 per 1000 livebirths (95%CI 4 to 9 per 1000 live births) at 4 to 8 years age (Himmelmann 2014). For children born at <37 weeks GA, the rate of any CP was 13 per 1000 live births (95%CI 10 to 16 per 1000 live births).

Low quality evidence from one study (sample size 46) showed that among children born at 32-36 weeks GA the rate of any CP was 4.6 per 1000 live births (95%CI 2.7 to 7.3 per 1000 live births) at age 4-7 years (Nordmark 2001).

4.1.4.4.11 Diplegia or tetraplegia

Moderate quality evidence from one study (sample size 94,466 live births) showed that among children born at <37 weeks GA the rate of bilateral spastic CP was 7.5 per 1000 livebirths (95%CI 5 to 10 per 1000 live births) at 4 to 8 years age (Himmelmann 2014).

4.1.4.4.12 Prevalence of intellectual disability

≤ 28 completed weeks of gestation

Moderate intellectual disability

Moderate to low quality from 4 studies (sample size range from 165 to 576) showed that among children born at a range of 23 to 27 weeks GA the prevalence of intellectual disability (BSIDIII -2SD to -3SD) ranged from 6.4 (95%CI 4.6 to 8.8) to 24% (95%CI 20 to 29) (Doyle 2011; Moore 2012; Anon 1997; Serenius 2013). One further low quality study (sample size 77) used the Dutch version of BSIDII, which showed that the prevalence of intellectual disability was 10.4% (95%CI 4.6 to 19.5) (MDI 55-69) (De Groote 2007).

Moderate quality evidence from two studies (sample size range from 75 to 1508) showed that among children born at 24-27 weeks GA or <27 weeks GA the prevalence of intellectual disability (K-ABC 55-69) was 14.9% (95%CI 10.5 to 20.2) and 10.7% (95%CI 4.7 to 19.9) at 5 years and 7-9 years respectively (Foix-Helias 2008; Stahlmann 2009).

Moderate quality from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of intellectual disability (IQ -2 to -3SD on K-ABC, GMDS or NEPSY) was 19.9% (95%CI 15.1 to 25.5) at 6 years (Marlow 2005).

Moderate quality evidence from one study (sample size 306) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (Full scale IQ WPPSI-R 55-70) was 4.9% (95%CI 2.8 to 8) at 5 years (Leversen 2011). Low quality evidence from one study (sample size 141) showed that the prevalence (WISC-IV -2SD to -3SD) was 8.5% (95%CI 4.4 to 14.1) in children born in the same gestational age range but assessed at 8 years (Roberts 2010).

Moderate to severe intellectual disability

Moderate to low quality evidence from 5 studies (sample size ranged from 19 to 1508) showed that among children born at GA range 24 to 28 weeks GA the prevalence of intellectual disability (MPC <70 or IQ <70 K-ABC) ranged from 17.6% (95%CI 12.8 to 23.2) to 41% (95%CI 18 to 67) at a range of 5-9 years (Beaino 2011; Foix-Helias 2008; Larroque 2008; Rieger-Fackeldey 2010; Stahlmann 2009).

Moderate to low quality evidence from 5 studies (sample size ranged from 77 to 3785) showed that among children born at a GA range 22-27 weeks GA the prevalence of intellectual disability (BSID <-2SD or MDI <70) ranged from 15.2%(95%CI 10.1 to 21.6) to 39% (95%CI 37 to 41) at 18-36 months (Doyle 2011; Moore 2012; Anon 1997; Vohr 2005; De Groote 2007).

Moderate to low quality evidence from two studies (sample size 203 to 1455) showed that among children born at 22-27 or <27 weeks GA the prevalence of intellectual disability (WPPSI-R IQ <70) was 11.8% (95%CI 6.2 to 19.7) and 5.6% (95%CI 3.3 to 8.8) respectively at 5 years (Mikkola 2005; Leversen 2011).

Low quality from one study (sample size 141) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (WISC-IV IQ <-2SD) was 14.6% (95%CI 9.3 to 21.4) at 8 years corrected age (Roberts 2010). Low quality evidence from one other study showed that the prevalence (using WISC-III <70) in 275 children born at <28 weeks GA was 5.1% (95%CI 2.8 to 8.4) (Anderson 2003).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks the prevalence of intellectual disability (IQ <-2SD [K-ABC, GMDS or NEPSY]) was 40.7% (95%CI 34.4 to 47.2) at 6 years (Marlow 2005).

Low quality evidence from one study (sample size 244) showed that among children born at <27 weeks GA the prevalence of intellectual disability (Griffiths <2SD) was 10.4% (95%CI 5.8 to 16.8) at 12 months corrected age (Sutton 1999).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence of intellectual disability (verbal, DAS II <=2SD) was 17% (95%CI 14.5 to 19.5) and 15% (95%CI 12.7 to 17.6) for non- verbal reasoning (DAS II <=2SD) at 10 years (Joseph 2016b).

Severe intellectual disability

Moderate quality evidence from two studies (sample size ranged from 75 to 1508) showed that among children born at <27 weeks or 24-27 weeks GA the prevalence of intellectual disability (IQ <55, K-ABC) was 14.7% (95% CI 7.6 to 24.7) and 2.7% (95%CI 1 to 5.8) at 5-9 years (Stahlmann 2009; Foix-Helias 2008).

Moderate to low quality evidence from 5 studies (sample size ranged from 77 to 576) showed that among children born at GA range 23 to 27 weeks the prevalence of intellectual disability (BSIDIII <-3SD or MDI <55) ranged from 3.6% (95%CI 1.4 to 7.8) to 18.2% (95%CI 10.3 to 28.6) across the studies (Moore 2012; Anon 1997; De Groote 2007; Serenius 2013; Doyle 2011).

Low quality evidence from one study (sample size 141) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (IQ <-3SD, WISC-IV) was 6.3% (95%CI 2.9 to 11.5) at 8 years corrected age (Roberts 2010).

Moderate quality evidence from one study (sample size 306) showed that among children born at 22-27 weeks GA the prevalence of intellectual disability (IQ <55, WPPSI-R) was 2.9% (95%CI 1.4 to 5.5) at 5 years (Leversen 2011).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of intellectual disability (IQ <-3SD, K-ABC, GMDS or NEPSY) was 20.8% (95%CI 15.8 to 26.4) at 6 years (Marlow 2005).

Moderate quality evidence from one study (sample size 142) showed that among children born at a mean GA of 27 weeks, the prevalence of intellectual disability (IQ <71 WPPSI) was 4.2% (95%CI 1.6 to 9.0%) At 4 years (Salakorpi 2001).

28-31 completed weeks of gestation

Moderate intellectual disability

Moderate quality evidence from one study (sample size 1508) showed that among children born at 28-32 weeks GA the prevalence of intellectual disability (MPC 55-69) was 8.7% (95CI 7.2 to 10.4) at 5 years (Foix-Helias 2008). In the same study, the prevalence in children born at 24-32 weeks GA was 9.6% (95%CI 8.2 to 11.2).

Moderate to severe intellectual disability

Moderate to low quality evidence from 4 studies (sample size ranged from 1455 to 1812) showed that among children born at a gestational age range of 28-32 weeks the prevalence of intellectual disability (MPC <70, K-ABC) was similar across the studies (range 8.9% (95%CI 7.3 to 10.7) to 12.1% (95%CI 10 to 14.4)) at 5 years (Beaino 2011; Marret 2007; Foix-Helias 2008; Larroque 2008).

(*A number of studies reported intellectual disability in children born at <32 weeks GA. One study of moderate quality in 3785 children born at 22-32 weeks GA found that the prevalence for intellectual disability (MDI <70, BSIDII) was 33.8% (95%CI 32.3 to 35.4) at 18-22 months corrected age (Vohr 2005).*

Low quality evidence from two studies (sample size ranged from 203 to 259) showed that among children at 23-32 weeks or mean GA 27.3 (2.1) the prevalence of intellectual disability (IQ<70, WISC-IV or DAS, or IQ<70, WPPSI-R) was 15.8% (95%CI 11.6 to 20.9) and 9.4% (95%CI 5.7 to 14.2) respectively at 5 years (Andrews 2008; Mikkola 2005).

Moderate quality evidence from two studies (sample size ranged from 1508 to 1812) showed that among children born at 24-32 weeks and <33 weeks GA the prevalence was the same (11.9% (95%CI 10.3 to 13.7)) at 5 years (Foix-Helias 2008; Larroque 2008).

Moderate quality evidence from one study (sample size 402) showed that among children born at <32 weeks GA/<1500g the prevalence of intellectual disability (IQ<-2SD, revised Amsterdam Child Intelligence Test) was 6.2% (95%CI 4.1 to 9) at 5 years (de Kleine 2003).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of intellectual disability (MDI <70, BSIDII) was 25.9% (95%CI 23.7 to 28.2) at 18-22 months corrected age (Vohr 2005). Another study reported a prevalence of 17% (95%CI 11 to 24) at <32 weeks GA (Cognitive delay, <2SD BSID) (Toome 2012).

Low quality evidence from one study (sample size 347) showed that among children born at <33 weeks GA the prevalence of intellectual disability (DQ <70, Brunet-Lezine) was 2.3% (95%CI 1 to 4.5) at 2 years (corrected age) (Charkaluk 2010).

Severe intellectual disability

Moderate quality from one study (sample size 1508) showed that among children born at 28-32 weeks GA the prevalence of intellectual disability (MPC <55) was 2.3% (95%CI 1.5 to 3.2) at 5 years (Foix-Helias 2008). In the same study, the prevalence in children born at 24-32 weeks GA was 2.3% (95%CI 1.6 to 3.2).

32-36 completed weeks of gestation

Moderate to severe intellectual disability

Low quality evidence from one study (sample size 646) showed that among children born at 32-34 weeks GA the prevalence of intellectual disability (MPC<70) was 7.6% (95%CI 5.7 to 9.9) at 5 years (Marret 2007).

4.1.4.4.13 *Prevalence of intellectual disability by week of gestational age*

Moderate intellectual disability

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate intellectual disability (BSIDII -2 to -3 SD) was 13.2% (95%CI 4.4 to 28.1%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate intellectual disability (full scale IQ 55-70, WPPSI-R) was 6.9% (95%CI 2.6 to 14.4%) compared to a prevalence of 2.6% (95%CI 0.7 to 6.6%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of intellectual disability (IQ -2 to -3 SD, KABC GMDS or NEPSY) was 33.3% (95%CI 15.6 to 55.3%) compared to a prevalence of 18.8% (95%CI 12.7 to 26.1%) in children born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate to severe intellectual disability

Low quality evidence from one study (sample size 244) showed that among children born at 23 weeks GA the prevalence of moderate to severe intellectual disability (major developmental delay, Griffiths <2SD) was 100% (95%CI 25 to 100%) compared to a prevalence of 3.9% (95%CI 0.81 to 11%) in children born at 26 weeks GA, assessed at 12 months corrected age (Sutton 1999).

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate to severe intellectual disability (cognitive impairment BSIDIII ≤-2SD) was 31.6% (17.5 to 48.7%) compared to a prevalence of 12.0% (95%CI 8.2 to 16.6%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Low quality evidence from one study (sample size 1503) showed that among children born at 24-26 weeks GA the prevalence of moderate to severe intellectual disability (MPC<70, KABC) was 15.7% (95%CI 9.2 to 24.2) compared to a prevalence of 8.9% (95%CI 6.2 to 12.0%) in children born at 31-32 weeks GA, assessed at 5 years (Beaino 2011).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate to severe intellectual disability (full scale IQ <70, WPPSI-R) was 9.2% (95%CI 4.1 to 17.3%) compared to a prevalence of 2.6% (95%CI 0.7 to 6.6%) in children born at 26-27 weeks GA, assessed at 5 years (Leversen 2011).

Moderate quality evidence from one study (sample size 1534) showed that among children born at 24-25 weeks GA the prevalence of moderate to severe intellectual disability (MPC <70, KABC) was 12.5% (95%CI 4.7 to 25.3%) compared to a prevalence of 10.7% (95%CI 7.5 to 14.6%) in children born at 32 weeks GA. However, the prevalence was higher in children born at 26 weeks GA (prevalence 21.1% (95%CI 11.4 to 33.9%), 27 weeks (prevalence 18.6% (95%CI 12.1 to 26.9%), and 28 weeks GA (prevalence 20.7% (95%CI 14.5 to 28%) (Larroque 2008).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe intellectual disability (MPC <70, KABC) was 9.9% (95%CI 6.5 to 14.3%) compared to a prevalence of 5.3% (95%CI 2.0 to 11.2%) in children born at 34 weeks GA, assessed at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate to severe intellectual disability (IQ≤-=2SD, KABC GMDS or NEPSY) was 58.3% (95%CI 36.6 to 77.9%) compared to a prevalence of 35.4% (95%CI 27.6 to 43.8%) in children born at 25 weeks GA, assessed at 5 years (Marlow 2005).

Severe intellectual disability

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of severe intellectual disability (cognitive impairment, BSIDIII <-3SD) was 18.4% (95%CI 7.7 to 34.3%) compared to a prevalence of 7.6% (95%CI 4.6 to 11.6%) in children born at 26 weeks GA, assessed at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of severe intellectual disability (full scale IQ <55, WPPSI-R) was 4.6% (95%CI 1.3 to 11.4%) in children born at 26-27 weeks GA, assessed at 5 years age (Leversen 2011).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of severe intellectual disability (IQ <-3SD, KABC, GMDS or NEPSY) was 25.0% (95%CI 9.8 to 46.7%) compared to a prevalence of 16.7% (95%CI 11 to 23.8%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

4.1.4.4.12 Prevalence of speech and/or language disorder

≤ 28 completed weeks of gestation

Moderate and severe speech and/or language disorder

Moderate quality evidence from one study (sample size 456) showed that among children born at <27 weeks GA the prevalence of moderate language impairment (-2 to -3SD BSIDIII) was 9.4% (95%CI 6.7 to 12.7) (Serenius 2013).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of moderate communication impairment (-2SD to -3SD BSIDIII) was 5.4% (95%CI 3.7 to 7.6) at 3 years age (Moore 2012). In the same study, there was a prevalence of 11.6% (95%CI 9.1 to 14.5) in children with moderate to severe impairment (<=2SD BSIDIII).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe speech/communication impairment ranged from 1.10% to 5.3% depending on whether they could communicate by a systemised method or not at 30 months (median) (Wood 2000).

Low quality evidence from one study (sample size 241) showed that among children born at <=25+6 weeks GA the prevalence for total severe impairment (PLS <2SD) was 15.6% (95%CI 10.8 to 21.4) at a median age of 6 years (Wolke 2008). However, the prevalence of severe communication impairment and severe language impairment in children (sample size ranged from 456 to 576) born at <27 weeks was lower in two studies of moderate to low quality (6.30% (95%CI 4.4 to 8.6) and 6.60% (95%CI 4.4 to 9.5) respectively) at the age of 2.5 to 3 years age (Serenius 2013; Moore 2012).

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate communication impairment (-2 to -3 SD BSID III) was 10.5% (95%CI 2.3 to 24.8) compared to 4.4% (95%CI 2.2 to 7.7) at 26 weeks GA (at the age of 3 years). A similar trend was observed when severe communication impairment was assessed (<-3SD BSIDIII), with prevalence increasing with decreasing gestational age by week. At 22-23 weeks GA, the prevalence was 15.8% (95%CI 6 to 31.3) (Moore 2012) compared to the prevalence at 26 weeks GA, which was 4% (95%CI 1.9 to 7.2) (Moore 2012).

For moderate to severe impairment, there was a similar trend, prevalence in the 22-23 GA group was 26.5% (95%CI 13.4 to 43.1) compared to 8.4% (95% CI 5.3 to 12.5) in the 26 weeks GA group (Moore 2012).

28-31 completed weeks of gestation

Low quality evidence from one study (sample size 155) showed that among children born at <32 weeks GA the prevalence of moderate language delay (<2SD BSIDIII) was 33% (95%CI 26 to 41) at 2 years (corrected age) (Toome 2012).

4.1.4.4.15 Prevalence of speech and language disorder by week of gestation

Moderate speech and language disorder

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate speech/language disability (communication impairment, BSIDII -2 to -3 SD) was 10.5% (95%CI 2.9 to 24.8%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate to severe speech and language disorder

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate to severe speech/language disability (communication impairment, BSIDII ≤-2 SD) was 26.3% (95%CI 13.4 to 43.1%) compared to a prevalence of 8.4% (95%CI 5.3 to 12.5%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Severe speech and language disorder

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of severe speech/language disability (communication impairment, BSIDII <-3 SD) was 15.8% (95%CI 6.0 to 31.3%) compared to a prevalence of 4.0% (95%CI 1.9 to 7.2%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

4.1.4.4.16 Prevalence of attention deficit hyperactivity disorder

≤ 28 completed weeks of gestation

Low quality evidence from two studies (sample size 205 to 219) showed that among children born at <26 weeks GA and x adolescents born at <28 weeks GA the prevalence of ADHD (including any type, DAWBA or ChIPs) was 11.5% (95%CI 7.3 to 17) at the age of 11 years and 14.6% (95%CI 10 to 20.2) at the age of 18 years respectively. In the same two studies, the prevalence of ADHD (combined) was 4.4% (95%CI 1.9 to 8.4) and 3.4% (95% CI 1.4 to 7) respectively at the ages of 11 years and at 18 years. Prevalence of ADHD (inattentive) in the two studies was 10.7% (95%CI 6.9 to 16) at the age of 11 years and 7.1% (95%CI 3.8 to 11.8) at the age of 18 years (Johnson 2010; Burnett 2014).

Low quality evidence from one study of (sample size 205) showed that among children born at <26 weeks GA the prevalence of ADHD (hyperactive/impulsive, ChIPs) was 0.5% (95%CI 0.01 to 2.7) at the age of 18 years (Burnett 2014).

4.1.4.4.17 Prevalence of autism spectrum disorder

≤ 28 completed weeks of gestation

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks GA the prevalence of ASD (any) was 8% (95%CI 4.6 to 12.6) at the age of 11 years. In the same study, the prevalence of autistic disorder was 6.5% (95%CI 3.5 to 10.8) and for atypical autism, the prevalence was 1.5% (95%CI 0.3 to 4.3) (Johnson 2010).

Moderate quality evidence from one study (sample size 857) showed that among children born at <28 weeks GA the prevalence of ASD (ADI-R and ADOS-2) was 9.2% (95%CI 7.4 to 11.4%) and 7.1% (95%CI 5.5 to 9.0) respectively at 10 years age (Joseph 2016a).

4.1.4.4.18 Prevalence of specific learning difficulty

≤ 28 completed weeks of gestation

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks GA the prevalence reading impairment (WIAT-II <-2SD) was 30.2% (95%CI 24.1 to 36.9) at the age of 11 years (Johnson 2011). However, in another study of low quality, 275 children who were born at <28 weeks GA had a lower prevalence of reading impairment (WRAT 3 <70) was lower (5.8% (95%CI 3.4 to 9.3)) when assessed at the age of 8 years (Anderson 2003). In the same two studies, there was a higher prevalence of arithmetic impairment (43.7% (95%CI 37 to 50.6)) in children born at <26 weeks GA compared with a prevalence of 6.6% (95%CI 4 to 10.2) in children born at <28 weeks GA (Johnson 2011; Anderson 2003)

Low quality evidence from one study (sample size 257) showed that among children born at <28 weeks GA the prevalence of spelling impairment was 2.5% (95%CI 1 to 5.2) assessed at the age of 8 years (Anderson 2003).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence of academic achievement (WIAT-III <=-2SD) was 14% (95%CI 11.7 to 16.5) for word reading, 16% (95%CI 13.7 to 18.6) for pseudoword decoding, 14% (95%CI 11.7-16.5) for spelling, and 17% (95%CI 14.5 to 19.6) for numeric operations when assessed at the age of 10 years (Joseph 2016b).

28-31 completed weeks of gestation

Low quality evidence from one study (sample size 135) showed that among children born at <33 weeks GA the prevalence of delayed numerical skills (TEDI-MATH <40) was 20% (95%CI 13.6 to 27.8) (at the age of 8 years (Kiechl-Kohlendorfer 2013).

4.1.4.4.19 Prevalence of developmental coordination disorder

≤ 28 completed weeks of gestation

Low quality evidence from one study (sample size 298) showed that among children born at 22-27 weeks GA the prevalence of DCD was higher in a cohort born in 1997 (16% (95%CI 10.1 to 23.3)) compared to a cohort born in 1991(sample size 298) (10% (95%CI 6.9 to 14.1)) (Roberts 2011).

28-31 completed weeks of gestation

Moderate to low quality evidence from two studies (sample size ranged from 280 to 402) showed that among children at <32 weeks GA the prevalence of DCD or motor delay was 22.3% (95%CI 18.3 to 26.7) at the age of 5 years and 30.7% (95%CI 25.4 to 36.5) at the age of 7-8 years. (de Kleine 2003; Foulder-Hughes 2003).

Moderate quality evidence from one study (sample size 168) showed that among children born between 24-31 weeks GA the prevalence of motor deficit was 17.9% (95%CI 12.4 to 24.5) at the age of 5 years (Agerholm 2011).

32-36 completed weeks of gestation

4.1.4.4.20 *Prevalence of mental and behavioural disorders*

≤ 28 completed weeks of gestation

Low quality evidence from one study (sample size 219) showed that among children born at <26 weeks GA the prevalence of emotional disorder (any) was highest among 11 year olds (9% (95%CI 5.4 to 13.6)), compared to conduct disorder (any), oppositional defiant disorder (5.5% (95%CI 2.9 to 9.4) and 5% (95%CI 2.5 to 8.8)), specific phobia (2.5% (95%CI 0.8 to 5.7)), or a number of disorders including specific phobia or social phobia, PTSD, generalised anxiety, disorder, childhood emotional disorder, and major depression (prevalence range from 0.5%(95%CI 0.01 to 2.8) to 2% (95%CI 0.5 to 5)) (DAWBA, Johnson 2011).

Low quality evidence from one study (sample size 205) showed that among children born at <28 weeks GA the prevalence of anxiety/mood disorder was highest (21% (95%CI 15.6 to 27.2)) in adolescents compared to mood disorder (16.1% (95%CI 11.4 to 22)), major depressive disorder (13.7% (95%CI 9.3 to 19.1)), anxiety disorder (BAI/CESD-R) (11.2% (95%CI 7.3 to 16.4)), co-morbid disorder (6.3% (95%CI 3.4 to 10.6)) and obsessive compulsive disorder (2% (95%CI 0.5 to 5)) (DSM-IV axis I, Burnett 2014).

4.1.4.4.21 *Prevalence of visual impairment*

≤ 28 completed weeks of gestation

Moderate quality evidence from one study (sample size 456) showed that among children born at <27 weeks GA the prevalence of visual impairment (any) was 3.7% (95%CI 2.2 to 5.9) at 2.5 years corrected age (Serenius 2013).

Moderate visual impairment

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of visual impairment (impaired but not blind) was 4.6% (95%CI 2.3 to 8) at 6 years age (Marlow 2005).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of visual impairment (functionally impaired vision) was 5.9% (95%CI 4.1 to 8;2) at 3 years age (Moore 2012).

Moderate quality evidence from one study (sample size 456) showed that among children born at <27 weeks GA the prevalence of visual impairment (moderate impairment) was 2.9% (95% CI 1.5 to 4.8) at 2.5 years corrected age (Serenius 2013).

Moderate to severe visual impairment

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-26 weeks GA the prevalence of unilateral blindness was 2.7% (95%CI 2 to 3.4) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 242) showed that among children born at <28 weeks GA the prevalence of moderate to severe visual deficiency (<3/10, one or both eyes) was 7% (95%CI 4.1 to 11) at 5 years age (Larroque 2008).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA the prevalence of visual impairment (impaired or blind) was 7.1% (95%CI 4.2 to 11.1) at 6 years age (Marlow 2005).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of impaired vision (blind or functionally impaired) was 6.9% (95%CI 5 to 9.3) at 3 years (Moore 2012).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of visual impairment (little useful vision) was 9.1% (95%CI 3.7 to 17.8) at 3 years age (de Groote 2007).

Low quality evidence from one study (sample size 88) showed that among children born at <28 weeks the prevalence of severe visual impairment (uni- or bilateral blindness or visual acuity <20/200 without glasses in at least one eye) was 12.5% (95%CI 6.4 to 21.3) at 11 years (Farooqi 2011).

Moderate quality evidence from two studies (sample size 306) showed that among children born at either 22-27 weeks GA or 23-25 weeks the prevalence for severe visual impairment was 0.3% (95%CI 0.01 to 1.8) and 1.2% (95%CI 0.03 to 6.2) respectively at 5 years (Leversen 2011).

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe visual impairment (blind or perceives light) was 2.5% (95%CI 1 to 5) at 30 months (median) (Wood 2000).

Moderate quality evidence from one study (sample size 411) showed that among children born at <27 weeks GA the prevalence of visual impairment (blind or able to only fixate and follow light binocularly) was 3.1% (95%CI 1.6 to 5.3) at 30 months corrected age (Holmstrom 2014).

Low quality evidence from one study (sample size 77) showed that among children born at <27 weeks GA the prevalence of visual impairment (no useful vision) was 2.6% (95%CI 0.9 to 9.1) at 3 years age (De Groote 2007).

Low quality evidence from two studies (sample size ranged from 189 to 219) showed that among children born at 23-27 weeks GA and 22-27 weeks GA the prevalence for blindness (<6/60 in both eyes) was 2.3% (95%CI 0.8 to 5.3) and 1.6% (95%CI 0.3 to 4.6) at 2 years and 8 years (corrected) respectively (Anon 1997; Anderson 2011).

Moderate to low quality evidence from three separate studies (sample size ranged from 306 to 373) showed that among children born at 22-27 weeks GA and also 23-25 weeks GA the prevalence for blindness was varied, ranging from 5.8% (95%CI 1.9 to 12.9) in the lower GA group (Leversen 2011), and 1.6% (95%CI ranged from 0.5 to 3.8) in the two 22-27 GA groups (Leversen 2010; Leversen 2011).

Moderate to very low quality evidence from 8 studies (sample size ranged from 19 to 3785) showed that among children born at various gestational ages (ranging from <26 weeks to <28 weeks) the prevalence of blindness was varied, ranging from 0.9% (95%CI 0.24 to 2.3) to 11% (95%CI 1.3 to 33) (Vohr 2005; Roberts 2010; Marlow 2005; Moore 2012; Hutchinson 2013; Serenius 2013; Anderson 2003; Rieger-Fackeldey 2010).

Low quality evidence from one study (sample size 1506) showed that among children born at <28 weeks GA the prevalence of severe visual impairment (functional blindness) was 0.8% (95%CI 0.3 to 1.7) at 10 years (Joseph 2016b).

28-31 completed weeks of gestation

Moderate to severe visual impairment

Low quality evidence from one study (sample size 1455) showed that among children born at 30-31 weeks GA the prevalence of visual impairment (visual acuity <3/10 in both eyes) was 1.5% (95%CI 0.7 to 2.8) at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA found that the prevalence of visual impairment (unilateral blindness) was 1.3% (95%CI 0.8 to 2) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 971) showed that among children born at 28-31 weeks GA the prevalence of moderate to severe visual deficiency (<3/10 in one or both eyes) was 2.1% (95%CI 1.3 to 3.2) at 5 years age (Larroque 2008).

Studies reporting vision impairment at <32 weeks GA

Moderate quality evidence from one study (sample size 3785) showed that among children born at 22-32 weeks GA the prevalence of unilateral blindness was 2.1% (95%CI 1.7 to 2.6) at 18-22 months corrected age (Vohr 2005).

Moderate quality evidence from one study (sample size 1697) showed that among children born at <33 weeks GA the prevalence of moderate to severe visual deficiency (<3/10 in one or both eyes) was 2% (95%CI 1.4 to 2.8) at 5 years (Larroque 2008).

Low quality evidence from one study (sample size 93) showed that among children born at <32 weeks GA the prevalence of visual impairment (worst eye blind or able to fixate torch) was 2.2% (95%CI 0.3 to 7.6) at 2.5 years corrected age (Hreinsdottir 2013).

Low quality evidence from one study with (sample size 155) showed that among children born at <32 weeks GA found that the prevalence of visual impairment (moderately reduced/blindness) was 0.64% (95%CI 0.02 to 3.5) at 2 years (corrected age) (Toome 2012).

Severe visual impairment

Moderate quality evidence from on study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of visual impairment (bilateral blindness) was 0.7% (95%CI 0.3 to 1.2) at 18-22 months corrected age (Vohr 2005). In the same study, the prevalence of bilateral blindness in children born at 22-32 weeks GA was 1.2% (95%CI 0.9 to 1.6) (Vohr 2005).

Low quality evidence from one study (sample size 93) showed that among children born at <32 weeks GA the prevalence of visual impairment (best eye blind or only able to fixate a torch) was 1.1% (95%CI 0.03 to 5.9) at 2.5 years corrected age (Hreinsdottir 2013).

32-36 completed weeks of gestation

Moderate to severe visual impairment

Low quality evidence from on study (sample size 1455) showed that among children born at 32-24 weeks GA the prevalence of visual impairment (visual acuity <3/10 in both eyes) was 1.7% (95%CI 0.9 to 3) at 5 years age (Marret 2007).

4.1.4.4.22 Prevalence of visual impairment by week of gestation

Moderate visual impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate visual impairment (functionally impaired vision) was 15.8% (95%CI 6.0 to 31.3%) compared to a prevalence of 3.2% (95%CI 1.4 to 6.2%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate visual impairment (visually impaired, not blind) was 8.3% (95%CI 1.0 to 27.0%) compared to a prevalence of 2.8% (95%CI 0.8 to 7.0%) in children born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Moderate quality evidence from one study (sample size 494) showed that among children born at 22-23 weeks GA the prevalence of visual impairment (any; best estimated visual acuity <20/40) was 23.8% (95%CI 12 to 40) compared to a prevalence of 13.4% (95%CI 6.9 to 22.7) at 24 weeks GA, prevalence of 7% (95%CI 3.4 to 12.6) at 25 weeks GA, and a prevalence of 5.1% (95%CI 2.1-1-.2) at 26 weeks GA (Hellgren 2016).

Moderate to severe visual impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate to severe visual impairment (functionally impaired vision) was 18.4% (95%CI 7.7 to 34.3%) compared to a prevalence of 4.4% (95%CI 2.2 to 7.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Low quality evidence from on study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe visual impairment (visual acuity <3/10 in both eyes) was 0.7% (95%CI 0.1 to 2.6) compared to a prevalence of 0.8% (95%CI 0.02 to 4.1%) in children born at 34 weeks GA. The prevalence was higher at GA 31 weeks (2.2% (95%CI 0.8 to 4.3%), and 33 weeks GA (2.3% (95%CI 0.5 to 6.5%), assessed at 5 years age (Marret 2007).

Moderate quality evidence from one study (sample size 1817) showed that among children born at 24-25 weeks GA the prevalence of moderate to severe visual impairment (<3/10 one or both eyes) was 9.3% (95%CI 3.1 to 20.3%) compared to a prevalence of 1.9% (95%CI 0.9 to 3.5%) in children born at 32 weeks GA, assessed at 5 years age (Larroque 2008).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate to severe visual impairment (visually impaired, or blind) was 16.7% (95%CI 4.7 to 37.4%) compared to a prevalence of 3.5% (95%CI 1.1 to 7.9%) in children born at 25 weeks GA, assessed at 6 years age (Marlow 2005).

Severe visual impairment

Moderate quality evidence from one study (sample size 411) showed that among children born at 22-23 weeks GA the prevalence of severe visual impairment (blind or able to only fixate and follow light binocularly) was 4.8% (95%CI 0.6 to 16.2%) compared to a prevalence of 1.4% (95%CI 0.2 to 4.8%) in children born at 26 weeks GA, assessed at 30 months corrected age (Holmstrom 2014).

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of visual impairment (blindness) was 2.6% (95%CI 0.1 to 13.8%) compared to a prevalence of 1.2% (95%CI 0.3 to 3.5%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of severe visual impairment (blindness) was 8.3% (95%CI 1.0 to 27.0%) compared to a prevalence of 0.7% (95%CI 0.02 to 3.8%) in children born at 25 weeks GA assessed at 6 years age (Marlow 2005).

Prevalence of visual impairment using per 1000 or 10,000 live births as denominator

Moderate to severe visual impairment (<28 weeks GA)

Very low quality evidence from one study (sample size 172, 584 livebirths) showed that among children born at <28 weeks GA the prevalence of moderate to severe visual impairment (<=6/18 in better eye or worse) was 182.5 cases per 10,000 livebirths (95%CI 102.5 to 299.1) at 12 years (Bodeau-Livinec 2007).

Moderate to severe visual impairment (28-31 weeks GA)

Very low quality evidence from one study (sample size 172, 584 livebirths) showed that among children born at 29-32 weeks GA the prevalence of moderate to severe vision impairment (<=6/18 in better eye or worse) was 37.1 cases per 10,000 livebirths (95%CI 14.9 to 76.2)at 12 years age (Bodeau-Livinec 2007).

Moderate to severe visual impairment (32-36 weeks GA)

Very low quality evidence from one study (sample size 172, 584 livebirths) showed that among children born at 33-36 weeks GA the prevalence of moderate to severe vision impairment (<=6/18 in better eye or worse) was 27 cases per 10,000 livebirths (95%CI 17.3 to 40.1) at 12 years age (Bodeau-Livinec 2007).

4.1.4.4.23 Prevalence of hearing impairment

≤ 28 completed weeks of gestation

Moderate hearing impairment

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks GA, the prevalence of hearing loss (corrected with hearing aids) was 2.9% (95%CI 1.2 to 5.9) when assessed at 6 years age (Marlow 2005).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of hearing loss (improved by aids) was 5.2% (95%CI 3.5 to 7.4) when assessed at 3 years age (Moore 2012).

Low quality evidence from one study (sample size 77) showed that among children born at < 27 weeks GA the prevalence of hearing impairment (but useful hearing) was 3.9% (95%CI 0.8 to 11) (De Groote 2007).

Moderate to severe hearing impairment

Low quality evidence from one study (sample size 141) showed that among children born at 22-27 weeks GA the prevalence of hearing impairment was 2.1% (95%CI 0.4 to 6) at 8 years corrected age (Roberts 2010).

Moderate quality evidence from one study (sample size 241) showed that among children born at <26 weeks the prevalence of moderate to severe hearing impairment was 5.8% (95%CI 3.2 to 9.6) at 6 years (Marlow 2005). In another study of low quality with 576 children born at <27 weeks GA the prevalence for severe hearing impairment was 5.4% (95%CI 3.7 to 7.6) at 3 years (Moore 2012).

Low quality evidence from one study (sample size 19) showed that among children born at mean 25.4 weeks GA the prevalence of hearing impairment (requiring hearing aid) was 11% (95%CI 1.3 to 33) at 5 years age (Rieger-Fackeldey 2010). Ten other studies (sample size ranged from 77 to 3785) of moderate to very low quality assessing hearing impairment or deafness (requiring hearing aids) in children born at a range of 22-28 weeks GA found that the prevalence was lower but varied, ranging from 0.7% (95%CI 0.14 to 2) to 5.7% (95%CI 1.9 to 12.8) (Farooqi 2011; Leversen 2011; Vohr 2005; Doyle 2011; Anderson 2011; De Groote 2007; Hutchinson 2013; Wood 2000; Serenius 2013; Anderson 2003).

Severe hearing impairment

Low quality evidence from one study (sample size 283) showed that among children born at 22-25 weeks GA the prevalence of severe hearing impairment (uncorrected without hearing aid) was 5.3% (95%CI 3.0 to 8.6) at 30 months (median) (Wood 2000).

Low quality evidence from one study (sample size 373) showed that among children born at 22-27 weeks GA the prevalence of deafness was 0.8% (95%CI 0.1 to 2.7) at 2 years (corrected age) (Leversen 2010). In another study (sample size 401) of low quality, the prevalence of deafness was 0.9% (95%CI 0.1 to 3.3) in children assessed at 2 years (Anon 1997). Prevalence of deafness was 0.2% (95%CI 0.01 to 1.2) in children (sample size 456) born at <27 weeks GA (moderate quality, Serenius 2013). At 5 years age, the prevalence of deafness was 1.0% (95%CI 0.2 to 2.8) in children (sample size 306) born at 22-27 weeks GA (moderate quality study, Leversen 2011).

Low quality evidence from one study (sample size 261) showed that among children born at <28 weeks GA the prevalence of severe hearing deficiency (>70 decibels in one or both ears or hearing aid) was 0.8% (95%CI 0.1 to 2.7) at 5 years age (Larroque 2008).

Low quality evidence from one study (sample size 576) showed that among children born at <27 weeks GA the prevalence of profound sensorineural hearing loss (not improved by aids) was 0.2% (95%CI 0.1 to 1) at 3 years age (Moore 2012). In another moderate quality study (sample size 241) children born at <26 weeks GA found that the prevalence of profound sensorineural hearing loss was 2.9% (95%CI 1.2 to 5.9) at 6 years age (Marlow 2005).

28-31 completed weeks of gestation

Moderate to severe hearing impairment

Moderate quality evidence from one study (sample size 3785) showed that among children born at 27-32 weeks GA the prevalence of permanent hearing loss (amplification in both ears) was 1.4% (95%CI 0.9 to 2.1) at 18-22 months corrected age (Vohr 2005).

Low quality evidence from one study (sample size 1455) showed that among children born at 30-31 weeks GA the prevalence for hearing loss >70 decibels was 0.30% (95%CI 0.04 to 1.1) at 5 years (Marret 2007).

Severe hearing impairment

Moderate quality evidence from one study (sample size 1020) showed that among children born at 28-31 weeks GA the prevalence for severe hearing deficiency (>70 decibels in one or both ears or hearing loss) was 0.5% (95%CI 0.2 to 1.1) at 5 years age (Larroque 2008).

4.1.4.4.24 *Prevalence of hearing impairment by week of gestation*

Moderate hearing impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate hearing impairment (hearing loss improved by aids) was 5.3% (95%CI 0.6 to 17.8%) compared to a prevalence of 5.2% (95%CI 2.8 to 8.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 241) showed that among children born at 24 weeks GA the prevalence of moderate hearing impairment was 2.7 (95%CI 0.3 to 9.6%) compared to a prevalence of 3.5% (95%CI 1.1 to 7.9%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

Moderate to severe hearing impairment

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of moderate hearing impairment (hearing loss improved by aids) was 7.9% (95%CI 1.7 to 21%) compared to a prevalence of 5.2% (95%CI 2.8 to 8.7%) in children born at 26 weeks GA, assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 306) showed that among children born at 23-25 weeks GA the prevalence of moderate to severe hearing impairment (hearing aid in both ears) was 2.3% (0.3 to 8.1%) compared to a prevalence of 1.3% (95%CI 0.2 to 4.7%) in children born at 26-27 weeks GA, assessed at 5 years (Leversen 2011).

Low quality evidence from one study (sample size 1455) showed that among children born at 30 weeks GA the prevalence of moderate to severe hearing impairment (hearing loss >70 decibels or aids in one or both ears) was 0.3% (95%CI 0.01 to 1.9%) compared to a prevalence of 1.5% (95%CI 0.2 to 5.3%) in children born at 34 weeks GA, assessed at 5 years (Marret 2007).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of moderate to severe hearing impairment was 4.2% (95%CI 0.1 to 21.1%) compared to a prevalence of 4.9% (95%CI 2.0 to 9.8%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

4.1.4.4.25 *Severe hearing impairment*

Low quality evidence from one study (sample size 576) showed that among children born at 22-23 weeks GA the prevalence of severe hearing impairment (profound sensorineural hearing loss not improved by aids) was 2.6% (95%CI 0.1 to 13.8%), assessed at 3 years (Moore 2012).

Moderate quality evidence from one study (sample size 1817) showed that among children born at 24-25 weeks GA the prevalence of severe hearing impairment (>70 decibels in one or both ears or hearing aid) was 1.7% (95%CI 0.04 to 9.2%) compared to a prevalence of 0.2% (95%CI 0.01 to 1.1%) in children born at 32 weeks GA, assessed at 5 years (Larroque 2008).

Moderate quality evidence from one study (sample size 241) showed that among children born at ≤23 weeks GA the prevalence of severe hearing impairment (profound sensorineural hearing loss) was 4.2% (95%CI 0.1 to 21.1%) compared to a prevalence of 1.4% (95%CI 0.1 to 4.9%) in children born at 25 weeks GA, assessed at 6 years (Marlow 2005).

4.2 Evidence to recommendations

4.2.1.1 Relative value placed on the outcomes considered

The Committee prioritised the following developmental outcomes: cerebral palsy, intellectual disability or global developmental delay, autism spectrum disorder, attention deficit/hyperactivity disorder, motor problems, speech, language and communication problems, executive function problems, and special educational needs.

These developmental disorders and problems were prioritised as they were considered to cause most concern among parents and carers and early identification and follow-up of these conditions have the greatest potential, once detected early and signposted to the appropriate services, to improve the outcomes for the child and family. These were also considered critical outcomes for which standardisation of clinical practice are needed, in view of significant variations in follow-up measuring these outcomes across the UK.

Other important outcomes considered in the reviews were: specific learning difficulty, developmental coordination disorder, mental disorders, social, emotional and behavioural problems, attention problems, visual impairment, hearing impairment, functional problems with feeding or eating, sleeping, and toileting, sensory sensitivity, problems specific to infancy including excessive crying, and irritability and poor self-regulation.

4.2.1.2 Consideration of clinical benefits and harms

Knowledge of risk factors for different development disorders and problems enables health care professionals to effectively identify babies and children born prematurely who are more likely to experience a developmental disorder or problems, and prioritise surveillance services accordingly. The Committee agreed that it was important to assess independent risk factors associated with each developmental disorder and problem, but appreciated that there was high degree of comorbidity in clinical practice and risk factors may not present independently.

The Committee recognised that while there was a large amount of evidence identified by the evidence review, there were several gaps in the evidence. These gaps included outcomes of interest, risk factors of interest, stratification by different gestational ages as well as the different ages at assessment. The gaps in evidence were due to both actual gaps in existing evidence and the relatively strict inclusion and exclusion criteria set out in the review protocol (Appendix D:). The recommendations should therefore be considered in the light of this absence of evidence. For example, if the only evidence found was among children born before 28 weeks' gestation, it does not necessarily mean children born at a more mature gestational age would not be at an increased risk of that outcome but rather that there is uncertainty due to the absence of evidence.

When deliberating about the evidence pertaining to risk factors and their associations with different disorders and problems, the Committee discussed:

- the magnitude of the risk estimate and whether the evidence from different studies reported, or largely reported, consistent findings regarding the direction of effect
- whether the evidence available was applicable to the UK setting
- circumstances where the study findings were inconsistent but conclusions could be drawn from well-conducted studies with robust findings
- circumstances where uncertainty remained after assessing the variations and heterogeneity across studies and no conclusions or recommendations could be made.

Specific developmental outcomes are discussed in the following sections along with the conclusions that the Committee reached when forming their recommendations.

Cerebral palsy

The Committee agreed that the evidence showed clearly that children born preterm were at an increased risk of cerebral palsy compared to children born at term. There was also clear evidence showing that the prevalence of cerebral palsy increased by decreasing gestational age with children born extremely preterm having a much higher prevalence of cerebral palsy compared to children born at later gestational ages.

In addition to gestational age at birth within the preterm population, the Committee considered the evidence on the association between cerebral palsy and different biological, neonatal, social, maternal and environmental risk factors.

The Committee concluded that evidence from several studies clearly indicated that grade 3 and 4 intraventricular haemorrhage and cystic periventricular leukomalacia were independent risk factors for cerebral palsy.

Regarding neonatal sepsis, the evidence largely showed an association between sepsis and cerebral palsy. There was some discrepancy in the evidence, which could be explained by sepsis being defined differently across the studies. The Committee agreed that culture-positive sepsis was shown to be an independent risk factor for cerebral palsy in children born preterm.

Evidence on the association between necrotising enterocolitis (NEC) and cerebral palsy was mixed. The Committee discussed that even though two publications showed an increased risk of cerebral palsy in the presence of NEC, the definition of NEC varied across studies and the evidence was too mixed for the Committee to confidently make a recommendation.

Regarding exposure to antenatal steroids, the evidence showed a largely protective effect of antenatal steroids on cerebral palsy even though not all effect estimates reached statistical significance. The Committee agreed that exposure to antenatal steroids was protective against cerebral palsy. The Committee was also aware that new evidence published after the re-run searches for the reviews showed a dose-dependent protective effect against neurodevelopmental impairment in children born extremely preterm which supports the recommendation made (Chawla 2016).

The Committee agreed that evidence on the association between postnatal steroid exposure and cerebral palsy was mixed or lacked statistical power. The Committee agreed that when considering postnatal steroid exposure, the dose and the duration of the steroids were important factors to consider. However, much of the evidence did not differentiate between doses and durations. One study reported stratified results in relation to the duration of administration of postnatal steroids and the results indicated long duration (>=57 days) increased the risk of cerebral palsy whereas shorter duration had no effect. However, in three studies where dosage or duration of steroid course was not specified, a significantly increased risk of cerebral palsy was shown. Two other studies also showed similar tendency even though they did not reach statistical significance. Based on these evidence, the Committee concluded that exposure to postnatal steroids increased the risk of cerebral palsy in children born preterm.

Regarding bronchopulmonary dysplasia, defined as requiring oxygen at 36 weeks' postmenstrual age, the Committee concluded that the evidence did not show an association with cerebral palsy. However, evidence from one large study showed a significantly increased risk of cerebral palsy (quadriparesis, diparesis) when the baby had bronchopulmonary dysplasia with need for continued mechanical ventilation at 36 weeks' postmenstrual age.

The Committee noted that it is known that there is a link between chorioamnionitis and cerebral palsy in the general population. However, the evidence among children born preterm did not show such an association. Evidence came from two studies, one of which showed no significant association between chorioamnionitis and cerebral palsy and the other showed a reduced risk of cerebral palsy in children born preterm with exposure to chorioamnionitis.

Although one study showed an increased risk of cerebral palsy among boys born preterm compared to girls, other studies did not find an association, therefore, no conclusions could be made.

Similarly, the evidence was mixed regarding being born small for gestational age and its link with cerebral palsy in children born preterm and no definite conclusion could be made.

The Committee discussed that generally, multiple pregnancy and young maternal age would be risk factors for cerebral palsy in preterm children, however, there was no clear evidence linking maternal age or multiple pregnancy to cerebral palsy.

Motor problems

Based on the evidence from three studies, the Committee concluded that children born preterm were at an increased risk of fine motor problems. The Committee discussed that the reason why one study did not find an association with preterm birth and fine motor problems could be due to the study population (the study excluded all children who needed tertiary care as neonates) as well the assessment method (such as the tool) and different cut-offs used. The assessment in this study was one-to-one with the child by a professional whereas the other studies relied on assessments completed by parents. The Committee noted that the proportion of children identified with problems using screening tools was typically higher than the proportion of children identified with problems using diagnostic assessments.

The Committee discussed the evidence on gross motor problems and concluded that children born before 32 weeks' gestation were at an increased risk of gross motor delay, as demonstrated in the evidence from several studies. The evidence among children born moderate to late preterm (32 to 36 weeks of gestation) did not reach statistical significance. The Committee concluded that the evidence showing an increased risk of delayed motor development among children born preterm supported the recommendations made on fine motor and gross motor problems.

Even though no evidence was found looking at the association between gestational age and developmental coordination disorder (DCD), there was some evidence that the prevalence of DCD is higher among children born preterm than children born at term. The prevalence of DCD among children born preterm ranged from 10% to 30%, which the Committee considered to be higher than what is typically observed in the general population according to their clinical knowledge and expertise.

The Committee reviewed the evidence on the relationship between brain lesions and motor delay, and concluded from 2 studies an increased risk of psychomotor developmental impairment (PDI <70) in children born preterm with grade 3-4 IVH. In an additional paper, a grade 3-4 IVH was significantly associated with overall neurodevelopmental impairment (defined as cerebral palsy, MDI<70, bilateral blindness or hearing impairment) but not with PDI <70 alone.

The Committee discussed the association between necrotising enterocolitis (NEC) and motor problems and concluded that there was evidence showing that preterm children with NEC requiring surgical treatment were at an increased risk of motor problems. The Committee discussed how the evidence on medically managed NEC did not show a clear association, thus, only the more severe form of NEC seemed to be associated with later motor delay.

The Committee discussed how the relationship between neonatal sepsis and motor delay depended on the definition and measurement of sepsis. The Committee noted that a diagnosis of sepsis could be made by clinical symptoms and signs augmented by culture positivity (blood, urine, or cerebrospinal fluid) requiring antibiotic treatment. The Committee agreed that the evidence available showed that children born extremely preterm with neonatal sepsis that was proven by culture and was treated with antibiotics for five or more days were at an increased risk of motor problems.

The evidence on the relationship between retinopathy of prematurity (ROP) and motor problems came from one study which looked at different levels of ROP and their associations with motor problems. The Committee discussed how even though the findings were mixed since the association between motor problems and some types of ROP did not reach statistical significance, there was indication that at least more severe levels of ROP were associated with motor problems.

The Committee discussed the evidence on antenatal exposure to steroids and its association with motor delay. Evidence was mixed with two studies showing a decreased risk of motor delay among very preterm children exposed to antenatal steroids while two studies found no significant association and one study found an increased risk of motor delay. The Committee discussed the discrepancy between the findings and noted that in the study that showed decreased risk of motor delay, highly intensive treatment was given to the children included. This could have potentially decreased the risk of motor delay independent of exposure to antenatal steroids. The Committee also discussed why the one study found an increased risk contrary to the other evidence available but could not find a reasonable explanation. The Committee discussed how generally it was thought that antenatal steroids were protective of developmental problems. In addition, the Committee were aware of a 2006 Cochrane review of randomized controlled trials that reported that antenatal steroids had a protective effect on developmental problems. This was not considered in the evidence review because randomised controlled trials were not included. Therefore, due to conflicting and unclear evidence, no conclusion could be made.

The evidence on postnatal exposure to steroids and its association with motor problems was scarce. The population in a study showing an increased risk of motor problems in the presence of exposure to postnatal steroids was considered to be somewhat selective since it only included children treated in neonatal units of a research network. Therefore, the Committee did not feel this was strong enough to make a recommendation. The Committee recognised that this was an area that was rapidly changing and further evidence was needed to draw conclusions.

The available evidence did not show a clear association between bronchopulmonary dysplasia (BPD) and motor problems, therefore, the Committee agreed that they were unable to make a recommendation.

The Committee discussed the evidence on motor delay in relation to the child’s sex, ethnicity, socioeconomic status, being born small for gestational age and maternal cocaine use but due to the limited evidence with mainly non-significant results and the low quality of the evidence that was available, the Committee decided that no conclusion could be made on the associations between those risk factors and motor problems.

Global developmental delay

Evidence from 5 studies was available on the association between gestational age and global developmental delay. Two studies found an increased risk of global developmental delay in children born before 32 weeks’ gestation and one study in children with very low birth weight children (mean gestational age 28.4 weeks). Evidence for an increased risk for global developmental delay following late and moderate preterm birth was mixed. A UK study of children born 32 to 36 weeks’ gestation found a clear association between late and moderate preterm birth and global developmental delay, however two other studies did not find a statistically significant association. These three studies used different assessment tools and employed different inclusion criteria for the term-born comparison group which made it difficult to directly compare the results. Despite conflicting evidence among children born moderate to late preterm, the Committee concluded that children born preterm appeared to be at increased risk of global developmental delay.

No evidence was identified regarding the association between brain abnormalities and global developmental delay. The Committee found this to be an unusual finding since there was clear evidence for the association between brain abnormalities and both intellectual disability and cerebral palsy.

The evidence on the association between being born SGA and global developmental delay among preterm children was scarce. However, the Committee agreed that the evidence showed that being born SGA was an independent risk factor for global developmental delay in children born preterm. The same study also showed an association between multiple birth and global developmental delay.

Regarding global developmental delay in relation to other biological and social factors, they concluded that evidence from 2 studies showed boys born preterm were at an increased risk of global developmental delay. Evidence from a UK study showed that children of non-white ethnicity and children from families with lower socioeconomic status who were born preterm were at an increased risk of developmental problems. The Committee discussed how these data were only among children born moderate to late preterm but concluded that it was appropriate to extrapolate these findings to children born at earlier gestational ages.

Intellectual disability

Evidence from several studies showed an increased risk in intellectual disability in children born preterm compared to children born at term. Furthermore, the evidence showed that prevalence of intellectual disability increased with decreasing gestation age at birth.

The Committee agreed that regarding the association between being born SGA and intellectual disability among children born preterm the evidence was mixed. It was noted that one study found no association between SGA and intellectual disability among children born between 24 and 28 weeks of gestation, however, the same publication reported a significant association between SGA and those born between 29 and 32 weeks of gestation. The Committee discussed that this may indicate that being born extremely pretermwas in itself such a severe risk factor for intellectual disability that being born SGA did not increase the risk additionally. In another publication with the same cohort, when the analysis was broken down by the severity of intellectual disability, it was found that SGA significantly increased the risk of severe intellectual disability among those born between 24 and 32 weeks of gestation. Acknowledging the limitations in evidence from subgroup analyses as such, and considering the evidence from another study that showed positive association between SGA and intellectual disability among those born at less than 27 weeks of gestation, the Committee considered SGA to be a risk factor for intellectual disability among those born at 24 to 32 weeks of gestation.

Regarding brain abnormalities and its association with intellectual disability among children born preterm, the evidence was mixed. The Committee noted that the types and severity of brain abnormalities considered in the different studies varied and intellectual disability was measured differently across the studies. However, the Committee concluded that there was enough evidence to show that more severe brain abnormalities, or more precisely grade 3 and 4 intraventricular haemorrhage and cystic periventricular leukomalacia, were associated with an IQ score less than 70 points regardless of the test. This was found in 6 studies, however, some of these findings did not reach statistical significance.

Evidence regarding neonatal sepsis proven by culture among children preterm (less than 28 weeks' gestation) from 3 studies showed an increased risk of intellectual disability.

Therefore, the Committee concluded that neonatal culture-positive sepsis increases the risk of intellectual disability among children born less than 28 weeks' gestation. Evidence among children born at later gestational ages was not available.

Evidence regarding the association between necrotising enterocolitis and intellectual disability among children born preterm was mixed. However, when looking at the more severe form of necrotising enterocolitis (grade II or more, requiring surgery, or perforated) the evidence clearly showed an increased risk of intellectual disability whereas medically managed necrotising enterocolitis or non-specified necrotising enterocolitis showed no association. Therefore, the Committee concluded that necrotising enterocolitis requiring surgery was an independent risk factor for intellectual disability among children born before 32 weeks of gestation. Evidence among children born at later gestational ages was not available.

Although findings were mixed, the Committee considered postnatal steroids to be an independent risk factor for intellectual disability. It was noted that the study that did not find a statistically significant association between postnatal steroids and intellectual disability had a selected population since children with cerebral palsy were excluded from the study. However, cerebral palsy and intellectual disability were closely associated. Therefore, the other studies that reported a significantly increased risk in intellectual disability in those who were exposed to postnatal steroids were considered to be more reliable and the conclusion was made based on them. The Committee agreed that their clinical experience did not contradict this finding. Evidence was only available among children born up to 32 weeks of gestation.

The evidence regarding the relationship between ROP and intellectual disability came from one study which looked at different levels of ROP and their associations with different levels of intellectual disability. The Committee discussed how even though not all comparisons reached statistical significance, the findings showed a clear trend that ROP was associated with intellectual disability in children born before 28 weeks' gestation. No evidence for later gestational ages was available.

The evidence for the association between BPD and intellectual disability was mixed. The Committee discussed how two studies showed an increased risk in intellectual disability (IQ score <70 points) with BPD among children born very preterm. One study found no association, however that study excluded children with cerebral palsy and as said, cerebral palsy and intellectual were known to be associated. Another study that found no association used a more strict cut-off for intellectual disability (score of <55). Based on these considerations, the Committee concluded that BPD was an independent risk factor for intellectual disability (defined as IQ score <70) in children born very preterm.

Evidence regarding the association between exposure to antenatal steroids and intellectual disability was mixed. Some studies found a protective effect of antenatal steroids on intellectual disability while some studies found no association. The Committee agreed that no firm conclusions could be made based on the available evidence.

Evidence regarding socioeconomic status and its association with intellectual disability were mixed but showed a clear tendency that preterm children from a disadvantaged background were at an increased risk of intellectual disability. The Committee also noted that preterm birth was known to be more common among mothers from socially disadvantaged backgrounds.

Evidence from three studies largely showed no association between chorioamnionitis and intellectual disability in children born preterm. The Committee considered it important how chorioamnionitis was defined, and determined it should be confirmed through histology or assessed clinically. The Committee concluded that there was no convincing evidence to show that chorioamnionitis would increase the risk of intellectual disability in children born preterm.

Evidence from 2 studies found no association between multiple birth and intellectual disability. The Committee noted that the two studies defined intellectual disability differently. They agreed that the evidence was very limited and no conclusions could be made.

Evidence regarding the association between maternal age and intellectual disability was scarce and the Committee was not able to draw any conclusions.

Special educational needs and educational attainment

The Committee agreed that the evidence underpinning the recommendations for special educational needs (SEN) should be from the UK only since educational settings varied considerably across countries. According to the UK evidence on special educational needs and educational attainment, the Committee concluded that all children born preterm were at an increased risk of special educational needs and the risk increased with decreasing gestational age. This conclusion was based on a large population-based study from Scotland and supported by other smaller studies. The Committee also discussed the risk of different subtypes of special educational needs among children born preterm (such as physical and motor SEN; language SEN; intellectual SEN; social, emotional or behavioural SEN) and whether any recommendations should be made on individual subtypes but since statistical power was considered low in some of the subtypes and statistical significance was not reached, the Committee decided to make a recommendation on global special educational needs only.

The evidence regarding educational attainment was also discussed. The Committee concluded that there was clear evidence from four UK studies on Foundation Stage and Key Stage 1 showing that children born preterm were at an increased risk of lower educational attainment during early school years compared to term children. The prevalence of low attainment increased with decreasing gestational age. The Committee were surprised that the evidence at key stage 2 to 4 showed no statistically significant association between prematurity and low attainment. There was also evidence showing that children born extremely preterm (before 26 weeks' gestation) had an increased risk of low attainment for reading and mathematics. Evidence on risk factors for low attainment was scarce but showed that intracranial haemorrhage and BPD were independently associated with delayed numeracy skills among children born before 32 weeks of gestation.

The Committee noted that no evidence was found on specific learning difficulties.

Evidence regarding risk factors associated with SEN was scarce so conclusions were difficult to reach. The Committee discussed that it was generally known that male sex was associated with SEN, however, evidence from only one study was available (from the UK) which showed that extremely preterm boys were more likely to SEN than extremely preterm girls. The same study also showed and association between brain abnormalities and SEN in children born extremely preterm. The same study was the only available study that looked at other risk factors in relation to SEN (NEC, antenatal steroids, postnatal steroids for chronic lung disease, chorioamnionitis, maternal ethnicity, maternal socioeconomic status, and maternal age) but found no statistically significant associations.

Autism spectrum disorder

Evidence regarding autism spectrum disorder (ASD) was found in two levels: symptoms suggestive of ASD assessed using screening tools and diagnosis of ASD assessed using diagnostic tools.

The evidence regarding symptoms suggestive of ASD was only available for children born before 28 weeks of gestation compared with term born children but showed a clearly increased risk when reported by both parents and teachers. The Committee, therefore, concluded that children born before 28 weeks of gestation were at an increased risk of symptoms suggestive of ASD. The Committee noted that evidence was not available for later gestational ages.

The evidence regarding ASD diagnosis among preterm came from two studies. One of these studies used parental reports of ASD diagnosis and the other study used data from an ASD register. Even though the Committee recognised that neither of these studies assessed the children using a diagnostic test, they concluded based on these evidence and their clinical experience that compared to children born at term, children born preterm were at an increased risk of ASD. There was also evidence showing that prevalence of ASD increased with decreasing gestational age.

The evidence regarding factors associated with ASD was relatively scarce. Intracranial haemorrhage was shown to increase the risk of ASD among children born before 34 weeks of gestation in 1 study which corresponded with the Committee's clinical experience.

The Committee also considered the evidence from 2 studies that showed boys born preterm to be at a significantly increased risk of ASD compared to girls born preterm. The Committee based the recommendation on this evidence in addition to their clinical knowledge and concluded that male sex was an independent risk factor for ASD.

Evidence on the association between ASD and neonatal sepsis, BPD, being born SGA, and ethnicity showed no association.

Attention, impulsivity and hyperactivity

Several studies using different screening tools found children born preterm were at an increased risk of attention problems, hyperactivity, and impulsivity. Some of the findings did not reach statistical significance likely due to relatively small sample sizes. The studies also used different instruments making it difficult to directly compare the results. However, the Committee agreed the evidence clearly showed that children born preterm were at an increased risk of symptoms of inattention, hyperactivity and impulsivity.

Evidence from two studies showed an increased risk of ADHD in children born before 28 weeks' gestation compared to term born children. A third study also showed an increased risk in all children born preterm compared to children born at term, however, this study relied on parent report of ADHD by asking the parent if a doctor had ever told them that the child had ADHD. The Committee considered this an unreliable way of measuring ADHD and gave this finding less weight. Another study among children born late preterm did not find an association and studies stratifying by different types of ADHD did not find an association. Therefore, the Committee concluded that there was an increased risk of ADHD (any type) among children born extremely preterm.

The Committee discussed that it was generally known that male sex was an independent risk factor for ADHD. As no evidence was available regarding the association between neonatal risk factors and ADHD, the Committee was unable to reach any conclusions.

Emotional and behavioural problems

The Committee discussed how many different criteria and tools were used across different studies to define and assess emotional and behavioural problems, making it difficult to compare the findings from different studies directly. The Committee, however, concluded that there was evidence to show that children born preterm were at an increased risk of behavioural problems, particularly internalising behaviours, compared with children born at term. The evidence showed that children born preterm had an increased risk of internalising behaviours, including anxiety, whereas the evidence on externalising behaviours was more mixed. An increased risk of hypoactivity or passivity was also found in preterm children at school age, both when observed by teachers and by parents.

Evidence regarding the association between different neonatal, biological, maternal, social factors and emotional and behavioural problems in children born preterm was relatively scarce. Evidence often came from one or two studies, however, these studies were well established cohort studies with moderate quality data. Evidence from a large French cohort study among children born preterm showed that major brain lesions increased the risk of behavioural problems. No association was found between minor or moderate brain lesions and behavioural problems. The Committee discussed how these findings indicated that only severe brain abnormalities (essentially cystic brain lesions) increased the risk of behavioural problems. The same study also found that a maternal age of less than 25 years (compared to 25 to 34 years) was a risk factor for total behavioural difficulties (as assessed by the SDQ) at both 3 and 5 years of age and that maternal self-report of poor mental wellbeing was associated with behavioural problems in the preterm child. A Dutch cohort study of moderate to late preterm children found that those from families of lower socioeconomic status to be at an increased risk of behavioural problems, especially internalising behaviours (assessed with CBCL) at preschool age. Evidence for sex of the child and being born SGA was either non-significant or equivocal and the Committee was not able to reach conclusions from these.

Speech, language and communication

Evidence on speech, language and communication problems was mixed. However, the Committee noted the studies that reported non-significant results on language and communication delay had wider exclusion criteria for their study populations. One study excluded all children requiring tertiary care as neonates and another study excluded all multiple births. The Committee pointed out that both can be common among children born preterm. In addition, in one of the studies only one component of communication problems was examined. Therefore, the Committee decided to make recommendations based on the other studies that examined global communication/language problems and showed a significant association between prematurity and speech, language and communication problems.

There was also some evidence on speech and language disorders among children born preterm. One study among children born extremely preterm found an increased risk of mild as well as moderate language impairment at 2.5 years of corrected age using the Bayley scales (mild -1 to -2 SD; moderate -2 to -3 SD). Another study among children born late preterm also found an increased risk of ICD-9 diagnosis of developmental speech or language delay in preschool age. Therefore, the Committee concluded that children born preterm were also at an increased risk of speech and language disorders.

Evidence on the association between different neonatal, biological, maternal, social factors and speech and language disorders and problems was scarce. However, the Committee agreed that evidence from a national cohort study from Estonia showing a significantly increased risk of language impairment in boys born preterm compared to girls born preterm at 2 years of age, and despite a relatively small sample size, this evidence was convincing enough to conclude that male sex was an independent risk factor for language delay (assessed with Bayley scales) among children born preterm.

The same Estonian study also showed an association between severe brain lesions and language delay (assessed with Bayley scales). Another study also found an association with severe periventricular-intraventricular haemorrhage and language delay (assessed with Bayley scales). Therefore, the Committee concluded that major brain lesions were independent risk factors for language disorder in children born preterm.

Feeding problems

Evidence on feeding problems was mixed. Although the majority of evidence showed no significant association between gestational age and feeding problems, the Committee noted that a significant association was found in two large studies among children born extremely preterm where feeding problems were defined as either total eating difficulties or oral motor problems. They thought the difference between these studies and the others, which showed no significant association, was mainly driven by motor problems which could be persistent. Therefore they concluded that among those born extremely preterm there was an increased risk of feeding problems which could persist until the age of 6 years.

The Committee also discussed the evidence for the effect of different biological, social, maternal and neonatal factors on the risk of feeding problems among children born preterm. Evidence on the effect of the child’s sex, ethnicity, socioeconomic status and being born SGA on feeding problems was inconclusive since only two low quality studies reported non-significant results. One study narratively reported an increased risk of feeding problems with brain lesions and with BPD, however, no effect estimates were given. Therefore, the Committee decided that they were unable to reach conclusions about the risk of feeding problems in relation to neonatal, biological, maternal and social factors.

Sleeping problems and sleep apnoea

The Committee discussed how they could not draw any conclusions on the risk of general sleeping problems among children born preterm since only one study with non-statistically significant results was included in the review. There was, however, evidence on an increased risk of sleep apnoea among children born preterm. The Committee discussed whether the risk of sleep apnoea increased by decreasing gestational age, but since there was only one study reporting on sleep apnoea, the Committee could not reach a definite conclusions on whether there was a dose-response relationship between sleep apnoea and gestational age.

Visual impairment

There was no evidence available on the association between preterm birth (versus term) and visual impairment. The Committee was surprised by this absence of evidence and was not able to conclude whether there was an increased risk of visual impairment in children born preterm. However, there was evidence on the prevalence rates of visual impairment in the population born preterm showing that the prevalence increased by decreasing gestational age. This could imply that there was an increased risk of visual impairment in children born preterm compared to term born children.

Evidence on the association between neonatal, biological, maternal, social factors and visual impairment was scarce.Although the majority of the available evidence showed that the association of visual impairment and risk factorsdid not reach statistical significance, there was some evidence that neonatal sepsis proven by culture and treated with antibiotics increased the risk of visual impairment in children born preterm. There was also evidence from the same cohort in another publication that suggested grade 3 and 4 intraventricular haemorrhage with shunt increased the risk of visual impairment. A national cohort study from Finland among children born very preterm found a considerably increased risk of visual impairment with ROP.

Hearing impairment

No evidence was identified regarding association between prematurity and hearing impairment. The Committee found the absence of evidence surprising and they were not able to reach conclusions as a result. However, there was evidence regarding prevalence levels of hearing impairment in the population born preterm. This evidence showed that the prevalence increased with decreasing gestational age, suggesting that the risk of hearing impairment may be increased in the population born preterm compared to children born at term.

Evidence regarding association between neonatal, biological maternal, social factors and hearing impairment was scarce and the existing evidence showed mostly non-statistically significant findings. The Committee was, therefore, unable to reach conclusions about these associations. However, there was evidence from one study showing that culture-proven sepsis with antibiotic treatment for 5 or more days significantly increased the risk of unilateral or bilateral hearing aid use in children born preterm. Therefore, the Committee concluded that culture-proven neonatal sepsis was an independent risk factor for hearing impairment in children born preterm.

Executive function problems

Evidence from three different studies showed that preterm birth was associated with executive function problems, specifically in planning, organisation, and working memory. The Committee noted that the findings were consistent between studies where outcomes were reported by either parents/teachers or assessed by a trained professional. The evidence was only available for children born before 32 weeks of gestation. The Committee thought because the evidence supported the significant association between essential components of executive functions such as planning, organising, and working memory, it could be concluded that there was an increased risk of executive function problems in children born before 32 weeks of gestation.

No evidence was found on the association between neonatal, biological, maternal, social factors and executive functions, therefore, the Committee was not able to reach conclusions on these potential associations.

4.2.2 Consideration of economic benefits and harms

A systematic review of the economic literature was conducted but no relevant studies were identified that were applicable to this review question.

Since the recommendations do not provide instructions for action, which makes the economic implications difficult to assess, the overall economic impact was considered unlikely to be significant. It is expected that increased awareness of prevalence and risk factors will lead healthcare professionals into taking action, such as a referral to specialist services for diagnostic assessment if parents and carers or health professionals have a concern or suspicion that there could be signs of a problem or disorder. While these actions would have cost implications, it is likely that the investigation costs would be outweighed by the potential cost and effectiveness offsets associated with earlier identification.

4.2.3 Quality of evidence

Overall, evidence on the risk and prevalence of developmental disorders and problems was of very low to moderate quality. The main reasons for downgrading the quality of the evidence were:

- limited description of the population and sample at hand
- high attrition (sometimes including failing to report the reasons for losses to follow-up and failing to report the characteristics of the ones lost to follow-up compared to the ones followed-up)
- insufficient description of the risk factors and the way they were assessed or measured (in the risk reviews)
- insufficient description of the outcome assessments
- high imprecision of point estimates (that is, wide confidence intervals) due to relatively low sample sizes
- insufficient or unclear adjustments for potential confounders (risk reviews).

The Committee also recognised that there was a large variation and heterogeneity across the studies in terms of:

- inclusion/exclusion criteria of participants
- gestational age of participants
- setting and year of measurement (for example, 1992 versus 2012)
- participant age at time of outcome assessment
- criteria or definitions of the outcome
- tools and scales used to assess the outcome
- level of outcome severity (for example, grade of intraventricular haemorrhages)
- criteria and definitions of risk factors (in the risk reviews)
- adjustments made in multiple variable analyses (in the risk reviews).

For these reasons, it was agreed that pooling of the findings using meta-analyses would not be appropriate.

4.2.4 Other considerations

Evidence on the prevalence of developmental disorders and problems in the population born preterm was used to guide the Committee in making recommendations about which populations were expected to benefit most from enhanced surveillance and support. The prevalence rates among children born preterm were not compared to prevalence rates among children born at term, therefore, the Committee could not reach an evidence-based conclusion that a prevalence of an outcome was increased in the population born preterm. However, for many outcomes, the prevalence rates in the general population were known and widely accepted by the Committee which made it clear when the evidence revealed an increased prevalence in the population born preterm. For example, the rate of cerebral palsy in the general population was known to be 1 to 2 per 1000 whereas the evidence among children born before 28 weeks' gestation showed a considerably higher prevalence which ranged between 5 to 25%. The Committee considered presenting the available evidence on prevalence of different developmental disorders and problems in the population born preterm, for example, in a table format. It was discussed that this could guide health care providers, parents and carers to understand the likelihood of the child developing specific developmental disorders and problems. However, due to the heterogeneity of the evidence and wide range of the estimates, it was concluded that the degree of uncertainty was sufficiently high that the presentation of prevalence estimates would not be clinically meaningful or helpful when counselling parents and carers.

The Committee discussed how it was important to recognise that developmental problems presented on a continuum with the severity of the problem ranging from a mild problem with limited effect on function to a severe disorder affecting all aspects of life. The Committee thought it was important that children born preterm who had been classified to have 'mild' problems were neither automatically considered to have problems nor automatically considered not to have problems. The Committee acknowledged that sometimes the distinction between a disorder and a milder problem was a difficult, or even artificial, distinction to make. The severity of the problem could have a significant effect on the life of the child and his or her family. However, the Committee also discussed how sometimes multiple mild problems could amount to a considerable functional problem for the child. For example, a child with a problem classified as mild with previously mild functional problems may face considerable difficulties or functional problems when entering school. The Committee also discussed how problems that were considered mild, for example, as determined by a result in an assessment of 1 to 2 standard deviations below the mean, may as well be considered in the normal range as they may not have an effect on day to day function. . The Committee concluded that these problems should not be over-medicalised and each child should always be considered individually.

The Committee discussed how in addition to parents and clinicians, it was crucial that professionals working in the education and social sectors were made aware of the developmental problems and challenges that the child born preterm was facing.

4.3 Recommendations

4.3.1 Risk and prevelance of developmental problems and disorders

1. Be aware that children born preterm are at increased risk of developmental problems and disorders.

2. Be aware that for recommendations in this section:

- for some developmental problems and disorders there was an absence of evidence about overall risk and prevalence in children born preterm, and some papers included specific gestational ages at birth from which the committee was unable to extrapolate to other gestational ages
- for some developmental problems and disorders the evidence was underpowered to detect an effect
- other gestational ages and other factors not listed here might also be associated with increased risk of developmental problems and disorders.

Cerebral palsy

3. Be aware that children born preterm are at increased risk of cerebral palsy, and that:

- the following are independent risk factors:
  - grade 3 or 4 intraventricular haemorrhage
  - cystic periventricular leukomalacia
  - neonatal sepsis
  - bronchopulmonary dysplasia for which mechanical ventilation was still needed at 36 weeks' postmenstrual age
  - antenatal steroids not given
  - postnatal steroids given to babies born before 32+0 weeks' gestation
- prevalence increases with decreasing gestational age.

See also the NICE guideline on cerebral palsy in children and young people under 25.

Motor problems

4. Be aware that children born preterm are at increased risk of motor problems, and that the following are independent risk factors:

- brain lesions (for example, grade 3 or 4 intraventricular haemorrhage, periventricular leukomalacia, infarct)
- necrotising enterocolitis that needed surgery
- neonatal sepsis

- severe retinopathy of prematurity.

5. Be aware that there is increased prevalence of developmental coordination disorder in children born preterm compared with the general population.

Intellectual disability

6. Be aware that children born preterm are at increased risk of intellectual disability, and that:

- the following are independent risk factors:
  - grade 3 or 4 intraventricular haemorrhage
  - cystic periventricular leukomalacia
  - neonatal sepsis in babies born before 28+0 weeks' gestation
  - necrotising enterocolitis that needed surgery in babies born before 33+0 weeks' gestation
  - bronchopulmonary dysplasia for which mechanical ventilation was still needed at 36 weeks' postmenstrual age in babies born before 28+0 weeks' gestation
  - severe retinopathy of prematurity in babies born before 28+0 weeks' gestation
  - small for gestational age
  - postnatal steroids given to babies born before 33+0 weeks' gestation
  - mother from a low-income or disadvantaged background

- prevalence increases with decreasing gestational age.

Special educational needs and educational attainment

7. Be aware that children born preterm are at increased risk of having special educational needs, and that the following are independent risk factors:

- brain lesions detected by ultrasound
- male sex.

8. Be aware that children born preterm are at increased risk of low educational attainment at the end of the early years foundation stage and at key stage 1, and that:

- prevalence of low educational attainment increases with decreasing gestational age
- there is increased risk of low attainment for reading and numeracy particularly in children born before 26+0 weeks' gestation
- the following are independent risk factors for delayed numeracy in children born before 32+0 weeks' gestation:
  - intracranial haemorrhage
  - bronchopulmonary dysplasia for which mechanical ventilation was still needed at 36 weeks' postmenstrual age.

Attention, impulsivity and hyperactivity

9. Be aware that children born before 33+0 weeks' gestation are at increased risk of symptoms of hyperactivity, impulsivity and particularly inattention at preschool and school ages.

10. Be aware that children born before 28+0 weeks' gestation are at increased risk of ADHD, and that male sex is an independent risk factor.

Autism spectrum disorder

11. Be aware that children born before 28+0 weeks' gestation are at increased risk of symptoms of social communication impairment, which may suggest a problem in the autism spectrum.

12. Be aware that children born preterm are at increased risk of autism spectrum disorder, and that:

- the following are independent risk factors:
  - intracranial haemorrhage in babies born before 34+0 weeks' gestation
  - male sex
- prevalence increases with decreasing gestational age.

Emotional and behavioural problems

13. Be aware that children born preterm are at increased risk of emotional and behavioural problems, particularly internalising behaviours and passivity, at preschool and primary school ages, and that the following are independent risk factors:

- major brain lesions (for example, periventricular leukomalacia, parenchymal lesions)
- mother with mental health problems
- mother younger than 25 years
- mother from a low-income or disadvantaged background.

Speech, language and communication

14. Be aware that children born preterm are at increased risk of speech, language and communication problems and disorders, and that the following are independent risk factors for language disorder:

- grade 3 or 4 intraventricular haemorrhage
- cystic periventricular leukomalacia
- male sex.

Feeding problems

15. Be aware that children born preterm are at increased risk of oro-motor feeding problems, and that this increased risk persists until at least 6 years of age in children born before 26+0 weeks.

Sleep problems

16. Be aware that children born preterm are at increased risk of sleep apnoea up to 6 years of age.

Visual impairment

17. Be aware that the prevalence of visual impairment increases with decreasing gestational age in children born preterm, and that the following are independent risk factors:

- grade 3 or 4 intraventricular haemorrhage with a shunt
- neonatal sepsis in babies born before 33+0 weeks' gestation
- retinopathy of prematurity requiring treatment.

Hearing impairment

18. Be aware that the prevalence of hearing impairment increases with decreasing gestational age in children born preterm, and that neonatal sepsis is an independent risk factor in babies born before 28+0 weeks' gestation.

Executive function problems

19. Be aware that children born before 32+0 weeks' gestation are at increased risk of executive function problems at preschool and school ages.

Developmental problems

20. Be aware that children born preterm are at increased risk of developmental problems, and that the following are independent risk factors:

- small for gestational age
- male sex
- mother from a low-income or disadvantaged background
- black, Asian or other minority ethnic group
- multiple pregnancy.

5 Information, support and developmental surveillance

5.1 Introduction

Few families are prepared for a premature birth and many are unaware of the possible consequences for the future development and health of their child. Families and carers may spend weeks or even months in the busy and often stressful environment of the NICU, having to share the care of their baby and make decisions in conjunction with a range of medical, nursing and ancillary staff. They often only have the opportunity to be with their baby for a continuous 24 hour care period immediately prior to discharge and only if the hospital accommodation supports them rooming-in. This means that many families and carers can feel ill-equipped to care for their baby following discharge and may experience high levels of anxiety after leaving the hospital. The evidence review on support (Section 5.1.2Table 35: Summary of included studies) aims to identify high quality support strategies for parents and carers, while the evidence review on information provision aims to establish what information should be available for families and carers to support the developmental needs of their child.

Detection of developmental problems and disorders in all children is achieved via the Healthy Child surveillance programme which incorporates screening programmes recommended by the National Screening Committee (NSC) and Public Health England. As children born preterm typically have a higher risk of developmental disorders and problems (see section), there may be delays in recognition of problems and disorders and access to required services which leads to poorer outcomes and higher costs in the long term. The evidence reviews on identification of developmental disorders and problems (see section 5.1.2.9) and delivering enhanced support and surveillance (see section 5.1.4) aim to identify which children are likely to benefit from additional developmental surveillance and determine what any additional assessments should include. The review on sharing information (see section 5.1.5) aims to identify what information should be shared between services involved in the developmental follow-up of children born preterm with a view to improving long-term outcomes for the child and effective planning of regional services.

5.1.1 Information provision

Review question:

What information about development and follow-up arrangements should be provided to parents and carers of preterm babies and to children and young people who were born preterm?

5.1.1.1 Description of clinical evidence

The aim of this review was to identify the information that should be provided to parents and carers about development and follow-up arrangements of babies, children and young people children who were born preterm.

Qualitative studies were selected relevant for inclusion for this review. We looked for studies that collected data using qualitative methods (such as semi-structured interviews, focus groups, and surveys with open-ended questions) and analysed the data qualitatively (including thematic analysis, framework thematic analysis, and content analysis etc.). Survey studies restricted to reporting that reported descriptive data that were analysed quantitatively were excluded.

Given the nature of qualitative reviews, categories/themes are were summarised from the literature and were not restricted to those identified as likely themes by the Committee.

For full details see review protocol in Appendix D:.

A total of 15 studies were identified for the inclusion in this review (Ardal 2011, Arockiasamy 2008, Brazy 2001, Brinchmann 2002, Doyle 2014, Gaucher 2011, Guillen 2012, Harvey 2013, Ignell Mode 2014, Keenan 2005, Nicolaou 2009, Niela-Vilén 2015, Padden 1997, Reyna 2006, Russell 2014).

The majority of included studies collected data by semi-structured interviews. One study presenteds the results of a workshop to discuss follow-up arrangements for preterm children (Doyle 2014).

Studies were carried out in the following countries:

- 4 studies were carried out in the UK (Harvey 2013, Nicolaou 2009, Padden 1997, Russell 2014)
- 4 studies were carried out in the USA (Brazy 2001, Guillen 2012, Keenan 2005, Reyna 2006)
- 3 studies were carried out in Canada (Ardal 2011, Arockiasamy 2008, Gaucher 2011)
- 1 study was carried out in Australia (Doyle 2014)
- 1 study was carried out in Sweden (Ignell Mode 2014)
- 1 study was carried out in Finland (Niela Vilén 2015)
- 1 study was carried out in Norway (Brinchmann 2002).
- 1 study was conducted during the antenatal period, to and obtained the views of pregnant women about the information they needed regarding preterm birth (Gaucher 2011).
- 3 studies were conducted whilst the infants were still admitted to in the neonatal unit (Ignell Mode 2014, Harvey 2013, Padden 1997). The time since the infants' admission ranged from 4 days to 53 days.
- 8 studies were conducted after the infants had been discharged from the neonatal unit (Ardal 2011, Brazy 2001, Brinchmann 2002, Guillen 2012, Keenan 2005, Nicolaou 2009 Niela-Vilén 2015, Reyna 2006). These were conducted at different times following the infants discharge:
- One 1 study was conducted 2-3 weeks after discharge (Reyna 2006).
- Two 2 studies were conducted during the first 3 months (Ardal 2011, Keenan 2005).
- One 1 study was conducted in the period up to around 1 year following discharge (Niela-Vilén 2015).
- Three 3 studies included children up to the age of around 24 months (Brazy 2001, Guillen 2012, Nicolaou 2009).
- One 1 study included infants who had been discharged (or died) between 1 and 8 years previously (Brinchmann 2002).
- Two studies included the parents of some infants who were still admitted, and some who had been discharged (Arockiasamy 2008, Russell 2014). One of these studies reports includeding infants between 6 weeks and almost 1 year of age (Russell 2014). The other does not report the age of the infants (Arockiasamy 2008).
- All studies included parents (or prospective parents) of preterm babies. Two of the studies also included healthcare professionals (Doyle 2014, Guillen 2012). Two studies included a small number of term babies with serious illness as well as preterm infants (Arockiasamy 2008: n=3/16 babies, Brinchmann 2002: n=3/26 babies).

Evidence from these are summarised in the clinical GRADE evidence profile below (Table 21). See also the study selection flow chart in Appendix F:, and exclusion list in Appendix G:.

5.1.1.2 Summary of included studies

Table 21: Summary of included studies for information provision to parents and carers of preterm infants

Study	Study design/methods	Participants/respondents	Aims of study	Comments/ Major limitations
Ardal 2011	Qualitative (semi-structured interviews)	n=8 mothers of preterm babies born at 24-29 weeks. All mothers spoke little or no English.	To determine the information needs of mothers who speak little/no English, and to assess their opinion of a “buddy” scheme (matching them with women who speak their language, who previously had a preterm baby admitted to the same NICU).	The relationship between the researcher and the study sample was not clearly described. It is unclear whether the researchers have managed their pre-understanding in relation to the analysis.
Arockiasamy 2008	Qualitative (semi-structured interview)	n=16 fathers of preterm (n=13) and term (n=3) infants admitted to NICU for >30 days Preterm infants ranged from 23 to 36 weeks gestation.	To describe the experiences of fathers regarding the care of their infant in a level III NICU	Sufficient data are not presented to support the findings. Unclear whether the researcher has managed his own pre-understanding in relation to the data analysis - the interviewer had been involved in the care of some of the study participants. 3 of the infants included in this study were babies born at term with serious illness.
Brazy 2001	Qualitative (semi-structured interview and questionnaire data with free text responses)	n=19 (n=15 mothers and n=4 fathers) Preterm infants born at 24 to 33 weeks gestation.	To discover how parents of premature babies obtain information and support. To identify the parents’ process of seeking information, the kind of information they sought and the resources they used to meet those needs.	The authors do not report whether data saturation was reached during the study. Data analysis, including coding and theme generation was not clearly described, nor validated by a second researcher. No quotations are used to support the findings. It is unclear whether the researcher managed their own pre-understanding in relation to the analysis.

Study	Study design/methods	Participants/respondents	Aims of study	Comments/ Major limitations
Brinchmann 2002	Qualitative (unstructured interview)	n=35 (4 mothers, 1 father and 15 couples) Parents of premature children (22-29 weeks) who had experience of life-and-death decisions regarding their infant	To generate knowledge about parents' participation in life-and-death decisions surrounding their premature infants	It is unclear whether data saturation was achieved with the sample included. It is unclear whether the analysis was independently verified. 3 of the 26 infants in this study were born at term with serious illnesses.
Doyle 2014	Qualitative (Workshop to discuss follow-up arrangements)	Not reported	To identify a framework for which children need follow-up, what outcomes should be assessed and how, where and when follow-up should be conducted.	Unclear how participants were selected for the discussion group. Only summary outcomes of the workshop are presented, not the discussion surrounding them.
Gaucher 2011	Qualitative (semi-structured interview)	n=5 Pregnant women admitted to hospital with threatened preterm labour at 26 to 30+2 weeks gestation.	To explore the concerns of mothers regarding premature labour, and the expectations that they have of the antenatal consultation with a neonatologist.	The relationship between the researcher and participants is not clearly described, and therefore it is unclear whether the researchers' pre-understanding has been managed appropriately. Insufficient data are presented to support the findings relevant to this review.
Guillen 2012	Qualitative (focus groups and semi-structured interviews)	n=31 clinicians n=30 parents of preterm babies born at <26 weeks.	To identify topics to discuss during antenatal counselling of prospective parents of a preterm infant, in order to develop a decision aid regarding delivery-room resuscitation	The method of sample selection for clinicians and parents is not clearly described. The relationship between the researcher and the study sample is not clearly described, and therefore it is unclear whether the researcher has managed their own pre-understanding.

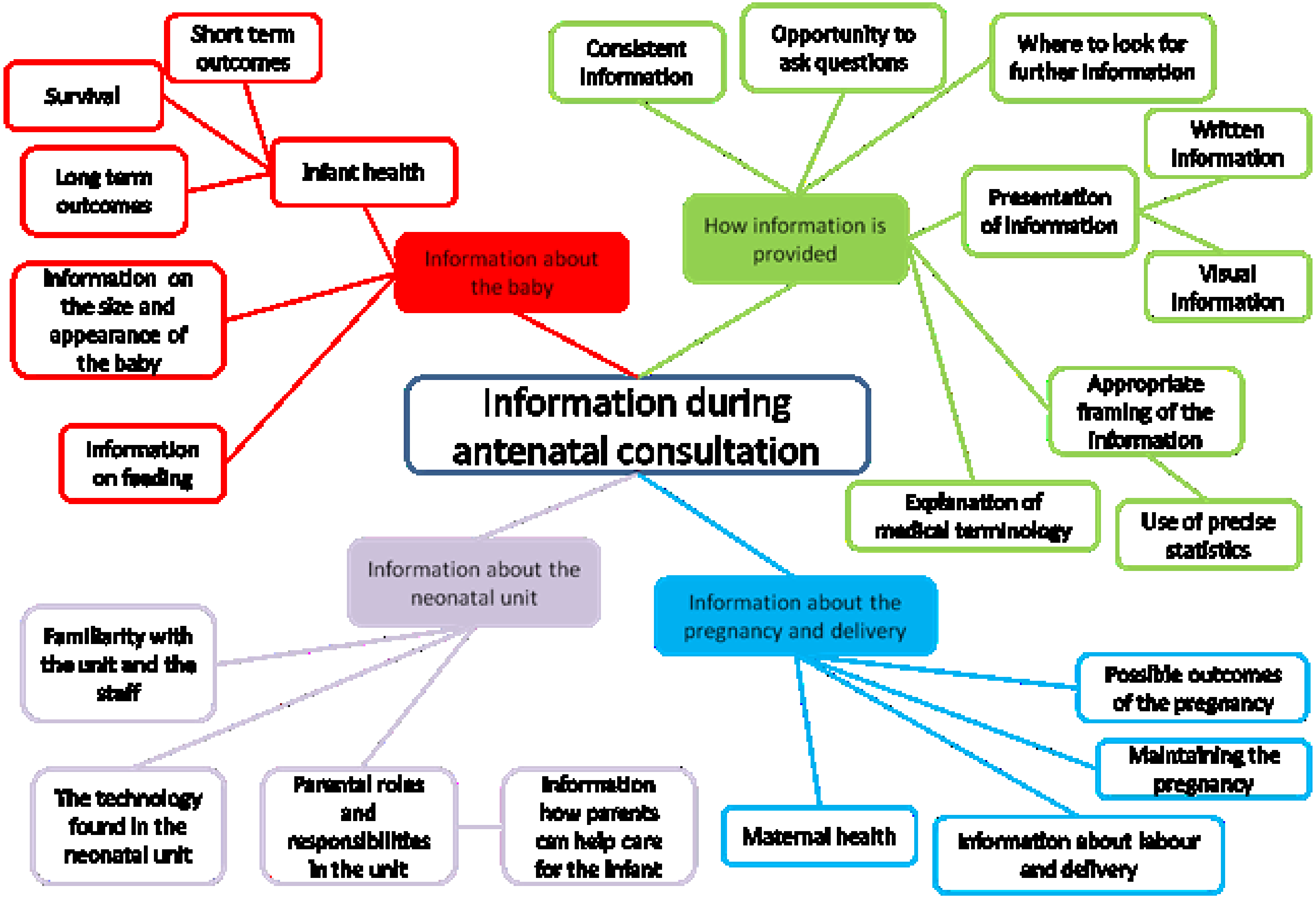
Study	Study design/methods	Participants/respondents	Aims of study	Comments/ Major limitations
				Insufficient data are presented to support the findings. Parents were interviewed many months after the birth of their preterm infant, and their perception of what information would have been useful during antenatal counselling is likely to have been affected by the experiences of their own child during their stay in NICU.
Harvey 2013	Qualitative (Semi-structured interview)	n=18 Parents of preterm babies born at 23 to 33 weeks.	To explore parental information needs during their baby's care in the neonatal unit.	The relationship between researcher and participants is unclear.
Ignell Mode 2014	Qualitative (semi-structured interview)	n=8 fathers of preterm babies born at 23 to 36 weeks.	To explore fathers' perceptions of the information they received while their infant was admitted to NICU	It is unclear whether data saturation has been achieved in terms of collection and analysis. Insufficient data are presented to support the findings of relevance to this review. The authors were involved in the care of the infants included in the study, and it is unclear whether they have managed their own pre-understanding in relation to analysis.
Keenan 2005	Qualitative (structured interview with some open-ended response questions)	n=15 mothers of preterm infants born at 23 to 28 weeks gestation.	To understand the views of mothers and counsellors regarding their roles in the decision making process for delivery-room resuscitation of premature infants.	The relationship between the researcher and participants is not clearly described therefore it is unclear whether their re-understanding has been appropriately managed. It is unclear whether data saturation has been achieved in terms of collection and analysis. The analysis of the qualitative data is not clearly described - it is unclear

Study	Study design/methods	Participants/respondents	Aims of study	Comments/ Major limitations
				how themes were generated and whether the analysis was independently verified. Insufficient data are reported to support the analysis.
Nicolaou 2009	Qualitative (semi-structured interview with some directive questions)	n=20 mothers of preterm infants born at 23 to 34 weeks.	To explore the early experiences of mothers regarding interaction with their premature infants. To identify information and support needs of mothers of premature infants.	Majority of participants were educated to degree level or higher, and may not be representative of the whole population.
Niela-Vilén 2015	Qualitative (analysis of posts on social media site)	n=30 mothers of preterm infants born at <35 weeks	To describe the perceptions, issues and problems relevant to mothers when they were breastfeeding their preterm infants.	It is unclear whether women would have posted comments about all issues that were of importance to them on this site, and relevant themes may have been missed. Due to the nature of the study it is difficult to determine whether data saturation has been achieved. The first author was also a midwife participating in the support group. It is unclear whether her pre-understanding has been managed appropriately when analysing the data.
Padden 1997	Qualitative (semi-structured interview)	n=36 mothers of preterm infants born at 27-34 weeks	To explore the subjective experiences of mothers of preterm infants in the early post partum period.	Unclear whether saturation in terms of data collection and analysis was achieved.
Reyna 2006	Qualitative (semi-structured interview)	n=27 mothers of preterm infants born at <32 weeks	To explore mothers' perceptions of their experiences in feeding their preterm infants in the early weeks after hospital discharge.	The relationship between the researcher and the study participants is unclear and therefore it is not clear whether the pre-understanding of the researchers was managed appropriately.

Study	Study design/methods	Participants/respondents	Aims of study	Comments/ Major limitations
				The authors state that no attempt was made to ensure data saturation.
Russell 2014	Qualitative (semi-structured interview)	n=39 parents of preterm babies born at <32 weeks	To explore parents views and experiences of the care of their very premature baby on NICU	It is unclear whether saturation has been achieved as this was an analysis of data collected for a different study. It is unclear whether the analysis has been independently verified.

1 Theme maps

Figure 1: Theme map 1


Information during antenatal consultation
Information about the baby
Infant health
Survival
Short term outcomes
Long term outcomes
Information on the size and appearance of the baby
Information on feeding
How information is provided
Consistent Information
Opportunity to ask questions
Where to look for further information
Presentation of information
Written information
Visual information
Appropriate framing of the information
Explanation of medical terminology
Use of precise statistics
Information about the neonatal unit
Familiarity with the unit and the staff
The technology found in the neonatal unit
Parental roles and responsibilities in the unit
Information how parents can help care for the infant
Information about the pregnancy and delivery
Possible outcomes of the pregnancy
Maintaining the pregnancy
Information about labour and delivery
Maternal health

1

1

Figure 2: Theme map 2

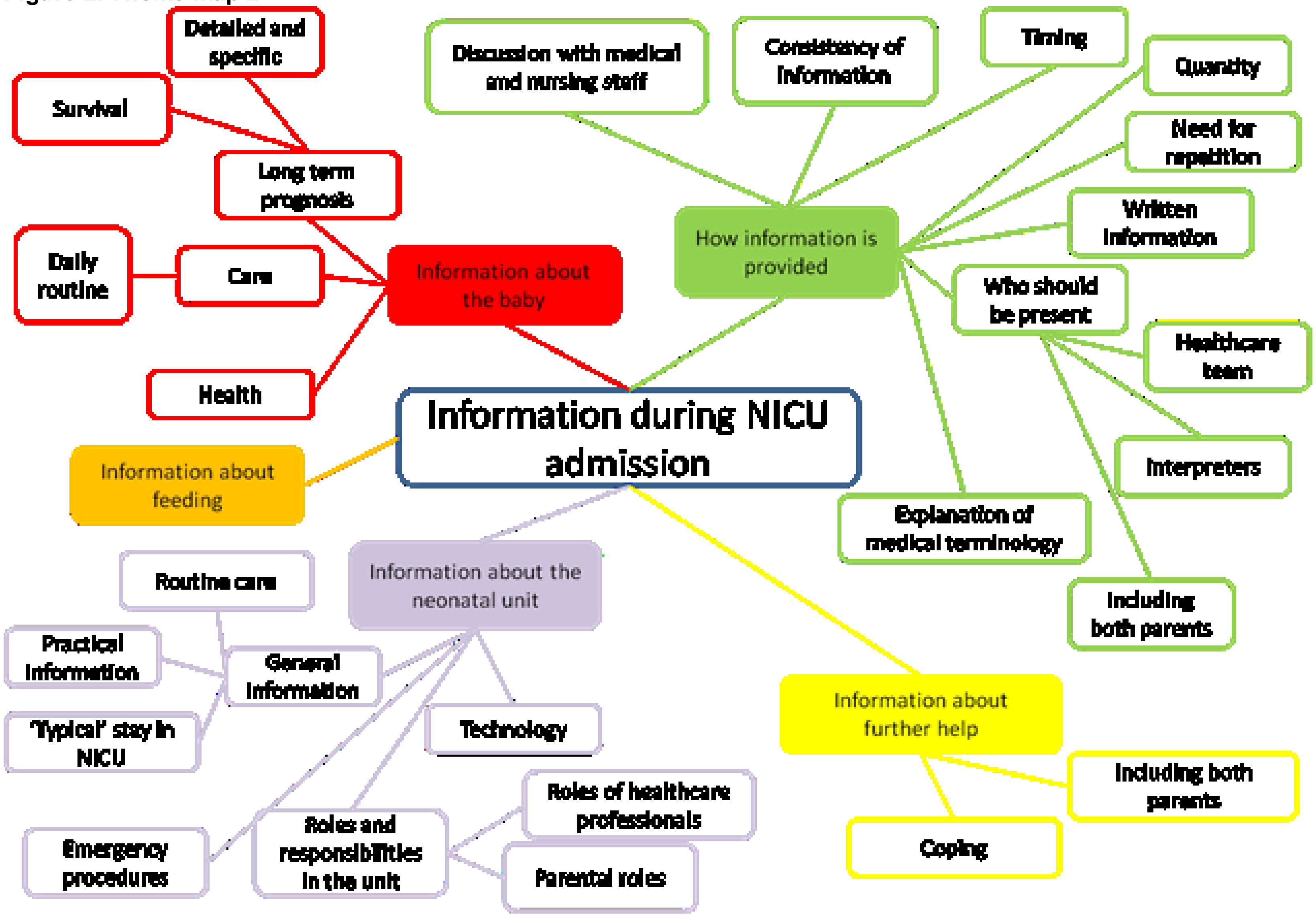
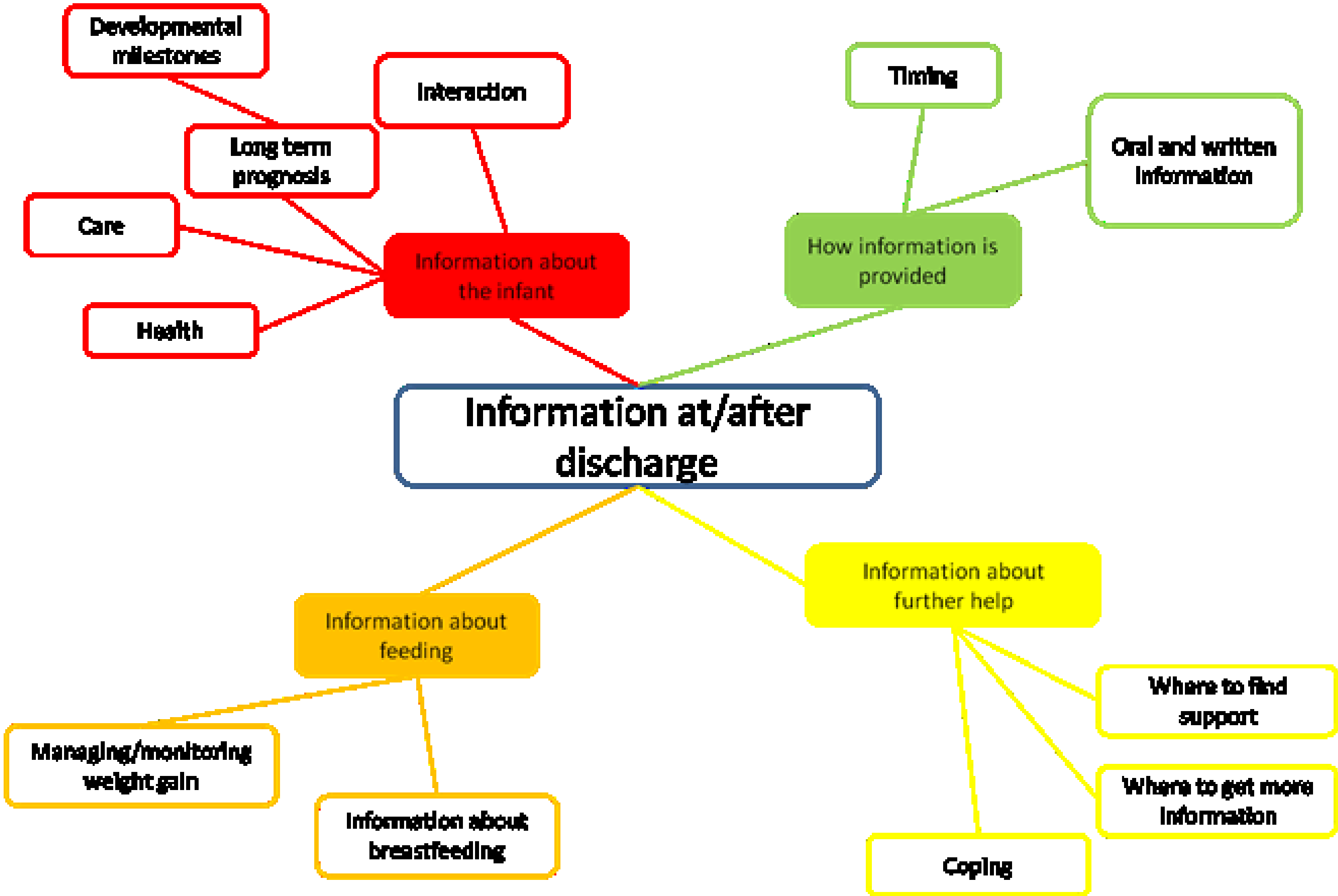


Figure 3: Theme map 3



5.1.1.3 Clinical evidence profiles

Information provision during antenatal consultation

Table 22: Theme 1: How information is provided

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: The format of information					
1 study (Guillen 2012)	Qualitative study (focus groups and interviews)	The authors identified that the majority of parents and clinicians felt that visual information (pictures, pamphlet or film) would be helpful when providing information. However, some parents worried that visual images may cause increased stress.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Keenan 2005)	Qualitative study (structured interviews with free response questions)	1 study found that mothers of preterm infants expressed a desire for written information.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Framing of information					
1 study (Guillen 2012)	Qualitative study (focus	The authors identified that the majority of parents and neonatal nurses thought that	Limitation of evidence	Major limitation	Low

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
	groups and interviews)	exact statistics should be provided regarding outcomes for premature infants. However, the majority of physicians felt that exact statistics should not be used.	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Subtheme 3: Terminology					
1 study (Keenan 2005)	Qualitative study (structured interviews with free response questions)	1 study found that mothers wanted less medical terminology. "When doctors would explain the words kept getting bigger and bigger; it would be helpful to have someone to break it down into more simple explanations"	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 4: Consistency of information from healthcare providers					
1 study (Gaucher 2011)	Qualitative study (semi-structured interviews)	1 study identified that pregnant women hospitalised for preterm labour wanted consistent information from healthcare providers	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Subtheme 5: Opportunity to ask questions					
1 study (Gaucher 2011)	Qualitative study (semi-structured interview)	1 study identified that pregnant women hospitalised for preterm labour wanted the opportunity to ask questions from the neonatologist. "Sometimes, I find it goes fast, that we don't have time to ask our questions. (...) It would	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
		only take the doctor an extra minute or two, but it would save us from being anxious and having unanswered questions"	Sufficiency or saturation	Sufficient	
Subtheme 6: Where to look for further information					
1 study (Brazy 2001)	Qualitative study (semi-structured interview)	The authors found that parents wanted to know where they could obtain further information if they required it.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Table 23: Theme 2: Information about the baby

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
Subtheme 1: Health of the infant					
1 study (Gaucher 2011)	Qualitative study (semi-structured interview)	1 study found that pregnant women hospitalised for threatened preterm labour wanted detailed, specific and precise information about short and long-term outcomes for their baby. Topics identified as being of importance were respiratory distress, neurological complications, sepsis, feeding difficulties and the possible length of hospitalisation	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Guillen 2012)	Qualitative study (focus	The authors identified that the majority of parents and clinicians felt that information	Limitation of evidence	Major limitation	Low

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
	groups and interviews)	should be provided about survival, short and long term morbidities. Specific topics included: lung disease and bronchopulmonary dysplasia retinopathy of prematurity sepsis intraventricular haemorrhage need for surgery for a patent ductus arteriosus	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Brazy 2001)	Qualitative study (semi-structured interview)	The authors found that parents wanted more information about the health of their infant.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Feeding					
1 study (Gaucher 2011)	Qualitative (semi-structured interviews)	1 study found that pregnant women hospitalised for preterm labour wanted information on breast feeding and feeding strategies for preterm infants.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Subtheme 3: The appearance of the baby					
1 study (Guillen 2012)	Qualitative study (focus groups and interviews)	The authors identified that parents and clinicians thought information on the expected size and appearance of a preterm infant would be useful.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

Table 24: Information about the pregnancy and the delivery

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 1: Maintaining the pregnancy					
1 study (Brazy 2001)	Qualitative study (semi-structure interview and questionnaire)	The authors found that parents were given information about how to continue with the pregnancy for as long as possible.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Possible outcomes of the pregnancy					
1 study (Brazy 2001)	Qualitative study (semi-structure interview and questionnaire)	The authors found that parents were given information about the possible outcomes of the pregnancy.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: The health of the mother					
1 study (Brazy 2001)	Qualitative study (semi-	The authors found that parents were given information about maternal health.	Limitation of evidence	Major limitation	Low

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
	structured interview)		Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 4: Information about labour and delivery					
1 study (Brazy 2001)	Qualitative study (semi-structured interview)	The authors found that parents would have liked more information about a "typical" labour and delivery for a premature baby.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Keenan 2005)	Qualitative study (structured interview with some free response questions)	Mothers explained that they appreciated explanations and knowing what would happen in the delivery room.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1

2 **Table 25: Information about the neonatal unit**

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Subtheme 1: Familiarity with the staff and the unit					
1 study (Ignell Mode 2014)	Qualitative study (semi-	1 study identified that fathers felt that the opportunity to visit the neonatal unit and	Limitation of evidence	Minor limitation	Moderate

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
	structured interviews)	meet the healthcare professionals before the infant was born was extremely useful. "What was fantastic was that we could meet a physician and a nurse from here already at the delivery unit, before the infant was born. That information was nearly the most valuable of it all"	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Appearance of the NICU					
1 study (Gaucher 2011)	Qualitative study (semi-structured interviews)	1 study identified that pregnant women hospitalised for threatened preterm labour would like information about the technology they could expect to see in the NICU.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Subtheme 3: Parental roles and responsibilities					
1 study (Gaucher 2011)	Qualitative study (semi-structured interviews)	1 study identified that mothers wanted information about what their roles and responsibilities would be if their baby was admitted to the NICU.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Gaucher 2011)	Qualitative study (semi-structured interviews)	1 study found that women wanted information on how they would be able to help care for their baby – whether they could touch or hold the infant.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

Information needs during the NICU stay

Table 26: Theme 1: How information is provided

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
Subtheme 1: The format of information					
1 study (Arockiasamy 2008)	Qualitative (semi-structured interview)	1 study identified that fathers would suggest having availability of written information about common medical conditions. One father also suggested having access to online material that he could discuss with the doctor.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Ignell Mode 2014)	Qualitative (semi-structured interview)	Fathers highlighted that written information was most useful when it was supported by oral information from the healthcare professionals.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Minor limitation	
			Sufficiency or saturation	Sufficient	
Subtheme 2: Explanation of terminology					

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
1 study (Ardal 2011)	Qualitative (semi-structured interview)	1 study conducted after discharge from NICU found that mothers (who did not speak English as a first language) found medical terminology confusing and difficult to understand. "But for me, no, the doctor never explained it in terms that I could understand. She used a lot of medical terminology, and for me that was the end of the world."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Ignell Mode 2014)	Qualitative (semi-structured interview	1 study identified that fathers viewed medical terminology as impeding the information flow	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Consistency of information from healthcare providers					
1 study (Arockiasamy 2008)	Qualitative (semi-structured interview)	1 study found that fathers wanted consistency in information provision and expressed a desire for a specific physician to be their primary contact, as well as an identified nurse or group of nurses.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Russell 2014)	Qualitative (semi-structured interview	One study reported that parents found conflicting information from different staff members was confusing and stressful. "Because you come in one day, say the day before, especially there was a guy there that, he promoted to hold her, literally whenever we was in, either of us, he	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
		would say, 'Hold her, it's the best thing you could do'. And then you'd come in the next day thinking 'oh yes, I get to hold her'. And you have a different nurse that says, 'no, no you've held her this week, you don't need to hold her for the rest of the week'… and then you'd almost feel devastated that you couldn't do that."	Sufficiency or saturation	Unclear	
1 study (Ignell Mode 2014)	Qualitative (semi-structured interview	1 study identified that fathers felt conflicting opinions and information from different staff members was confusing. A specific example was conflicting information about limits for alarms on medical equipment.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 4: Repetition of information					
1 study (Padden 1997)	Qualitative (semi-structured interview)	1 study found that mothers reported needing to ask questions repeatedly before they were certain of what was said. "We must have asked a hundred times what each machine does, and they always tell us again and again"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 5: Giving the right amount of information					
1 study (Harvey 2013)	Qualitative (semi-structured interviews)	1 study identified that parents varied with regard to how much information they wanted. "Too much knowledge can give you too many sleepless nights. She's in the right place, with the right care. I don't need to know anything else."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
			Sufficiency or saturation	Sufficient	
1 study (Arockiasamy 2008)	Qualitative (semi-structured interview)	1 study identified that some fathers felt they were given too much information. "There were times when it was too much information"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Ignell Mode 2014)	Qualitative (semi-structured interview	1 study identified that fathers felt a large quantity of information given at once was difficult to understand, as they were unable to identify which pieces of information were relevant to them.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Russell 2014)	Qualitative (semi-structured interview	One study identified that parents had difficulty taking in all the information that they were being given "I guess they do explain it to you when you first come in but they don't... you can't remember, you can't take stuff in. I think that follow-up explanation of everything… cos it took me ages to ask…"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Brinchmann 2002)	Qualitative (unstructured interview)	One study found that parents felt they should be asked how much information they wanted (with regard to life-and-death decisions about their infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
		"I think that on certain occasions the doctors should perhaps take the initiative to work out an agreement with parents such as: 'Shall I bother you with all the details that worry me, or shall I not say anything, or shall we try to find a good middle ground about what I tell you?' I had more than enough problems without having to worry about all the things that could go wrong."	Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 6: Who should provide information					
1 study (Padden 1997)	Qualitative (semi-structured interviews)	1 study found that many mothers received sufficient and good information from the nurses on the unit. "The nurses explain so you can understand" "They manage to put some time aside for small talk, they give us lots of information often even before we ask" However, others wanted more communication with the doctors. "I know there's nothing to worry about, but it would be nice occasionally, even once, to sit down and discuss things with his doctor"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 7: Who should be present when information is provided					
1 study (Ignell Mode 2014)	Qualitative (semi-structured interview)	Fathers felt that the daily medical round was a useful source of information. One father thought that the entire care team should be present to be updated about the infant's condition. "I think that information is the best, when the round is there....then everyone in the room knows what the physician said and the plan for the care...."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Padden 1997)	Qualitative (semi-	1 study found that mothers wanted a time to be set for both parents to meet the doctor together.	Limitation of evidence	Minor limitation	Moderate

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
	structured interviews)		Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Ardal 2011)	Qualitative (semi-structured interviews)	1 study conducted after discharge from NICU found that mothers who did not speak English as a first language found that serious misunderstanding could occur without appropriate use of interpreters. "I went outside the unit and called my husband to tell him that Michael was dying. Only after a nurse who speaks [my language] arrived, she helped me . . . As a result, I knew that Michael only had a minor infection."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Subtheme 8: Timing of information provision					
1 study (Brinchmann 2002)	Qualitative (unstructured interview)	1 study found that parents felt they needed adequate time to received important information (about life-and-death decisions regarding their infant) and also that this information should be provided to them when they are ready to receive such important news. "It was very important for us to get some time with these very busy doctors." "He just stood there and asked us whether we had thought about whether, should she get worse, she should be put on a respirator. He showed humility and asked in a pleasant manner, but I still felt that it was an awful imposition. I mean, if they are going to ask you whether to let your baby die, I think that they should have asked us to discuss it with them, asked if we wanted to talk about it."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1

1 **Table 27: Theme 2: Providing information about the NICU**

2

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Information about the unit and routine NICU care					
1 study (Ignell Mode 2014)	Qualitative study (Semi-structured interview)	Some fathers wanted written information about the NICU and neonatal intensive care, so that they would have an idea about what would happen during the rest of their infants stay.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Ardal 2011)	Qualitative study (Semi-structured interview)	Mothers (who did not speak English as a first language) used their linguistically matched parent-buddies for information about the NICU itself, such as where they could eat, use of the family room and where they could use the breast pump. “She told me that transferring to Level II means that the baby is improving. This is useful.”	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Harvey 2013)	Qualitative study (Semi-structured interview)	1 study identified that parents wanted information about when routine investigations were carried out on the unit. "I found a blob of ultrasound jelly on her head on Saturday, I said what's this? Ah Saturday, we do the routine scans"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
Subtheme 2: Information about technology					
1 study (Ignell Mode 2014)	Qualitative study (semi-	1 study found that several fathers wanted a complete introduction to the infants care	Limitation of evidence	Minor limitation	Moderate

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
	structured interviews)	space, as it was the natural location for discussions to occur, and could be perceived as frightening. They suggested a demonstration of some of the technical equipment and information about acceptable values on the monitoring equipment would make them feel less anxious.	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Brazy 2001)	Qualitative (semi-structured interview s)	1 study identified that parents wanted more technical information whilst their infants were acutely unwell in the NICU.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Information about emergencies					
1 study (Ignell Mode 2014)	Qualitative study (semi-structured interviews)	1 study identified that fathers wanted information about guidelines for emergencies to help reduce the anxiety they felt about not knowing what could be done.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 4: Information on roles and responsibilities					
1 study (Ardal 2011)	Qualitative study (Semi-structured interview)	Mothers (who did not speak English as a first language) used their linguistically matched parent-buddies for information about the structure of the NICU and what happens there.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
		"[The parent-buddy] gave us a lot of information... She explained to me what a primary nurse was, how the neonatologists work . . ."	Sufficiency or saturation	Sufficient	
1 study (Ignell Mode 2014)	Qualitative study (semi-structured interview)	One father wanted information about what the staff would expect from him whilst he was at the unit "If possible, more spontaneous information about what is expected from parents when they are here"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Table 28: Theme 3: Providing information about the infant

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: The health of the infant					
1 study (Brazy 2001)	Qualitative study (semi-structured interview)	1 study identified that parents of preterm infants wanted more information on the health of their infant during the NICU admission.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 2: Care of the infant					

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
1 study (Harvey 2013)	Qualitative study (semi-structured interview)	1 study identified that parents also valued information about day-to-day aspects of routine care for their baby. "...the information you want as a Mum, did he go through the night? Did he have all his feed? Was he whinging? The little things, which the staff don't think is important"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	
1 study (Russell 2014)	Qualitative study (semi-structured interview)	1 study identified that parents appreciated information about the baby's daily routine "And I think they were really, you know, explained everything. Every time we went to the incubator, whoever the nurse was on looking after her, you know, always explained how she'd been doing, how she'd been…they talked…it was really lovely".	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Brazy 2001)	Qualitative study (semi-structured interview)	1 study identified that parents of preterm infants wanted more information on the care of their infant during the NICU admission.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Long term prognosis					
1 study (Harvey 2013)	Qualitative study (semi-structured interview)	1 study identified that parents wanted detailed and specific information about how their baby was progressing, and the longer term prognosis. "To say she's fine doesn't really tell me anything at all"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
			Sufficiency or saturation	Sufficient	
1 study (Ignell Mode 2014)	Qualitative study (semi-structured interview)	1 study identified that fathers wanted early information about the care of their infant, and the possible course of events, to help them view the situation in the long term and bond with the baby. "I mean, the kind of information you want, will they survive or will they die, and that is probably difficult to answer...."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Table 29: Theme 4: Providing information about feeding

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
1 study (Russell 2014)	Qualitative study (semi-structured interviews)	1 study identified that parents wanted more information about breast feeding, and the facilities available. "I kept asking, when do I start expressing.... and it was about day 4, I think before they said to me, oh yea, here's a kit, go and express".	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1 **Table 30: Providing information about support**

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Coping					
1 study (Brazy 2001)	Qualitative study (semi-structured interviews)	1 study identified that parents of preterm infants wanted more information about coping.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 2: Including fathers					
1 study (Arockiasamy 2008)	Qualitative (semi-structured interviews)	1 study identified that fathers wanted to be involved when information was provided about support (such as social services), so that they too could access these facilities.	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

2

Information provision at or after discharge from NICU

Table 31: Theme 1: How to provide information

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
Subtheme 1: Providing information in different formats					
1 study (Doyle 2014)	Qualitative (workshop of health care professionals and parents)	The authors conclude that parents should be given information about the likely prognosis for their child, including written information.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Providing information at the right time					
1 study (Doyle 2014)	Qualitative (workshop of health care professionals and parents)	1 study identified that information (about prognosis) needs to be provided at the appropriate time to enable: decision making for life events (for example school choices, deferred or delayed school entry) screening and assessment for developmental disorders (for example Autism Spectrum disorder) monitoring for less visible medical conditions (for example hypertension).	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Table 32: Information about the infant

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
Subtheme 1: Infant health					
1 study (Brazy 2001)	Qualitative study (semi-		Limitation of evidence	Major limitation	Low

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
	structured interviews)	1 study identified that parents of preterm infants wanted more information about the health of their infant.	Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Infant care					
1 study (Brazy 2001)	Qualitative study (semi-structured interviews)	1 study identified that mothers wanted more information about the care of their infant.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Information about longer term prognosis and development					
1 study (Doyle 2014)	Qualitative (workshop of health care professionals and parents)	1 study identified that parents feel there is a lack of long term information about their infant.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Nicolaou 2009)	Qualitative (semi-structured interviews with some directive questions)	1 study identified that mothers wanted more information about developmental milestones, and how they differ for preterm infants.	Limitation of evidence	No limitation	High
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 4: Information about interacting with preterm infants					
1 study (Nicolaou 2009)	Qualitative (semi-structured interviews with some directive questions)	1 study identified that information given at the time of transition to home focussed on medical issues, rather than interaction. Mothers reported wanting more information on developmental play, appropriate toys and interaction. “We were given information but it was all very medical.....in terms of actually how to care for him and what to do when we got home there wasn’t really anything” “..we did have a resuscitation course. But that was pretty well it....I think that’s probably one of the things I found the hardest, the limited amount of information available regarding dealing with premature babies.”	Limitation of evidence	No limitation	High
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Sufficient	

Table 33: Information about feeding

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Managing feeding and weight gain appropriately					
1 study (Niela-Vilén 2015)	Qualitative (analysis of posts on a peer support group website)	1 study found that mothers wanted information on transitioning from bottle to breast feeding. "In what phase have you transferred from bottle to breast? Is there any age/weight-based guideline when you can try breastfeeding only? It is so much easier with a bottle, when you know for sure how much the baby is eating. Nevertheless, you can't perform test weighing at home, so how can I manage?"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
1 study (Reyna 2006)	Qualitative (semi-structured interviews)	1 study found that mothers wanted specific information on increasing feeds. "...basically how much to give him. When I should give it to him and if I feed him and he's still hungry should I give him more? How much more should I give him? How do I know when he's not hungry anymore, or if he's not hungry did he get enough milk in his feeding?" "They gave me instruction as every 3 to 4 hours ad lib. I didn't ask that right now she's on 2 ounces, when do I take her to 3 or 2.5 ounces?"	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: General guidance on breastfeeding					
1 study (Niela-Vilén 2015)	Qualitative (analysis of posts on a peer support group website)	1 study found that mothers felt unprepared for managing breast feeding at home and wanted individual support from the neonatal nurses. "...I was hoping for more information especially about how to manage at home, when the baby is used to the bottle, and what kind of problems may exist and how to manage them." "They didn't provide much support or instructions for home.'You can breastfeed once a day for a start'. That was the only advice I got."	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1 **Table 34: Finding further help**

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Coping					
1 study (Brazy 2001)	Qualitative study (semi-	1 study identified that parents wanted more information about coping.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
	structured interviews)		Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Availability of support					
1 study (Doyle 2014)	Workshop of health care professionals and parents	1 study identified that parents should be given information about where they can find longer term support for their child after discharge from the NICU.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Availability of further information					
1 study (Doyle 2014)	Workshop of health care professionals and parents	1 study identified that parents should be given specific website addresses to use for further information, to avoid the need to search the internet extensively.	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1

5.1.1.4 Economic evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

5.1.1.5 Evidence statements

5.1.1.5.1 During antenatal consultation

How information is provided

Format of information

Low quality evidence from one study (using semi-structured interviews and focus groups) identified that many parents of preterm infants and clinicians thought visual information (pictures, pamphlets or film) would be helpful during the antenatal consultation. However, some parents were concerned that visual images may cause increased stress. Low quality evidence from a second study, using a structured interview design, also found that mothers wanted written information to be provided.

Framing of information

Low quality evidence from a single study, using focus groups and semi-structured interviews, identified that parents of preterm infants and neonatal nurses thought that exact statistics should be provided on possible outcomes. In contrast, most physicians felt that exact statistics should not be used and favoured using statements such as “many” or “about a half”.

Terminology

Low quality evidence from a single study using structured interviews identified that mothers of preterm infants wanted better explanation of medical terminology.

Consistency

Moderate quality evidence from a single study using semi-structured interviews identified that pregnant women hospitalised for possible preterm labour wanted consistent information from healthcare providers.

Asking questions

Moderate quality evidence from a single study using semi-structured interviews identified that pregnant women hospitalised for possible preterm labour wanted the time and opportunity to ask questions.

Further information

Low quality evidence from a single study using semi-structured interviews identified that parents of preterm infants wanted to know where they could obtain further information.

Information about the baby

Health of the infant

Moderate quality evidence from one study using semi structured interviews found that pregnant women who were hospitalised for possible preterm labour wanted detailed, specific and precise information about the short- and long-term outcomes for their baby. In particular, women wanted information about respiratory distress, neurological complications, sepsis, feeding difficulties and the possible length of hospitalisation. Low quality evidence from a second study using focus groups and interviews of healthcare professionals and parents of preterm infants identified the important areas to discuss as survival and short- and long-term outcomes. Specific topics felt to be important were lung disease and bronchopulmonary dysplasia, retinopathy of prematurity, sepsis, intraventricular haemorrhage and the need for surgery for a patent ductus arteriosus. Low quality evidence from a third study using semi-structured interviews found that parents of preterm infants wanted more information about the health of their baby.

Feeding

Moderate quality evidence from a single study using semi-structured interviews of women hospitalised for threatened preterm labour identified that women wanted information on breast feeding, and feeding strategies for preterm infants.

The appearance of the baby

Low quality evidence from a single study (using semi-structured interviews and focus groups) found that healthcare professionals and parents thought that information on the anticipated size and appearance of a preterm baby would be helpful.

Information about the pregnancy and delivery

Maintaining the pregnancy

Low quality evidence from a single study using semi-structured interviews found that parents of preterm babies were given information on how to continue with the pregnancy for as long as possible during the antenatal consultation.

Possible outcomes of the pregnancy

Low quality evidence from a single study (using semi-structured interviews) found that parents of preterm infants had been given information on the possible outcomes of the pregnancy during the antenatal consultation.

Maternal heath

Low quality evidence from a single study using semi-structured interviews found that parents of preterm infants had been given information about maternal health during the antenatal consultation.

Labour and delivery

Low quality evidence from one study using semi-structured interviews found that parents of preterm infants would have liked more information about a “typical” labour and delivery for a premature baby. Very low quality evidence from a second study using structured interviews found that mothers of preterm infants valued explanations and knowing what would happen in the delivery room.

Information about the neonatal unit

Familiarity with the staff and the unit

Moderate quality evidence from a single study using semi-structured interviews found that fathers of preterm infants valued the opportunity to visit the neonatal unit and meet some of the staff before the birth of their baby.

Appearance of the NICU

Moderate quality evidence from a single study using semi-structured interviews found that pregnant women hospitalised for threatened preterm labour would like more information about the sort of technology they could expect to see on the neonatal unit.

Parental roles and responsibilities

Moderate quality evidence from a single study using semi-structured interviews found that pregnant women hospitalised for threatened preterm labour wanted information about what would be expected of them on the neonatal unit. The same study also found that women wanted information on caring for their infant and whether they could touch or hold the baby.

5.1.1.5.2 *During NICU admission*

How information is provided

Format of information

Moderate quality evidence from one study using semi-structured interviews found that fathers of preterm infants would have liked written information about common medical conditions affecting preterm infants. Access to online material was also suggested to be of possible use. Moderate quality evidence from a second study using semi-structured interviews highlighted that fathers of preterm infants viewed written information as being most valuable when it was supported by oral information.

Terminology

Moderate quality evidence from one study using semi-structured interviews found that mothers (of preterm infants) who spoke little or no English found medical terminology very confusing and difficult to understand. Moderate quality evidence from one further study using semi-structured interviews also reported that fathers of preterm infants viewed medical terminology as impeding the provision of information.

Consistency

Moderate quality evidence from one study using semi-structured interviews found that fathers of preterm infants wanted consistency in the information they received from different healthcare professionals. The same study found that fathers would have liked a specific physician and nurse (or group of nurses) to be identified as their primary contact. Moderate quality evidence from two further studies (both using semi-structured interviews) found that parents and fathers of preterm infants found conflicting advice and opinions from different healthcare professionals was confusing and stressful.

The need to repeat information

Moderate quality evidence from a single study using semi-structured interviews found that mothers of preterm infants needed to ask questions repeatedly before they could feel sure about what had been said.

The amount of information

Moderate quality evidence from four studies (all using semi-structured interviews) found that parents of preterm infants had difficulty taking in a lot of information at once, and sometimes felt that they received too much information. Moderate quality evidence from one further study (using unstructured interviews) identified that parents who had experience of life-and-death decision regarding their preterm infants felt that they should be asked how much information they wanted to receive.

Who should provide information

Moderate quality evidence from a single study using semi-structured interviews found that many mothers of preterm infants were happy with the information they received from the nurses on the unit. However, others wanted more communication with the doctors, especially if their baby was receiving medical assistance or there were other concerns.

Who should be present when information is provided

Moderate quality evidence from one study using semi-structured interviews found that fathers of preterm infants felt the daily medical round was a useful source of information, and that the entire care team should be present to be updated about the infant's condition. Moderate quality evidence from a second study using semi-structured interviews found that mothers of preterm infants wanted an opportunity for both parents to receive information together. Moderate quality evidence from one study using semi-structured interviews found that mothers who spoke little or no English could suffer serious misunderstandings unless an interpreter was present when information was provided.

Timing of information provision

Moderate quality evidence from a single study using semi-structured interviews found that parents who had experience of life-and-death decision regarding their preterm infants felt that adequate time most be allocated to receive such important information. The same study identified that parents needed to be prepared to hear difficult news, and finding the right time to do this was important.

Information about the neonatal unit

Information about the unit and routine NICU care

Moderate quality evidence from one study using semi-structured interviews found that fathers of preterm infants wanted written information about the neonatal unit and the process of neonatal intensive care, so that they would know what might happen during the time that their baby was in hospital. Moderate quality evidence from a second study using semi-structured interviews, found that mothers (of preterm infants) who spoke little or no English valued receiving practical information about the NICU (such as where to eat, where to express milk) as well as information about how their baby would progress through the different levels of care (from intensive care to intermediate care etc.). Moderate quality evidence from one further study using semi-structured interviews found that parents of preterm infants wanted information about when routine investigations/procedures would be carried out.

Information about technology

Moderate and low quality evidence from two studies using semi-structured interviews found that parents of preterm infants wanted more information about the technical equipment on the neonatal unit.

Information about emergencies

Moderate quality evidence from a single study using semi-structured interviews found that fathers of preterm infants would like to have information about emergency procedures to help them manage their anxiety.

Roles and responsibilities

Moderate quality evidence from a single study using semi-structured interviews found that mothers (of preterm babies) who spoke little or no English valued information about the roles of different healthcare professionals working on the neonatal unit. Moderate quality evidence from one further study using semi-structured interviews found that fathers of preterm infants wanted information about their own role, and the expectations that staff would have of them.

Information about the infant

The health of the infant

Low quality evidence from one study using semi-structured interviews identified that parents of preterm infants wanted more information on the health of their baby.

Care of the infant

Moderate quality evidence from two studies using semi-structured interviews found that parents of preterm infant valued having information about their baby's daily routine, and day-to-day care (such as how they were feeding and sleeping). Low quality evidence from one further study using semi-structured interviews identified that parents of preterm infants wanted more information on the care of their baby.

Long term prognosis

Moderate quality evidence from one study using semi-structured interviews found that parents of preterm infants wanted detailed and specific information about their baby's progress and long term prognosis. Moderate quality evidence from a second study using semi-structured interviews found that fathers of preterm infants wanted early information about the possible course of events for their baby.

Information about feeding

Moderate quality evidence from a single study using semi-structured interviews found that parents of preterm infants wanted more information about breastfeeding, and the facilities available.

Information about support

Information about coping

Low quality evidence from a single study using semi-structured interviews identified that parents of preterm infants wanted more information about coping when their infant was admitted to the NICU.

Including fathers

Moderate quality evidence from one study using semi-structured interviews identified that fathers of preterm infants wanted to be included in discussions about ongoing support services.

5.1.1.5.3 At or after discharge from neonatal intensive care unit

How to provide information

Format of information provision

Low quality evidence from a single study (reporting on a workshop comprising both healthcare professionals and parents of premature infants) identified that parents should be given oral and written information about the likely prognosis for their child.

Timing of information provision

Low quality evidence from a single study (reporting on a workshop comprising both healthcare professionals and parents of premature infants) identified that information about prognosis must be provided at the appropriate time during follow-up.

Information about the infant

Infant health and care

Low quality evidence from a single study using semi-structured interviews identified that parents of preterm infants wanted more information about the health and care of their child.

Information about longer term prognosis and development

Low quality evidence from a single study (reporting on a workshop comprising both healthcare professionals and parents of premature infants) identified that parents feel there is a lack of long-term information about their infant. High quality evidence from one further study using semi-structured interviews found that mothers of preterm infants wanted more information on developmental milestones and how they differ for preterm infants.

Interaction

High quality evidence from a single study using semi-structured interviews identified that mothers of preterm infants wanted more information at discharge on developmental play and interaction with their infant, rather than the sole focus to be on medical information.

Feeding

Managing feeding and weight gain appropriately

Moderate quality evidence from a single study (which analysed social media posts) found that mothers of preterm infants wanted more information on transitioning from bottle to breast feeding and how they would know if their infant was taking enough milk. Moderate quality evidence from a second study using semi-structured interviews found that mothers of preterm infants wanted specific information on how to increase feeds appropriately.

Breastfeeding

Moderate quality evidence from a single study (which analysed social media posts) found that mothers of preterm infants felt unprepared for managing breastfeeding at home and wanted more advice and support before leaving hospital.

Finding further help

Coping

Low quality evidence from a single study using semi-structured interviews found that parents of preterm infants wanted more information about how to cope after discharge from hospital.

Support

Low quality evidence from a single study (which reported on a workshop comprising healthcare professionals and parents of preterm infants) concluded that parents should be given information about where they can find longer term support for the child after discharge from the neonatal unit.

Further information

Low quality evidence from a single study (which reported on a workshop comprising healthcare professionals and parents of preterm infants) concluded that parents should be given information about helpful websites where they can access further information.

5.1.1.6 Economic evidence statement

A literature review of published cost-effectiveness analyses did not identify any relevant studies and no economic modelling was undertaken for this question.

5.1.1.7 **Evidence to recommendations**

5.1.1.7.1 Relative value placed on the outcomes considered

The aim of this review was to identify what information should be provided to the parents and carers of children who were born preterm and themes in qualitative reviews are driven by the included evidence. Whereas all the themes identified were considered important, some were more relevant than other for this particular guideline. The themes identified in the evidence that the Committee considered most important were the timing and format of information provision, consistency of information provision, and content of information provided (topics, level of detail, terminology). These themes were considered important because it was known that the engagement and involvement of parents and carers improves outcomes for the child and because providing information reduces confusion and unnecessary stress and anxiety among parents and carers, which in turn can also improve the outcomes for the child. The Committee agreed that the most crucial time points for providing information are the different transition points, for example, when the child is transferred between units, discharged from hospital, or when the child is entering education services.

5.1.1.7.2 Consideration of clinical benefits and harms

Information about prematurity and the potential consequences it may have on the development of a child should be provided to parents and carers according to their individual needs. The evidence showed that it iwas important for parents to receive information about the possible prognosis for their child regarding developmental outcomes. The Committee agreed that the evidence for risk of developmental disorders and problems in children born preterm identified in the guideline should be made available to parents where this was available and explained in the context of their child. The Committee acknowledged the imprecision, uncertainty and lack of evidence available on the developmental disorders and problems among children born preterm (please see section 4.2) and discussed the potential harms that providing information that is uncertain might have on the parents and carers. The Committee agreed that it was important to maintain a balance in providing information that was factual and honest but that would not cause unnecessary worry and anxiety.

The Committee also discussed how the level of detail that parents and carers would like to have may differ as shown by the evidence. The Committee agreed that information provided to the parents and carers should be tailored according to their individual needs taking into account their level of education, potential language barrier, cultural and spiritual needs. Some parents and carers may, for example, wish to receive information containing detailed, exact statistics and medical terminology whereas some parents may find this unhelpful and confusing. This was also highlighted in the evidence. The health care professionals should also have a consistent message when discussing with parents and carers in order to avoid confusion among parents and carers. This should include information on discharge which should be shared consistently amongst healthcare professionals. This highlighted the importance of good communication between healthcare professionals caring for the family. The Committee discussed how information about developmental follow-up should primarily be given by healthcare professionals who have expertise in developmental follow-up of children born preterm.

The discharge plan should be developed and shared with the parents and carers. Before discharge, the parents and carers should be given information to learn techniques and skills to care for their child at home, including feeding, sleeping, play and interaction with the child. The evidence showed that some parents might feel isolated and anxious with the preterm, particularly for reasons such as fear of infection once discharged. The Committee agreed that the parents and carers should be reassured and given information and support regarding the risk of infection. The Committee also discussed that the parents should also be explained that corrected age should be used when assessing the development of the child for the first two years. This was because after 2 years the impact of weeks of prematurity would become less important.

The Committee agreed that was important to provide parents and carers with sufficient information and support regarding caring for the child at home. The evidence in this review showed that parents and carers would have liked to have advice on daily activities with the preterm child after discharge, for example, on feeding and interaction with the child (please see section 5.1.2.8).

The Committee agreed that the parents and carers should be clearly informed about the enhanced support and surveillance programme, what it entailed and why it was needed for their child, including information about the process for arranging the follow-up. They should be given information about a point of contact whom they could ask questions about follow-up or any other concerns. The Committee also agreed that the parents and carers should be given information about the routine postnatal care and the Healthy Child Programme. The Committee considered this important because parents and carers of children born preterm who were still in neonatal services may not be aware that they still fell within the remit of other routine services.

Information about opportunities for peer support should be made available for the parents and carers, which could include local peer support groups or online-based groups. The evidence in this review brought up the importance of peer support and the Committee discussed that it was harder to establish peer support connections with a preterm baby because participation at antenatal and postnatal support groups or mother and baby groups may not be possible, therefore providing information about possibilities of peer support was important at this stage.

The evidence in this review emphasised the importance of having different types of information available, including written information, visual information and oral information. The Committee discussed that there were various types of information that could be provided to parents and carers depending on preference. The Committee members discussed how online information was thought to be generally more useful than printed leaflets although printed information should also be made available. The Committee also agreed how important it was that patient information was found easily online and kept up to date. The Committee agreed that the health care providers should be prepared to provide information orally during visits or on the phone.

5.1.1.7.3 *Consideration of economic benefits and harms*

A systematic review of the economic literature was conducted but no relevant studies were identified which were applicable to this review question.

The economic implications of this topic were considered but not thought to be substantial. The provision of information does have resource implications as it requires time to be spent by the health care professionals providing it. However, the majority of the recommendations made reflect current best practice and so the recommendations are not expected to require a substantial increase in resources.

There is the potential for inconsistency in practice though with the information that parents receive varying across service providers. Therefore, it is possible that there could be increased costs for service providers that are not currently providing the information outlined in the recommendations.

Any increase in the time spent by clinicians in providing information as a result of the recommendation was thought likely to be cost-effective as the increased costs would be offset by potential cost savings and effectiveness gains. There could be cost savings associated with educating parents upfront perhaps meaning that they would be less likely to require additional support when concerns arise. There could be effectiveness gains too as parents become better educated and are able to recognise problems when they arise leading to earlier identification and management. There could also be effectiveness gains associated with reducing parent anxiety and providing reassurance.

5.1.1.7.4 *Quality of evidence*

Low to high quality qualitative evidence was included in the review. The main reasons for downgrading of evidence included likely bias in the selection of participants, lack of saturation in the data analysis, unclear relationship of the investigator with the participants, insufficient data to support findings and unclear hypothesis/model generated. A variety of the themes regarding information provision that parents and carers reported as helpful or unhelpful were reported across the studies, however, due to uncertainty in data saturation or sufficiency in some findings, this evidence should be interpreted with caution.

5.1.1.7.5 *Other considerations*

The Committee discussed the concept of high quality information provision (with a view to empowering shared decision-making) as also outlined in the guideline on Patient experience in adult NHS services: improving the experience of care for people using adult NHS services. They noted how the themes were consistent with those found in this review (individualised approach, variety of formats, sensitive to cultural, spiritual or religious belief, timing of information, need for consistent message, and promotion of shared decision-making). The Committee agreed that the principles were relevant to this population and should be considered when providing information to children, parents and carers.

The Committee also considered how communication needs may differ according to the English language comprehension skills of the children, families and carers. They discussed how it is important to establish effective ways of communicating and explore different ways of improving communication (for example, using communications aids, or involving an interpreter).

5.1.1.7.6 *Key conclusions*

The Committee reviewed the themes identified by the evidence review and concluded that information provision should be tailored to individual family circumstances, taking into account the child's potential developmental needs, the need for consistency in information sharing among healthcare professionals, their level of education, any social care needs they have, as well as any cultural, spiritual or religious beliefs. They stressed how this aspect of care could greatly influence the overall experience of the family and support the aim of early detection of developmental problems or disorders. Lastly, the Committee recognised that the engagement and involvement of parents and carers was crucial because it can improve developmental outcomes for the child.

5.1.1.8 Recommendations

See Section 5.2.

5.1.2 Support of children who are born preterm

Review question:

What support do parents and carers report was or would have been helpful to them in the care of infants who were born preterm both at discharge and during subsequent follow-up?

5.1.2.1 Description of clinical evidence

Qualitative studies were eligible for inclusion in this review. We looked for studies that collected data using qualitative methods (such as semi-structured interviews, focus groups, or surveys with open-ended questions) and analysed data qualitatively (including thematic analysis, framework thematic analysis or content analysis). Survey studies that analysed descriptive data quantitatively were excluded.

Categories and/or themes were obtained from the literature.

For full details see review protocol in Appendix D:.

A total of 20 studies (Benzies 2015; Chiu 2012; Frisman 2012; Garel 2006; Harrison 1997; Lasby 2004; Lee 2009; Lee 2013; Little 2015; May 1997; Neu 2008; Nicolau 2009; Niela-Vilen 2015; Philips-Pula 2013;; Reyna 2006; Sommer 2015; Thomas 2009; Turner 2013; Vasquez 1995; Whittingham 2014) were identified for the inclusion in this review.

The majority of studies obtained data via semi-structured interviews or focus groups. The most common data analysis method employed across studies was thematic analysis.

Studies were carried out in the following countries:

- 1 in the UK (Nicolaou 2009)
- 3 in Canada (Chiu 2012; Harrison 1997; Thomas 2009)
- 7 in the USA (Benzies 2014; Little 2015; May 1997; Neu 2008; Philips-Pula 2013; Reyna 2006; Vasquez 1995)
- 2 in Taiwan (Lee 2009; Lee 2013)
- 1 in Finland (Niela-Vielen 2015)
- 1 in New Zealand (Sommer 2015)
- 1 in France (Garel 2006)
- 1 in Sweden (Frisman 2012)
- 2 in Australia (Turner 2013; Whittingham 2014)

Studies were carried out in the following settings:

- 7 studies were at NICU discharge (Harrison 2009; Lasby 2004; Little 2015; Nicolaou 2009; Sommer 2015; Turner 2013; Whittingham 2015).
- 1 study was after NICU discharge to low risk unit (Sommer 2015)
- 15 studies were at home, after NICU discharge (Benzies 2014; Chiu 2012; Frisman 2012; Garel 2006; Lasby 2004; Lee 2013; Little 2015; May 1997; Neu 2008; Nicolau 2009; Niela-Vilen 2015; Philips-Pula 2013;; Thomas 2009; Turner 2013; Whittingham 2014).

Evidence from these are summarised in the clinical GRADE evidence profile below (Table 35). See also the study selection flow chart in Appendix F:, and exclusion list in Appendix G:.

5.1.2.2 Summary of included studies

Table 35: Summary of included studies

Study	Study design/methods	Participants/respondents	Aims of study	Comments
Interviews/focus-groups				
Benzies 2014 (USA)	Qualitative (Semi-structured Interview)	n=85 (fathers from one centre) Infants born at 35 weeks GA	To explore the father's perceptions of the positive and negative aspects of his experiences that influence interactions with his infant and his perceived needs for support in his role.	Relationship between the researcher and the selected sample was not clearly described Unclear achievement of data saturation Unclear how categories/themes derived for thematic analysis Unclear saturation in terms of analysis Unclear validation of independent validation Unclear hypothesis/theory/model generated
Chiu 2012 (Canada)	Qualitative (Videotape/interview)	n=12 mother-infant dyads Infants born at <37 weeks GA (24-36 weeks range)	To explore the changes in mother-infant interaction of preterm infants and their mothers who received home care occupational therapy	Unclear relationship between the researcher and the selected sample Unclear role of researcher Unclear achievement of saturation (data collection or analysis) Unclear independent validation of the analysis
Frisman 2012 (Sweden)	Qualitative (Interview)	n=11 women who were grandmothers to preterm infants who were born at 25 to 34 weeks GA	To explore and describe the experience of becoming a grandmother to a preterm infant, and balancing their involvement with care of the infant	Unclear saturation in data collection Unclear if a theory or model was generated
Garel 2006 (France)	Qualitative (Semi-structured interview)	n=20 mothers of children born preterm between 26-32 weeks GA	To assess qualitatively mothers' physical and psychological health, their perception of their child's health and development, and their	Unclear saturation during data collection or analysis

Study	Study design/methods	Participants/respondents	Aims of study	Comments
			difficulties with childcare from 2 months post discharge to 1 year after a very preterm delivery	
Harrison 2009 (Canada)	Qualitative (in depth interview)	n=20 women who were mothers of a preterm infant born at ≤35 weeks GA	To explore women's perceptions of barriers to support during family caregiving in a Canadian setting	The study compared two groups, women caring for adults with cognitive impairment compared with women who were caring for infants born preterm Saturation of data was not clearly described, as well as saturation of analysis The analysis was not clearly described Unclear if the process of analysis was thematic Unclear if data sufficient to support findings Unclear if the analysis was validated independently Unclear hypothesis or theory or model generated
Lasby 2004 (Canada)	Qualitative component of a randomised controlled trial (Focus group interviews of a convenience sample of mothers from the trial)	n=14 mothers of infants who weighed <1250g	To explore the experiences of mothers who received support from the neonatal transition care programme, after discharge of their infants from hospital	The study was a qualitative component of a randomised trial for NTCP compared with PHN support Method of selection was not clearly described The relationship between the researcher and the selected sample was not clearly described Data collection procedure was not described Roles of the researchers are not clearly described Unclear if saturation had been achieved

Study	Study design/methods	Participants/respondents	Aims of study	Comments
				Analysis method not clearly described Unclear how categories/themes derived Unclear if sufficient data was presented to support findings Unclear if saturation in terms of analysis was achieved Unclear if researcher managed own pre-understanding in relation to analysis Unclear if analysis was independently validated
Lee 2009 (Taiwan)	Qualitative (in depth interview)	n=31 mothers of very low birth weight infants born between 23-33 weeks GA	To report the breastfeeding experience of mothers with very low birth weight babies	The data collection procedure was described, but not according to a theoretical framework Unclear if data saturation was achieved in the analysis Unclear hypothesis, theory or model generated from the results
Lee 2012 (Taiwan)	Qualitative (in depth interview)	n=19 parents (11 mothers, 8 fathers) of infants born very low birth weight ranging from 620-1470g	To explore the perceptions and experiences of Taiwanese parents in coping with the unfolding evidence of a disability, their response to the official diagnosis, and their views about their child's developmental disability	Unclear saturation during data collection Unclear if analysis was independently validated
Little 2015 (USA)	Qualitative (Focus groups and interviews)	n=44 parents (10 focus groups at 5 sites (each group with 3 to 7 participants)	To explore existing barriers and challenges to early intervention referral, enrolment, and service provision for very low birth weight (<1500g) infants	Unclear if data collection saturation was achieved Insufficient data presented to support findings
May 1997 (USA)	Qualitative (Semi-structured interview)	n=14 mothers of infants born preterm between 23-34 weeks GA	To explore the process mothers use to seek help in providing care to low birth weight infants	Roles of the researchers were not clearly described regarding analysis of data

Study	Study design/methods	Participants/respondents	Aims of study	Comments
				Unclear if data collection saturation was achieved Unclear if a theory or hypothesis was generated from the results/findings
Neu 2008 (USA)	Qualitative (Interview)	n=12 adolescent mothers of infants born between 32 to 35 weeks GA	To examine early adaptation challenges and strengths of young mothers with preterm infants	Unclear saturation during data collection Unclear if sufficient data presented supported findings Unclear saturation during data analysis Unclear if analysis was validated independently
Nicolau 2009 (UK)	Qualitative (Interview)	n=20 mothers who met the inclusion criteria and volunteered to participate in the study, whose infants were born between 23 to 34 weeks GA	i.To explore thoughts and experiences of mothers concerning their early interactions with their preterm infants ii. To explore the perceived support and information needs of mothers of preterm infants	Unclear saturation in terms of analysis Unclear if analysis has been independently validated
Niela-Vielen 2015 (Finland)	Qualitative (Interview)	n=30 mothers of preterm infants born at < 35 weeks GA, and 3 peer supporters	To explore mothers views and perceptions of issues and problems that were relevant to them when they were breastfeeding their preterm infants	Unclear if data collected according to a theoretical framework Unclear saturation was achieved during data collections Unclear if saturation was achieved during data analysis Unclear if the analysis was independently validated The first author was a midwife participating in the peer-support group, which may have some influence on her perception of breastfeeding.

Study	Study design/methods	Participants/respondents	Aims of study	Comments
Philips-Pula 2013 (USA)	Phenomenological (In depth interview)	n=8 mothers of preterm infants born between 24 to 34 weeks GA	To examine the experiences of mothers of preterm infants during the first 6 months at home following discharge from NICU	The relationship between the researcher and the selected sample is unclear Unclear if data saturation achieved during data collections Not enough data to support findings Unclear if data saturation achieved during the analysis Unclear hypothesis, theory or model generated from findings
Reyna 2006 (USA)	Qualitative (Interview)	n=27 mothers of preterm infants born at 35 weeks GA	To explore mothers' perception of their experiences in feeding their preterm infants in the early weeks after hospital discharge	Unclear if saturation was achieved during data collection Insufficient data to support results/findings Unclear if saturation was achieved during data analysis Unclear if the analysis was independently validated Unclear hypothesis, theory or model generated
Sommer 2015 (New Zealand)	Qualitative (Interview)	n=6 parents (5 mothers and 1 father) of preterm infants born between 23+6 to 29 weeks GA	To investigate parents' perceptions of preterm infants transfer, to provide neonatal clinicians with insights to facilitate optimal service provision	Unclear relationship between researcher and selected sample Data collection procedure not clearly described and not according to a theoretical framework Roles of the researchers are not clearly described Unclear saturation during data collection Analysis description is vague Partial explanation of thematic analysis used Saturation during data analysis unclear

Study	Study design/methods	Participants/respondents	Aims of study	Comments
				Unclear if researcher managed pre-understanding in relation to the analysis Unclear if data was independently validated in the analysis
Thomas 2009 (Canada)	Qualitative (Interview)	n=5 fathers of very low birthweight infants	To explore the factors that first time fathers of a very low birth weight infant perceive to influence their parenting self-efficacy beliefs	The relationship between the researcher and selected sample not clearly described Roles of the researcher not clearly described Unclear saturation during data collection Unclear saturation during data analysis Unclear if analysis validated independently
Turner 2013 (Australia)	Qualitative (Interview)	n=9 mothers who consented to first interview at NICU and second after discharge, infants were born at 24 to 31 weeks GA	To explore emotional reactions during the transition to home from the NICU for parents who participated in a support group	Unclear saturation during data analysis Unclear if researcher managed own pre-understanding in relation to analysis Unclear if analysis was independently validated
Vasquez 1995 (USA)	Qualitative (Interview)	n=14 parents of very low birth weight infants of <1500g	To describe parents' method of adaptation to the problems of caring for a very low birth weight infant at home	The relationship between researcher and selected sample not clearly described Roles of the researcher not clearly described Unclear if saturation achieved during data collection Unclear if sufficient data supported findings Unclear if saturation achieved during data analysis

Study	Study design/methods	Participants/respondents	Aims of study	Comments
				Unclear if analysis independently validated
Whittingham 2014 (Australia)	Qualitative (Focus groups)	n=18 parents of children born very preterm (≤32 weeks gestation)	i. To identify from the parents' own perspective the unique aspects of parenting an infant born very preterm ii. To asses parental preferences for support including opinions of a new tailored parenting intervention	Unclear if saturation achieved during data collection Unclear if analysis was independently validated

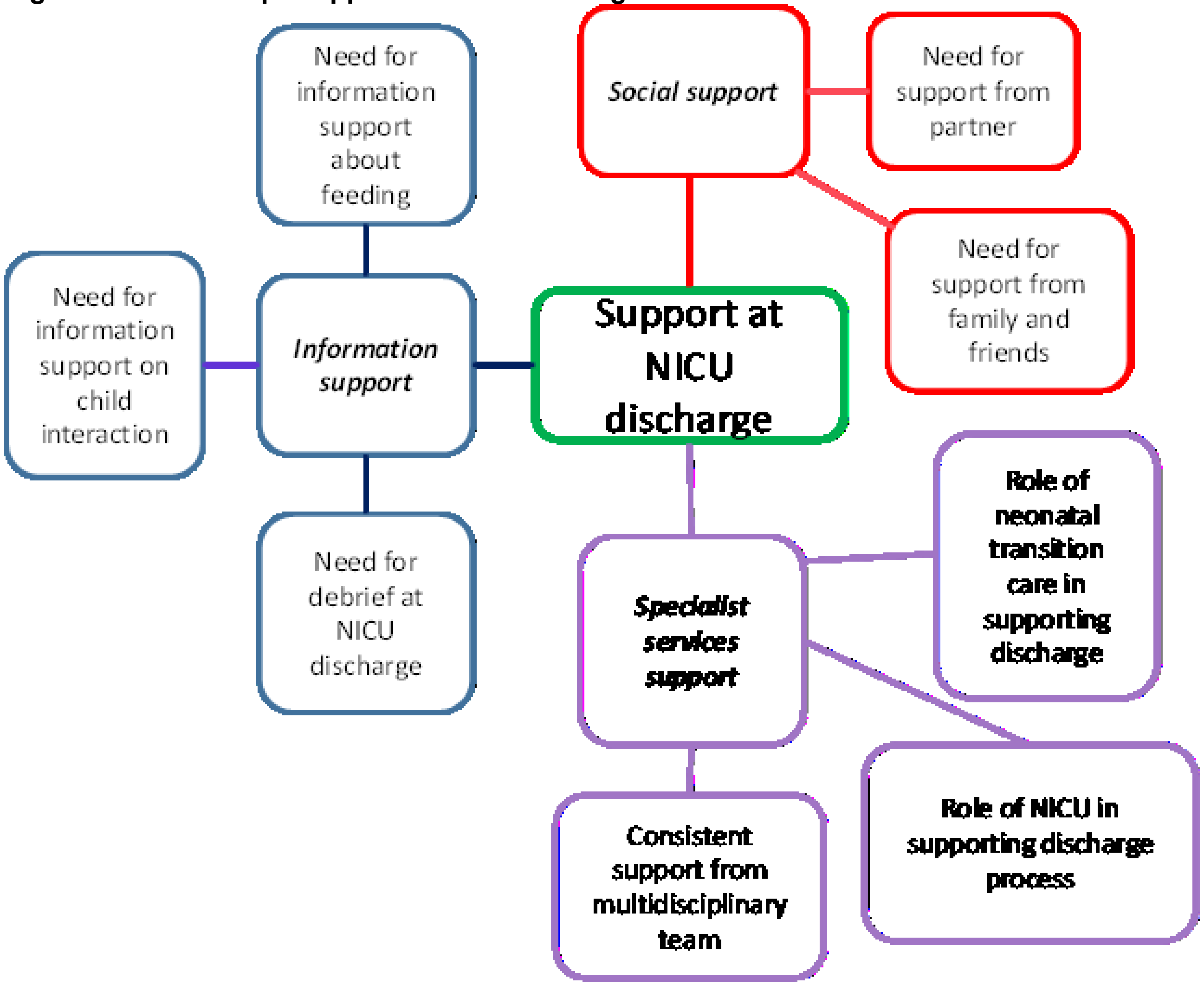
5.1.2.3 Theme maps

Two theme maps were generated according the settings where the studies were carried out:

Figure 4: Theme map: Support at NICU discharge - upport needs perceived by parents/carers at NICU discharge

Figure 5: Theme map: Support after NICU discharge - needs perceived by parents/carers after NICU discharge

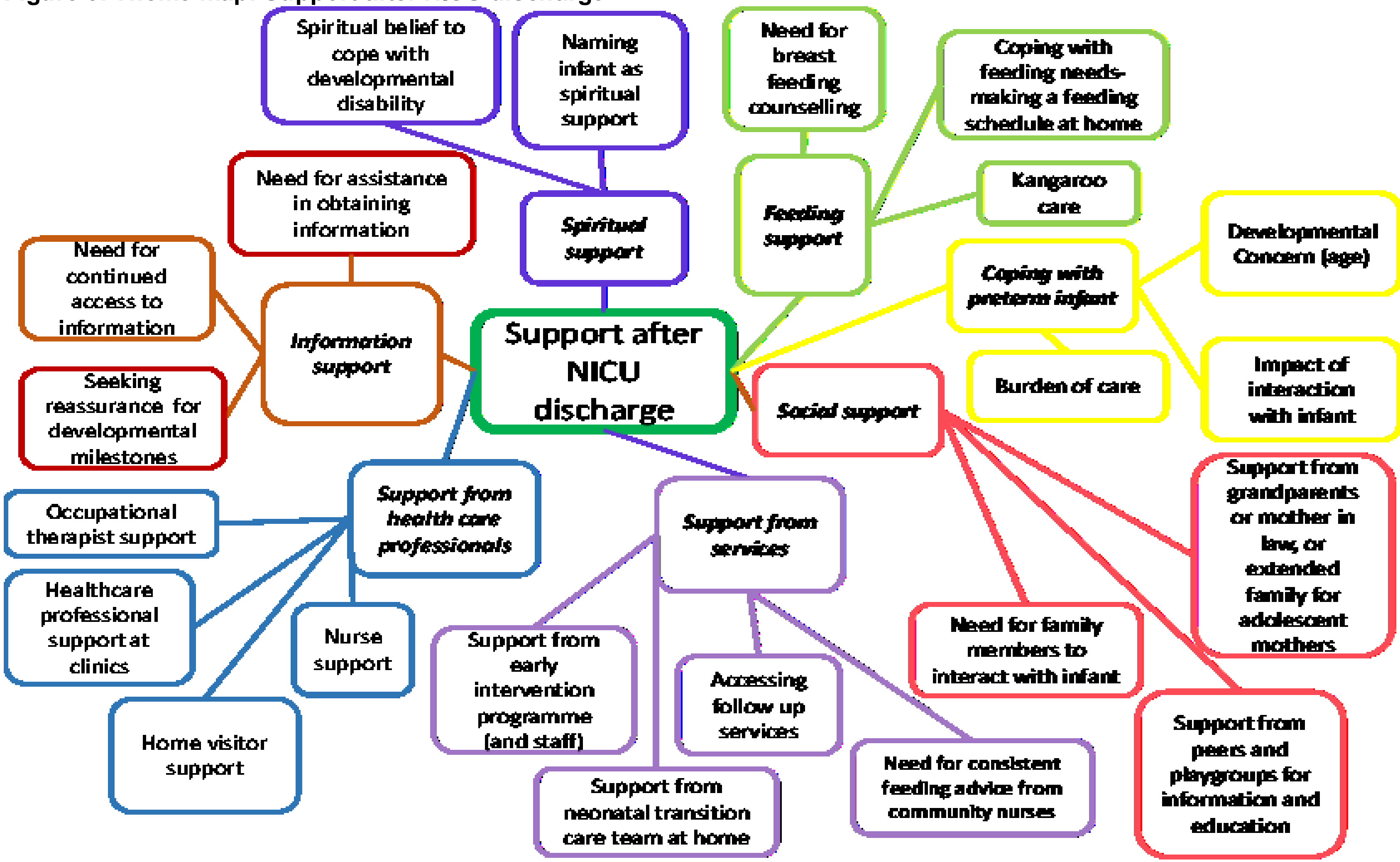
Figure 4: Theme map: Support at NICU discharge



1

1

Figure 5: Theme map: Support after NICU discharge



5.1.2.4 Clinical evidence profiles

Table 36: Theme 1: Social support

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Need for support from partner					
1 study (Harrison 2009)	In depth interview	1 study conducted shortly after discharge from NICU (Canada) showed that mothers frequently excused their husbands from providing help with household duties: "If he's lying on the couch with a very sleepy look on his face and says 'don't worry dear, I'll clean it up', I'll say 'don't worry about it', because I know his heart is not in it" (mother)	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2; Need for support from friends and family					
1 study (Turner 2013)	Semi-structured interview	1 study conducted at discharge from NICU (Australia) found that parents had an increased distress due to family and friends input about their concerns regarding the infants' health: "..one of my girlfriends was bombarding me the day before we actually picked her up..My head was spinning..I got in the car and said to my partner, 'I'm not going to cope. This is too much' " (mother)	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Table 37: Theme 2: Specialist services support

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Role of NICU in supporting discharge process					

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
1 study (Sommer 2015)	Semi-structured interview	1 study conducted in parents' home or workplace (New Zealand) reflected on the anxiety they experienced regarding transfer from NICU to another regional unit: "feeling like you're kind of whisked out a back door and it's like that abandonment" or "it would have been reassuring to know that NICU hadn't washed their hands completely" (parent)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
1 study (Turner 2013)	Semi-structured interview, support group	1 study (Australia) reported that mothers' anxious experience of discharge at NICU was difficult to cope with due to lack of assistance in providing support with how to manage complications at home: "they taught us..[cardiopulmonary resuscitation] CPR and stuff like that..and in my head it was like 'well what if something goes wrong and I don't know how to do the CPR?' "	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Role of neonatal transition care in supporting discharge					
1 study (Lasby 2004)	Focus group interview	1 study conducted at discharge from hospital (Canada) found that mothers felt anxious about taking their infant home, but found support from the transition care programme helpful: "The first week I was nervous, but once I had [the nurse] coming and I knew to expect her...it made it so much easier for me to just tend to [my baby] and to get over any apprehensions I had of having him home and not having a full staff of nurses there and learn that I was his full caregiver and whatever we did was ok" (mother)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1 **Table 38: Inforrmation support**

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Subtheme 1: Need for further Information and support on child interaction					
1 study (Nicolaou 2009)	Interview	1 study conducted at the time of discharge from NICU identified that mothers of preterm infants found a lack of information given to them about interacting with their infant: *"we were given preparation but it was all very medical. We had booklets and discussions about RSV, meningitis, all the things he could pick up, but in terms of how to actually care for him and what to do when we got him home there really wasn't anything" (mother of preterm infant)*	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Need for debrief at NICU discharge					
1 study (Whittingham 2014)	Interview; Survey	1 study conducted in NICU prior to discharge (Australia) identified that parents felt it would be important to debrief close to time of discharge: *"I felt emotionally I don't think that I would take it in at that stage. Maybe at the special care or close to the end..to be in the ICU and have that emotional weight [parenting support] would just be an extra weight added..." (parent of preterm infant)*	Limitation of evidence	Minor limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Unclear	
			Sufficiency or saturation	Unclear	
Subtheme 3: Need for information support on feeding support at discharge from NICU					
1 study (Harrison 2009)	In depth interview	1 study conducted at discharge from NICU (Canada) showed that mothers' fear of refusal of support, fear of exposure, or fear of failing to care for their infant, was a barrier to requesting support:	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
		"I don't like to ask other people to do things for me. I will do them on my own if it kills me. So that leads to all kinds of problems" (mother) "When you're asking for support, a lot of times you've got to tell them the reason why and go into great depth about it. You can't just say, would you do this for me" (mother) "when you can't manage on your own, you feel like somehow you've failed, and so if you 're a failure, you hate to point this out to someone else and ask for help" (mother)	Sufficiency or saturation	Unclear	
Subtheme 4: Support during NICU discharge about feeding schedule					
1 study (Reyna 2006)	Interview	1 study conducted at discharge from NICU (USA) showed that mothers were anxious and apprehensive about their infants after discharge, especially with feeding: *"the only concern I have is, I don't want them to choke, I'm fearful of choking" (mother)* In the same study, mothers found difficulties with understanding discharge instructions and feeding schedule and were hesitant to liberalise their infant's intake after discharge: *"I'm afraid of missing a feeding...the hardest part is when she's 3 hours this time and then she doesn't eat for 4 hours the next time, and I'm thinking I'm late, I didn't feed her" or "they gave me instructions as every 3 to 4 hours ad lib. I didn't ask that right now she's on 2 ounces, when do I take her to 3 or 2.5 ounces"* (mother)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1 **Table 39: Theme 3a: Information support**

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Need for further Information and support on child interaction					
1 study (Nicolaou 2009)	Interview	1 study conducted at the time of discharge from NICU identified that mothers of preterm infants found a lack of information given to them about interacting with their infant: *"we were given preparation but it was all very medical. We had booklets and discussions about RSV, meningitis, all the things he could pick up, but in terms of how to actually care for him and what to do when we got him home there really wasn't anything" (mother of preterm infant)*	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Need for debrief at NICU discharge					
1 study (Whittingham 2014)	Interview; Survey	1 study conducted in NICU prior to discharge (Australia) identified that parents felt it would be important to debrief close to time of discharge: *"I felt emotionally I don't think that I would take it in at that stage. Maybe at the special care or close to the end..to be in the ICU and have that emotional weight [parenting support] would just be an extra weight added..." (parent of preterm infant)*	Limitation of evidence	Minor limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Unclear	
			Sufficiency or saturation	Unclear	
Subtheme 3: Need for information support on feeding support at discharge from NICU					
1 study (Harrison 2009)	In depth interview	1 study conducted at discharge from NICU (Canada) showed that mothers' fear of refusal of support, fear of exposure, or fear of failing to care for their infant, was a barrier to requesting support:	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
		"I don't like to ask other people to do things for me. I will do them on my own if it kills me. So that leads to all kinds of problems" (mother) "When you're asking for support, a lot of times you've got to tell them the reason why and go into great depth about it. You can't just say, would you do this for me" (mother) "when you can't manage on your own, you feel like somehow you've failed, and so if you 're a failure, you hate to point this out to someone else and ask for help" (mother)	Sufficiency or saturation	Unclear	
Subtheme 4: Support during NICU discharge about feeding schedule					
1 study (Reyna 2006)	Interview	1 study conducted at discharge from NICU (USA) showed that mothers were anxious and apprehensive about their infants after discharge, especially with feeding: *"the only concern I have is, I don't want them to choke, I'm fearful of choking" (mother)* In the same study, mothers found difficulties with understanding discharge instructions and feeding schedule and were hesitant to liberalise their infant's intake after discharge: *"I'm afraid of missing a feeding...the hardest part is when she's 3 hours this time and then she doesn't eat for 4 hours the next time, and I'm thinking I'm late, I didn't feed her" or "they gave me instructions as every 3 to 4 hours ad lib. I didn't ask that right now she's on 2 ounces, when do I take her to 3 or 2.5 ounces"* (mother)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1

1 **After discharge from NICU**

2 **Table 40: Theme 1: Coping with preterm infant, interactionwith infant, and developmental concerns after discharge from NICU**

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Coping with a very low birth weight infant (<1500g)					
1 study (Vasquez 1995)	Interview	1 study conducted after hospital discharge (USA) found that parents were protective to their infants from germs, strangers, friends and close family members, and also isolated: "when people come over...mostly relatives...I did tell them that they couldn't touch the baby" or "we didn't go to restaurants until 3 months after discharge...we didn't take him out much those first couple of months. And we still don't go out much" (parents)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Coping with preterm infant-burden of care					
1 study (May 1997)	Interview	1 study conducted after discharge from hospital (USA) found that mothers expressed burden of care of their infants at home, physical and emotional strain and changes to lifestyle: "I think an important time for people to be reached when they have premature children is in the first week, because you're terrified and you have no idea" (mother of preterm infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Impact of interaction with father on infant development					
1 study (Benzies 2014)	Interview	1 study conducted after discharge from NICU (USA) found that fathers interacting with their infants was a positive aspect for their infant's development:	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
		Spending time with infant "I love when I can spend the whole day with the baby" or "getting on the floor and watching them play" or "taking the baby for walks in the park" Many fathers liked "playing in the bathtub" or "putting him to bed" Watching the infant grow and learn One father stated that he "looked forward to each new step and each new development" Being recognised by the infant Some fathers stated that their child's recognition and excitement contributed to joys of fatherhood: "I enjoy that he smiles at me, that I make him happy, and that he knows who I am" (father)	Applicability of evidence Sufficiency or saturation Coherence of findings Applicability of evidence Sufficiency or saturation	Applicable Unclear Coherent Applicable Unclear	
Subtheme 4: Parents developmental concern with infant age					
1 study (Vasquez 1995)	Interview	1 study conducted after discharge from hospital (USA) found that parents were concerned with the infants actual age: "..we were talking about celebrating her birthday. When she turns 1..will she really be 1? Developmentally, she will be a little behind. We'll just do it on her real birthday, the day she should have been born" (parents)	Limitation of evidence Coherence of findings Applicability of evidence Sufficiency or saturation	Major limitation Coherent Applicable Unclear	Low

1

2 **Table 41: Social support**

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
Subtheme 1: Family support-grandmother/mother –in- law					

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
3 studies (Frisman 2012; Philips-Pula 2013; Thomas 2009)	Interviews	1 study conducted at home or at hospital after discharge from NICU found that grandmothers acknowledged that the parents of the preterm infant needed their support with regard to housework and shopping: "Having an infant in the neonatal ward made them isolated from the world. So in that way they needed more practical help than otherwise" (grandmother of a preterm infant) 1 study conducted after discharge at home or another choice of place (USA) identified that mothers of preterm infants found that support from their mothers was helpful: "whenever I get tired my mom will say 'bring him to me and go take a nap or something' and that helps" (mother of preterm infant) 1 study conducted after discharge from NICU (Canada) fathers found that their mother in law's support was helpful in caring for their infant: "she's extremely capable..feeding, teaching my mother tongue [language] and manners, how to handle a baby physically..in some ways through her caring for our baby, it was for us a kind of training" (father of VLBW infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Family support- extended family					
1 study (Neu 2008)	Interview	1 study conducted at home after discharge from hospital (USA) found that support from extended family members was helpful to adolescent mothers of preterm infants:	Limitation of evidence	Major limitation	Moderate
			Coherence of findings	Coherent	

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
		"I have lots of cousins who live very close. In the evening we get together and play with our babies and just talk" (mother)	Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 3: Difficulties of family interaction with infant					
1 study (Vasquez 1995)	Interview	1 study conducted after discharge from NICU (USA) found that parents were angered by some remarks that family members made and did not interact with the infant because they were afraid: "they're afraid of him, some people are afraid to touch him...he's so small. I'm talking about relatives, the people that I expect to love him. They love him...but don't show it. They haven't celebrated his birth yet..it's been 7 months" (parents)	Limitation of evidence	Major limitation	Low
			Coherence of findings	coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 4: Peer support group					
1 study (Turner 2013)	Semi-structured interview	1 study conducted after discharge from NICU (New Zealand) found that parents attending a baby playgroup was helpful for them to reconnect with other parents to gain support for their infant's care: " the support is carrying on now...having a kid who's..nearly 6 months old, but only 4 months corrected...I'm starting to think about solids...and that's something that I'll..go to the playgroup" (parents of preterm infant) In the same study, parents also found the support group helpful in providing information and educational support: "..definitely [found information and educational content useful]" (parent of preterm infant)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 5: Need for peer support					
1 study (Garel 2006)	Semi-structured interview	1 study conducted at home 2 months after discharge from NICU (France) and 1 year after delivery identified that mothers of preterm infants found: "the need for contacts and meetings with other parents of very preterm babies and written information" (mother)	Limitation of evidence	Minor limitation	Very low
			Coherence of findings	Coherent	
			Applicability of evidence	Unclear	
			Sufficiency or saturation	Unclear	

1

2 **Table 42: Theme 3: Spiritual support**

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 1: Spiritual support of parents to cope with developmental disability					
1 study (Lee 2013)	Interview	1 study conducted at home (at 6 to 12 months follow-up) (Taiwan) found that parents personal belief helped them to cope with the developmental disability of their preterm infant: "I was very disappointed at first because I planned to teach him to play tennis when he was older...Now I consider my son's condition [possible permanent disability] as a tough trial God gave me...Ever since I knew the possible prognosis related to his physical functioning, I have more empathy when seeing other handicapped children. I think God is fair. I appreciate that my son's current condition is not as severe as the one shown on TV" (Christian father)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Naming of infant as spiritual support					

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
1 study (Chiu 2012)	Interview/videotape	1 study conducted at home after discharge (Canada) found that naming their infants after ancestors was supportive for mothers to come to terms with their preterm status: "[The name is] from one of those that leads you to the Promised Land..because he was premature..birth was so risky and when he came out and I heard him cry. It's like, he made it!..So I gave him the name" (mother) Mothers of Canadian caucasian background told how naming their infant after an ancestors gave the baby the strength to survive: "That was our first kind of leap of faith after she was born because the chances of her making it, weren't 100%...when she was born..we always wanted to name our baby after our mothers..she's going to make it and we gave her the real name" (mother)	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1 **Table 43: Theme 4: Information support**

Study information			Quality assessment		
Number of studies	**Design**	**Description of theme or finding**	**Criteria**	**Rating**	**Overall**
Subtheme 1: Assistance in obtaining Information, assessment, treatment respite caregiving, and support					
1 study (May 1997)	Interview	1 study conducted at home after NICU discharge (USA) stated that mothers recognised the need for assistance in obtaining information, assessment, treatment, respite caregiving and support: "One thing is that I wish there were more resources to rely on, to fall back on. I wish there were more studies done and more statistics" (mother of preterm infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Need for continued information support					

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
1 study (Benzies 2014)	Interview	1 study conducted after discharge (USA) found that fathers would have liked to receive continued access to information regarding: "suggestions or links to resources for further learning" (father of preterm)	Limitation of evidence	Major limitation	Very low
			Coherence of findings	Coherent	
			Applicability of evidence	Unclear	
			Sufficiency or saturation	Unclear	
Subtheme 3: Seeking reassurance for developmental milestones concerns					
1 study (Benzies 2014)	Interview	1 study conducted after discharge from hospital (USA) found that Fathers' were aware of their infant's development regarding developmental milestones. One parent sought information from the home visitor with concerns: "..some of his cousins are the same age and walking-should he be walking?" (father)	Limitation of evidence	Major limitation	Low

1

2 **Table 44: Feeding support**

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Subtheme 1: Need for breast feeding support					
1 study (Niela-Vielen 2015)	Facebook breast feeding peer-support group	1 study conducted at home (Finalnd) found that mothers experience of breast feeding counselling at NICU did not facilitate their needs at home: "after discharge, we tried to practice breast feeding by ourselves. It didn't work out at all..the baby's latch wasn't right.." or "they said no breast feeding at all before the weight is clearly increasing. Well, after a	Limitation of evidence	Minor limitation	Low
			Coherence of findings	Unclear	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
		few weeks, the baby refused to suckle the breast and he only accepted the bottle" (mother of preterm after discharge from NICU) In the same study, mothers also stated that they wished for individual support and equal guidance and support (and counselling) from all nurses in order to maintain breast feeding and its potential challenges at home: "..I was hoping for more information especially about how to manage at home, when the baby is used to the bottle, and what kind of problems may exist and how to manage them. Your are not able to ask all relevant questions in hospital when you are worried about the health of your baby and the main issue is that the baby is getting food, one way or another. In hindsight, I would have acted differently when we got home, but then, as a novice, I ruined my opportunity to exclusively breast feed" (mother of preterm infant) In the same study, some mothers who were able to kangaroo care for their infant did not need to practice at home: "..we were able to kangaroo..they really encouraged us to do it. Both nurses and doctors..we hardly ever practiced kangaroo at home" (mother of preterm infant)			
Subtheme 2: Support with infant feeding needs after discharge					
1 study (Lee 2009)	Interview	1 study conducted after discharge from (Taiwan) found that mothers became familiar with the infant’s feeding needs in a positive manner:	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
		"when my baby came home, I made a time schedule listing what I should do, and recorded what I did and how much I fed her. It took me one month to get familiar with her and learn way to take care of her" (mother) In the same study, some mothers who bottle feed their breast milk found it difficult to feed their infant: "Every day feeding occupied the majority of my time. I fed her every 3 hours. The nurse told me to express even at night to supply efficiently. I felt my sleep was dissected into several segments" (mother)	Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

1

2 **Table 45: Healthcare professional support**

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 1: Health care professional support at clinics					
1 study (Philips-Pula 2013)	Interview	1 study conducted after NICU discharge (at home or at another choice of place) (USA) identified that at least one person who worked with mothers with their preterm infants was helpful: "The NP at the apnea clinic was amazing..the best..she understood everything" or "the nurses and neonatologists always talked to us like humans" (mother of preterm infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Home visitor support					
1 study	Interview	1 study conducted at home after discharge from NICU (USA) identified that fathers	Limitation of evidence	Major limitation	Low

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
(Benzies 2014)		liked having a health care professional as the home visitor One father found: "found comfort in knowing he could ask questions regarding the baby" (father of preterm infant) In the same study, fathers also stated that the home visits were helpful: "A full year of visits would be great..like having a teacher come once a month to help guide" (father of preterm infant) "it was good to have outside confirmation that I am a good dad" (father of preterm infant)	Coherence of findings	Coherent	
1 study (Nicolaou 2009)	Interview	In another study, mothers expressed that they would have liked more support in the early days when they took their infant home: "Hospital is probably the place that knows that we're all mums with new babies. It would have been great if we could have had a support group" (mother)	Applicability of evidence	Applicable	Moderate
			Sufficiency or saturation	Unclear Minor limitation Coherent Applicable Unclear sufficiency or saturation	
Subtheme 3: Nurse support					
1 study (May 1997)	Interview	1 study conducted after discharge from NICU (USA) stated that mothers found they could seek help with assessment and treatment when at home from the nurse at the follow-up clinic: "I'd call the home health nurses and say 'can you stop by today? I think he's got a cold in his lungs. Am I hearing things or do I need to take him to the doctors?' She would come out" (mother of preterm infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
Subtheme 4: Occupational therapist support					
1 study (Chiu 2012)	Interview/audio tape/focus group	1 study conducted at home (Canada) identified that mothers of preterm infants showed appreciation for the OT as a mentor and trusted expert: "They know what they (the babies) should be doing, and showing me what to do with her..it's amazing. If I didn't have that, I really wouldn't know..'what would she be doing?' Probably wouldn't even get her attention for 5 minutes..because I've worked with her every week and it gives us something different to do besides sitting there and playing with toys all day. The exercises are something we can do for an hour.." (one mother) In the focus group, mothers expressed that the OT helped with learning to play with their infant and facilitated positive interaction and motor development of the infant: "..we don't feel anxiety about the baby because we've had that (OT in the home)...it's been huge, and she's made great progress...the OT has taught us a lot..we know how to play with her in ways that are more therapeutic" (one mother) "the OT also gave me extra help with how I can massage him as he grows..and also taught me how to use the beach ball...since I did all that, I saw a very big improvement in my child...he is two times more active than before" (Tamil mother) Mothers in the focus group stated that OT support once a week was helpful:	Limitation of evidence	Major limitation	Low
			Coherence of findings	Coherent	
			Applicability of evidence	Applicability	
			Sufficiency or saturation	Unclear	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
		"..having the OT come in every week, was helpful, not only for exercises, she helps me, just by talking to me and telling me that my child is progressing, and that's positive, because the OT is quick to compliment and quick to let you know that you're doing a good job" (one mother)			

1

2 **Table 46: Theme 7: Specialist services support**

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
Subtheme 1: Coming to terms with having a preterm infant and impact on accessing to follow-up services					
1 study (Lee 2013)	Interview;	1 study conducted at home after discharge from NICU (Taiwan) found that mothers hesitated to apply for social welfare programmes, which affected their infant's follow-up: "I could not accept he was 'severely handicapped' at first, especially when I saw the doctor write down the term on his report...he needs to be evaluated after three years. So I told myself if we worked harder [at rehabilitation], maybe he would be normal or become mildly disabled" (mother of VLBW infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 2: Expectations of parents from early intervention services					
1 study (Lee 2013)	Interview;	1 study conducted at home after discharge from NICU (Taiwan) found that parents expected that early intervention services would stop functional deterioration of their VLBW infant and also impairment would disappear or become less obvious:	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
		" ..I believed if she continued her physical therapy, then one day she would walk like a normal child. No one would know she had been a premature baby with impairment" (mother)	Sufficiency or saturation	Unclear	
Subtheme 3: Early intervention support-understanding the infants needs					
1 study (Little 2015)	Focus group; interview	1 study conducted at three hospitals and local early intervention programmes (USA) identified that early intervention support was helpful for parents to understand medical and developmental needs of the infant when they could not understand the doctor: " sometimes we don't really understand the doctor, and then the EI provider comes and explains it" (parent of VLBW infant In the same study, EI staff also helped parents to recognise their infants medical and developmental problems: "My wife says how...she didn't notice...my daughter's problem, her neck. Early intervention did. And then I started to notice it too. So she had 2 therapists, one for the neck and one to help her play" (parent) EI support also helped with keeping parents engaged with their infants care: "The EI therapist writes what we did and what needs to be worked on and what was the improvement. And I get a copy of that at every visit" (one parent) EI staff explained their role in making observations about the infant's development and also the family's social situation:	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Study information			Quality assessment		
Number of studies	Design	Description of theme or finding	Criteria	Rating	Overall
		"EI is the eyes and ears for paediatricians and school systems and everybody" (EI local coordinator)			
Subtheme 4: Early intervention staff supporting parents during follow-up visits					
1 study (Little 2015)	Focus group; interview	1 study conducted at three hospitals and local early intervention programmes (USA) identified the EI staff support during doctors' visits to facilitate parents in receiving correct information: "we go as support systems, and.to make sure we have information correct. A lot of our families' educational levels make it hard for them to...talk about what their doctor explained" (local EI coordinator)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 5: Early intervention support-encouraging parents to attend follow-up clinics					
1 study (Little 2015)	Focus group; interview	1 study conducted at three hospitals and local early intervention programmes (USA) identified that EI was supportive in prompting parents to come back to NICU for follow-up after discharge: "EI has helped us out a lot..in terms of prompting parents to come back to the NICU follow-up clinic" (parents of VLBW infant)	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	
Subtheme 6: Role of neonatal transition care support					
1 study	Interview	1 study conducted after discharge from hospital (Canada) found that regular in-	Limitation of evidence	Major limitation	Very low

Study information		Description of theme or finding	Quality assessment		
Number of studies	**Design**		**Criteria**	**Rating**	**Overall**
(Lasby 2004)		home contact and prompt pager support from the NTCP nurses, and telephone contact with the dietician enhanced their maternal confidence and decreased the need to take their infant outside of the home for weight checks, routine assessments, and vaccinations: "It helps you gain confidence [The NTCP] are there for you at every intense time" or "I can't imagine what it would be like without them [NTCP]" (mother) In the same study, NTCP support impacted positively on mothers at home with their infants: "they [NTCP] are the hope because they've seen babies like ours-very small and they've grown up to be well-and it's the stories they [NTCP] tell. I can now give that future hope. whereas before I didn't look past this day, this week, or this month" (mother)	Coherence of findings	Unclear	
			Applicability of evidence	Applicability	
			Sufficiency or saturation	Unclear	
Subtheme 8: Community services- need for consistent feeding advice					
1 study (Whittingham 2014)	Focus group discussion	1 study conducted in a hospital setting (Australia) found that parents were confused by variation of support provided by community nurses compared with NICU: "my community nurse at the community health clinic told me I should be starting her on solids at her six months real age and then I rang special care and they said probably, we normally go corrected age but whatever the baby wants so I gave up and just ant with whatever she told me. But when I went back to the community nurse a couple of months later she was into me because this baby should be on mashed	Limitation of evidence	Minor limitation	Moderate
			Coherence of findings	Coherent	
			Applicability of evidence	Applicable	
			Sufficiency or saturation	Unclear	

Study information		Description of theme or finding	Quality assessment		
Number of studies	Design		Criteria	Rating	Overall
		...and you should fast track this baby through all of this and I just went you know, how am I supposed to know what I'm supposed to do?" (parent)			

5.1.2.5 Economic evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

5.1.2.6 Evidence statements

5.1.2.6.1 At discharge from NICU

Social support

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers frequently excused their husbands from providing help because they felt that their husbands were not willing to help.

Specialist services support

Multidiciplinary teams/NICU support

Moderate quality evidence from one study carried out among mothers using interview or support group discussion, showed that mothers found it difficult to cope at the time of discharge from NICU due to lack of assistance in providing support with how to manage complications at home.

Low quality evidence from one study carried out among parents using interview design, showed that parents found that lack of support felt like they were being abandoned at the time of discharge from NICU to another regional unit.

Neonatal transition care support

Low quality evidence from one study carried out among mothers using focus group or interview design, showed that mothers were anxious about taking their infants home from NICU but found support from the transition care programme nurses helpful immediately after discharge with regards to looking after their infant.

Information support

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers found that the information given to them at the time of discharge from NICU was not helpful for interacting with their infant after discharge from NICU.

Low quality evidence from one study carried out among parents using interview design, showed that parents would have found debriefing information support helpful and important at the time of discharge from NICU.

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers were fearful of requesting or refused support from others, which was not helpful in the care for their infants immediately at discharge from NICU.

Low quality evidence from one study carried out among mothers using interview design, showed that mothers were anxious and apprehensive about feeding their infants at discharge from NICU. Mothers in the same study found difficulties with discharge instructions and feeding schedules.

5.1.2.6.2 *After discharge from NICU*

Coping with low birth weight/preterm infants

Low quality evidence form one study carried out among parents of very low birth weight infants using interview design, showed that parents were protective to their infants from germs, strangers, friends and close family members. Parents were also restricted as they could not take their infant out for the first few months after discharge from NICU. In another moderate quality study carried out among mothers of preterm infants using interview design, mothers felt the burden of care in the first week at home as they were not prepared for the changes to lifestyle, physical and emotional strain.

Impact of interaction with father on infant development

Low quality evidence from one study carried out among fathers of preterm infants, using interview design, showed that fathers interacting (spending time) with their infants was a positive aspect for their infant's development and also recognition of the bond between the father and the infant.

Parents concerns with infant age

Low quality evidence from one study carried out among parents, using interview design, showed parents were concerned about their infants being behind developmentally at 1 years' age.

Social support

Family support (grandmother or parents of preterm infant)

Moderate quality evidence from three separate studies carried out among grandmothers or parents of preterm infants using interview design, showed that grandmothers were practically supportive to parents of the preterm infant with regards to help at home (house work and shopping) as well as caring for the infant. In one study carried out among fathers of preterm infants, showed that fathers found their mother-in-law supportive to them in feeding, and teaching them how to handle their infant physically.

Family support (extended family of adolescent mothers)

Moderate quality evidence from one study carried out among adolescent mothers using interview design, showed that mothers found support from extended family members to be helpful in playing with their infants and also to talk to other mothers about their infants.

Family interaction with infant

Low quality evidence from one study carried out among parents using interview design, showed that family members were reluctant to interact with the infant which angered parents of the infant.

Peer support

Low quality evidence from one study carried out among parents using interview design, showed that parents found attending a baby playgroup was helpful for them to reconnect with other parents to gain support for their infant's care. In the same study, parents found that the peer support group was also useful for accessing information and educational.

Very low quality evidence from one study carried out among mothers using interview design, showed that mothers found the need for contacts and meetings with other parents of very preterm infants and also written information.

Spiritual support

Moderate quality evidence from one study carried out among fathers using interview design, s showed that fathers' personal religious beliefs helped them to cope with developmental disability of their preterm infant.

Low quality evidence from one study carried out among mothers using interview design, showed that mothers' religious, or cultural beliefs (naming the infant) helped them to come to terms with their preterm infant.

Information support

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers found there was a need for assistance in obtaining information, assessment, treatment, respite caregiving and support at home.

Very low quality evidence form one study carried out among fathers using interview design, showed that fathers would have liked to receive continued access to information regarding suggestions or links to resources for further learning

Low quality evidence from one study carried out among fathers, using interview design, showed that fathers sought information from home visitors regarding concerns about developmental milestones of their infant, such as walking,

Feeding support

Breast feeding support

Low quality evidence from one study carried out among mothers using peer-support group design, showed that mothers experience of breast feeding counselling at NICU did not facilitate their needs at home, and wished for individual support and equal guidance (and counselling) from all nurses in order to maintain breast feeding and its potential challenges at home. In the same study, mothers who had been provided kangaroo care in NICU were able to provide this care for their infant at home.

Coping with infants feeding needs

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers were able to cope with their infants feeding needs in a positive manner after discharge from NICU but making a time schedule listing what they should do, and recorded how much they fed their infant. In the same study, some mothers also complained of exhaustion from feeding their infant as feeding was taking up majority of their time.

Health care professional support

Home visitor as the health care professional and support at home

Low quality evidence from one study carried out among fathers using interview design, showed that fathers preferred the home visitor to be a health care professional as this gave them more comfort to ask questions about their infant. In the same study, fathers found that frequent home visits were helpful, and provided confirmation for fathers regarding their parenting skills.

Occupational therapist support at home

Low quality evidence from one study carried out among mothers using interview design, showed that the occupational therapist was helpful in mentoring and was a trusted expert for mothers in showing how to interact with their infant. In the same study, among mothers using a focus group design, mothers found the occupational therapist supportive in providing help with learning to play with their infant and facilitated positive interaction as well as motor development of the infant. Mothers in the focus group expressed that support from the occupational therapist once a week was helpful to them in the care of their infant.

Health care professional support at follow-up clinics

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers found support from at least one health care professional (nurse or neonatologist) during follow-up clinics.

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers found the nurse to be supportive in help with assessment and treatment at follow-up clinics.

Specialist services support

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers hesitated to apply for social welfare programmes for their infant's follow-up after discharge from NICU.

Moderate quality evidence from one study carried out among mothers using interview design, showed that mothers expected that early intervention services would halt deterioration and impairment of their VLBW infant after discharge from NICU.

Early intervention service support

Moderate quality evidence from one study carried out among parents using focus group or interview design, showed that mothers found early intervention support provider helpful to understand medical and developmental needs of their infant. In the same study, mothers found that early intervention supported them to recognise their infants' medical and developmental needs, and also with continued engagement with their infant.

Moderate quality evidence from one study carried out among early intervention staff using focus group or interview design, showed that early intervention staff were supportive for making observations about infant development and the family's social situation. In the same study, early intervention local coordinator was supportive for paediatricians and school systems. Early intervention staff were supportive during parents' visits to doctors to facilitate them in receiving the correct information and also encouraging to prompt parents to come back to NICU for follow-up appointments after discharge.

Neonatal transition care support

Very low quality evidence from one study carried out among mothers using interview design, showed that mothers found regular in-home contact and prompt pager support from the neonatal transition care nurses helpful. In the same study, mothers found that telephone contact with the dietician also increased their confidence in caring for their infant at home.

Community support

Moderate quality evidence from one study carried out among parents using focus group discussion showed that parents found feeding advice given to them by community nurses conflicting compared to advice given to them in NICU.

5.1.2.7 Economic evidence statement

A literature review of published cost-effectiveness analyses did not identify any relevant studies and no economic modelling was undertaken for this question.

5.1.2.8 Evidence to recommendations

5.1.2.8.1 *Relative value placed on the outcomes considered*

The aim of this review was to understand how different support strategies were perceived from the perspective of the parents and carers of infants, children and young people born pre-term. Because this was a review of qualitative data, there were no pre-specified outcomes. Based on the evidence the Committee identified the following as important themes in relation to support: social support, information support, support from health care professionals and specialist services, support with feeding, and support with coping with a child born preterm. In addition, psychological support for parents and carers due to high levels of anxiety was considered important even though this was not identified in the evidence.

5.1.2.8.2 *Consideration of clinical benefits and harms*

The Committee agreed that the evidence largely reflected their experiences as parents or grandparents of a child born preterm or as health care professionals working closely with families of children born preterm. They noted that anxiety surrounding discharge from hospital was highlighted in the evidence and that this was a common sentiment experienced by the parents and carers of children born preterm. The Committee discussed that no matter how well the discharge was planned, it was natural that an element of anxiety is present. However, in order to minimise the anxiety surrounding discharge, planning and execution of discharge is crucial and should begin during admission to the neonatal unit. The plan for discharge should include educating the parents and transferring skills on how to take care of a child born preterm in the home environment, planning and providing support in establishing routines at home and having a point of contact for the parents and carers to rely on when concerns or questions arise. The need to provide support with discharge and establishment of daily routines at home was also supported by the evidence in this review.

All support for the families should be planned on an individual basis and take into account the needs of each family. Consideration should be given to the educational and socio-economic status of the family, the presence of a language barrier, or the spiritual needs of the parents.

The Committee acknowledged that the advice, support and skills transfer that parents receive currently varied across service providers. They agreed that it was important that the messages to the parents and carers are consistent in order to avoid additional stress and confusion. The Committee highlighted the importance of clear communication between healthcare professionals and service providers to ensure consistent information was provided to the parents (please see section 5.1.1.7). Although it was not highlighted in the evidence in this review, the Committee discussed that the consistency and communication between service providersis especially important for families of children born preterm because they are often engaged with several different service providers, including neonatal services, local health services, as well as social and educational services. In order to avoid duplication and potential inconsistency, the Committee discussed that the different professionals should be aware of each other, their roles in the care of the child and the information they have of the child. The Committee recognised the importance of getting consent from the parents or carers when sharing information between health services, social services and education services (see section 5.1.5.6).

The importance of having someone to ask questions or get reassurance from was seen as important, therefore, the Committee agreed that families of children born preterm most at risk of developmental problems and disorders should be assigned a key point of contact to which they could rely on when concerns or questions arise (see section 5.1.4). The Committee agreed that this person should be someone with experience of the needs of a premature children and therefore should be organised through the neonatal services.

The Committee discussed that the topics that cause concern in the immediate phase after discharge from hospital often relate to functional issues such as feeding, breathing, crying and sleeping. They felt that a telephone contact would be very useful especially during the immediate phase after discharge in order to have a chance to ask questions and to get reassurance that the child is doing well. In their experience the transition from the hospital environment with continuous supervision to the home environment can cause unexpected worry in parents, for example, the sound of the child breathing may sound different in the quiet home environment compared to the busy hospital unit, which may cause concern in the carer.

The Committee agreed that the key transitions are the most crucial time points during which parents and carers need support. These may include transition from hospital to home, from neonatal intensive care unit to a different hospital or unit, from specialised services to community services, from home to nursery care (and parent's return to work), or eventually to the education services. The Committee emphasised that parents and carers may need support when making decisions during these times. The Committee noted that these themes were partly reflected in the evidence which considered support during the first and second year after discharge, although evidence was mainly found on the immediate time during or after discharge. Issues surrounding the transition to education services, a key point in the child's and family's life, were not identified in the evidence review.

The Committee discussed that there were two distinct phases after discharge: 1. Immediately following discharge where the parent(s) concern and worry is acute and all effort is put into making sure that the child survives; 2. Sometime after discharge when the acute constant worry dissipates and the parents start concentrating on the longer-term development of the child. The support required from professional experienced in the needs of children born preterm would be different during both phases but needed at both times none the less.

A theme that frequently arose in the evidence was the need for support in relation to feeding. The Committee agreed that this was very important and frequently raised by the parents and carers in their experience as well. The Committee agreed that feeding impacts on many aspects of development, including growth, brain development, speaking, and interaction and providing support with it is essential. The health care professionals providing postnatal care and support in the community to the family after discharge should, therefore, have expertise in feeding issues.

One of the themes that came up in the evidence was peer support. The Committee was of the opinion that peer support is important and can be helpful for parents and carers. Members of the Committee expressed that even though formal peer support might be offered to them, it was sometimes difficult to utilise this because of frequent other appointments for the child. The parents and carers of children born preterm may also miss out on peer support opportunities that would normally be available to them, for example, antenatal and postnatal support groups, because of having to be in neonatal care. The Committee agreed that peer support, whether organised formally or informally by the parents themselves, was very important and helpful. They also said that digital or online peer support groups could be very useful bringing more flexibility in relation to the location and timing of this type of support.

Even though the need for psychological support was not found in the evidence, the Committee agreed that it was an important issue to consider and service providers should be aware of the parents' potential need for psychological support.

The evidence highlighted the importance of the father or partner and the extended family members (including grandparents) in the care of the child. There was some evidence showing that extended family members might be insecure or reluctant to take care of the child born preterm. The Committee agreed that it was very important that the grandparents and other members of the support system are included and where possible, skills transfer should be provided to them as well. However, the Committee recognised that not all parents have partners or families and it should always be up to the parent(s) to determine who should be involved, what kind of support they require and from whom.

5.1.2.8.3 *Consideration of economic benefits and harms*

A systematic review of the economic literature was conducted but no relevant studies were identified which were applicable to this review question.

The economic implications of this topic were considered but not thought to be substantial. The provision of support does have resource implications as it requires time to be spent by the health care professionals providing it. However, the majority of the recommendations made reflect current best practice and so the recommendations are not expected to require a substantial increase in resources.

There is thought to be inconsistency in practice though with the advice, support and skills transfer that parents receive varying across service providers. Therefore, it is possible that there could be increased costs for service providers that are not currently providing the support outlined in the recommendations.

Any increase in the time spent by clinicians in providing support as a result of the recommendation was thought likely to be cost-effective as the increased costs would be offset by potential cost savings and effectiveness gains. There could be cost savings associated with educating parents upfront perhaps meaning that they would be less likely to require additional support when concerns arise. There could also be effectiveness gains associated with reducing parent anxiety, which was identified as a key theme in the evidence review.

5.1.2.8.4 *Quality of evidence*

The evidence in this review ranged from very low to moderate quality. The Committee identified gaps in the evidence mainly in relation to psychological support and transition to education services. Many of the included studies came from countries other than the UK and therefore generalisation to UK settings should be undertaken with caution. However, the Committee agreed that the principles, if not the details, of the evidence were applicable to the UK context and reflected their experiences well.

5.1.2.8.5 *Other considerations*

The Committee discussed gestational age at birth being an important factor in relation to support needed. The needs of the families of a child born at 35 weeks of gestation and a child born at 25 weeks of gestation might vary considerably.

The Committee also recognised that currently there may be a disconnection between the neonatal and community services. The reality is that the community health visitors' expertise does not always cover prematurity which can be both frustrating for the parents and the health visitor.

5.1.2.8.6 *Key conclusions*

The guideline developers concluded that on-going access to support is essential for the families of children born preterm. The support provided should be specific to the needs of each child and their family. The support and advice provided by service providers should be consistent and the different service providers should engage with each other to provide the best possible support for the child and the family.

5.1.2.9 Recommendations

See Section 5.2.

5.1.2.10 Research recommendations

	What support do parents and carers report was helpful to them in the care of children who were born preterm at the time of transfer to education services?
Population	Parents or carers of children born less than 37 weeks' gestation
Intervention	Current support in relation to transfer to education services
Outcome	Parent and carer experiences
Study design	Qualitative study (for example, focus groups)
Timeframe	No follow-up required
Why this is needed	
Importance to 'patients or the population'	There is now a 'local offer',for children with Special Educational Needs and Disbability (SEND) and a process of Education, Health and Social Care plans that aim to be inclusive and prepare for transition to education services.
Relevance to NICE guidance	This study will provide valuable insights on the practical and qualitative aspects of support that may be used to guide future updates.
Relevance to the NHS	A positive impact in terms of parent satisfaction and engagement will promote more seamless public-NHS partnerships in health care. It will seek views from parents, carers and families (who are key stakeholders) and thus inform evaluation and improvement of care.
National priorities	Preterm births are one of the top 10 priorities identified nationally by the James Lind Alliance, specifically providing information of packages of care at or after discharge http://www.jla.nihr.ac.uk/priority-setting-partnerships/preterm-birth/top-10-priorities/ Developing an understanding of parental needs in delivering a developmental support and surveillance for children born preterm is an important component. The 2010 inquiry into the quality of general practice in England by the King's Fund highlighted the need for patient engagement (in this case, parents, carers and families of the child born and preterm) https://www.kingsfund.org.uk/projects/gp-inquiry/patient-engagement-involvement
Current evidence base	There are no data about the impact of a developmental surveillance programme in the UK.

	What support do parents and carers report was helpful to them in the care of children who were born preterm at the time of transfer to education services?
	There is currently a lack of 'end-user' contribution (parental, carer or family voice) in the evaluation of such programmes.
Equality	No specific equality issues were identified other than those relating to language and communication. Appropriate support, tools and techniques (for example, interpreters and translation of questionnaires) that enable communication should be employed.
Feasibility	No barriers to feasibility were identified.
Other comments	No other comments.

5.1.3 Identification of problems and disorders

Review question:

What is the usefulness of the following screening strategies in the identification of children and young people born preterm with intellectual disability, speech and language disorder, specific learning difficulty, social, emotional and mental health, and developmental co-ordination disorder:

- **healthy child programme (including plus/enhanced)**
- **parental observation/concern**
- **teachers observation/concern**
- **formal screening tests?**

5.1.3.1 Description of clinical evidence

This review aimed to identify methods leading to recognition of the neurodevelopmental disorders of intellectual disability, speech and language disorder, specific learning disorders, developmental co-ordination disorder and social, emotional and mental health disorders.

The purpose of the review was to look for approaches and simple screening tools that might be widely used to recognise those requiring a formal diagnostic assessment. The objectives of the review were to:

- assess the usefulness (diagnostic value) of the above approaches at identifying probable developmental disorders and problems in children and young people born preterm at different time points in order to initiate referral for specialist diagnostic assessment.
- inform a national programme of enhanced surveillance in children born preterm.

For full details see review protocol in Appendix D:.

A total of 13 studies (Blaggan 2012; Cuttini 2012; Dewey 2011; Halbwachs 2013; Indredavik 2005; Johnson 2008; Johnson 2010; Johnson 2014; Martin 2013; Schonhaut 2013; Simard 2012; Skellern 2001; Woodward 2011) were included in this review, including 12 diagnostic studies which assessed the diagnostic value of screening tools and 1 prognostic study (Johnson 2010) in which the association between earlier screening assessment and future diagnoses of a disorder was assessed.

Regarding settings, 4 studies were carried out in the UK (Blaggan 2014; Johnson 2008, 2010, 2014), 2 were from Canada (Dewey 2011; Simard 2012), 2 from Australia ((Martin

2013; Skellern 2011), and 1 each from France (Halbwachs 2013), Chile (Schonhaut 2013), Italy (Cutti 20123), and Norway (Indreadvik 2005), and the USA (Woodward 2011).

For screening strategies, we looked for studies that assessed the diagnostic value of the following:

- standard healthy child programme (including plus/enhanced health child programme)
- parental observation/concern
- teacher's observation/concern
- formal screening tests, including
  - Ages and stages questionnaire (ASQ)
  - Strength and Difficulties Questionnaire (SDQ)
  - Ages and stages questionnaire (ASQ) Social and Emotional
  - Developmental Coordination Disorder Questionnaire (DCDQ)
  - Parent report of children's abilities revised (PARCA-R), and
  - Schedule of Growing Skills

Evidence on all formal screening tests was found except for the Schedule of Growing Skills. No evidence was found on the standard healthy child programme (including plus/enhanced health child programme), parental observation/concern, or teacher's observation/concern.

The following evidence was considered in this review:

- 5 studies assessed the diagnostic value of ASQ compared to Wechsler Preschool & Primary Scale of Intelligence (WPPSI) or Bayley Scales of Infant and Toddler Development (BSID), and 4 studies assessed the diagnostic value of PARCA-R compared to BSID in identifying intellectual disability in preterm children, respectively.
- 1 study assessed the diagnostic value of PARCA-R compared to BSID in correctly identifying speech and language disability in preterm children;
- 2 studies assessed the diagnostic value of SDQ compared to Development and Well-Being Assessment (DAWBA) or a clinical diagnosis in correctly identifying emotional or conduct disorder; and
- 1 study assessed the diagnostic value of DCDQ compared to Movement ABC in correctly identifying developmental coordination disorder (DCD).
- 1 prognostic study (Johnson 2010) assessed the association between pervasive attentional and conduct problems, which were measured by SDQ among preterm children aged 6 years, and the diagnosis of any psychiatric disorder by DAWBA when the preterm children reached the age of 11 years.

Evidence from these are summarised in the clinical GRADE evidence profile in Section 5.1.3.3. See also the study selection flow chart in Appendix F:, forest plots in Appendix J:, study evidence tables in Appendix K and exclusion list in Appendix G:.

The feasibility of combining study data using meta-analysis was assessed. Due to the limited amount of evidence and the following differences between studies, it was not possible to pool the results:

- cut-off points used for index tests and reference standards
- gestational age at birth of participants
- ages of participants at the time of assessment.

5.1.3.2 Summary of included studies

Table 47: Summary of included studies

Study	Index test and reference standard	Population	Outcomes	Comments
Intellectual disability				
ASQ				
Halbwachs 2013 (France)	Index test: ASQ score 270; ASQ score 285 Reference standard: IQ < 70 on WPPSI-III; IQ < 85 on WPPSI-III	Children born before 36 weeks' gestation assessed at 5 years of age	Sens; Spec; LR+; LR-	
Simard 2012 (Canada)	Index test: ASQ < 1 SD; ASQ < 1.5 SD; ASQ < 2 SD Reference standard: Bayley MDI < 85; Bayley PDI < 85	Children born between 29 and 36 weeks' gestation assessed at 12 months' and 24 months' corrected age	Sens; Spec; LR+; LR-	
Skellern 2011 (Australia)	Index test: ASQ < 1SD Reference standard: Bayley MDI < 1 SD	Children born before 31 weeks' gestation assessed at 18 months' corrected age	Sens; Spec; LR+; LR-	
Schonhaut 2013 (Chile)	Index test: ASQ-3 Psychometric Values (< 2 SD) Reference standard: Bayley III ≥ 1 SD below the mean	Children born between 32 and 36 weeks' gestation; children born before 32 weeks' gestation or with birthweight <1500 g assessed at 8 months', 18 months', and 30 months' corrected age	Sens; Spec; LR+; LR-	
Woodward 2012 (USA)	Index test: ASQ > 1 SD;	Children born between 23 and 31 weeks' gestation assessed at 18-22 months' corrected age	Sens; Spec;	Participants were recruited from an early RCT

Study	Index test and reference standard	Population	Outcomes	Comments
	ASQ > 2 SD Reference standard: BSID MDI or PDI > 2SD below the mean; BSID MDI or PDI > 1SD below the mean		LR+; LR-	
PARCA-R				
Blaggan 2014 (UK)	Index test: PARCA-R < 73 Reference standard: Bayley III-MDI < 70	Children born between 32 and 36 weeks' gestation assessed at 25 months' corrected age	Sens; Spec; LR+; LR-	
Cutti 2012 (Italy)	Index test: PARCA < 44; PARCA < 46; PARCA < 68 Reference standard: BSID-II < 70; BSID-II MDI < 70; BSID-II MDI < 85	Children born between 22 and 31 weeks' gestation assessed at 2 years' corrected age	Sens; Spec; LR+; LR-	
Johnson 2008 (UK)	Index test: PARCA < 44; PARCA < 49 Reference standard: BSID-II MDI < 70	Children born before 32 weeks' gestation assessed at 2 years' corrected age	Sens; Spec; LR+; LR-	Participants were recruited from an early RCT
Martin 2013 (Australia)	Index test: PARCA < 19 (on the cognitive component) Reference standard: Bayley III cognition score < 70	Children born preterm (median gestational age at birth 27 weeks) assessed at 24 months' corrected age	Sens; Spec; LR+; LR-	Participants were recruited from an early RCT
Speech and language disorder				
PARCA-R				

Study	Index test and reference standard	Population	Outcomes	Comments
Martin 2013 (Australia)	Index test: PARCA < 23 (on language component) Reference standard: Bayley III language score < 70	Children born preterm (median gestational age at birth 27 weeks) assessed at 24 months' corrected age	Sens; Spec; LR+; LR-	Participants were recruited from an early RCT
Social, emotional and mental health				
SDQ				
Indreadvik 2005 (Norway)	Index test: SDQ > 90th centile (mother); SDQ > 90th centile (father); SDQ > 90th centile (teacher) Reference standard: Any psychiatric diagnosis	Children born between 24 and 36 weeks' gestation assessed at 14 year of age	Sens; Spec; LR+; LR-	
Johnson 2014 (UK)	Index test: Parent abnormal SDQ score according to published norms; Teacher abnormal SDQ score according to published norms Reference standard: Conduct disorder measured by DAWBA; Emotional disorder measured by DAWBA	Children born before 26 weeks' gestation assessed at 11 years of age	Sens; Spec; LR+; LR-	
Developmental coordination order (DCD)				
DCDQ				
Dewey 2011 (Canada)	Index test: DCDQ < 15th percentile Reference standard:	Children born between 24 and 35 weeks' gestation assessed at 5 years of age	Sens; Spec; LR+; LR-	

Study	Index test and reference standard	Population	Outcomes	Comments
	Movement ABC < 15th percentile			
Prognostic study				
Johnson 2010 (UK)	Prognostic factor assessed: pervasive attentional problems measured by SDQ at 6 years: Pervasive conduct problems measured by SDQ at 6 years Outcome: psychiatric disorders (assessed by DAWBA)	Children born before 26 weeks' gestation assessed at 11 years of age	Adjusted ORs	Prospective study where adjusted ORs were reported

Ages and stages questionnaire (ASQ); Development and Well-Being Assessment (DAWBA); Developmental Coordination Disorder Questionnaire (DCDQ); Gestational age (GA); Odds ratio (OR); Parent report of children's abilities revised (PARCA-R); Strength and Difficulties Questionnaire (SDQ); sensitivity (sens); specificity (spec); Positive likelihood ratio (LR+); Negative likelihood ratio (LR-).

5.1.3.3 Clinical evidence profiles

Evidence was summarised in the adapted GRADE Tables (Table 48,Table 49,Table 50,Table 51) When assessing the diagnostic values of screening tools, we focused on sensitivities, specificities, positive likelihood ratios, and negative likelihood ratios.

The following definitions have been used when summarising the levels of sensitivity and specificity:

- High – 90% and above
- Moderate – 75% to 89%
- Low – 74% or below

The following terms have been used when summarising the positive and negative likelihood ratios

Positive likelihood ratio (LR +):

- Very useful – > 10
- Moderately useful – ≥ 5 to 10
- Not useful – < 5

Negative likelihood ratio (LR -):

- Very useful – < 0.1
- Moderately useful – ≥ 0.1 to 0.2
- Not useful – > 0.2

Table 48: Clinical evidence profile: Diagnostic accuracy of screening tool (ASQ) in correctly identifying intellectual disability

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Sens**	**Spec**	**LR+**	**LR-**	**Quality**	**Importance**
Developmental delay screened by ASQ-3 Psychometric values (< 2 SD) among children born preterm (GA 32-36wks) at 8 months, 18 months, and 30 months corrected age; Diagnosis tool: Bayley III ≥ 1 SD below the mean												
1 (Schonhaut 2013)	Cross sectional	Serious risk of bias [1]	N/A	no serious indirect-ness	serious imprecision [2]	none	0.80 (0.61-0.91)	0.73 (0.63-0.81)	2.9 (2.0-4.3)	0.27 (0.1-0.6)	Low	Critical
Developmental delay screened by ASQ-3 Psychometric values (< 2 SD) among children born preterm (GA <32wks) at 8 months, 18 months, and 30 months corrected age Diagnosis tool: Bayley III ≥ 1 SD below the mean												
1 (Schonhaut 2013)	Cross sectional	Serious risk of bias 1	N/A	no serious indirect-ness	serious imprecision [2]	none	0.86 (0.60-0.96)	0.86 (0.73-0.93)	6.0 (2.9-12.3)	0.17 (0.05-0.6)	Low	Critical
Screening: ASQ > 1 1SD below the mean, among children born at 25.4 weeks GA (range: 23.0-31.0 weeks), at 18-22 months corrected age; Diagnosis: BSID-II (> 2SD below the mean, either MDI or PDI)												
1 (Woodward 2012)	Follow-up of RCT, cross sectional study	Serious risk of bias [3]	N/A	No serious indirectness	No serious imprecision	None	0.94 (0.89-1.00)	0.32 (0.23-0.40)	1.39 (1.21-1.60)	0.16 (0.05-0.49)	Moderate	Critical
Screening: ASQ > 2 SD below the mean, among children born at 25.4 weeks GA (range: 23.0-31.0 weeks), at 18-22 months corrected age; Diagnosis: BSID-II (> 2SD below the mean, either MDI or PDI)												
1 (Woodward 2012)	Follow-up of RCT, cross	Serious risk of bias 3	N/A	No serious indirectness	No serious imprecision	None	0.73 (0.60-0.84)	0.65 (0.55-0.73)	2.05 (1.58-2.76)	0.42 (0.27-0.65)	Moderate	Critical

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens	Spec	LR+	LR-	Quality	Importance
	sectional study											
Screening: ASQ > 2 SD below the mean, among children born at 25.4 weeks GA (range: 23.0-31.0 weeks), at 18-22 months corrected age; Diagnosis: BSID-II (> 1SD below the mean, either MDI or PDI)												
1 (Woodward 2012)	Follow-up of RCT, cross sectional study	Serious risk of bias 3	N/A	No serious indirectness	No serious imprecision	None	0.63 (0.53-0.72)	0.76 (0.64-0.85)	2.47 (1.58-3.86)	0.50 (0.38-0.67)	Moderate	Critical
ASQ < 2 SD, among children born at < 31 weeks' GA, assessed at 18 months corrected age; Diagnosis: Bayley MDI < 1SD												
1 (Skellern 2001)	Cross sectional study	Serious risk of bias 4	N/A	No serious indirectness	Serious imprecision 2	None	0.50 (-0.19-1.19)	0.91 (0.79-1.03)	5.5 (0.81-37.2)	0.55 (0.14-2.2)	Low	
ASQ < 1SD; among children born at 29-36 weeks' GA, assessed at 12 months corrected age; Diagnosis: BSID-II MDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk	N/A	No serious indirectness	Serious imprecision 2	None	0.60 (0.39-0.81)	0.68 (0.59-0.77)	1.83 (1.17-2.87)	0.60 (0.36-1.01)	Low	Critical
ASQ < 1.5 SD; among children born at 29-36 weeks' GA, assessed at 12 months corrected age; Diagnosis: BSID-II MDI < 85												

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Sens**	**Spec**	**LR+**	**LR-**	**Quality**	**Importance**
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.45 (0.23-0.67)	0.78 (0.71-0.87)	2.25 (1.23-4.11)	0.68 (0.46-1.01)	Moderate	Critical
ASQ < 2 SD; among children born at 29-36 weeks' GA, assessed at 12 months corrected age; Diagnosis: BSID-II MDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.20 (0.02-0.38)	0.88 (0.82-0.95)	1.50 (0.53-4.21)	0.93 (0.75-1.15)	Moderate	Critical
ASQ < 1SD; among children born at 29-36 weeks' GA, assessed at 12 months corrected age; Diagnosis: BSID-II PDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Very serious imprecision [5]	None	0.52 (0.38-0.67)	0.90 (0.83-0.96)	5.04 (2.46-10.3)	0.53 (0.38-0.74)	Low	Critical
ASQ < 1.5 SD; among children born at 29-36 weeks' GA, assessed at 12 months corrected age; Diagnosis: BSID-II PDI < 85												

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens	Spec	LR+	LR-	Quality	Importance
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Very serious imprecision [5]	None	0.39 (0.24-0.53)	0.96 (0.92-1.00)	7.33 (2.62-20.5)	0.65 (0.51-0.83)	Low	Critical
ASQ < 2 SD; among children born at 29-36 weeks' GA, assessed at 12 months corrected age; Diagnosis: BSID-II PDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Very serious imprecision [5]	None	0.25 (0.12-0.38)	0.97 (0.94-1.00)	9.85 (2.29-42.4)	0.76 (0.64-0.91)	Low	Critical
ASQ < 1SD; among children born at 29-36 weeks' GA, assessed at 24 months corrected age; Diagnosis: BSID-II MDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.92 (0.81-1.00)	0.558 (0.45-0.66)	2.07 (1.59-2.69)	0.14 (0.04-0.53)	Moderate	Critical
ASQ < 1.5 SD; among children born at 29-36 weeks' GA, assessed at 24 months corrected age; Diagnosis: BSID-II MDI < 85												

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens	Spec	LR+	LR-	Quality	Importance
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.88 (0.74-1.00)	0.72 (0.63-0.82)	3.34 (2.27-4.90)	0.16 (0.05-0.46)	Moderate	Critical
ASQ < 2 SD; among children born at 29-36 weeks' GA, assessed at 24 months corrected age; Diagnosis: BSID-II MDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.75 (0.58-0.92)	0.78 (0.69-0.87)	3.46 (2.17-5.51)	0.33 (0.17-0.63)	Moderate	Critical
ASQ < 1SD; among children born at 29-36 weeks' GA, assessed at 24 months corrected age; Diagnosis: BSID-II PDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.50 (0.31-0.69)	0.73 (0.64-0.83)	1.82 (1.09-3.03)	0.69 (0.47-1.02)	Moderate	Critical
ASQ < 1.5 SD; among children born at 29-36 weeks' GA, assessed at 24 months corrected age; Diagnosis: BSID-II PDI < 85												

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Sens**	**Spec**	**LR+**	**LR-**	**Quality**	**Importance**
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.50 (0.31-0.69)	0.73 (0.64-0.83)	1.82 (1.09-3.03)	0.69 (0.47-1.02)	Moderate	Critical
ASQ < 2 SD; among children born at 29-36 weeks' GA, assessed at 24 months corrected age; Diagnosis: BSID-II PDI < 85												
1 (Simard 2012)	Follow-up of longitudinal study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.31 (0.13-0.49)	0.92 (0.86-0.98)	3.95 (1.51-10.36)	0.77 (0.59-0.97)	Moderate	Critical
ASQ score 270, among children born at ≤ 35 week GA, assessed at 5 years; Diagnosis: IQ lower than score 70 on WPPSI-III												
1 (Halbwachs 2013)	Cross sectional study	No serious risk of bias of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.85 (0.68-0.94)	0.81 (0.77-0.85)	4.46 (3.47-5.7)	0.18 (0.07-0.45)	Moderate	Critical
ASQ score 280, among children born at ≤ 35 week GA, assessed at 5 years; Diagnosis: IQ lower than score 85 on WPPSI-III												
1 (Halbwachs 2013)	Cross sectional study	No serious risk of bias	N/A	No serious indirectness	No serious imprecision	None	0.80 (0.71-0.87)	0.54 (0.48-0.60)	1.74 (1.50-2.02)	0.37 (0.24-0.56)	High	Critical

1 *Sens: sensitivity; Spec: specificity; LR+: positive likelihood ratio; LR-: negative likelihood ratio;*

1. Evidence was downgraded by 1 level because of the selection bias in the sample recruited;
2. Evidence was downgraded by 1 level due to the wide confidence intervals on some accuracy estimates;
3. Evidence was downgraded by 1 level because of the selection bias of the sample recruited (follow-up study of an earlier RCT)
4. Evidence was downgraded by 1 level because the study did not clearly report whether diagnosis outcome assessors were blinded to the screening results.
5. Evidence was downgraded by 1 level due to the very wide confidence intervals on some accuracy estimates

Table 49: Clinical evidence profile: Diagnostic accuracy of screening tool (PARCA-R) in correctly identifying intellectual disability

Quality assessment							**Summary of findings: diagnostic accuracy**					
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Sens**	**Spec**	**LR+**	**LR-**	**Quality**	**Importance**
Developmental delay screened by PARCA-R < 44, among children born at 22-31wks GA assessed at 2 years (corrected age) Diagnosis tool: BSID-II MDI < 70												
1 (Cuttini 2012)	Cross sectional study	Serious risk of bias [1]	N/A	No serious indirectness	Serious imprecision [2]	None	0.64 (0.35-0.92)	0.79 (0.71-0.87)	3.01 (1.69-5.36)	0.46 (0.21-1.01)	Low	Critical
Developmental delay screened by PARCA-R < 46, among children born at 22-31wks GA assessed at 2 years (corrected age) Diagnosis tool: BSID-II MDI < 70												
1 (Cuttini 2012)	Cross sectional study	Serious risk of bias 1	N/A	No serious indirectness	Very serious imprecision [3]	None	0.73 (0.46-0.99)	0.77 (0.69-0.85)	3.17 (1.92-5.22)	0.35 (0.13-0.93)	Very low	Critical
Developmental delay screened by PARCA-R < 68, among children born at 22-31wks GA assessed at 2 years (corrected age) Diagnosis tool: BSID-II MDI < 70												
1 (Cuttini 2012)	Cross sectional study	Serious risk of bias [1]	N/A	No serious indirectness	Very serious imprecision on 3	None	0.85 (0.71-0.98)	0.64 (0.54-0.73)	2.34 (1.71-3.20)	0.24 (0.09-0.60)	Very low	Critical
Development delay screened by PARCA-R < 44, among children born at 22-31wks GA assessed at 2 years (corrected age) Diagnosis: BSID-II MDI < 70												

1 (Johnson 2008)	Cross sectional study	Serious risk of bias 4	N/A	No serious indirectness	Serious imprecision 2	None	0.85 (0.58-0.96)	0.87 (0.81-0.92)	6.72 (4.16-10.8)	0.18 (0.05-0.63)	Low	Critical
Development delay screened by PARCA-R < 49, among preterm children (GA < 32wks) assessed at 2 years (corrected age); Diagnosis: BSID-II MDI < 70												
1 (Johnson 2008)	Cross sectional study	Serious risk of bias [4]	N/A	No serious indirectness	Serious imprecision [2]	None	0.85 (0.58-0.960	0.83 (0.77-0.88)	5.11 (3.36-7.82)	0.18 (0.05-0.66)	Low	Critical
Developmental delay screened by PARCA-R < 73, among children born at 32-36wks GA assessed at 25 months corrected age; Diagnosis: BSID-III MDI < 70												
1 (Blaggan 2014)	Follow-up of a cohort study (cross sectional study)	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.90 (0.77-1.03)	0.76 (0.70-0.82)	3.73 (2.80-4.97)	0.13 (0.04-0.49)	Moderate	Critical
Developmental delay screened by PARCA ≤19 (cognitive component), among children born at median 27wks GA assessed at 5 years Diagnosis: Bayley-III cognition score < 70												
1 (Martin 2013)	Follow-up of an RCT	Very serious risk of bias [1, 4,]	N/A	No serious indirectness	Serious imprecision [2]	None	0.89 (0.68-1.00)	0.89 (0.84-0.94)	8.25 (5.18-13.14)	0.12 (0.02-0.79)	Very low	Critical
Developmental delay screened by PARCA ≤ 23 (language component), among children born at median 27wks GA assessed at 5 years Diagnosis: Bayley-III cognition score < 70												
1 (Martin 2013)	Follow-up of an RCT	Very serious risk of bias [1, 4,]	N/A	No serious indirectness	Serious imprecision [2]	None	0.75 (0.54-0.96)	0.79 (0.74-0.85)	3.62 (2.42-5.30)	0.32 (0.13-0.74)	Very low	Critical

Sens: sensitivity; Spec: specificity;
LR+: positive likelihood ratio;
LR-: negative likelihood ratio;
N/A: not applicable;
1. Evidence was downgraded by 1 level because the study did not clearly report whether the outcome assessors were blinded to the screening test results
2. Evidence was downgraded by 1 level due to the wide confidence intervals on some accuracy estimates;

3. Evidence was downgraded by 1 level due to the very wide confidence intervals on some accuracy estimates
4. Evidence was downgraded by 1 level because of the selection bias of the sample recruited (follow-up study of an earlier RCT)

Table 50: Clinical evidence profile: Diagnostic accuracy of screening tool (SDQ) in correctly identifying emotional and mental health disorder

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Sens	Spec	LR+	LR-	Quality	Importance
Any psychiatric disorder screened by SDQ > 90th percentile (mother's report), among children born at 24-36 weeks GA assessed at 14 years Diagnosis tool: psychiatric diagnosis by interview												
1 (Indredavik 2005)	Cross sectional	Serious risk of bias [1]	N/A	No serious indirectness	Serious imprecision 2	None	0.85 (0.67-1.00)	0.58 (0.42-0.74)	2.04 (0.32-3.12)	0.25 (0.06-0.92)	Low	Critical
Any psychiatric disorder screened by SDQ > 90th percentile (father's report), among children born at 24-36 weeks GA assessed at 14 years Diagnosis tool: psychiatric diagnosis by interview												
1 (Indredavik 2005)	Cross sectional	Serious risk of bias [1]	N/A	No serious indirectness	No serious imprecision	None	0.50 (0.24-0.76)	0.75 (0.61-0.90)	2.06 (0.93-4.59)	0.66 (0.38-1.15)	Moderate	Critical
Any psychiatric disorder screened by SDQ > 90th percentile (teacher's report), among children born at 24-36 weeks GA assessed at 14 years Diagnosis tool: psychiatric diagnosis by interview												
1 (Indredavik 2005)	Cross sectional	Serious risk of bias 1	N/A	No serious indirectness	Serious imprecision 2	None	0.57 (0.31-0.83)	0.88 (0.78-0.98)	4.80 (1.88-12.28)	0.49 (0.26-0.90)	Low	Critical
Emotional disorder screened by abnormal parent SDQ, among children born at <26wks GA assessed at 11 years Diagnosis tool: DWABA												
1 (Johnson 2014)	Cross sectional	No serious risk of bias	N/A	No serious	Serious imprecision [2]	None	0.67 (0.43-0.85)	0.80 (0.78-0.820	3.29 (2.13-5.09)	0.41 (0.22-0.80)	Moderate	Critical

				indirectness								
Emotional disorder screened by abnormal teacher SDQ, among children born at <26 weeks GA assessed at 11 years Diagnosis tool: DWABA												
1 (Johnson 2014)	Cross sectional	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.29 (0.12-0.53)	0.90 (0.88-0.93)	2.37 (1.01-5.52)	0.81 (0.61-1.09)	Moderate	Critical
Conduct disorder screened by abnormal parent SDQ, among children born at <26 weeks GA assessed at 11 years Diagnosis tool: DWABA												
1 (Johnson 2014)	Cross sectional	No serious risk of bias	N/A	No serious indirectness	Very serious imprecision [3]	None	0.67 (0.37-0.88)	0.90 (0.89-0.92)	6.91 (3.84-12.41)	0.37 (0.16-0.82)	Low	Critical
Conduct disorder screened by abnormal teacher SDQ, among children born at <26 weeks GA assessed at 11 years Diagnosis tool: DWABA												
1 (Johnson 2014)	Cross sectional	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [2]	None	0.33 (0.12-0.60)	0.95 (0.94-0.97)	5.89 (2.48-19.16)	0.70 (0.47-1.05)	Moderate	Critical

Sens: sensitivity; Spec: specificity;
LR+: positive likelihood ratio;
LR-: negative likelihood ratio;
N/A: not applicable;
1. Evidence was downgraded by 1 level because the study did not clearly report whether the outcome assessors were blinded to the screening test results
2. Evidence was downgraded by 1 level due to the wide confidence intervals on some accuracy estimates;
3. Evidence was downgraded by 1 level due to the very wide confidence intervals on some accuracy estimates

Table 51: Clinical evidence profile: Diagnostic accuracy of screening tool (DCDQ) in correctly identifying developmental coordination disorder (DCD)

Quality assessment							Summary of findings: diagnostic accuracy					
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Sens**	**Spec**	**LR+**	**LR-**	**Quality**	**Importance**
DCDQ ≤ 15th percentile, among children born at 24-35wks GA assessed at 5 years Diagnosis: Movement ABC												
1 (Dewey 2011)	Cross sectional	No serious risk of bias	N/A	No serious indirectness	Serious imprecision [1]	None	0.37 (0.25-0.48)	0.91 (0.83-1.00)	4.49 (1.45-13.9)	0.69 (0.56-0.85)	Moderate	Critical

Sens: sensitivity; Spec: specificity;
LR+: positive likelihood ratio;
LR-: negative likelihood ratio;
N/A: not applicable;
1. Evidence was downgraded by 1 level due to the wide confidence intervals on some accuracy estimates;

5.1.3.4 Economic evidence

The identification of problems and disorders that might arise during the developmental follow-up of preterm children is of considerable economic concern. It is important – both clinically and economically – to begin to manage conditions as early as possible, but screening and surveillance tools require resource input from the NHS. On the other hand, many problems and disorders are almost impossible to identify with great certainty early, and early misdiagnosis may have economic and human costs (such as increasing parental anxiety). The ideal identification strategy uses as few tests as possible to diagnose children as accurately as possible, and therefore the competing use of available resources makes the question very appropriate for economic modelling.

No economic evaluations of the identification of problems and disorders during follow-up of pre-term infants was found.

5.1.3.4.1 Methods

A Markov decision analytic model was developed in Microsoft Excel® to assess the cost effectiveness of various surveillance strategies.

A conventional health economic model on the most accurate identification tools would require cost and QALY inputs from three cohorts of children in order to produce ICERs:

- with the condition and treated
- without the condition and treated
- with the condition and not treated

However, for obvious ethical reasons, there is no evidence available on children who were confirmed as having a condition and then not treated. Consequently the conventional model structure of a cost-effectiveness analysis was not selected, and instead a cost-consequence analysis was chosen; keeping the main outputs (percentage of children diagnosed) in their natural units.

The model was designed to consider the costs of identifying developmental conditions with various combinations of testing schedules and instruments. The tables below detail the testing schedules and screening instruments that were considered in the model.

Table 52: Summary characteristics of testing schedules

Schedule name	Source	Characteristics	Notes
'Screen and forget'	Assumption based	One test before 1 year, no subsequent tests	Included as a baseline – not intended as a realistic option
Southampton Protocol	University Hospital Southampton (communication with Committee member)	Seven contacts in first and second year, eight contacts in year four	Noted by the Committee as a high-quality UK-based service
Nottingham Protocol	Marlow et al (2005)	Three contacts in first year, six in second year and three in third year	
Old Canadian	Synnes et al (2006)	Four contacts in first year, two in second year and one contact in third and fifth year	Canada had 19 protocols; the model assumes a weighted average of these protocols

Schedule name	Source	Characteristics	Notes
New Canadian	Canadian Government	Six contacts each in year one, two and three	Noted by the Committee as a high quality service (although not based in the UK)
Healthy Child Programme	UK Government	Six contacts in first year, one contact in second year and two contacts in third and fifth year	
Healthy Child Program + Recs	Guideline Committee	As Healthy Child Program, plus two additional contacts in second year, and one additional contact at third and fifth years	Intended to approximately represent recommendations made by the Committee

Table 53: List of screening instruments included in economic model

Instrument	Sensitivity	Specificity	Source
Never offer reference standard[b]	0.00	1.00	Assumption
Always offer reference standard[b]	1.00	0.00	Assumption
PARCA-R <49 cutoff	0.35	0.90	Blaggan et al. 2014
PARCA-R <44 cutoff	0.35	0.94	Blaggan et al. 2014
PARCA-R <73 cutoff	0.90	0.76	Blaggan et al. 2014
DCDQ <15% cutoff	0.37	0.91	Dewey et al. 2011
ASQ 285 (for IQ <85)	0.80	0.54	Halbwachs et al. 2013, based on accuracy data at 5 years of age in children born at ≤35 weeks.
ASQ 270 (for IQ <70)	0.85	0.81	Halbwachs et al. 2013, based on accuracy data at 5 years of age in children born at ≤35 weeks.
VLBW, mother's SDQ >90% and in-depth interview	0.85	0.58	Indredavik et al. 2005
VLBW, father's SDQ >90% and in-depth interview	0.50	0.75	Indredavik et al. 2005
VLBW, teacher's SDQ >90% and in-depth interview	0.57	0.88	Indredavik et al. 2005
<26wk GA, diagnosed psychiatric disorder, parent SDQ score (conduct disorder)	0.67	0.90	Johnson et al. 2014
<26wk GA, diagnosed psychiatric disorder, teacher SDQ score (conduct disorder)	0.33	0.95	Johnson et al. 2014
ASQ-3 <2SD below mean	0.59	0.87	Schonhaut et al. 2013
ASQ <1SD (BSID-II PDI <85)	0.60	0.68	Simard et al. 2012, based on accuracy data at 12 months

Instrument	Sensitivity	Specificity	Source
			corrected age in children born at 29-36 weeks.
ASQ <1.5SD (BSID-II PDI <85)	0.45	0.78	Simard et al. 2012, based on accuracy data at 12 months corrected age in children born at 29-36 weeks.
ASQ <2SD (BSID-II PDI <85)	0.20	0.88	Simard et al. 2012, based on accuracy data at 12 months corrected age in children born at 29-36 weeks.

(a) Source is clinical review for all except where noted below. ASQ = Ages and Stages Questionnaire, PARCA-R = Parent Report of Children's Abilities-Revised, DCDQ = Developmental Coordination Disorder Questionnaire, SDQ = Strengths and Difficulties Questionnaire, VLBW = Very Low Birth Weight, GA = Gestational Age
(b) By definition

The model identifies the cost of diagnosing a condition given a certain identification strategy followed in the population of children born preterm (subdivided by degree of prematurity). To make the model more relevant to clinical practice, the base case assumes that one instrument is used to identify multiple conditions; for example one instrument can be used to identify moderate intellectual disability and severe intellectual disability. Long term costs of treating conditions are not considered in the model due to data limitations, but it is assumed that the treatment of these conditions is cost-effective and therefore society would prefer a diagnosis to no diagnosis.

The model assumes that children will contact the healthcare service a number of times, and at each contact there is a probability that the care professional will notice something concerning about their development, or the parent or carer will raise a concern. When a concern is raised, it is assumed a screening or identification instrument is offered such as the Ages and Stages Questionnaire. If this instrument also indicates concern, an appointment for a 'reference standard' diagnostic test is made – for example the Bayley scales of infant and toddler development. It is assumed this reference standard is perfectly sensitive and specific, so once a concern is escalated to this level it is impossible for an incorrect diagnosis to be made.

Each developmental problem and disorder is specified with an age at which it becomes 'obvious', meaning that there is no question that a disorder exists, or potentially 'existed in the past'. Since there is no health economic evidence considering lifetime costs following on from diagnosis at different ages, the model runs only until the age at which the last condition becomes 'obvious', which the Committee agreed was likely to be around 18.

Where possible, costs were based on an NHS and Personal Social Services perspective as outlined in the NICE Reference Case (The guidelines manual, NICE October 2014). The price year for costs was 2016.

Table 54 The table below shows the estimated costs for the use of each screening instrument in the model. Note that on top of the cost of actually administering the test, it is assumed that there is a cost associated with explaining the results of the test. This cost is likely to be higher where the test indicates cause for concern and lower where the test does not. A flat cost of £27 (a GP telephone appointment) is added to all tests to be indicative of the cost of a 'no concern' discussion, and the additional cost of a 'reason for concern' discussion is added to the cost of the reference standard instrument.

Table 54: Estimated cost of screening instruments

Instrument	Estimated total / Test	License fee / Test	Salary cost / Test	Notes
No Test	£0.00	N/A	N/A	

Instrument	Estimated total / Test	License fee / Test	Salary cost / Test	Notes
ASQ	£28.50	£0.00	£1.50	No fee per test, assumed to be set by parent (free) and scored by practice nurse for 2.5 minutes at £36 / hour
SDQ (parent)	£28.64	£0.14	£1.50	£0.14 fee per test (converted to UK currency from US fee of $0.20), set by parent (free) and scored by practice nurse for 2.5 minutes at £36 / hour
SDQ (teacher)	£38.64	£0.14	£11.50	£0.14 fee per test (converted to UK currency from US fee of $0.20), assumed to be set by teacher for 15 minutes at £40 / hour[b] and scored by practice nurse for 2.5 minutes at £36 / hour
DCDQ	£27.61	£0.01[b]	£0.60	Nominal fee per test, assumed to be set and scored by parent (free) with practice nurse providing one minute of advice at £36 / hour
PARCA-R	£28.50	Unknown	Unknown	No information found, assumed to be similar to ASQ / SDQ
In-depth interview	£95.50	N/A	£68.50	Assumed to be 30 minutes with consultant

(a) ASQ = Ages and Stages Questionnaire, PARCA-R = Parent Report of Children's Abilities-Revised, DCDQ = Developmental Coordination Disorder Questionnaire, SDQ = Strengths and Difficulties Questionnaire
(b) Cost of teachers' salary potentially falls outside the scope of NICE Reference Case as it is not an NHS / PSS cost. However it is thought teacher time represents an opportunity cost to the NHS in the case of preterm infants, so there is at least a reason to consider teacher time as a relevant cost even if taking a very strict definition of payer perspective. However this should likely not be the full cost of the teacher's time to the state.

As the simulation runs over a time period of greater than one year, a discount rate of 3.5% for both costs and QALYs is employed as per the NICE Reference Case.

5.1.3.4.2 Results

Table 55 demonstrates the main schedule of results. The costs describe the total cost over 18 years to identify one case of a developmental problem. It demonstrates that for any given test or screening strategy, there is always some other test or screening strategy with a lower cost for at least some population. This means that there is no 'dominated' test or schedule. However in general the Healthy Child Programme (HCP) appears to offer the cheapest screening strategy. Note that the HCP+Recs strategy significantly outperform the HCP alone for some screening strategies, whereas the HCP alone only slightly outperforms the HCP+Recs across the board.

As with the schedules of screening, there is no instrument which universally dominates, although PARCA-R <73 cutoff, ASQ at any cutoff and parent-scored SDQ in combination with a diagnosed psychiatric disorder all performed well in general. No test or no screening is an order of magnitude more expensive than even the worst screening test; these should be avoided if at all possible. Always offering a reference standard test performs well given its unsophisticated nature; clinicians who are uncertain of how to use screening instruments might consider a referral without too much risk of making a cost-ineffective decision.

Table 55: Main schedule of results: total cost over 18 years to identify one case of a developmental problem

	'Screen and Forget'	Southampton	Nottingham	Old Canadian	New Canadian	HCP	HCP + Recs
No tests	£23,113	£23,782	£23,469	£23,342	£23,662	£23,436	£23,561
Always Test	£23,387	£1,179	£1,255	£991	£1,119	£922	£1,030
PARCA-R <49 cutoff	£23,241	£1,474	£2,451	£3,318	£1,577	£2,140	£1,686
PARCA-R <44 cutoff	£23,239	£1,451	£2,438	£3,309	£1,558	£2,128	£1,670
PARCA-R <73 cutoff	£23,367	£904	£1,164	£948	£905	£815	£864
DCDQ <15% cutoff	£23,244	£1,393	£2,276	£2,997	£1,483	£1,934	£1,565
ASQ 285 (for IQ <85)	£23,358	£1,073	£1,323	£1,075	£1,060	£933	£1,002
ASQ 270 (for IQ <70)	£23,354	£897	£1,190	£972	£906	£825	£869
VLBW, mother's SDQ >90% & in-depth interview	£23,499	£1,432	£1,556	£1,269	£1,376	£1,172	£1,284
VLBW, father's SDQ >90% & in-depth interview	£23,418	£1,713	£2,107	£2,120	£1,705	£1,647	£1,648
VLBW, teacher's SDQ >90% & in-depth interview	£23,440	£1,574	£1,913	£1,772	£1,564	£1,451	£1,500

	'Screen and Forget'	Southampton	Nottingham	Old Canadian	New Canadian	HCP	HCP + Recs
<26wk GA, diagnosed psychiatric disorder, parent SDQ score (conduct disorder)	£23,312	£952	£1,353	£1,164	£982	£927	£954
<26wk GA, diagnosed psychiatric disorder, teacher SDQ score (conduct disorder)	£23,249	£1,595	£2,688	£3,713	£1,716	£2,413	£1,859
ASQ-3 <2SD below mean	£23,296	£1,036	£1,482	£1,362	£1,071	£1,036	£1,046
ASQ <1SD (BSID-II PDI <85)	£23,308	£1,133	£1,526	£1,375	£1,150	£1,077	£1,109
ASQ <1.5SD (BSID-II PDI <85)	£23,270	£1,275	£1,876	£2,127	£1,322	£1,451	£1,325
ASQ <2SD (BSID-II PDI <85)	£23,208	£2,870	£5,347	£7,569	£3,306	£5,427	£3,879

5.1.3.4.3 Sensitivity analyses

Numerous one-way sensitivity analyses were conducted to assess the consequences of the uncertainty around the key input parameters (see Appendix H: for the full results). The model behaved as expected and the results were generally thought to be robust. A particularly notable sensitivity analysis was that conducted on the cost of the reference standard, which showed that always testing is preferred if the reference test is <£100, but that otherwise some kind of screening protocol is preferred.

5.1.3.5 Evidence statements

5.1.3.5.1 ASQ

Among preterm children (GA 32-36wks) assessed at age 8-month, 18-month, and 30-month:

Low quality evidence from 1 study investigating the diagnostic value of ASQ on intellectual disability found that a cut-off of ASQ < 2SD below the mean gave a moderate sensitivity, low specificity, and not useful positive or negative likelihood ratio for the reference diagnosis of developmental delay defined as Bayley-III ≥ 1SD below the mean for this population.

Among extremely preterm children (GA 32 weeks) assessed at age 8-month, 18-month, and 30-month:

Low quality evidence from 1 study investigating the diagnostic value of ASQ on intellectual disability found that a cut-off of ASQ < 2SD below the mean gave a moderate sensitivity and specificity, and moderately useful positive or negative likelihood ratio for the reference diagnosis of developmental delay defined as Bayley-III ≥ 1SD below the mean for this population.

Among preterm children (mean GA 25.4wks) assessed at age 18-22 months corrected age:

Moderate quality evidence from 1 study investigating the diagnostic value of ASQ on intellectual disability found that a cut-off of ASQ > 1 SD below the mean gave a high sensitivity, low specificity, and not useful positive or negative likelihood ratio on intellectual disability compared to the reference standard of BSID-II > 2 SD below the mean in this population. When the cut-off of ASQ > 2 SD was assessed, it gave a low sensitivity and specificity, and not useful positive or negative likelihood ratio (moderate quality evidence). The same was found when the ASQ cut-off > 2 SD was assessed and when the reference standard was BSID-II > 1 SD below the mean (moderate quality evidence).

Among preterm children (GA < 31wks) assessed at 18 months corrected age:

Low quality evidence from 1 study investigating the diagnostic value of ASQ on intellectual disability found that a cut-off of ASQ < 2 SD gave a low sensitivity, high specificity, moderately useful positive likelihood ratio and not useful negative likelihood ratio when compared to the reference standard of Bayley MDI < 1 SD.

Among preterm children (GA 29-36wks) assessed at 12 months corrected age:

Moderate to low quality evidence from 1 study investigating the diagnostic value of ASQ on intellectual disability found that a cut-off of ASQ <1SD gave a low sensitivity and specificity, and not useful positive and negative likelihood ratio when compared to the reference standard of BSID-II MDI < 85. The same study found that the cut-offs of ASQ < 1.5 SD and ASQ < 2 SD each gave a moderate specificity, however both gave a low sensitivity and not useful positive or negative likelihood ratios (moderate quality evidence).

When the reference standard was BSID-II PDI < 85, moderate to low quality evidence from the same study found that the cut-offs of ASQ < 1 SD, < 1.5 SD, and < 2 SD all gave a high specificity and moderately useful positive likelihood ratios. However all cut-offs assessed gave a low specificity, and not useful negative likelihood ratios.

Among preterm children (GA 29-36wks) assessed at 24 months corrected age:

Moderate to low quality evidence from 1 study investigating the diagnostic value of ASQ on intellectual disability found that a cut-off of ASQ <1 SD gave a high sensitivity, low specificity, not useful positive likelihood ratio but moderately useful negative likelihood ratio when the reference standard was BSID-II MDI < 85 in this population. The same study found that the cut-offs of ASQ < 1.5 SD and ASQ < 2 SD each gave a moderate specificity, moderate or close to moderate specificity, and not useful positive likelihood ratio. The cut-off of ASQ < 1.5 SD was found to give a moderately useful negative likelihood ratio on compared to BSID-II MDI < 85 in this population.

When the reference standard was BSID-II PDI < 85, moderate quality evidence from the same study found that the cut-offs of ASQ < 1 SD, < 1.5 SD both gave a low sensitivity and specificity, and not useful positive and negative likelihood ratio. The cut-off of ASQ < 2 SD in this study gave a high specificity but low sensitivity and not useful positive or negative likelihood ratio when compared to the reference standard of BSID-II PDI < 85.

Among preterm children (GA ≤ 35wks) assessed at age 5 years:

Moderate quality evidence from 1 study found that ASQ score < 270 gave a moderate sensitivity or specificity, not useful positive likelihood ratio, and moderately useful negative likelihood ratio compared to IQ score < 70 on WPPSI-III in this population.

Moderate quality evidence from the same study found that ASQ score < 280 gave a moderate sensitivity, low specificity, and not useful positive or negative likelihood ratio compared to IQ score < 85 on WPPSI-III in this population.

5.1.3.5.2 *PARCA-R*

Among preterm-children (GA 22-31 weeks) assessed at 2 years (corrected age):

Low quality evidence from 1 study investigating the diagnostic value of PARCA-R on intellectual disability found that a PARCA-R score < 44 gave a low sensitivity, moderate specificity, and not useful positive and negative likelihood ratio for diagnosis of developmental delay defined as BSID-II MDI < 70 (reference standard) in this population. The same was found for the cut-off of PARCA-R score < 46. When the cut-off of PARCA-R score < 68 was assessed, it was found to give a moderate sensitivity, low specificity and not useful positive and negative likelihood ratio. The reference standard was BSID-II MDI < 70 for all of the comparisons.

Among preterm-children (GA < 32 weeks) assessed at 2 years (corrected age):

Low to moderate quality evidence from 1 study investigating the diagnostic value of PARCA-R on intellectual disability found that a cut-off of PARCA-R score < 44 and < 49 gave moderate sensitivity and specificity, a moderately useful positive and negative likelihood ratios when compared to the reference standard of BSID-II MDI < 70.

Among pre-term children (GA 32-36 weeks) assessed at 25 months (corrected age):

Moderate quality evidence from 1 study investigating the diagnostic value of PARCA-R on development delay found that a cut-off of PARCA-R score < 73 gave a high sensitivity, moderate specificity, not useful positive likelihood ratio, and moderately useful negative likelihood ratio when compared to the reference standard of BSID-III MDI < 70.

Among pre-term children (median GA 27 weeks) assessed at 24 months corrected age:

Very low quality evidence from 1 study assessing the diagnostic value of PARCA-R on cognitive impairment found that a cut-off PARCA score ≤ 19 (cognitive component) gave a moderate sensitivity and specificity, moderately useful positive and negative likelihood ratio when the reference standard was BSID-II cognition score < 70 in this population.

Very low quality evidence from 1 study assessing the diagnostic value of PARCA-R on cognitive impairment found that a cut-off PARCA score ≤ 23 (language component) gave a moderate sensitivity and specificity, and not useful positive and negative likelihood ratio when the reference standard was BSID-II cognition score < 70 in this population.

5.1.3.5.3 *SDQ*

Among preterm children (GA 24-36 weeks) assessed at 14 years:

Low quality evidence from 1 study assessing the diagnostic value of SDQ on psychiatric disorders found that a cut-off SDQ score > 90th percentile (mother's report) gave a moderate sensitivity, low specificity, and not useful positive and negative likelihood ratio compared to clinical diagnosis of a psychiatric disorder based on psychiatric interviews in this population. SDQ score > 90th percentile (both the father's and teacher's reports) gave a low sensitivity, moderate specificity, and not useful positive and negative likelihood ratio when the reference standard was clinical diagnosis based on psychiatric interviews.

Among preterm children (GA < 26 weeks) assessed at 11 years:

Moderate to low quality evidence from 1 study assessing the diagnostic value of SDQ on emotional disorder found that abnormal SDQ score (as defined by the instrument developers) reported by either parents or teachers gave a low sensitivity, moderate specificity, and not useful positive and negative likelihood ratio when the reference standard was diagnosis by DAWBA.

Moderate to low quality evidence from the same study assessing the diagnostic value of SDQ on conduct disorder found that abnormal SDQ score (as defined by the instrument developers) reported by either parents and teachers gave a low sensitivity, moderate specificity, moderately useful positive likelihood ratio, and not useful negative likelihood ratio when the reference standard was diagnosis by DAWBA.

5.1.3.5.4 DCDQ

Among preterm children (GA 24-35 weeks) assessed at 5 years:

Moderate to low quality evidence from 1 study assessing the diagnostic value of DCDQ on developmental coordination disorder (DCD) found that a cut-off of DCDQ score ≤ 15th percentile gave a low sensitivity, high specificity, and not useful positive and negative likelihood ratio when the reference standard was Movement ABC score ≤ 15th percentile.

The association between SDQ measured at 6 years and the diagnosis of psychiatric disorder (DWAB) at 11 years:

Moderate quality evidence from 1 study found that pervasive attentional problems measured by SDQ at age 6 years was positively associated with the risk of psychiatric disorders when the preterm children reached 11 years of age. The same positive association was found between the pervasive conduct problems measured by SDQ at age 6 years and the diagnosis of psychiatric disorders made at age 11 years using DAWBA.

5.1.3.6 Economic evidence statement

A literature review of published cost-effectiveness analyses did not identify any relevant studies. The economic modelling undertaken for this question demonstrated that for any given schedule, there is always some other strategy with a lower cost for at least some populations. This means that there is no 'dominated' schedule. However the results do provide an indicated that the HCP and the HCP+Recs may be preferred as they perform well in most populations.

Similarly, it was found that there is no instrument which universally dominates, although PARCA-R <73 cutoff, ASQ at any cutoff and parent-scored SDQ in combination with a diagnosed psychiatric disorder all performed well in general.

5.1.3.7 Evidence to recommendations

5.1.3.7.1 Relative value placed on the outcomes considered

The aim of the review was to assess the value and accuracy of different tools to identify developmental problems and disorders in children born preterm. The Committee focused on the sensitivity, specificity, positively likelihood ratios, and negative likelihood ratios when considering the value of tools for identifying children born preterm who were at risk of developmental disorders.

The Committee considered the relative importance of having a high false positive and high false negative result from the screening and the consequences for the child and family. They agreed that screening tools should have a high sensitivity at this stage to identify as many children born preterm who were at risk as possible so that they could be referred for further assessment and treatment/intervention. Although specificity was generally more important at diagnosis, the Committee noted that a high specificity at a surveillance level was also important because in a test with high specificity, a positive result would indicate that one could be fairly sure that a child who screened positive had the problem or disorder and the child would not be subject to unnecessary further testing. The Committee recognised that the prevalence of the condition under consideration was an important factor and agreed that likelihood ratios were the most important measures of the value of the tools because they do not vary according to prevalence and can be used to determine post-test probabilities of the condition. The positive likelihood ratio reports how many times more likely children born preterm with disorder were to have a positive screening result compared with those who did not have the disorder. The higher the value, the more likely it was that a child with a positive test has the disorder.

5.1.3.7.2 Consideration of clinical benefits and harms

The Committee discussed how identifying evidence for the effectiveness of isolated screening measures was important for directing professionals in their practice, but that such screening measures should never be used in isolation. Professionals should always endeavour to gather and triangulate several sources of information about a child's development when forming any view about potential developmental problems or disorders.

In this review the accuracy of different screening tools were compared to diagnostic tests which are used in clinical practice and considered as golden standards for the identification of different developmental disorders or problems in children born preterm.

Screening tools for identifying global developmental delay/intellectual disability:

The Committee discussed the evidence on the accuracy of the ASQ and PARCA-R as tools to identify global developmental delay/intellectual disability compared with standardised tests.

The Committee noted that the evidence on ASQ from 4 studies used different diagnostic cut-offs and the ages of assessment varied among children born preterm (12 months, 18 months, 24 months and 25 months corrected age, and 5 years). The evidence showed mixed results of the accuracy of the ASQ compared to diagnostic tests considered as gold standards in current practice. Somewhat to their surprise, the Committee was not convinced of the usefulness of ASQ as a tool to identify global developmental delay/intellectual disability among children born preterm. Therefore, the Committee agreed not to recommend ASQ for screening children born preterm in the enhanced surveillance programme. The Committee agreed that more research was needed on the predictive value of ASQ at different ages.

Regarding PARCA-R, 3 studies carried out among preterm children at 2 years (corrected age) reported moderate to high sensitivity, and positive and negative likelihood ratios showing the PARCA-R to be a moderately useful test at identifying global developmental delay/intellectual disability at 2 years of age when compared to the standardised BSID test (see also section 5.1.4.6).

The Committee agreed that the evidence was more strongly in favour of the PARCA-R compared to the ASQ as a screening tool to identify children who may have global developmental delay/intellectual disability at 2 years of age and therefore recommended

PARCA-R as the tool to be used at the 2 year assessment among children born preterm in the enhanced surveillance programme.

The Committee discussed the need to set a PARCA-R cut-off which is suggestive of global developmental delay/intellectual disability (when compared with BSID mental developmental index score of <70). They based this on moderate quality evidence which found a score of <44 to provide moderate sensitivity and specificity, a moderately useful positive and negative likelihood ratios.

The Committee noted that since PARCA-R was completed by parents, it could be administered electronically or through post. It didnot require a trained professional to administer it and therefore, would not have a significant resource impact when introduced as part of the enhanced surveillance programme. The Committee recognised the potential problem of poor return rates but also noted that since these families were expected to come for a clinic visit, although not ideal, the questionnaire could be filled in during the visit as well. The scoring of PARCA-R was considered easy and relatively quick to do. However, the Committee recognised that since the PARCA-R was not age-standardised, it could only be used in a limited time frame between 22 and 26 months of age (corrected for children born preterm). As with any other parent-filled questionnaire, a potential language barrier was also considered. Therefore, the Committee agreed that when the PARCA-R was not appropriate, a suitable alternative should be used and this should be selected by the healthcare professional depending on the needs of the child.

Screening tools for identifying DCD/motor problems:

Evidence from 1 study carried out among children born preterm at age 5 years assessing the diagnostic value of DCDQ reported a high specificity, but low sensitivity for identifying DCD/motor problems. The positive and negative likelihood ratios did not indicate DCDQ to be a useful tool in identifying DCD or motor problems in children born preterm. The Committee noted that the high specificity could be useful to correctly rule in children born preterm at risk of motor problems because in a test with high specificity, one could be fairly sure that a positive screening test result indicated that the child may have difficulties in this area. However, overall the scarce evidence that was available did not show DCDQ to be a very useful tool in identifying DCD or motor problems in children born preterm, the Committee agreed that its use should not be recommended for inclusion the enhanced surveillance programme for children born preterm.

Due to the lack of evidence, the Committee recommended that further research should be carried out on the value of DCDQ or other screening tools for identifying motor problems among pre-school children born preterm.

Screening tools for identifying social, emotional and behavioural problems:

The accuracy of the SDQ compared to diagnostic tests (clinical diagnosis or DAWBA) in identifying emotional disorder and conduct disorder in children born preterm was assessed in 2 studies among children aged 11 and 14 years. The evidence showed that SDQ had a high specificity for emotional disorders, high specificity and moderately useful positive likelihood ratio for conduct disorder when administered by either parents or teachers. Even though the evidence did not show SDQ to be a very useful test for identifying emotional or conduct disorders, the Committee noted SDQ would likely flag up concerns or problems in relation to behavioural and emotional development. Therefore, the Committee agreed to recommend the use of SDQ at 4 years of age to the children born preterm in the enhanced surveillance programme (see Section 5.1.4.6).

Screening tool for identifying specific learning disorder:

The Committee noted that no evidence was found on screening tools identifying specific learning disorders in this review.

5.1.3.7.3 Consideration of economic benefits and harms

A systematic review of the economic literature was conducted but no relevant studies were identified that were applicable to this review question.

The Committee's considerations around the economic benefits and harms was informed by the results of the economic analysis. It wasnoted that identifying developmental disorders and problems carried a benefit in terms of providing an opportunity to offer earlier intervention and support which could reduce the costs later on. However, the Committee also noted that there could be considerable 'hidden' costs in the form of anxiety caused to parents and over treatment of infants who did not actually have a developmental disorder or problem. In addition, there was a direct cost of offering identification tests, which varied by the intensity of the test.

The Committee were aware that there was no 'perfect' identification strategy, and that the mathematically optimal screening strategy depended on characteristics of the target population and assumptions about problem / disorder onset. In particular, the costs associated with particular strategies may vary in different healthcare geographies (especially relating to whether a particular healthcare geography had the resources to monitor 'false positives' and access to sufficient expertise to perform reference standard tests where appropriate).

A key point discussed by the Committee was the anxiety that unnecessary referrals could cause to parents. However, the Committee agreed that failing to diagnose a condition which really did exist almost always led to more difficulties compared to raising concerns about one which did not exist, and that the relative difference between these two possibilities depended on details of the circumstances. The Committee argued that in the context of a supportive healthcare team, being told an infant was being monitored for a minor condition could be more reassuring than a correct non-identification. The Committee also emphasised that in practice identification of risk rarely relied simply upon one screening test and that other features such as level of concern, severity, persistence and pervasiveness were all features to be taken into consideration in deciding about referral to a diagnostic pathway. Nevertheless, the Committee did not substantially change the overall conclusions regarding the effectiveness of early identification and importance of highly accurate tests.

5.1.3.7.4 Quality of evidence

Low to moderate quality evidence was found in the review. The main reason for downgrading of evidence was the fact that several studies were follow-ups of earlier randomised controlled trials and therefore subject to selection bias. Some studies did not clearly report whether the diagnostic assessment was performed without knowledge of, or blinded to the results of, the screening test results. A few studies did not report sufficient data therefore the 2x2 tables for diagnostic accuracy calculations could not be constructed. As the diagnostic accuracy estimates on some outcomes had wide confidence intervals, the evidence was downgraded for imprecision.

5.1.3.7.5 Other considerations

The Committee noted that possible early motor signs suggestive of cerebral palsy are addressed in recommendation 1.3.3 in the Cerebral Palsy in Children guideline (expected publication January 2017). This guideline refers the same recommendation, however, dystonia was removed because it was not considered relevant to children born preterm.

The Committee also noted that for the identification of autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD), NICE guidance on identification of ASD and ADHD should be used, specifically recommendation 1.3.3.

5.1.3.7.6 Key conclusions

5.1.3.8 Based on evidence, the Committee recommended the use of PARCA-R to identify if a child is at risk of global developmental delay, early intellectual disability or language problems at 2 years (corrected age) and SDQ at 4 years of age to check for social, attentional, emotional and behavioural problems. Recommendations

See Section 5.2.

5.1.3.9 Research recommendations

Research Question	1. **What is the accuracy of the parent-completed Parent Report of Children's Abilities-Revised (PARCA-R) questionnaire for predicting intellectual disability, language impairment and special educational needs at age 4 years for children born preterm?**
Population	Children born less than 37 weeks of pregnancy
Intervention	Parent Report of Children's Abilities-Revised (PARCA-R) questionnaire completed by parents when the child is 2 years (corrected age).
Comparator	Age appropriate gold standard tests of cognitive and language development at 4 years chronological age.
Outcome	The prognostic accuracy of the PARCA-R for predicting special educational needs in children born preterm at school age. Diagnostic accuracy: sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio
Study design	Prognostic and diagnostic study
Timeframe	2 to 3 years' follow-up
Why is this needed?	
Importance to 'patients' or the population	Parent-completed questionnaires such as the PARCA-R are used to identify children at risk of developmental problems and disorders. Although the PARCA-R has good diagnostic accuracy for identifying children at risk of concurrent developmental problems at age 2 years (corrected age), its accuracy at 4 years for predicting risk of intellectual disability, language impairment and learning difficulties that require special educational provision at school is not known. If the PARCA-R is able to accurately identify children at risk, then preventive intervention or enhanced surveillance may be offered during the preschool years and the results could be used to inform early years provision. Improved identification and provision of interventions is expected to lead to a reduced prevalence of intellectual disability at school age and improved developmental outcomes for children born preterm.

Relevance to NICE guidance	Developmental screening using the PARCA-R at age 2 (corrected) is recommended for children born preterm having enhanced developmental surveillance. Evidence on the prognostic accuracy of the PARCA-R at 4 years would further strengthen its use as a screening tool for identifying children at risk. The guidance may be updated if more optimal PARCA-R cut-off scores were identified.
Relevance to the NHS	Early identification of children at risk for later cognitive, language and learning disorders would enable the provision of intervention to reduce the risk of preterm children developing disorders, promote cognitive and language development over the early years and facilitate performance at school. Ultimately this may improve general health and well-being and reduce the prevalence of intellectual and learning disorders in this population, thereby reducing demands on the NHS for long term healthcare provision.
National priorities	National neonatal data collection and NHS commissioning arrangements have prioritised the need to collect and record developmental follow-up data for children born preterm who are at risk of developmental problems and disorders at age 2 (corrected).
Current evidence base	The PARCA-R was shown to have optimal diagnostic accuracy for identifying preterm children at risk of delayed cognitive development 2 years (corrected age) compared with other developmental screening tools. However, there was no evidence relating to the predictive validity of the PARCA-R in the preterm population at 4 years. Therefore the prognostic accuracy of the PARCA-R for later cognitive, language and learning disorders is not known.
Equality	The prognostic accuracy of the PARCA-R should be explored in children whose parents do not speak English.
Feasibility	This research should be feasible with adequate funding as the population of children born preterm is sufficiently large and it may be possible to conduct this research within the population of preterm children enrolled on the enhanced developmental surveillance pathway using routinely collected follow-up data to 4 years of age where the guideline is implemented. A difficulty may lie in recruiting a sample that is representative of the preterm population as a whole in terms of socio-economic and demographic characteristics.

Research question	**2. What is the accuracy of the parent-completed Ages and Stages Questionnaire, 3rd edition (ASQ-3)**

	for detecting concurrent intellectual disability and motor impairment between the ages of 2 years (corrected) and 4 years in children born preterm?
Population	Children born less than 37 weeks of pregnancy
Intervention	ASQ-3 completed by parents of children born preterm at the ages of 2 (corrected) and 4 years.
Comparator	Age-appropriate gold standard test for intellectual disability and motor impairment.
Outcome(s)	Diagnostic accuracy: • sensitivity • specificity • positive predictive value • negative predictive value • positive likelihood ratio • negative likelihood ratio
Study design	Diagnostic study
Timeframe	2 to 3 years' follow-up
Why is this needed	
Importance to 'patients' or the population	The ASQ is widely used to identify children at risk of developmental problems and disorders, and there are many versions of the questionnaire that span the preschool years. If the ASQ-3 was found to have sufficient predictive accuracy for detecting intellectual disability and motor impairment between the ages of 2 years (corrected age) and 4 years, this developmental check could be considered for use in enhanced developmental surveillance.
Relevance to NICE guidance	If the ASQ-3 was found to have acceptable diagnostic accuracy at age 4 years then it could be considered as an option for first-line developmental screening for preterm children and potentially reduce the number of children requiring a full standardized diagnostic assessment at 4 years of age.
Relevance to the NHS	Use of the ASQ-3 as a first-line developmental screening tool in the preterm population would reduce the resources and costs needed to provide enhanced developmental surveillance for preterm babies at risk of developmental problems and disorders compared with standard care.
National priorities	National neonatal data collection and NHS commissioning arrangements have prioritised the need to collect and record developmental follow-up data for higher risk preterm children at the age of 2 years (corrected age).
Current evidence base	The Committee considered evidence relating to the diagnostic accuracy of the ASQ-3 in order to determine whether this is an appropriate developmental screening tool in the preterm population. Low to moderate quality evidence from 4 studies provided mixed results regarding the diagnostic accuracy of the ASQ-3, with most

	reporting positive and negative likelihood ratios that were assessed as 'not useful' or, at best, 'moderately useful'. As such the evidence did not support a recommendation for use of the ASQ-3 to identify children at risk for intellectual disability and motor impairment at 2 years of age (corrected). In addition, only 1 study reported the diagnostic utility of the ASQ-3 at 5 years of age for which cut-offs produced positive and negative likelihood ratios that were, for the most part, not useful. There were no studies of the diagnostic accuracy of the ASQ-3 in identifying developmental problems in preterm children at 4 years of age, and no evidence of the prognostic accuracy of the ASQ-3 when completed by parents of preterm children at any age.
Equality	The diagnostic accuracy of the ASQ-3 should be explored in children whose parents are unable to speak English and where there is no validated appropriate translation available.
Feasibility	This research should be feasible with adequate funding as the population of children born preterm is sufficiently large. Care should be taken to ensure a representative sample is recruited, although this may be challenging in terms of socio-economic and demographic characteristics of the population.

1

Research question	**3. What is the accuracy of the parent-completed Strengths and Difficulties Questionnaire (SDQ) for predicting social, attentional, emotional and behavioural problems in children born before 28+0 weeks' gestation?**
Population	Children born less than 37 weeks of pregnancy
Intervention	Strengths and Difficulties Questionnaire (SDQ) completed by parents when the child is age 4 years
Comparator	Diagnostic reference standard
Outcome	•Prognostic accuracy up to 16 years: diagnosis of social, attentional, emotional and behavioural problems (including diagnoses of Autism Spectrum Disorder (ASD), Attention-Deficit/Hyperactivity Disorder (ADHD), Anxiety Disorders, Depressive Disorders, Obsessive-Compulsive and Related Disorders, Feeding and Eating Disorders and Disruptive, Impulse-Control and Conduct Disorders) by age 16. •Diagnostic accuracy as assessed at age 4: sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio
Study design	Prognostic and diagnostic accuracy study
Timeframe	Up to 12 years' follow-up
Why is this needed	

Importance to 'patients or the population'	Social, attentional, emotional and behavioural problems in children born preterm may go unnoticed, yet can have an adverse impact on a child's health and wellbeing, quality of life and school performance, as well as on their family. Identifying children at risk of these problems will enable appropriate intervention and family support to be provided in order to reduce their impact. In particular, identifying problems before school entry will support education planning and promote social and emotional development and attainment at school.
Relevance to NICE guidance	Screening for social, attentional, emotional and behavioural problems using the parent completed SDQ at age 4 years is recommended for children born before 28 weeks' gestation as part of the enhanced developmental surveillance. Information about the prognostic accuracy of the SDQ may inform revision of the guideline in terms of the choice of screening tool or the most appropriate cut-offs in this population.
Relevance to the NHS	The increased risk for mental health disorders in children born before 28 weeks gestational age places increased demands on paediatric and child and adolescent mental health services. Early identification of children at risk for disorders followed by intervention is expected to improve long term outcomes and reduce the prevalence of disorders in this population.
National priorities	1 in 10 children aged 5 to 16 years have a diagnosable mental health problem. The five year forward view for mental health report published by the independent Mental Health Taskforce to the NHS in England in 2016 highlighted the promotion of mental health and prevention of poor mental health as one of the key priority actions for the NHS by 2020/2021.
Current evidence base	The SDQ is widely used in UK clinical settings and by education professionals to screen for mental health disorders in children and young people, and its validity, reliability and diagnostic utility for use in the general population is well established. However there is a lack of evidence about the diagnostic or prognostic accuracy of the SDQ in identifying children born before 28 weeks' gestation who are at risk of mental health disorders. Two studies were included in the guideline which assessed the utility of the SDQ in this population at the age 11 and 14 years. There was a lack of evidence on the diagnostic and prognostic accuracy of the SDQ when used at 4 years of age in this population.
Equality	The prognostic accuracy of the SDQ should be explored in children whose parents do not speak English and where an appropriate validated translation is not available.
Feasibility	This research is feasible as the SDQ is widely available for use and will be used to screen

	extremely preterm children at age 4 years as part of enhanced developmental surveillance. Care should be taken to ensure a representative sample of preterm children born before 28 weeks' gestational age is recruited.
Other comments	The prognostic accuracy of other parent-completed behavioral screening tools may be also be considered to inform this guideline.

1

Research question	**4. What is the accuracy of the Preschool Language Scales 5th edition (PLS-5), completed by parents together with a speech and language therapist, for detecting speech and language problems at 2 years (corrected age) in children born preterm?**
Population	Children born less than 37 weeks of pregnancy
Intervention	Pre Language Scales 5th edition (PLS -5) completed by parents or carers in conjunction with a speech and language therapist when the child is aged 2 years (corrected).
Comparator	Children aged 2 years (corrected) who were not born preterm.
Outcome(s)	•Prognostic accuracy (speech, language and communication difficulties at later ages) •Diagnostic accuracy (sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio)
Study design	Diagnostic and prognostic study
Timeframe	3-5 years' follow-up
Why is this needed?	
Importance to 'patients or the population'	The PLS-5 may provide information about speech and language at 2 years (corrected age) which is not identified by the PARCA-R questionnaire. Identification of speech, language and communication problems at this age may allow early intervention that will help children when they move into early years education, as well as during their school years. It may also help to prevent other problems in the future, such as mental health problems and conduct disorders.
Relevance to NICE guidance	Screening for speech and language problems and disorders at age 2 (corrected) is currently done using the PARCA–R. The PLS–5 may provide information about receptive and expressive function which is not covered by the PARCA–R. Information about the prognostic accuracy of the PLS - 5 may also help inform future updates of the guideline.
Relevance to the NHS	Early identification of speech, language and communication problems may enable greater

	success when children transition into early years education services and later into school, and prevent other additional problems such as mental health or conduct disorders emerging. This would, in turn, reduce the financial burden on the NHS for services in these areas.
National priorities	Over 1 million children and young people in the UK have long term speech, language and communication difficulties. Of this group, between 50% and 90% are at risk of developing literacy problems, which in turn affects access to the curriculum. Over 65% of 7 to 14 year olds with conduct disorders have speech, language and communication difficulties. The Bercow Report (2008) highlighted how over 77% of parent respondents did not receive sufficient information about communication support or information. In addition, it was found that front-line staff may not understand speech, language and communication problems, and were therefore not able to advise families effectively. It is important to identify speech, language and communication problems early to promote effective parent to child interaction, identify appropriate sources of support such as speech and language therapy early on, and ensure strategies are in place to maximize access to the curriculum in school, reduce the development of conduct disorders and reduce the risk of exclusion.
Current evidence base	This guideline identified an inverse relationship between decreasing gestational age and the risk of speech, language and communication problems. In addition, there were increased risks for hyperactivity, impulsivity and particularly inattention, Autism Spectrum Disorder, hearing impairment and Intellectual Disabilities, all of which have additional speech, language and communication difficulties. The PLS is currently used in research as well as by health care practitioners such as speech and language therapists. It has been validated in English and Spanish populations. There is little evidence at present which describes the PLS – 5 as being useful in the identification of children born preterm, although some studies highlight the value of using the PLS with children who have developing features of autism.
Equality	To ensure equality of access, issues related to English as an additional language for families will need to be considered.
Feasibility	This research is feasible as the PLS 5 is currently used widely by clinicians working with pre-school populations.
Other comments	It may be useful to compare the identification of speech, language and communication difficulties with the Macarthur Bates questionnaire in the PARCA–R and the BSID–III.

1

Research Question	**5. What is the accuracy of a Wechsler Preschool and Primary Scale of Intelligence 4th Edition (WPPSI-IV) assessment at age 4 years for predicting later educational difficulties in children of primary school age who were born before 28^{+0} weeks' gestation?**
Population	Children born before 28+0 weeks' gestational age.
Intervention	Wechsler Preschool and Primary Scale of Intelligence 4th Edition (WPPSI-IV) administered at age 4 years.
Comparator	•Age appropriate gold standard tests of cognitive ability and academic achievement administered at ages 5 to 11 years •Identification of special educational needs at ages 5 to 11 years •Key Stage 2 national attainment tests at 11 years of age
Outcome	•Prognostic accuracy of diagnosis of cognitive impairment by 11 years •Diagnostic accuracy as assessed at age 11: sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio and negative likelihood ratio •Receiver operating characteristic curves as assessed at age 11 to determine the cut-off scores that have optimal accuracy
Why is this needed?	
Study design	Prognostic and diagnostic study.
Timeframe	Up to 7 years' follow-up
Why is this needed	Children born before 28+0 weeks' gestation are at increased risk of intellectual disability, which may have an adverse impact on their learning and achievement at school but may not be apparent at the 2-year developmental assesmnet. Determining the predictive accuracy of a WPPSI-IV assessment is key to providing parents or carers with accurate information about their child's likely development, so that educational support can be provided in order to reduce the risk of long-term intellectual disability.
Importance to 'patients' or the population	Children born before 28 weeks' gestation are at risk for cognitive deficits which may have an adverse impact on their learning and achievement at school. Learning difficulties may become apparent or exacerbated during early childhood as schooling places increasing cognitive demands on the child. Performing a cognitive assessment at 4 age years, prior to school entry can be used to inform parents of their child's risk for learning difficulties in order that support can be put in place from the outset of schooling. It is important to identify not only

	the children who have ongoing problems, but also those children who are likely to have problems later in school. Determining the prognostic accuracy of a WPPSI-IV assessment is key to providing parents with accurate information about their child's ongoing development and risk for later difficulties in order that appropriate support and educational provision can be put in place to reduce the risk of long term intellectual or learning disability.
Relevance to NICE guidance	A diagnostic assessment using the WPPSI-IV is recommended for children born less than 28+0 weeks' gestational at age 4 years as part of enhanced developmental surveillance. The 4 year assessment will enable early identification of intellectual disabilities in order to facilitate educational planning and special educational provision, if needed, from the outset of schooling. Understanding the prognostic accuracy of the WPPSI-IV assessment at 4 years of age could be used to determine whether this is the most appropriate measure to predict which preterm children are likely to have difficulties at school.
Relevance to the NHS	Improved educational outcomes may reduce the prevalence of learning disabilities and improve the general health and well-being of this vulnerable population of children.
National priorities	None identified.
Current evidence base	The WPPSI-IV is a standardized test of cognitive development for use in children of preschool and school age. It is considered the current gold standard in diagnostic cognitive assessment for children aged 2 to 7 years and is used clinically to assess children for cognitive delays and intellectual disabilities. The prognostic accuracy of the WPPSI-IV for identifying children born before 28 weeks' gestational age who are at risk for later learning disorders and special educational needs is not known.
Equality	The value of carrying out a WPPSI-IV assessment should be explored in children who do not speak English, in those with neurodevelopmental disorders, speech and language disorders or neuromotor impairments which are common among extremely preterm children, all of which may affect the validity of the test.
Feasibility	This research should be feasible as children born before 28 weeks' gestation will routinely be assessed using the WPPSI-IV as part of enhanced developmental surveillance and data relating to special educational needs and performance in Key Stage 2 attainment tests are routinely recorded by the Department for Education. Care should be taken to ensure that a representative sample of preterm children is recruited.

Other comments	The prognostic accuracy of other standardized preschool cognitive tests could be explored in to inform the best choice of preschool cognitive assessment for this guideline.

5.1.4 Delivering enhanced support and surveillance

Review question:

What is the most effective setting and staffing model for the follow-up for the identification of developmental problems and disorders and support of babies, children and young people born preterm?

5.1.4.1 Description of clinical evidence

No relevant clinical studies were identified for this review but a total of 10 expert commentary papers or reports, or developmental follow-up models from different experts or institutions were included (Adams 2014; BAPM 2008; Doyle 2014; Frisk 2011; Gong 2015; Hussey-Gardner 2002; Marshall and Zolotor 2003; Salt and Redshaw 2006; Toome et al 2008; Vollmer 2012). Three publications from the United Kingdom were included (BAPM 2008; Salt and Redshaw 2006; Vollmer 2012). The included publications and models were identified either through the literature search or through the assistance of the Guideline Committee members.

Further detail on the evidence can be found in section 1.6 and Appendix K:.

5.1.4.2 Summary of included studies

A summary of the publications that were included in this review are presented in Table 56.

Table 56: Summary of included studies

Publication	Country	Key content
Adams 2014	Switzerland	A publication describing the recommendations of the Swiss Society of Neonatology, the Swiss Society of Developmental Pediatrics and the Swiss Society of Neuropediatrics on follow-up assessment of high-risk newborns (including children born at <32 weeks of gestation) in Switzerland.
BAPM 2008	UK	A report presenting the work of the BAPM and Royal College of Paediatrics and Child Health working group on the classification of health status at 2 years as a perinatal outcome.
Doyle 2014	Australia/New Zealand	A publication summarising the discussions and recommendations made by an expert panel in Australia during a 2-day workshop on long-term developmental follow-up of high risk children.
Frisk 2011	Canada	A poster of a developmental follow-up program for children born preterm in parts of Ontario, Canada.
Gong 2015	USA	A report summarising practices in developmental follow-up of NICU survivors across seven centres in Texas, USA, and the conclusions made by experts during a one-day summit aiming to standardise follow-up care.
Hussey-Gardner 2002	USA	A publication describing Maryland's Premature Infant Developmental Enrichment (PRIDE) program which is a collaborative practice between service providers at NICU, NICU follow-up program and early intervention program.
Marshall and Zolotor 2003	USA	A commentary outlining the care needs of the NICU graduate during the first few years after discharge.
Salt and Redshaw 2006	UK	A commentary on the neurodevelopmental follow-up of children born preterm after 2 years of age. Discusses the areas that should be assessed, the assessment instruments, the assessment timings and the professional groups that should be involved.
Toome 2008	Estonia	A developmental follow-up program for children born preterm in Estonia.
Vollmer 2012	UK	A neurodevelopmental follow-up program for high risk infants (including children born preterm) in University Hospital Southampton NHS Foundation Trust.

5.1.4.3 Economic evidence

A systematic review of the economic literature was conducted but no relevant studies were identified that investigated the resource implications of enhanced surveillance methods for children born preterm.

Surveillance strategies are used to monitor children born preterm in order to identify any developmental problems and disorders that might arise. More intensive surveillance strategies enable problems and disorders to be identified earlier but increasing the frequency or number of surveillance strategies can have significant resource implications. Therefore, there is a need to balance the clear benefits of earlier detection against the costs of surveillance when deciding upon the optimal surveillance strategy.

The analysis aimed to estimate the resource impact associated with an enhanced surveillance strategy for children born preterm. For the full technical report (see Appendix I:).

5.1.4.3.1 Methods

A resource impact analysis was developed in Microsoft Excel®. The analysis was conducted from the perspective of the NHS and Personal Social Services (PSS) as outlined in the NICE Reference Case (The guidelines manual, NICE October 2014).

The analysis focuses on two assessment points, which were identified by the guideline Committee as time points where changes could be made to current practice to enhance the overall surveillance strategy:

1. Assessment at 2 years of age for children born before 30^{+0} weeks' gestation
2. Assessment at 4 years of age for children born before 28^{+0} weeks' gestation

Note that, for the assessment at 2-years of age, the analysis focused only on children born before 30^{+0} weeks gestation. Children born before 36^{+6} weeks of gestation with specific risk factors were also identified by the Committee as a group that would benefit from enhanced surveillance. However, it was not possible to include this group in the analysis as the size of the population could not be reliably estimated due to a lack of sufficient evidence on the proportion of children with risk factors.

At each time point, an 'enhanced surveillance' and 'current practice' strategy is compared. At the 2-year assessment, current practice was assumed to be a structured, face-to-face diagnostic assessment, which would involve a clinical psychologist, neonatologist or paediatrician with expertise in neonatology, an occupational therapist or physiotherapist and a nurse. In the enhanced surveillance strategy, a screening test was assumed to be used instead of the structured assessment meaning that a clinical psychologist would not be required. At the 4-year assessment, it was assumed that routine assessments were not undertaken in current practice. In the enhanced surveillance strategy, it was assumed that assessments would be undertaken involving a clinical psychologist and a paediatrician.

The number of children born preterm who would be assessed at 2 and 4 years of age was estimated using data on the number of live births by gestational age and the number of infant deaths in England and Wales from the Office for National Statistics (ONS). To estimate the number of children that would be alive at the assessment time points, mortality rates were applied to the live birth data. Infant mortality rates were estimated from the total number of infant deaths (occurring up to one year after birth) in England and Wales in 2013 from the ONS. Mortality from other causes in years two, three and four was estimated using ONS life tables 2013-15, which give an estimate of the annual probability of death given a person's age and gender.

The tables below show the estimated population that would be assessed at two and four years of age.

Table 57: Estimated number of preterm children born before 30^{+0} weeks' gestation assessed at 2-years of age

Gestational age (weeks)	Total live births	Infant mortality rate	Estimated deaths in year 1	Estimated deaths in year 2	Estimated population in year 2
≤ 22	462	88%	408	0	54
23	293	70%	205	0	88
24	465	41%	189	0	276
25	534	24%	127	0	407
26	560	17%	96	0	464
27	735	11%	80	0	655
28	959	9%	84	0	874
29	1,119	5%	51	0	1,068
Total	5,127	-	1,240	1	3,886

Table 58: Estimated number of preterm children born before 28^{+0} weeks' gestation assessed at 4-years of age

Gestational age (weeks)	Total live births	Infant mortality rate	Estimated deaths in year 1	Estimated deaths in year 2-4	Estimated population in year 4
≤ 22	462	88%	408	0	54
23	293	70%	205	0	88
24	465	41%	189	0	276
25	534	24%	127	0	407
26	560	17%	96	0	464
27	735	11%	80	0	655
Total	3,049	-	1,105	1	1,943

The costs associated with assessments were estimated using relevant staff costs from NHS Reference Costs 2014/15 and the Unit Costs of Health and Social Care 2015.

The cost of an assessment by a neonatologist or paediatrician was estimated to be £165.20 and £192.99, respectively based on outpatient costs associated with 'Neonatology' and 'Paediatrics' from NHS reference costs 2014/15. The costs of an assessment by a clinical psychologist was estimated to be £201.38 based on the outpatient cost associated with 'clinical psychology' from NHS reference costs 2014/15.

The costs of a nurse visit was estimated to be £94.91 based on the cost associated with 'Nursing services for children' from the community health services section of NHS reference costs 2014/15. The costs of an occupational therapist and physiotherapist visit was estimated to be £131.72 and £91.71, respectively based on the cost associated with 'Occupational therapist, child, one to one' and 'Physiotherapist, child, one to one' from the community health services section of NHS reference costs 2014/15.

The overall costs for the assessments at two and four years of age under current practice and enhanced surveillance scenarios are shown in the Table 59.

Table 59: Assessment costs of surveillance strategies for children born before 30^{+0} weeks' gestation at age 2

Surveillance strategy and assessments	Estimated costs	Source
Current practice		

Surveillance strategy and assessments	Estimated costs	Source
Paediatrician / neonatologist†	£179.09	NHS reference costs 2014/15 - outpatient costs for 'Neonatology' and 'Paediatrics'
Clinical psychologist	£201.38	NHS reference costs 2014/15 - outpatient costs for 'Clinical psychology'
Nurse	£94.91	NHS reference costs 2014/15 – 'Nursing services for children' (community health services)
Occupational therapist / physiotherapist‡	£111.71	NHS reference costs 2014/15 – 'Occupational therapist, child, one to one' and Physiotherapist, Child, One to One (community health services)
Total cost for assessment	**£587.10**	
Enhanced surveillance		
Paediatrician / neonatologist†	£179.09	NHS reference costs 2014/15 - outpatient costs for 'Neonatology' and 'Paediatrics'
Nurse	£94.91	NHS reference costs 2014/15 – 'Nursing services for children' (community health services)
Occupational therapist / physiotherapist‡	£111.71	NHS reference costs 2014/15 – 'Occupational therapist, child, one to one' and Physiotherapist, Child, One to One (community health services)
Total cost for assessment	**£385.72**	

†Average cost estimated assuming weighting of 50% for paediatricians (£192.99) and 50% for neonatologists (£165.20). Alternative scenarios are explored in sensitivity analysis.
‡Average cost estimated assuming weighting of 50% for occupational therapists (£131.72) and 50% for physiotherapists (£91.71). Alternative scenarios are explored in sensitivity analysis.

Table 60: Assessment costs of surveillance strategies at children born before 28^{+0} weeks' gestation at 4-years of age

Surveillance strategy and assessments	Estimated costs	Source
Current practice		
No assessment	£0.00	
Enhanced surveillance		
Paediatrician	£192.99	NHS reference costs 2014/15 - outpatient costs for 'Neonatology' and 'Paediatrics'
Clinical psychologist	£201.38	NHS reference costs 2014/15 - outpatient costs for 'Clinical psychology'
Total cost for assessment	**£394.36**	

5.1.4.3.2 Results

Base case results

The estimated population and resource impact of the surveillance programs at the assessments undertaken at 2 and 4 years of age are shown in the tables below.

At the 2-year assessment, it can be seen that total cost of the enhanced surveillance programme is estimated to be £1,498,962 whereas current practice is estimated to cost £2,281,534. Thus, the enhanced surveillance programme for children being assessed at two years of age was estimated to result in a cost saving of £782,572.

At the 4-year assessment, it can be seen that total cost of the enhanced surveillance programme is estimated to be £766,426 whereas there is no cost associated with current practice (since assessments at four years are not currently undertaken as part of routine practice). Therefore, the additional cost of the enhanced surveillance programme at the 4-year assessment timepoint is estimated to be £766,426.

Taking account of the costs at the 2- and 4-year assessment points, the enhanced surveillance programme was estimated to result in a cost saving of £16,146.

Table 61: Estimated costs of enhanced surveillance for children born before 30^{+0} weeks' gestation at 2 years of age

Gestational age (weeks)	Estimated population size at 2-year assessment	Estimated costs		
		Current practice	Enhanced surveillance	Difference
≤ 22	54	£31,711	£20,834	-£10,877
23	88	£51,414	£33,779	-£17,635
24	276	£161,981	£106,421	-£55,560
25	407	£238,888	£156,949	-£81,939
26	464	£272,488	£179,024	-£93,464
27	655	£384,744	£252,776	-£131,968
28	874	£513,406	£337,306	-£176,099
29	1,068	£626,901	£411,872	-£215,029
Total	3,886	£2,281,534	£1,498,962	-£782,572

Table 62: Estimated costs of enhanced surveillance for children born before $28+^{0}$ weeks' gestation at 4 years of age

Gestational age (weeks)	Estimated population size at 4-year assessment	Estimated costs		
		Current practice	Enhanced surveillance	Difference
≤ 22	54	£0	£21,296	£21,296
23	88	£0	£34,529	£34,529
24	276	£0	£108,784	£108,784
25	407	£0	£160,433	£160,433
26	464	£0	£182,998	£182,998
27	655	£0	£258,387	£258,387
Total	1,943	£0	£766,426	£766,426

Sensitivity analysis results

Various deterministic sensitivity analyses were conducted to assess the areas of uncertainty. The results of the sensitivity analysis for the assessment at 2 and 4 years of age are shown in the table below (Table 63: Sensitivity analysis results for assessment of children born before 30^{+0} weeks' gestation at two years of age. Particularly noteworthy were the alternative scenarios where changes were made to the surveillance scenario assumed to represent current practice. When assuming that a proportion of places were already following the enhanced surveillance programme, the cost saving at two years and the cost increase at four years was found to diminish.

Table 63: Sensitivity analysis results for assessment of children born before 30^{+0} weeks' gestation at two years of age

Modelled scenario	Total estimated cost difference	
	Assessment at 2 years of age (corrected)	Assessment at 4 years of age
Population increased by 25%	-£978,215	£958,032
Population decreased by 25%	-£586,929	£574,819
25% of places already following enhanced surveillance strategy	-£586,929	£574,819
50% of places already following enhanced surveillance strategy	-£391,286	£383,213
75% of places already following enhanced surveillance strategy	-£195,643	£191,606

5.1.4.3.3 *Conclusion*

The results of the analysis showed that the enhanced surveillance programme is likely to lead to cost savings for children born before 30^{+0} weeks gestation at 2 years of age and a cost increase for children born before 28^{+0} weeks gestation at 4 years of age. When considering the changes at 2 years of age and 4 years of age together, the enhanced surveillance program was found to result in a modest cost saving (£16,146).

This redistribution of resources from the assessment at 2 years to the assessment at 4 years should achieve improvements in the detection of developmental problems and disorders in a cost-effective manner.

5.1.4.4 Evidence statements

5.1.4.4.1 *UK publications*

A publication by Vollmer 2012 introduced the standardised neurodevelopmental follow-up program for high risk newborns (including children born at less than 31 weeks of gestation) at the University Hospital Southampton NHS Foundation Trust. The document included the enrolment criteria for the neonatal neurodevelopmental follow-up programme, the referral pathways to the follow-up programs, the organisation of the follow-up, and the timing and content of the follow-up program. Once a child has been enrolled in the program, she or he is followed up at the following intervals (with assessments done at these ages in parenthesis):

- 3 months corrected age (Hammersmith Infant Neurological Examination [HINE]; Abnormal Involuntary Movement Scale [AIMS])
- 12 months corrected age (HINE; AIMS; Gross Motor Function Measure [GMFM] for children with cerebral palsy; Ages and Stages Questionnaire [ASQ]; sleep questionnaire)
- 24 months corrected age (For infants born at less than 28+6 weeks of gestation, and infants with HIE or focal lesions: with Bayley III for cognitive, language and motor assessment; Health Status Classification System – Preschool [HSCS-PS]; Child

Behaviour Checklist [CBCL] 1.5-5; sleep questionnaire, Quantitative Checklist for Autism in Toddlers [Q-CHAT]. For infants born at more than 28+6 weeks of gestation: postal questionnaires ASQ; HSCS-PS; CBCL 1.5-5; sleep questionnaire)

- 4 years of chronological age (modified Touwen; M-ABC; Wechsler Preschool and Primary Scale of Intelligence [WPPSI]; visuo-motor test [VMI]; HSCS/Health Utilities [HUI]; CBCL 1.5-5; Behaviour Rating Inventory of Executive Function – Preschool [BRIEF-P]; sleep questionnaire. For infants born at more than 28+6 weeks of gestation: postal questionnaires ASQ; HSCS/HUI; CBCL 1.5-5; BRIEF-P; sleep questionnaire).

Medical history, growth parameters and nutrition matters are evaluated and measured and a general examination is done at each follow-up.

Additional assessments are done at 6 months corrected age for those babies with concerns at 3 months corrected age or for those with unilateral brain lesions. Additional assessments are done at 18 months corrected age for children with concerns at 12 months corrected age.

A report published by the British Association of Perinatal Medicine (BAPM, 2008) presented the work of the BAPM and Royal College of Paediatrics and Child Health working group on the classification of health status at 2 years as a perinatal outcome. The working group recommended that all children born at less than 31 weeks of gestation or with birth weight less than 1000 grams should receive a follow-up evaluation at 2 years of age. Furthermore, the working group recommended that service providers would consider including children born at less than 32 weeks of gestation or with a birth weight of less than 1500 grams into their 2-year follow-up services. The report defined the neurodevelopmental outcomes of interest at 2 years of age which included motor function, cognitive function, hearing, speech and language, and vision. The report also recommended specific definition criteria for moderate and severe neurodevelopmental disability and instruments for the assessment of some of these neurodevelopmental outcomes. These details are outlined in (Table 3 in Appendix L:)

Another publication from the UK (Salt & Redshaw, 2006) discussed the neurodevelopmental follow-up of children born preterm after 2 years of age. The authors listed the neurodevelopmental areas that should be assessed, the timings of follow-up, and the individuals that should be part of the follow-up (see Table 4 in Appendix L:), and the instruments that could be used to assess cognitive ability, speech and language, behavioural adjustment, and motor development (see Table 5 in Appendix L:). The authors concluded that formal follow-up of children born preterm after two years of age should be carried out. The follow-up should include assessments of cognitive ability, neuropsychological functioning (including executive functioning, non-verbal learning, visual-motor skills, speech and language and sensory impairment), academic achievement, behavioural adjustment, motor development, disability, quality of life and social skills and adjustment. The instruments used should be standardised and validated instruments whenever possible. A mixture of assessment with trained professionals as well as parent or child report is recommended. After two years of age, the authors recommended follow-up at 3-4 years of age, and at school age with choice of assessment at school entry (5-6 years), when established at school (7 years), early adolescence (12 years) and later (around 15 years). Children identified as having neurodevelopmental problems should be followed-up and assessed in more detail.

5.1.4.4.2 *Follow-up models from other countries*

Another publication from Switzerland (Adams 2014) described the recommendations of the Swiss Society of Neonatology, the Swiss Society of Developmental Pediatrics and the Swiss Society of Neuropediatrics on follow-up assessment of high-risk newborns (including children born at <32 weeks of gestation) in Switzerland. These children are followed up in developmental paediatric or neuropaediatric units (follow-up centres) which are specialised and experienced in developmental assessments and use validated and standardised instruments to assess the child. The follow-up model includes recommendations for assessments at 18 to 24 months' corrected age and at 5 to 6 years chronological age. At 18 to 24 months' corrected age, Bayley III, neurological, visual and hearing examinations should be performed. At 5 to 6 years of age K-ABC II, neurological, motor, visual, and hearing examinations and behaviour assessment should be performed (see Table 6 in Appendix L:). Assessments between 3 and 15 months' corrected age and at 3 to 4 years should be organised according to the individual centres' strategies and the needs of the child. The paper recommended the following steps to ensure highest possible follow-up rate:

- Families should be made aware of the importance of follow-up during the initial hospitalisation.
- Neonatologist should arrange the first follow-up examination directly or send a copy of the discharge report to the follow-up centre closest to the child's home.
- First contact between the follow-up centre and the family should be established via a secretary or a physician, accompanied by a written invitation.
- Twin/triplets should be invited to follow-up examinations simultaneously.
- If parents refuse follow-up or do not show up to a follow-up examination, the responsible paediatrician should be informed in order for the paediatrician to contact the parents directly.
- One publication (Doyle 2014) summarised the discussions and conclusions made by experts mainly from Australia and New Zealand in a two-day workshop in Australia and introduced a suggested scheme for follow-up of high risk children (including children born preterm) from the early neonatal period until adulthood. The discussions included the following areas: who should be followed-up, why should they be followed-up, what outcomes should be assessed during the follow-up, when should the children be followed-up, who should be involved in the follow-up, and what assessment tools should be used. The outcomes of interest in the follow-up of high risk children were grouped into four broad domains: physical health, learning and cognition, mental health, and quality of life. In addition, family outcomes such as parental mental health and carer-child interaction should be assessed. The model presented includes the following ages of assessment: 2-6 weeks; 3-4 months; 8 months; 12 months; 15-18 months; 24 months; 36 months; 4-5 years (preschool age); 6-8 years (1-2 years after starting school); 12-14 years; when transitioning to adulthood; and adulthood. Table 7 in Appendix L:provides a detailed summary of the recommendation for follow-up including the timing and the relative importance of each outcome of interest. Table 8 in Appendix L: summarises the tools for assessment recommended by Doyle 2014.
- Another document introduced the Infant and Child Development Services (ICDS) and the Preterm Pathways within the ICDS which is a model of developmental follow-up of children born preterm in Central West and Durham Regions in the Canadian state of Ontario (Frisk 2011). The model includes details about the referral criteria for the ICDS and Preterm Pathways, the levels of service for different preterm children, the care pathways for different preterm children depending on the underlying risk factors for developmental problems, the screening intervals, and the developmental areas that are being assessed, the instruments used, as well as the timing and place of these assessments.
- The referral criteria for the ICDS from birth until 18 months corrected age include the following:
- all children born preterm with very low birth weight (less than 1500 grams)
- children born preterm with low birth weight (1500 grams or more) with developmental delays, feeding or tone issues
- very low birth and low birth weight children who had one or more of the following risk factors:
    - abnormal cranial ultrasound scans

  - ventilation or oxygen treatment at 36 weeks postmenstrual age
  - neonatal seizures
  - being one of multiples
  - significant psychosocial issues
  - family history of learning problems, hearing impairment, developmental delay, language disorders, ADHD, ASD, fetal alcohol spectrum disorder, developmental disabilities.
- The referral criteria for the ICDS after 18 months corrected age include the following:
- very low birth weight and low birth weight children with developmental delays, feeding or tone issues
- very low birth weight and low birth weight children with typical development who had one or more of the following risk factors:
  - abnormal cranial ultrasound scans
  - bronchopulmonary dysplasia
  - ventilation or oxygen treatment at 36 weeks postmenstrual age
  - microcephaly
  - significant psychosocial issues
- very low birth weight or low birth weight children referred by neonatal follow-up staff because of other concerns.

After referral and initial consultation, there are three levels of service: monitoring (for all preterm children) which includes screening at regular intervals, education and counselling in relation to prematurity issues, referral facilitation, access to group programs and presentations, provision of contact number if problems arise; drop in or home consultation for preterm children with mild motor delay, mild feeding problems, or mild tone problems which includes the same services as the monitoring service and monthly 1 hour consultation including progress update and programming suggestions; and finally home visiting for preterm children with severe feeding problems, significant delays, deteriorating pattern of development, failure to make progress, or poor psychosocial situations including the same services as in the monitoring service and 1.5 hour consultation every 1-6 weeks including progress update and programming suggestions.

The children in the Preterm Pathway program are placed on one of nine pathways based on birth weight, presence and severity of medical complications and risk factors, family history of developmental problems and psychosocial issues. For children born preterm with minor medical conditions risk factors, birth weight of more than 1500 grams and no additional risk factors the preterm pathway continues until 36 months corrected age. For children born preterm with minor medical conditions or risk factors, birth weight of more than 1500 grams and family history of developmental problems the pathway continues until 36 months corrected age with optional preschool screening at 54 months chronological age. For all other children referred to the model, the preterm pathway continues until 54 months chronological age.

The screening is done at the following intervals (with the screening instruments in parenthesis):

- 4 months corrected age (Ages and Stages Questionnaire Third Edition [ASQ-3]; Alberta Infant Motor Scale [AIMS])
- 8 months corrected age (ASQ-3; AIMS)
- 12 month corrected age (ASQ-3; AIMS; Receptive-Expressive Emergent Language Third Edition [REEL-3]; Sensory Motor Screen Toddler [SMST])
- 18 months corrected age (ASQ-3; REEL-3; SMST; Modified Checklist for Autism in Toddlers [M-CHAT])

- 24 months corrected age (ASQ-3; REEL-3; SMST; M-CHAT)
- 30 months corrected age (ASQ-3; REEL-3; SMST)
- 36 months corrected age (ASQ-3; REEL-3)48 months chronological age (ASQ-3)
- 54 months chronological age (Early Screening Profiles [ESP] Cognitive, Language & Self-Help Social Profiles; Brigance Expressive Language composite; ASQ Fine Motor, Gross Motor, Problem-solving scales; Gradied Reading Assessment and Diagnostic Evaluation – Preschool [GRADE-P] Phonological composite; Bracken School Readiness Assessment Third Edition [BSRA-3]; Behaviour Rating Inventory of Executive Function – Preschool [BRIEF-P].

During some of the screenings parental well-being is also been screened using Edinburgh Postnatal Depression Scale or Parent Health Questionnaire 9.

Hearing and vision will be tested from all the children in the preterm pathway at regular intervals. For children who fail screenings or for whom there are concerns of developmental problems or delays further testing is done.

All data is collected into a database which can be used for example to refine the pathways in the model depending on the percentage of children with given developmental problems and the nature of those problems.

The Estonian guideline for developmental follow-up of very preterm infants provided a summary for follow-up assessment in the first and second year of life (Toome 2008). In the first year, assessments by a paediatrician and a physiotherapist in the follow-up clinic occur at 40 weeks postmenstrual age and at 2, 4, 6, 9 and 12 months corrected age. Assessment by a neurologist and a psychologist or a speech therapist occurs at 12 months corrected age, or earlier if required (decided by paediatrician). Hearing and vision are assessed at 40 weeks postmenstrual age. Retinopathy of prematurity (ROP) can be assessed further if required (decided by ophthalmologist or paediatrician). An orthopaedist assesses hips at 2 months corrected age by ultrasound (US) and at 4 months by X-ray and can be assessed further if required. A family practitioner assessment occurs at 2 months corrected age, and follow-up to 12 months corrected age. At 18 and 24 months corrected age, the infant is assessed further by a neurologist if abnormalities are present at 12 or 18 months corrected age, or if the infant is referred by a paediatrician. At both 18 and 24 months corrected age, follow-up assessment includes physiotherapy and hearing screening. At 24 months corrected age, the infant is also assessed by a clinical psychologist using BSID-III and a speech therapist using Reynell-III as well as assessed by a paediatrician. Table 9 and 10 in Appendix L: summarise the follow-up for very preterm babies for the first two years in Estonia (Toome 2008).

5.1.4.4.3 *Other relevant publications*

One publication (Gong 2015) reviewed practices in developmental follow-up of NICU survivors across seven centres in Texas, USA, with the aim to plan a standardised best practice programme that would facilitate and improve growth and feeding outcomes, developmental delay, and secondary social, emotional, or behavioural outcomes. The paper summarised the conclusions made by the involved experts during a one-day summit. The paper concluded that a quality comprehensive follow-up care for NICU survivors should include the following components:

- Personnel should include a multidisciplinary team including physicians, psychologists, nurses, social workers, physical, occupational, speech, and respiratory therapists, nutritionists, lactation consultants, case managers, and early intervention collaborators.
- NICU follow-up programme should provide support for case management and include home visits.
- There should be a standardised, uniform, evidence-based guidelines for developmental follow-up of NICU survivors.

- Processes need to be established to engage effectively with neonatologists, community paediatricians, and other primary care providers, including data sharing.
- Mechanisms for tracking during and after discharge from clinic should be established including follow-up at school age, adolescence and adulthood.
- A database for tracking and research should be established.
- Family support groups should be established.
- Educational programs as well as a website with resources should be provided to families, service providers and the community.
- There should be an appropriate space for the follow-up clinic.

One paper (Hussey-Gardner 2002) introduced the Maryland's Premature Infant Developmental Enrichment (PRIDE) program which is a collaborative practice between service providers from a neonatal intensive care unit (NICU), a NICU follow-up program and an early intervention program. The aim of the PRIDE program is to allow the families streamlined access to early intervention, eliminate duplication of evaluations, and facilitate timely acquisition of services. Infants at high risk of developmental problems are enrolled from NICU and NICU follow-up clinic to PRIDE. Eligibility is assessed by a multidisciplinary team consisting of neonatologist, developmental paediatricians, nurse, psychologist, occupational therapist, speech and language pathologist, physical therapist, and an onsite PRIDE service co-ordinator.

The PRIDE program includes an evaluation by a developmental specialist and potential referrals, service coordinator who acts as a liaison between services and the family. The PRIDE service co-ordinator works with the family from enrolment to the program at NICU or NICU follow-up clinic until the child turns three years old through frequent home visits, phone calls or follow-up clinic visits. The service co-ordinator facilitates creating and updating the individualised family service plan (IFSP) according to the needs of the child and the family. The service co-ordinator acts as a liaison and facilitates communication between the family, NICU and NICU follow-up staffs and early intervention service providers as well as other community resources. When the child turn two years, the service co-ordinator begins to facilitate the transition from NICU follow-up clinic and early intervention program to community services or school-based special education program.

The authors concluded that there are three key components for successfully replicating Maryland's PRIDE program: having a liaison between the hospital and the local early intervention program; having an onsite service co-ordinator to ensure communication between service providers as well as the family and to facilitate in creating an individualised service plan for the child and the family; and finally, ensuring that hospital staff is educated and advised on the importance and functions of an early intervention program.

One paper (Marshall and Zolotor, 2003) outlined the care needs of the NICU graduate during the first few years after discharge, which is influenced by the infant's medical history and risk factors for future sequelae. Recognition of growth failure, nutritional deficiencies, or neurosensory abnormalities identifies the infant with ongoing medical issues and requirements for further evaluation.

The authors suggested that evaluation should include periodical assessment of development of gross motor, fine motor, cognitive and communicative skills, behavioural or learning problems (which may not be detected until school age). Formal developmental screening by specialised, multidisciplinary clinics enables early detection of abnormalities and referral for interventional services. Neuromuscular assessment should be provided to identify abnormalities in tone, movement and posture to help diagnose cerebral palsy. Thorough assessment of communication skills include assessment of language comprehension and expression, interaction, attachment and use of gestures. Language deficits usually become apparent in preschool age. In addition, hearing and vision assessments are essential in the first years of life. Infants should receive on-going monitoring of hearing throughout the first 3 years of age, and those infants who have additional risk factors, for example, family history of permanent childhood hearing loss, suspicion of syndromes associated with hearing loss, congenital infections, or neonatal risk factors, should be monitored every 6 months until pre-school age. In addition, infants who have a history of retinopathy of prematurity require follow-up. Annual examinations to assess visual impairment should occur throughout early years as uncorrected poor vision may contribute to developmental delay. Overall, the authors concluded that careful management, subspecialist collaboration, community resources, and family support reduce morbidity and improve the overall outcome for the premature infant.

5.1.4.5 Economic evidence statement

A literature review of published cost-effectiveness analyses did not identify any relevant studies. The economic analysis undertaken for this question demonstrated that the enhanced surveillance programme is likely to lead to cost savings for children born before 30^{+0} weeks gestation at 2 years of age and a cost increase for children born before 28^{+0} weeks gestation at 4 years of age.

When considering the changes at 2 years of age and 4 years of age together, the enhanced surveillance program was found to result in a modest cost saving (£16,146). This redistribution of resources from the assessment at 2 years to the assessment at 4 years should achieve improvements in the detection of developmental problems and disorders in a cost-effective manner

5.1.4.6 Evidence to recommendations

5.1.4.6.1 *Relative value placed on the outcomes considered*

The aim of the review was to identify the most effective setting and staffing model for developmental follow-up for children born preterm. The Committee agreed that the most important outcomes to be considered were identification of developmental disorders and problems; early intervention; parental satisfaction and experience, parental support and audit information.

5.1.4.6.2 *Consideration of clinical benefits and harms*

The publications included in this review provided examples of development surveillance practice in different settings but did not provide evidence-based information to guide the development of the recommendations. Therefore, the recommendations were largely based on evidence about the risk of developmental disorders and problems at different gestational ages (see section 4), independent risk factors (see section 4), tools used for identification (see Section 5.1.2.9), and the Committee's clinical knowledge and expertise.

In formulating the recommendations on the developmental follow-up of children born preterm, the Committee considered the following:

- Which children born preterm should receive developmental follow-up?
- At what timepoints should the follow-up take place?
- What should be assessed during the follow-up visits?
- What screening and diagnostic tools should be used during assessments?
- Where should the assessments take place?
- Which professionals should be involved?

Currently, all children in the UK are eligible for enrollment in the national Healthy Child Programme. The Committee carefully considered which children born preterm should be eligible to receive developmental follow-up in addition to the Healthy Child Programme. The terms 'enhanced developmental support' and 'enhanced developmental surveillance' were used to describe these programmes of support and monitoring for developmental problems and disorders.

The Committee discussed how the risk of developing a problem or disorder is impacted by gestational age and other underlying antenatal, perinatal and neonatal factors. Bearing this in mind, they agreed that eligibility for both enhanced support and enhanced surveillance should take account of gestational age and other risk factors in order to target the children who are most likely to develop problems and disorders and benefit from early intervention. Targeting preterm children who are likely to be at higher risk of developing a problem or disorder will reduce the number of false positives which in turn reduces the burden on the child, parents and carers as well as the health system. The Canadian follow-up model was discussed as an example that utilised risk factors to determine which children received follow-up, but the Committee agreed that a simpler model would be better suited to the UK context.

The Committee discussed possible criteria for entering children into follow-up pathways. When considering gestational age at birth as a risk factor, they agreed that the evidence regarding the risk and prevalence of developmental disorders and problems did not often provide clear thresholds for degree of risk according to gestational age at birth; the risk and prevalence of various problems and disorders were approximately continuously distributed without clear evidence of a 'cliff' effect. However, the Committee noted that children born at less than 28 weeks' gestation were at an increased risk of not only cerebral palsy and moderate to severe intellectual disability but also presented with a range of special educational needs. For example, evidence from a large study from the UK showed that the prevalence of special educational needs increased with decreasing gestational age, with a clear increase in prevalence evident at 27 to 28 weeks' gestational age (MacKay 2010). Therefore, the Committee agreed that children born before 28 weeks' gestation should receive developmental support in the first two years of life and surveillance up to 4 years of age (uncorrected).

The Committee considered how children born between 28 and 30^{+0} weeks of gestation were likely to have been received specialist neonatal care and therefore have some risk of developmental problems and disorders. Based on their clinical experience and the evidence on risk of developmental problems and disorders, they agreed that all of these children should also be eligible enhanced support and enhanced surveillance through follow-up to 2 years of age (corrected).

Children born between 30 and 36^{+6} weeks of gestation who present with specific risk factors for developmental problems and disorders were also considered likely to benefit from enhanced support and surveillance programme up to 2 years (corrected age). There are substantially more children born between 30^{+0} and 36^{+6} gestational age compared to those born at lower gestatational ages. Because this is a large group and the children are considered, as a group, to be at lower overall risk of developmental problems and disorders because of their gestational age, any other factor or combination of factors that is considered sufficient to make them eligible for enhanced surveillance needed to be sufficiently robust. The specific risk factors were discussed at length and agreed by the Committee based on the evidence that these factors can independently have on developmental outcomes, together with their clinical knowledge and expertise. They include: grade 2 or 3 hypoxic ischaemic encephalopathy; a brain abnormality on neuroimaging, for example, grade III or IV intraventricular haemorrhage, cystic periventricular leukomalacia; neonatal bacterial or herpetic meningitis proven by culture. These factors were chosen because they were identified with developmental problems and disorders in the evidence base and would always require enhanced developemental support and surveillance. The Committee also discussed how there are a wide range of other risk factors that may increase the likelihood of developmental disorders and problems in children born between 30^{+0} to 36^{+6} weeks gestation, but that clinical judgement, taking into account the prevalence and severity of these risk factors, should be used. The Committee also discussed how it was important that the eligibility criteria allowed flexibility to enrol a child in the enhaced support and surveillance programme if they presented with considerable and obvious risk factors other than those listed, for example, a genetic abnormality which may be associated with learning difficulties.

The Committee discussed how all children born preterm, not just those children born between 30 and 36^{+6} weeks' gestation without risk factors, would be followed-up through the usual established national Healthy Child Programme and that this should provide an additional route for identifying developmental disorders and problems. The Committee highlighted that there needed to be a level of flexibility between the follow-up and care pathways. Should a problem arise at any point, the child should be transferred to an investigative pathway accordingly.

When considering enhanced support, the Committee decided that a single point of contact from whom to seek advice should be available for all parents and carers of eligible preterm children. This contact could be by telephone, e-mail, or other messaging service, or face-to - face (including home visits), depending on individual need. This service should be organised by the neonatal network of care as an outreach post-discharge service and have professionals who are experts in preterm development. The Committee agreed that more frequent contact with neonatal services could be helpful in the immediate period after discharge from the hospital by aiming to reduce anxiety surrounding the care of the child and support identification and management of early developmental problems.

The Committee agreed that children receiving enhanced developmental surveillance should be reviewed regularly during the first 2 years of life (that is, up to the age of 2 [corrected age]) and that this should include at least 3 visits for developmental review, one of which would be at 2 years of age (corrected age). They debated specifying ages for all of the visits but decided to leave this to local providers, recognising that many children born preterm would also be receiving follow-up for medical reasons. Committee members suggested, for example, a follow-up regime might include a visit at 3 to 4 months of age (corrected age), 9 to 12 months of age (corrected age) and 2 years (corrected age).

The Committee considered what assessments to include in the 2-year visit and in particular which specific tools should be used to check for developmental disorders and problems. The following disorders were considered important outcomes for the first 2 years: cerebral palsy, global developmental delay (intellectual disability), autism spectrum disorder, persistent feeding problems and communication and language delays. In general, the Committee agreed that the main objective of the assessment at 2 years of age should be to identify severe developmental impairment and cerebral palsy, if present. Identifying cerebral palsy at the earliest opportunity should lead to improved outcomes for the child and their families.

The Committee considered the evidence showing that PARCA-R (see section 5.1.3.5.1), when used at age 22 to 26 months, was found to be a reliable screening check for global developmental delay when compared with the BSID. PARCA-R is inexpensive and easy to administer because it is a parent-filled questionnaire. The Committee recognised the possibility of poor return rate of the PARCA-R questionnaire, however, they noted that the families were expected to attend a clinic appointment and, although not ideal, the PARCA-R could potentially be filled in during the visit and the scoring of it can be done relatively quickly. The Committee considered this made it a good tool to identify concern in relation to global developmental delay and, therefore, recommended the use of PARCA-R, at a minimum, for all children having enhanced surveillance. The Committee agreed than when PARCA-R is was suitable, for example due to a language barrier or the assessment was done outside the validated time frame of PARCA (which is 22 to 26 months), a recognised alternative should be used. The recommendations were considered to be the minimum set of checks that should be offered.

The Committee noted the importance of referring to the NICE guideline on autism spectrum disorder in under 19s: recognition, referral and diagnosis. They noted that there were no specific screening checks recommended for ASD but refer the reader to recommendation 41 on recognising signs and symptoms of possible autism spectrum disorder.

The Committee emphasised that at each contact, parents and carers should be actively queried about any concerns they might have about the development of their child. These developmental concerns should be taken seriously and assessed. The results of this assessment and any formal screening checks, together with any other aspects of development should be discussed with the parents and carers. If significant concerns about the child are identified then they should be referred into local pathways for diagnosis and intervention (including Early Years education).

When considering the developmental assessment for children born less than 28 weeks' gestational age at the age 4, the Committee anticipated that although most severe developmental disorders and cerebral palsy would be identified by 2 years of age, significant problems were often missed or could not be reliably assessed at that age (for example, behavioural problems) that could have a negative impact on the child, particularly school-based learning. In addition, problems and disorders of a lesser severity (including milder forms of cerebral palsy and neurodevelopmental disorders) may only become evident at the later age. While it was recognised that assessments at 4 years were taking place potentially very close to the time of school entry, when decisions about school admissions may have already been made, the assessment at this age did have advantages in terms of better understanding the child's overall development. The timing of the assessment would act as a 'safety net' and an entry point for neurodevelopmental pathway for those not identified earlier. The Committee also discussed how the assessment at age 4 years was unlikely to inform decision about choice of school but rather inform the educational staff of potential special needs.

The Committee considered what assessments to include in the 4-year visit and in particular which specific tools should be used to check for developmental disorders and problems in children who did not raise concern at 2 years of age. They prioritised the following outcomes, all of which were considered to have considerable impact on school readiness: intellectual development; emotional, attention and social behaviour; fine and gross motor development; speech, language and communication; hearing; and vision. Ideally, the assessment at 4 years would be completed in collaboration with the educational services but at a minimum should provide educational services with a developmental report of the child that can be used to inform educational plans for the child (see also section 5.1.5.6).

The Committee considered whether to recommend a screening tool or a standardised test for the assessment of intellectual ability at 4 years of age (uncorrected). Because of the evidence on significantly increased risk and prevalence of intellectual disability in this group of children born extremely preterm, the Committee decided to recommend the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) because it was a recognised and commonly used test in the NHS for this age group and because it was thought that naming a specific test would enable uniformity of data to inform national audit. The Committee agreed that in case WPPSI is not possible to administer for example, due to the child's motor or sensory impairment, a suitable alternative should be used.

The Committee discussed how behavioural problems were challenging to assess but the assessmentis essential for families because behavioural problems and disorders could have considerable implications on daily life, social life and educational attainment. The Committee agreed that there was evidence to support the use of the Strengths and Difficulties Questionnaire (SDQ) (see also section 5.1.3.5.1).

In general, the Committee agreed that, like at the 2-year assessment, good clinical practice was to use the results of the different tools in conjunction with parental concern and all available other sources of information including previous assessment results, when making decisions about the significance of the results.

Following the 4 year assessment, the results of the assessment should be collated with any other other concerns or observations into a comprehensive report of the child's strengths and difficulties in a format that could be used by parents and carers, as well as professionals (such as educational staff to inform the support an individual child is likely to need when starting school). The clinician involved in this appointment should have the appropriate skills in order to integrate the information, communicate with parents and reach decisions about further referral.

The Committee agreed that the enhanced support and surveillance program should be provided as an integral part of a neonatal service working together with local health services as appropriate. Although the Committee found no evidence on this point, they agreed that those engaged in enhanced support and surveillance should comprise a multidisciplinary team of professionals with expertise in the following areas: neonatal care; child development in children born preterm; support of parents (including feeding support); administration and interpretation of screening tools and standardised tests; integration of information from a wide range of sources to formulate a comprehensive report on the child's strengths and difficulties; and local care pathways and early years education. In order for the enhanced support and surveillance programme to work practically, the Committee agreed that the following professionals were essential to the core multidisciplinary team: neonatologist or paediatrician with expertise in neonatal care; occupational therapist or physiotherapist (during the first two years, for example, to assess movement); nurse with experience in neonatal care and neonatal post discharge care (for example, to provide enhanced support); clinical or educational psychologist (at 4 year assessment, for example, to administer the WPPSI); paediatrician with expertise in neurodevelopment (at 4 year assessment). The Committee agreed that there were a range of other professionals who are also key to the assessment of developmental problems and disorders at various ages and that easy access to these people when needed should be made available: occupational therapist, physiotherapist, speech and language therapist, community nurse, paediatric neurologist, and dietitian.

The Committee discussed how the early birth of a child increased the child's overall risk of future developmental problems and disorders and the importance of sharing this information with healthcare professionals and education staff. Thus, their preterm birth can be considered together with any concerns that might arise during developmental surveillance within the Healthy Child Programme. However, this information should only be disclosed with parental permission.

In addition to the follow-up pathways, the Committee discussed the importance of collecting national neonatal audit data for benchmarking purposes. The collection of neonatal audit data can be used to inform neonatal services and support changes to practice. The Committee discussed that neonatal audit data should be collected from a tightly defined population and the Committee agreed that audit data should be collected from children born before 28 weeks of gestation. Including children born before 28 weeks of gestation allows the monitoring of children cared for in the local neonatal units as well as those transferred into neonatal intensive care units. The Committee agreed that the neonatal audit should include data on the following conditions at 2 years of age: diagnosis of cerebral palsy and the Gross Motor Function Classification System (GMFCS) if cerebral palsy is present; PARCA-R score; impairments in hearing; vision; speech or language; or motor skills. At 4 years of age the following information should be recorded for audit purposes: diagnosis of cerebral palsy including the GMFCS; WPPSI fullsacle IQ score and subscale scores for verbal comprehension index, visual spatial, fluid reasoning, working memory and processing speed; SDQ total difficulty scores, subscale scores and impact scale; impairment in hearing; results of the national prthoptic screening test; any formal diagnosis of a developmenytal disorder (for example ASD); and epilepsy currently receiving treatment.

5.1.4.6.3 Consideration of economic benefits and harms

A systematic review of the economic literature was conducted but no relevant studies were identified that were applicable to this review question.

The Committee recognised that there are large economic implications if enhanced developmental follow-up is offered to all children born preterm. However, as the risk and prevalence of developmental disorders and problems is lower among the children born moderate to late preterm compared with extremely preterm and very preterm births, it was not considered appropriate to automatically place all children born preterm on an enhanced support and surveillance pathway. The Committee decided that children born before 30 weeks should receive enhanced support and surveillance up to 2 years (corrected age) and children born before 28 weeks up should receive surveillance up to 4 years of age. However, all children born before 36+6 weeks of gestation with certain risk factors (see above) for developmental disorders or problems would be included in the enhanced surveillance follow-up pathway. It was noted that the vast majority of preterm births occur between 32+0 and 36+6 weeks of gestation.

In the Committee's estimation, offering the PARCA-R screening test rather than a structured face-to-face diagnostic assessment (such as the BSID) at the 2 year (corrected) assessment was a significant deviation from current practice. They discussed the economic implications of using PARCA-R over BSID; PARCA-R was considered to be much cheaper than BSID as it was administered by parents and can be completed via postal questionnaire, online, or on a tablet during clinic visit whereas the BSID was a face-to-face assessment conducted by a healthcare professional, most likely a clinical psychologist. Based on NHS Reference Costs 2014/15, an assessment by a clinical psychologist was estimated to cost £201.38. The recommendation was estimated to affect 3,597 births per year (based on data from children born at less than 30 weeks gestation in 2014 minus estimated mortality by age 2 from the Office for National Statistics). Thus, the recommendation was estimated to result in a cost saving of £724,286 per year.

The Committee thought that outcomes from the assessment would not be impaired by use of the PARCA-R instead of the BSID as evidence showed the PARCA-R to be a reliable screener for global developmental delay/intellectual disability that correlates well with BSID. Furthermore, the Committee agreed that the clinical decision-making following an abnormal score on either the PARCA-R or BSID would be similar. However, the downside of PARCA-R was that it was not age-standardised and could only be used in a narrow age spectrum of 22 to 26 months of (corrected) age. The Committee also discussed how PARCA-R may not identify mild to moderate intellectual disability so well. Also, unlike BSID, PARCA-R does not identify specific areas with problems but rather flags that there are general problems or concerns which can then be explored in a more detailed assessment. Despite these limitations, the Committee concluded that PARCA-R was the most cost-effective tool to be used to assess developmental delay.

The recommendation to offer developmental assessment at the age of 4 years to children born at less than 28 weeks was also likely to be a significant deviation from current practice. In comparison to current practice, this recommendation required two additional consultations at the 4-year assessment: an assessment by a psychologist and a consultation with a paediatrician. Based on NHS Reference Costs 2014/15, an assessment by a psychologist was estimated to cost £201.38 while a consultation with a paediatrician was estimated to cost £192.18. The recommendation was estimated to affect 2,438 children per year (based on data from children born at less than 28 weeks gestation in 2014 minus estimated mortality by age four from the Office for National Statistics) and thus estimated to cost an additional £696,260 per year.

While the 2-year assessment recommendation was associated with a substantial resource impact, the Committee believed that it was likely to be a cost-effective use of resources. The enhanced surveillance assessment should lead to the early identification of children with problems or disorders as a result of targeted referral for diagnosis and appropriate management. Thus, effectiveness gains and potentially downstream cost savings would make it likely that the recommendation was cost-effective in cost per QALY terms.

Overall, it was estimated that the increased cost of the assessment at 4 years would be more than offset by the cost savings at 2 years of age. When considering the changes in assessment at 2 and 4 years together, the differences from current practice were estimated to result in a modest cost-saving of £28,026. Thus, the changes effectively represent a redistribution of resources from the assessment at 2 years to the assessment at 4 years. By using more resources on a higher risk group at a timepoint where disorders and problems may be more evident, it was hoped that the changes will achieve improvements in the detection of developemental problems and disorders in a cost-effective manner.

5.1.4.6.4 *Quality of evidence*

No clinical evidence was identified for this review. The grey literature provided examples or suggestions of models for developmental follow-up of children born preterm and commentaries in relation to the topic. As these publications were not research evidence, even though some of them might be evidence-based approaches, they could not be evaluated formally. Due to the absence of clinical evidence and the obvious limitations of the 'evidence' included in this review, the Committee relied largely on their clinical knowledge and expertise when forming the recommendations.

5.1.4.6.5 *Other considerations*

Assessments and follow-up of cerebral palsy should be conducted in line with the NICE guidance on cerebral palsy in children (expected publication January 2017). The Committee also noted that for the identification of autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD), NICE guidance on identification of ASD and ADHD should be used, specifically recommendation 12.

The Committee discussed the possibility of conducting the 2 and 4 year assessments by telephone interview for children who are geographically remote. It is possible to administer the PARCA-R over the phone in conjunction with general developmental enquiry. Parents and carers could be asked if they have concern about hearing and motor development and their feedback can be used to determine whether the child needs to be seen in person. However, the Committee considered that this is not ideal. Because of the nature of the 4 year assessment, they agreed that this should be conducted in person.

The Committee considered the needs of children, parents and carers who are travellers and/or live in temporary accommodation. As such, the recommendations require that enhanced developmental support be tailored to take account of individual preferences and needs.

5.1.4.6.6 *Key conclusions*

The Committee concluded that enhanced support and surveillance up to 2 years of age (corrected) should be available to children who are born before 30 weeks' gestation, and those born between 30 and 36 weeks' gestation who have specific risk factors for developmental problems and disorders. Children born before 28 weeks' gestation should also receive surveillance at 4 years. Taken together, these changes to current practice are expected to be a cost-effective use of resources. All children born preterm should receive developmental surveillance as part of the Healthy Child programme in conjunction with any enhanced support and surveillance they may be receiving.

5.1.4.7 Recommendations

See Section 5.2.

5.1.4.8 Research recommendations

	6. **Does enhanced developmental support and surveillance improve outcomes for the parents and carers of children born preterm?**
Population	Parents or carers of children born less than 37 weeks' gestation
Intervention	Enhanced developmental support and surveillance
Comparator	Current practice
Outcome	Parent reported outcome measures (PROM), for example, psychological well-being of parents or carers (depression and anxiety) at key time-points during enhanced support and surveillance, quality of life Parent reported experience measures (PREM), for example, experience of services and satisfaction Adherence to enhanced surveillance
Study design	Prospective qualitative study
Timeframe	2 year follow-up
Why this is needed	
Importance to 'patients or the population'	Ehanced developmental support and surveillance up to age 4 years for children born preterm who fulfil the necessary criteria is expected to increase the detection of developmental problems and disorders and improve outcomes for these children However the acceptability of this approach to parents, carers and families also needs to be taken into consideration. A study that looks at the impact of wnhanced developmental support and surveillance on parents and carers (for outcomes such as experience of services, satisfaction and anxiety) may help to identify where improvements can be made to future support and surveillance.
Relevance to NICE guidance	This study will provide valuable insights on the practical and qualitative aspects of enhanced support and surveillance and further guide updates. It will also in part audit the utility, success or failure of this Guideline, and in so doing, strengthen the concept of follow-up in this high-risk group.
Relevance to the NHS	A positive impact in terms of parent satisfaction and engagement will promote more seamless public-NHS partnerships in health care. It will seek views from parents or carers (who are key stakeholders) and thus inform evaluation and improvement of care.
National priorities	Preterm births are one of the top 10 priorities identified nationally by the James Lind Alliance, specifically providing information of packages of care at or after discharge

	http://www.jla.nihr.ac.uk/priority-setting-partnerships/preterm-birth/top-10-priorities/ Developing an understanding of parental needs in delivering a developmental support and surveillance for children born preterm is an important component. The 2010 inquiry into the quality of general practice in England by the King's Fund highlighted the need for patient engagement (in this case, parents, carers and families of the child born and preterm) https://www.kingsfund.org.uk/projects/gp-inquiry/patient-engagement-involvement
Current evidence base	There are no data about the impact of a developmental surveillance programme in the UK. There is currently a lack of 'end-user' contribution (parental, carer or family voice) in the evaluation of such programmes.
Equality	No specific equality issues were identified other than those relating to language and communication. Appropriate support, tools and techniques (for example, interpreters and translation of questionnaires) that enable communication should be employed.
Feasibility	No barriers to feasibility were identified.
Other comments	No other comments.

5.1.5 Sharing information

Review question:

What information should be shared between those delivering NHS commissioned care and also between the NHS and the educational sector on the developmental follow-up of babies, children and young people born preterm?

5.1.5.1 Description of clinical evidence

One study (Johnson 2015) was included in this review. This survey study from the UK assessed the knowledge and information needs of teaching staff and educational psychologists on prematurity in order to determine how prepared they feel to support children born preterm in schools.

5.1.5.2 Summary of included studies

Table 64: Summary of included studies for sharing information

Study	Aim of the study	Study type	Population	Comments
Johnson (2015) UK	To assess the knowledge and information needs of education professionals to determine how prepared they feel to support the growing number of preterm children entering schools today.	A national survey, Preterm Birth-Knowledge Scale (PB-KS)	N=585 teachers N=212 educational psychologists	validated scale low response rate respondents not representative of the target population

5.1.5.3 Economic evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

5.1.5.4 Evidence statements

Low quality evidence from 1 study of teaching staff (n=585) and educational psychologists (n=212) in the UK on their knowledge in relation to the developmental and educational consequences of preterm birth showed that teaching staff had a mean accuracy of 45% (SD 17%) in the Preterm Birth-Knowledge Scale (PB-KS) while educational psychologists had a mean accuracy of 52% (SD 15%) on the same scale. Twelve percent of the teaching staff has an accuracy of less than 25% and 5.2% of the educational psychologists had an accuracy of less than 25%. The teaching staff had significantly lower scores than the educational psychologists. For both groups, the greatest accuracy of responses to the PB-KS were on items about neurosensory outcomes, such as cerebral palsy, and the need for assistance with daily activities. Only 11 to 18% of all the respondents knew that children born very preterm might be at a higher risk of being inattentive and have poorer peer relationship skills than their peers born at term. Only 8% of the teaching staff knew that difficulties in mathematics is a particular deficit that children born preterm might have.

The same study also looked at the information needs of teaching staff and educational psychologists in relation to preterm birth. The evidence from the study showed that 38% of the teaching staff felt that they were adequately equipped to support a child born preterm while 14% of the teaching staff felt that had received sufficient training on issues relating to prematurity. Over 80% of all respondents requested more information about preterm birth. Around 75% of the respondents felt that disclosure of preterm birth status would be beneficial for the child and would not lead to negative labelling of the child.

5.1.5.5 Economic evidence statement

A literature review of published cost-effectiveness analyses did not identify any relevant studies and no economic modelling was undertaken for this question.

5.1.5.6 Evidence to recommendations

5.1.5.6.1 Relative value placed on the outcomes considered

The aim of the review was to identify what information should be shared and how it should be shared between those delivering NHS commissioned care, and between NHS and the educational services on the developmental follow-up of babies, children and young people born preterm. The Committee agreed that the most important outcomes to consider were communication between NHS organisations and between NHS and educational organisations, parent and carer satisfaction, and benchmarking data.

5.1.5.6.2 Consideration of clinical benefits and harms

Limited evidence was identified for this review. The only study provided evidence on the knowledge level and information needs of the teaching staff and educational psychologists in relation to prematurity. As no other evidence was found the recommendations made were largely based on the knowledge and expertise of the Committee.

The Committee discussed how effective information sharing between service providers was required to streamline service, avoid duplication of work and prevent confusion among the families of the child born preterm.

- The Committee agreed that the most important time points for information sharing were the transition phases such as discharge from the hospital, transition from the neonatal outreach services to the community services, and transition to early year's education or school. These were considered the most important time's points for information sharing because at these transitions different individual or groups of professionals become involved in the care and follow-up of the child and there could be a risk of missing valuable information or duplication of services. The members of the Committee who were parents or grandparents of children born preterm also raised how the limitations regarding information sharing between service providers was most apparent during the transition phases.
- The Committee agreed that the neonatal discharge summary should always include information about antenatal, obstetric and neonatal risk factors for developmental problems and disorders. The discharge summary should be shared with parents as well as the primary and secondary healthcare teams.
- The evidence in this review found that three quarters of educational professionals considered that knowing about a child's premature status would be beneficial and would not lead to negative labelling. However, the Committee discussed how it was not necessary to disclose this information if no problems or disorders had been identified. On the other hand, sometimes problems unrelated to the child's premature birth could arise later, for example when a child starts school. A large UK study showed social factors may account for more than the premature status when considering the outcome of educational attainment among children born late preterm (Quigley 2012). However, if developmental problems or disorders were identified at any time point, the Committee agreed that this information should be shared with other service providers, including primary and secondary healthcare teams, and if needed, with education services and social services. For education services and social services, the consent of the parents or carers was always required.
- The Committee recommended a comprehensive developmental assessment at 4 years of age for children born before 28+0 weeks' gestation (see section 5.1.4.6). They discussed how ideally, the educational services should be involved because one aim of the assessment is to establish the educational needs that the child may have when entering school. However, including the education services in the assessment was considered oftenunrealistic, therefore, a comprehensive summary of the strengths and difficulties of the child according to the assessment at 4 years should be developed using a range of information from different sources, including parents and carers, health care services and early years education services, where appropriate. This comprehensive summary should be shared (with consent from the parent or carer) with the education services in order for the educational services to develop a comprehensive educational plan for the child according to the child's individual needs. The Committee agreed that the parents or carers should be given a copy of the information that had been shared if they so request.
- The Committee recognised the importance of collecting information for the neonatal audit for two reasons: 1) to have national data on developmental outcomes for children born preterm in the enhanced surveillance program and 2) in order for individual neonatal units to benchmark against other units. If a unit is found to be an outlier the reasons can be explored and practices can be altered in order to improve outcomes for the children. The outcomes included in the neonatal audit collection were chosen based on their clear diagnostic criteria or measurement based on a score that is objective and allows comparison.

5.1.5.6.3 Consideration of economic benefits and harms

A systematic review of the economic literature was conducted but no relevant studies were identified which were applicable to this review question.

The economic implications of this topic were considered but not thought to be substantial. The sharing of information does have resource implications as it requires time to be spent by the individuals sharing it (in health care and educational services). However, any increase in time is not expected to be significant especially since the majority of the recommendations reflect current best practice. Therefore the recommendations are not anticipated to require a substantial increase in resources. There is the potential for inconsistency in practice though and it is therefore possible that there could be increased costs for service providers that are not currently sharing the information outlined in the recommendations, such as the comprehensive summary of the strengths and difficulties of the child according to the developmental assessment at 4 years.

Any increase in costs as a result of an increase in time spent sharing information was thought likely to be cost-effective as effective information sharing between service providers could streamline the service, avoid duplication of work and prevent confusion among the families of the children born preterm.

5.1.5.6.4 Quality of evidence

The one study included in this review was considered to be of low quality. The sample was not representative of the overall population and the response rate was low. However, the questionnaire used was validated for use in the population.

5.1.5.6.5 Other considerations

The evidence suggests that some educational staff may have limited understanding of the effect that preterm birth may have on the child's development and school attainment. Therefore, the Committee discussed how it was important to ensure educational staff were aware of the risks that prematurity could have on the child's educational attainment, motor and behavioural features.

5.1.5.6.6 Key conclusions

The guideline developers concluded that:

- Sharing information between service providers was essential in order to effectively plan the care for the child and avoid duplication of services and confusion among parents and carers.
- The neonatal discharge summary, including information about antenatal, perinatal and neonatal risk factors for developmental problems and disorders, should be shared with the appropriate primary and secondary health care providers.
- At any point, if there were concerns or if developmental disorders or problems were identified, the information should be shared between the tertiary, secondary, and primary healthcare teams involved in the care of the child.
- When appropriate, information about the child's developmental concerns or problems should be shared with the education services, including early year's education services, with consent from the parent or carer.
- After the 4-year assessment of children born before 28+0 weeks' gestation, a comprehensive summary of the child's strengths and difficulties should be written bringing together information from a range of sources. If a problem or a disorder is identified, this summary should be shared, with consent from the parent or carer, with the education services in order to facilitate the development of a comprehensive plan for education considering the needs of the child.

- It was important to collect neonatal audit data on clearly defined outcomes in order to benchmark and improve the practice of neonatal units.

5.2 Recommendations

5.2.1 Information and support for parents and carers of all preterm babies

Providing information and support

21. Provide information about the risk and prevalence of developmental problems and disorders to parents or carers of preterm babies, and discuss this with them.

22. Provide information to parents or carers of preterm babies that is tailored to their individual circumstances, taking into account:

- their child's potential developmental needs
- their level of education
- any social care needs they have
- any cultural, spiritual or religious beliefs.
- the need for consistency in information sharing among healthcare professionals

23. Follow the principles in the NICE guideline on patient experience in NHS services in relation to communication (including different formats and languages), information and shared decision-making.

24. Provide emotional and psychological support as needed to parents or carers of preterm babies.

25. Provide information to parents or carers of preterm babies about opportunities for peer support.

Information and support leading up to and on discharge

26. Before discharging a preterm baby:

- agree a discharge plan with the parents or carers
- ensure that the discharge plan includes clear information about any antenatal and perinatal risk factors for developmental problems and disorders (see section 4.3.1)
- share the discharge plan with parents or carers and with primary and secondary healthcare teams.

27. Help parents or carers to gain the knowledge, skills and confidence they need to look after their baby at home and support the baby's developmental needs, taking into account that they are likely to be anxious about managing their baby's care after discharge. This may relate to:

- interaction with the baby
- managing feeding

- patterns of sleeping
- impact on day-to-day living, such as social isolation because of fear of infection.

28. Involve the social support networks (which may include partners, grandparents or other family members) of parents and carers of a baby born preterm when planning discharge and during follow-up.

29. Explain to parents and carers at the time of discharge that their child's developmental (corrected) age, which is calculated from their original due date (and not the date they were born), will be used for the first 2 years when assessing their functional and developmental skills (such as walking and talking).

30. Inform parents or carers of all preterm babies about the Healthy Child Programme, which includes national recommendations for all children about screening (for example, newborn hearing screening) and surveillance (including social, emotional, behavioural and language development).

Care, support and follow-up after discharge

31. Inform parents or carers about the routine postnatal care and support available as described in the NICE guideline on postnatal care up to 8 weeks after birth.

32. Healthcare professionals providing postnatal care and support in the community for babies born preterm should have the skills and knowledge to recognise and manage problems in these babies, including:

- providing feeding support
- addressing concerns about sleeping
- facilitating interaction between the parents or carers and the baby.

5.2.2 Enhanced developmental support and surveillance

Criteria for enhanced developmental support and surveillance up to 2 years (corrected age)

33. Provide enhanced developmental support and surveillance by a multidisciplinary team (see section 5.2.3) up to 2 years (corrected age) for children born preterm who have a developmental problem or disorder, or are at increased risk of developmental problems or disorders based on the following criteria:

- born before 30+0 weeks' gestation **or**
- born between 30+0 and 36+6 weeks' gestation and has or had 1 or more of the following risk factors:
  - a brain lesion on neuroimaging likely to be associated with developmental problems or disorders (for example, grade 3 or 4 intraventricular haemorrhage or cystic periventricular leukomalacia)
  - grade 2 or 3 hypoxic ischaemic encephalopathy in the neonatal period
  - neonatal bacterial meningitis
  - herpes simplex encephalitis in the neonatal period

34. Consider providing enhanced developmental support and surveillance by a multidisciplinary team (see section 5.2.3) up to 2 years (corrected age) for children born between 30^{+0} and 36^{+6} weeks' gestation who do not have any of the risk factors listed in recommendation 33 but are thought, using clinical judgement, to be at increased risk of developmental problems or disorders in the first 2 years of life and taking into account the presence and severity of risk factors (see recommendations 3-20)

35. Inform parents or carers of preterm babies who meet the defined criteria about the arrangements for enhanced developmental support and surveillance for their child.

Enhanced developmental support

36. Provide parents or carers of a preterm baby having enhanced developmental support with a single point of contact within the neonatal service for outreach care after discharge.

37. Use a range of approaches when providing enhanced developmental support and tailor the support to take account of individual preferences and needs. Approaches may include:

- face-to-face meetings, in clinics or in the home
- a telephone helpline
- electronic communication, for example by text message or email.

Enhanced developmental surveillance

38. For all children born preterm who are having enhanced developmental surveillance, provide:

- a minimum of 2 face-to-face follow-up developmental visits in the first 2 years of life **and**
- a developmental assessment at 2 years (corrected age) (see recommendation 43).

39. At each visit for a child born preterm who is having enhanced developmental surveillance:

- ensure that this is conducted by professionals with appropriate skills (see recommendation and 49 and 50)
- ask parents or carers whether they have any concerns about their child's development
- include checks for developmental problems and disorders (See recommendation 40)
- carefully assess and review any developmental concerns arising either from parent or carer report or at the visit itself
- correct for gestational age up to 2 years (corrected) when assessing development
- discuss any concerns with parents or carers
- consider further investigation or referral if a developmental problem or disorder is suspected or present
- refer the child to the appropriate local pathway if needed.

Checking for developmental problems and disorders

40. At each visit for a child born preterm who is having enhanced developmental surveillance up to 2 years (corrected age), and at the 4-year assessment (for children born before 28+0 weeks; see recommendation 46), check for signs and symptoms of problems and disorders as appropriate, such as:

- cerebral palsy (see recommendation 41)
- global developmental delay
- autism spectrum disorder (See recommendation 42)
- visual impairment
- hearing impairment
- feeding problems
- sleep problems
- speech, language and communication problems
- motor problems
- attention, impulsivity and hyperactivity
- emotional and behavioural problems
- executive functional problems
- special educational needs

41. Recognise the following as possible early motor signs of cerebral palsy:

- delayed motor milestones, such as late sitting, crawling or walking (correcting for gestational age)
- unusual fidgety movements or other abnormalities of movement, including asymmetry or paucity of movement
- abnormalities of tone, including hypotonia (floppiness) or spasticity (stiffness)
- persisting feeding difficulties.

See also the NICE guideline on cerebral palsy in children and young people under 25.

42. For guidance on recognising signs and symptoms of possible autism spectrum disorder, see the NICE guideline on autism spectrum disorder in under 19s: recognition, referral and diagnosis.

Developmental assessment at 2 years (corrected age)

43. Provide a developmental assessment at 2-years (corrected age) for children born preterm who are having enhanced developmental surveillance. This assessment should include:

- all aspects listed in recommendation 39
- at a minimum, use the Parent Report of Children's Abilities - Revised (PARCA-R) to identify if the child is at risk of global developmental delay, early intellectual disability or language problems:
  - if the PARCA-R is not suitable (for example, because of poor English language comprehension or the child being outside the validated age range of 22 to 26 months), use a suitable alternative.
- ensuring that checks of vision and hearing have been carried out in line with national recommendations.

44. If findings from the developmental assessment at 2 years (corrected age) or 4 years (see recommendation 46) suggest any developmental problems or disorders:

- refer the child to an appropriate local pathway, which may involve child health and education services
- share information with:
  - parents or carers
  - primary and secondary healthcare teams
- ask parents or carers for permission to share information with:
  - education services
  - social care services as approptiate.

Discharge from enhanced surveillance at 2 years

45. After the developmental assessment at 2 years (corrected age):

- advise parents or carers of all children that their child should continue to be followed-up in the healthy child programme **and**
- advise parents or carers of children born before 28+0 weeks' gestation that the child will also be offered a further developmental assessment at 4 years .

Developmental assessment at 4 years for children born before 28+0 weeks' gestation

46. Provide a developmental assessment at 4 years for all children born before 28+0 weeks' gestation. This assessment should:

- be conducted by professionals with appropriate skills (see recommendations 49 and 50)
- take into account information provided by parent or carers (see recommendation 39)
- include a review of previous assessments and information from all other relevant sources
- include checks for developmental problems and disorders (see recommendation 40)
- use:
  - the Strengths and Difficulties Questionaire (SDQ) to check for social, attentional, emotional and behavioural problems
  - as a minimum, the Wechsler Preschool and Primary Scales of Intelligence 4th Edition (WPPSI) test, including subscales for verbal comprehension, visual spatial skills, fluid reasoning, working memory and processing speed:
    - ◊ if the WPPSI is not suitable (for example, because of sensory or motor impairment), use a suitable alternative
- include ensuring that the child has been offered orthoptic vision screening as recommended by the National Screening Committee.

47. Provide a comprehensive summary of the child's strengths and difficulties, including any developmental problems and disorders, after the 4-year assessment that:

- is in a format that is accessible to parents and carers

- if needed, informs the development of a plan for intervention and support, including educational support.

See also recommendation 44 about referral and information sharing

5.2.3 Delivering enhanced developmental support and surveillance

48. Enhanced developmental support and surveillance for children born preterm who meet the defined criteria (see recommendations 33, 34 and 46) should:

- be provided as an integral part of a neonatal service working together with local health services
- empower parents and carers to be involved in decisions about their child's care
- be delivered by a multidisciplinary team with the necessary skills (see recommendation 49)
- record outcomes at specified time points for national audit (see section 5.2.4)
- be monitored by checking adherence to the recommendations in this guideline, including follow-up rates and outcomes, as part of the routine provision of neonatal care by neonatal operational delivery networks and commissioners

49. Multidisciplinary teams delivering enhanced developmental support and surveillance for children born preterm should include professionals with knowledge and expertise in the following areas:

- neonatal care
- development of children born preterm, including developmental problems and disorders (see recommendation 40)
- providing support in the community, for example for feeding problems
- administering and interpreting results from questionnaires and standardised tests (such as the PARCA-R, SDQ and WPPSI)
- collating information from a range of sources to facilitate decision making and writing reports
- local care pathways, including Early Years education.

50. Multidisciplinary teams delivering enhanced developmental support and surveillance for children born preterm should include the following professionals:

- for follow-up to 2 years (corrected age):
    - neonatologist or paediatrician with expertise in neonatal care
    - occupational therapist or physiotherapist
    - outreach nurse or nurse with expertise in neonatal care
- for the assessment at 4 years (see recommendation 46):
    - clinical or educational psychologist
    - paediatrician with expertise in neurodevelopment.

51. Multidisciplinary teams delivering enhanced developmental support and surveillance for children born preterm should have access to the following professionals:

- community nurse

- occupational therapist
- physiotherapist
- paediatric neurologist
- paediatrician with expertise in neurodevelopment
- dietitian
- speech and language therapist.

5.2.4 Neonatal audit

52. Record the following information, as applicable, in the National Neonatal Research Database for every child born preterm who has enhanced developmental surveillance:

- whether the child had specialist neonatal care and details of discharge
- at the assessment at 2 years (corrected age) (see recommendation 43)
  - diagnosis of cerebral palsy
  - Gross Motor Function Classification System (GMFCS) score if cerebral palsy is present
  - PARCA-R score
  - epilepsy that is currently being treated
  - impairments of hearing, vision, speech and language, and motor skills[1]
- at the assessment at 4 years (see recommendation 46)
  - diagnosis of cerebral palsy
  - GMFCS score if cerebral palsy is present
  - WPPSI full scale IQ score, and subscale scores for verbal comprehension, visual spatial skills, fluid reasoning, working memory and processing speed
  - SDQ total difficulty score, subscale scores and impact score
  - any formal clinical diagnoses of a developmental disorder (for example, autism spectrum disorder)
  - epilepsy that is currently being treated
  - the presence of a hearing impairment, defined as profound deafness or impairment severe enough to need hearing aids or cochlear implant
  - results of national orthoptic vision screening (see recommendation 46).

53. Record routine educational measures at key stage 2 (including special educational needs and disability [SEND]) on an operational delivery network-wide basis, to allow educational outcomes at 11 years to be linked to neonatal information.

[1] As defined in Figure 3 in Classification of health status at 2 years as a perinatal outcome, report of a BAPM/RCPCH working group, version 1.0, 8 January 2008.

References

Adams 2014

Adams, M., Borradori-Tolsa, C., Bicle-Graz, M., Grunt, S., Weber, P., Capone Mori, A., Bauder, F., Hagmann, C., Natalucci, G., Pfister, R., Latal, B., Follow-up assessment of high-risk newborns in Switzerland, Paediatrica, 25, 6-10, 2014

Adams-Chapman 2008

Adams-Chapman, I., Hansen, N. I., Stoll, B. J., Higgins, R., Neurodevelopmental outcome of extremely low birth weight infants with posthemorrhagic hydrocephalus requiring shunt insertion, Pediatrics, 121, e1167-e1177, 2008

Agerholm 2011

Agerholm, H., Rosthoj, S., Ebbesen, F., Developmental problems in very prematurely born children, Danish Medical Bulletin, 58, A4283, 2011

Allred 2014

Allred, E. N., Capone Jr, A., Fraioli, A., Dammann, O., Droste, P., Duker, J., Gise, R., Kuban, K., Leviton, A., O'Shea, T. M., Paneth, N., Petersen, R., Trese, M., Stoessel, K., Vanderveen, D., Wallace, D. K., Weaver, G., Retinopathy of prematurity and brain damage in the very preterm newborn, Journal of AAPOS, 18, 241-247, 2014

Ambalavanan 2012

Ambalavanan, N., Carlo, W. A., Tyson, J. E., Langer, J. C., Walsh, M. C., Parikh, N. A., Das, A., Van Meurs, K. P., Shankaran, S., Stoll, B. J., Higgins, R. D., Generic, Database, Subcommittees of the Eunice Kennedy Shriver National Institute of Child, Health, Human Development Neonatal Research, Network, Outcome trajectories in extremely preterm infants, Pediatrics, 130, e115-25, 2012

Ancel 2006

Ancel, P. Y., Livinec, F., Larroque, B., Marret, S., Arnaud, C., Pierrat, V., Dehan, M., N'Guyen, S., Escande, B., Burguet, A., Thiriez, G., Picaud, J. C., Andre, M., Breart, G., Kaminski, M., Cerebral palsy among very preterm children in relation to gestational age and neonatal ultrasound abnormalities: The EPIPAGE cohort study, Pediatrics, 117, 828-835, 2006

Andersen 2011

Andersen, G. L., Romundstad, P., Cruz, J. D. L., Himmelmann, K., Sellier, E., Cans, C., Kurinczuk, J. J., Vik, T., Cerebral palsy among children born moderately preterm or at moderately low birthweight between 1980 and 1998: A European register-based study, Developmental Medicine and Child Neurology, 53, 913-919, 2011

Anderson & Doyle 2003

Anderson, P., Doyle, L. W., Neurobehavioral Outcomes of School-age Children Born Extremely Low Birth Weight or Very Preterm in the 1990s, Journal of the American Medical Association, 289, 3264-3272, 2003

Anderson & Doyle 2004

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

Anderson 2003

Anderson, P., Doyle, L. W., Victorian Infant Collaborative Study, Group, Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s, JAMA, 289, 3264-72, 2003

Anderson 2011

Anderson, P. J., De Luca, C. R., Hutchinson, E., Spencer-Smith, M. M., Roberts, G., Doyle, L. W., Attention problems in a representative sample of extremely preterm/extremely low birth weight children, Developmental Neuropsychology, 36, 57-73, 2011

Andrews 2008

Andrews, W. W., Cliver, S. P., Biasini, F., Peralta-Carcelen, A. M., Rector, R., Alriksson-Schmidt, A. I., Faye-Petersen, O., Carlo, W., Goldenberg, R., Hauth, J. C., Early preterm birth: association between in utero exposure to acute inflammation and severe neurodevelopmental disability at 6 years of age, American Journal of Obstetrics & Gynecology, 198, 466.e1-466.e11, 2008

Ardal 2011

Ardal,F., Sulman,J., Fuller-Thomson,E., Support like a walking stick: parent-buddy matching for language and culture in the NICU, Neonatal network : NN, 30, 89-98, 2011

Arnaud 2007

Arnaud, C., Daubisse-Marliac, L., White-Koning, M., Pierrat, V., Larroque, B., Grandjean, H., Alberge, C., Marret, S., Burguet, A., Ancel, P. Y., Supernant, K., Kaminski, M., Prevalence and associated factors of minor neuromotor dysfunctions at age 5 years in prematurely born children: The EPIPAGE study, Archives of Pediatrics and Adolescent Medicine, 161, 1053-1061, 2007

Arockiasamy 2008

Arockiasamy,V., Holsti,L., Albersheim,S., Fathers' experiences in the neonatal intensive care unit: a search for control, Pediatrics, 121, e215-e222, 2008

BAPM 2008

BAPM, RCPCH, Report of a BAPM/RCPCH Working Group: Classification of health status at 2 years as a perinatal outcome, 2008

Beaino 2010

Beaino, G., Khoshnood, B., Kaminski, M., Pierrat, V., Marret, S., Matis, J., Led, ESert B., Thiriez, G., Fresson, J., Roz, E. J. C., Zupan-Simunek, V., Arnaud, C., Burguet, A., Larroque, B., Br, EArt G., Ancel, P. Y., Predictors of cerebral palsy in very preterm infants: The EPIPAGE prospective population-based cohort study, Developmental Medicine and Child Neurology, 52, e119-e125, 2010

Beaino 2011

Beaino, G., Khoshnood, B., Kaminski, M., Marret, S., Pierrat, V., Vieux, R., Thiriez, G., Matis, J., Picaud, J. C., Roze, J. C., Alberge, C., Larroque, B., Breart, G., Ancel, P. Y., Epipage Study Group, Predictors of the risk of cognitive deficiency in very preterm infants: the EPIPAGE prospective cohort, Acta Paediatrica, 100, 370-8, 2011

Benzies & Magill-Evans 2015

Benzies, K. M., Magill-Evans, J., Through the eyes of a new dad: experiences of first-time fathers of late-preterm infants, Infant Mental Health Journal, 36, 78-87, 2015

Blaggan 2014

Blaggan, S., Guy, A., Boyle, E. M., Spata, E., Manktelow, B. N., Wolke, D., Johnson, S., A parent questionnaire for developmental screening in infants born late and moderately preterm, Pediatrics, 134, e55-62, 2014

Bodeau-Livinec 2007

Bodeau-Livinec, F., Surman, G., Kaminski, M., Wilkinson, A. R., Ancel, P. Y., Kurinczuk, J. J., Recent trends in visual impairment and blindness in the UK, Archives of Disease in Childhood, 92, 1099-1104, 2007

Bolisetty 2014

Bolisetty, S., Dhawan, A., Abdel-Latif, M., Bajuk, B., Stack, J., Lui, K., Intraventricular hemorrhage and neurodevelopmental outcomes in extreme preterm infants, Pediatrics, 133, 55-62, 2014

Brazy 2001

Brazy,J.E., Anderson,B.M., Becker,P.T., Becker,M., How parents of premature infants gather information and obtain support, Neonatal Network - Journal of Neonatal Nursing, 20, 41-48, 2001

Brinchmann 2002

Brinchmann, B. S., Forde, R., Nortvedt, P., What matters to the parents? A qualitative study of parents' experiences with life-and-death decisions concerning their premature infants, Nursing Ethics, 9, 388-404, 2002

Brown 2014

Brown, H. K., Speechley, K. N., Macnab, J., Natale, R., Campbell, M. K., Mild prematurity, proximal social processes, and development, Pediatrics, 134, e814-24, 2014

Burguet 1999

Burguet,A., Monnet,E., Pauchard,J.Y., Roth,P., Fromentin,C., Dalphin,M.L., Allemand,H., Maillet,R., Menget,A., Some risk factors for cerebral palsy in very premature infants: importance of premature rupture of membranes and monochorionic twin placentation, Biology of the Neonate, 75, 177-186, 1999

Burnett 2014

Burnett, A., Davey, C. G., Wood, S. J., Wilson-Ching, M., Molloy, C., Cheong, J. L. Y., Doyle, L. W., Anderson, P. J., Extremely preterm birth and adolescent mental health in a geographical cohort born in the 1990s, Psychological medicine, 44, 1533-44, 2014

Burnett 2014a

Burnett, A., Davey, C. G., Wood, S. J., Wilson-Ching, M., Molloy, C., Cheong, J. L., Doyle, L. W., Anderson, P. J., Extremely preterm birth and adolescent mental health in a geographical cohort born in the 1990s, Psychological Medicine, 44, 1533-44, 2014

Carlo 2011

Carlo, W. A., McDonald, S. A., Fanaroff, A. A., Vohr, B. R., Stoll, B. J., Ehrenkranz, R. A., Andrews, W. W., Wallace, D., Das, A., Bell, E. F., Walsh, M. C., Laptook, A. R., Shankaran, S., Poindexter, B. B., Hale, E. C., Newman, N. S., Davis, A. S., Schibler, K., Kennedy, K. A., Sanchez, P. J., Van Meurs, K. P., Goldberg, R. N., Watterberg, K. L., Faix, R. G., Frantz, I. D., 3rd, Higgins, R. D., Eunice Kennedy Shriver National Institute of Child, Health, Human Development Neonatal Research, Network, Association of antenatal corticosteroids with mortality and neurodevelopmental outcomes among infants born at 22 to 25 weeks' gestation, JAMA, 306, 2348-58, 2011

CASP 2013

Critical Appraisal Skills Programme (CASP). Qualitative checklist. 10 questions to help you make sense of qualitative research. CASP UK. Oxford, 2013.

Chan & Quigley 2014

Chan, E., Quigley, M. A., School performance at age 7 years in late preterm and early term birth: a cohort study, Archives of Disease in Childhood Fetal & Neonatal Edition, 99, F451-7, 2014

Charkaluk 2010

Charkaluk, M. L., Truffert, P., Fily, A., Ancel, P. Y., Pierrat, V., Neurodevelopment of children born very preterm and free of severe disabilities: The Nord-Pas de Calais Epipage cohort study, Acta Paediatrica, International Journal of Paediatrics, 99, 684-689, 2010

Chawla 2016

Chawla, S., Natarajan, G., Shankaran, S., Pappas, A., Stoll, B.J., Carlo, W.A., Saha, S., Das, A., Laptook, A.R., Higgins, R.D., Association of neurodevelopmental outcomes and neonatal morbidities of extremely premature infants with differential exposure to antenatal steroids, JAMA Pediatr, 170(12), 1164-1172, 2016.

Chiu 2012

Chiu, T. M. L., Wehrmann, S., Reid, D., Sinclair, G., Transforming mother-infant interaction within cultural and caregiving contexts: Home-based occupational therapy for preterm infants, Hong Kong Journal of Occupational Therapy, 22, 17-24, 2012

Chyi 2008

Chyi, L. J., Lee, H. C., Hintz, S. R., Gould, J. B., Sutcliffe, T. L., School Outcomes of Late Preterm Infants: Special Needs and Challenges for Infants Born at 32 to 36 Weeks Gestation, Journal of Pediatrics, 153, 25-31, 2008

Curtis & Burns 2015

Curtis, L., Burns, A., Unit Costs of Health and Social Care 2015. Personal Social Services Research Unit, University of Kent. UK, 2015.

Cuttini 2012

Cuttini, M., Ferrante, P., Mirante, N., Chiandotto, V., Fertz, M., Dall'Oglio, A. M., Coletti, M. F., Johnson, S., Cognitive assessment of very preterm infants at 2-year corrected age: performance of the Italian version of the PARCA-R parent questionnaire, Early Human Development, 88, 159-63, 2012

Davis 2007

Davis, N. M., Ford, G. W., Anderson, P. J., Doyle, L. W., Developmental coordination disorder at 8 years of age in a regional cohort of extremely-low-birthweight or very preterm infants, Developmental Medicine and Child Neurology, 49, 325-330, 2007

De Groote 2007

De Groote, I., Vanhaesebrouck, P., Bruneel, E., Dom, L., Durein, I., Hasaerts, D., Laroche, S., Oostra, A., Ortibus, E., Roeyers, H., Van Mol, C., Outcome at 3 years of age in a population-based cohort of extremely preterm infants, Obstetrics and Gynecology, 110, 855-864, 2007

De Jesus 2013

De Jesus, L. C., Pappas, A., Shankaran, S., Li, L., Das, A., Bell, E. F., Stoll, B. J., Laptook, A. R., Walsh, M. C., Hale, E. C., Newman, N. S., Bara, R., Higgins, R. D., Outcomes of small for gestational age infants born at <27 weeks' gestation, Journal of Pediatrics, 163, 55-60.e3, 2013

De Jong 2015

De Jong, M., Verhoeven, M., Lasham, C. A., Meijssen, C. B., van Baar, A. L., Behaviour and development in 24-month-old moderately preterm toddlers, Archives of Disease in Childhood, 100, 548-53, 2015

De Kleine 2003

De Kleine, M. J., den Ouden, A. L., Kollee, L. A., Nijhuis-van der Sanden, M. W., Sondaar, M., van Kessel-Feddema, B. J., Knuijt, S., van Baar, A. L., Ilsen, A., Breur-Pieterse, R., Briet, J. M., Brand, R., Verloove-Vanhorick, S. P., Development and evaluation of a follow-up assessment of preterm infants at 5 years of age, Archives of Disease in Childhood, 88, 870-5, 2003

Delobel-Ayoub 2006

Delobel-Ayoub, M., Kaminski, M., Marret, S., Burguet, A., Marchand, L., N'Guyen, S., Matis, J., Thiriez, G., Fresson, J., Arnaud, C., Poher, M., Larroque, B., Behavioral outcome at 3 years of age in very preterm infants: The EPIPAGE study, Pediatrics, 117, 1996-2005, 2006

Delobel-Ayoub 2009

Delobel-Ayoub, M., Arnaud, C., White-Koning, M., Casper, C., Pierrat, V., Garel, M., Burguet, A., Roze, J. C., Matis, J., Picaud, J. C., Kaminski, M., Larroque, B., Behavioral problems and cognitive performance at 5 years of age after very preterm birth: The EPIPAGE study, Pediatrics, 123, 1485-1492, 2009

DoH, 2009

Department of Health. Healthy Child Programme: Pregnancy and the first five years of life. Department of Health. UK, 2009.

DoH 2015

Department of Health. National Reference Costs 2014-15. National schedule of reference costs - NHS trusts and NHS foundation trusts. UK, 2015.

Dewey 2011

Dewey, D., Creighton, D. E., Heath, J. A., Wilson, B. N., Anseeuw-Deeks, D., Crawford, S. G., Sauve, R., Assessment of developmental coordination disorder in children born with extremely low birth weights, Developmental Neuropsychology, 36, 42-56, 2011

Dixon-Woods 2005

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., Sutton, A., Synthesising qualitative and quantitative evidence: a review of possible methods, Journal of Health Services & Research Policy, 10, 45-53, 2005

Downey 2015

Downey,L.C., O'Shea,T.M., Allred,E.N., Kuban,K., McElrath,T.F., Warner,D.D., Ware,J., Hecht,J.L., Onderdonk,A., Leviton,A., Antenatal and early postnatal antecedents of parent-reported attention problems at 2 years of age, Journal of Pediatrics, 166, 20-25, 2015

Doyle 2011

Doyle, L. W., Anderson, P. J., Callanan, C., Carse, E., Charlton, M. P., Davey, M. A., Davis, N., Duff, J., Hunt, R., De Luca, C., Hayes, M., Hutchinson, E., Kelly, E., McDonald, M., Opie, G., Roberts, G., Stewart, M., Ung, L., Watkins, A., Williamson, A., Woods, H., Changing long-term outcomes for infants 500-999 g birth weight in Victoria, 1979-2005, Archives of Disease in Childhood: Fetal and Neonatal Edition, 96, F443-F447, 2011

Doyle 2014

Doyle, L. W., Anderson, P. J., Battin, M., Bowen, J. R., Brown, N., Callanan, C., Campbell, C., Chandler, S., Cheong, J., Darlow, B., Davis, P. G., DePaoli, T., French, N., McPhee, A., Morris, S., O'Callaghan, M., Rieger, I., Roberts, G., Spittle, A. J., Wolke, D., Woodward, L. J., Long term follow-up of high risk children: who, why and how?, BMC Pediatrics, 14, 279, 2014

Drummond & Colver 2002

Drummond,P.M., Colver,A.F., Analysis by gestational age of cerebral palsy in singleton births in north-east England 1970-94, Paediatric and Perinatal Epidemiology, 16, 172-180, 2002

Faebo Larsen 2013

Faebo Larsen, R., Hvas Mortensen, L., Martinussen, T., Nybo Andersen, A. M., Determinants of developmental coordination disorder in 7-year-old children: a study of children in the Danish national birth cohort, Developmental Medicine and Child Neurology, 55, 1016-1022, 2013

Farooqi 2007

Farooqi, A., Hagglof, B., Sedin, G., Gothefors, L., Serenius, F., Mental health and social competencies of 10- to 12-year-old children born at 23 to 25 weeks of gestation in the 1990s: a Swedish national prospective follow-up study, Pediatrics, 120, 118-33, 2007

Farooqi 2011

Farooqi, A., Hagglof, B., Sedin, G., Serenius, F., Impact at age 11 years of major neonatal morbidities in children born extremely preterm, Pediatrics, 127, e1247-57, 2011

Farooqi 2013

Farooqi, A., Hagglof, B., Serenius, F., Behaviours related to executive functions and learning skills at 11 years of age after extremely preterm birth: A Swedish national prospective follow-up study, Acta Paediatrica, International Journal of Paediatrics, 102, 625-634, 2013

Farooqi 2016

Farooqi, A., Adamsson, M., Serenius, F., Hagglof, B., Executive functioning and learning skills of adolescent children born at fewer than 26 weeks of gestation, PLoS ONE, 11 (3) (no pagination), 2016

Fevang 2016

Fevang, S. K. E., Hysing, M., Markestad, T., et al.,, Mental Health in Children Born Extremely Preterm Without Severe Neurodevelopmental Disabilities, Pediatrics, 2016

Foix-L'Helias 2008

Foix-L'Helias, L., Marchand, L., Theret, B., Larroque, B., Ancel, P. Y., Blondel, B., Garel, M., Maillard, F., Missy, P., Sehili, F., Supernant, K., Durand, M., Matis, J., Messer, J., Treisser, A., Burguet, A., Abraham-Lerat, L., Menget, A., Roth, P., Schaal, J. P., Thiriez, G., Leveque, C., Marret, S., Marpeau, L., Boulot, P., Picaud, J. C., Donadio, A. M., Ledesert, B., Andre, M., Fresson, J., Hascoet, J. M., Arnaud, C., Bourdet-Loubere, S., Grandjean, H., Rolland, M., Leignel, C., Lequien, P., Pierrat, V., Puech, F., Subtil, D., Truffert, P., Boog, G., Rouger-Bureau, V., Roze, J. C., Ancel, P. Y., Breart, G., Kaminski, M., Du Mazaubrun, C., Dehan, M., Zupan-Simunek, V., Vodovar, M., Voyer, M., Impact of the use of antenatal corticosteroids on mortality, cerebral lesions and 5-year neurodevelopmental outcomes of very preterm infants: The EPIPAGE cohort study, BJOG: An International Journal of Obstetrics and Gynaecology, 115, 275-282, 2008

Foulder-Hughes & Cooke 2003

Foulder-Hughes, L. A., Cooke, R. W., Motor, cognitive, and behavioural disorders in children born very preterm, Developmental Medicine & Child Neurology, 45, 97-103, 2003

Frisk 2011

Frisk, V., Montgomery, L., Boychyn, E., Young, R., McLachlan, D., Neufeld, J., Van Ryn, E., Providing the right service at the right time to the right children born prematurely using evidence-based care pathways, 2011

Frisman 2012

Frisman,G.H., Eriksson,C., Pernehed,S., Morelius,E., The experience of becoming a grandmother to a premature infant - a balancing act, influenced by ambivalent feelings, Journal of Clinical Nursing, 21, 3297-3305, 2012

Garel 2007

Garel, M., Dardennes, M., Blondel, B., Mothers' psychological distress 1 year after very preterm childbirth. Results of the EPIPAGE qualitative study, Child: Care, Health & Development, 33, 137-43, 2007

Gaucher & Payot 2011

Gaucher,N., Payot,A., From powerlessness to empowerment: Mothers expect more than information from the prenatal consultation for preterm labour, Paediatrics and Child Health, 16, 638-642, 2011

Germa 2012

Germa,A., Marret,S., Thiriez,G., Rousseau,S., Hascoet,J.M., Paulsson-Bjornsson,L., Soderfeldt,B., Ancel,P.Y., Larroque,B., Kaminski,M., Nabet,C., Neonatal factors associated with alteration of palatal morphology in very preterm children: the EPIPAGE cohort study, Early Human Development, 88, 413-420, 2012

Glinianaia 2011

Glinianaia, S. V., Rankin, J., Colver, A., Cerebral palsy rates by birth weight, gestation and severity in North of England, 1991-2000 singleton births, Archives of Disease in Childhood, 96, 180-185, 2011

Goldstein 2013

Goldstein, R. F., Cotten, C. M., Shankaran, S., Gantz, M. G., Poole, W. K., Influence of gestational age on death and neurodevelopmental outcome in premature infants with severe intracranial hemorrhage, Journal of Perinatology, 33, 25-32, 2013

Gong 2015

Gong, A., Johnson, Y. R., Livingston, J., Matula, K., Duncan, A. F., Newborn intensive care survivors: a review and a plan for collaboration in Texas, Maternal Health, Neonatology and Perinatology, 1, 24, 2015

Green 1991

Green, S.B., How many subjects does it take to do a regression analysis, Multivariate Behav Res, 26(3), 499-510, 1991.

Guellec 2011

Guellec, I., Lapillonne, A., Renolleau, S., Charlaluk, M. L., Roze, J. C., Marret, S., Vieux, R., Monique, K., Ancel, P. Y., Epipage Study Group, Neurologic outcomes at school age in very preterm infants born with severe or mild growth restriction, Pediatrics, 127, e883-91, 2011

Guillen 2012

Guillen,U., Suh,S., Munson,D., Posencheg,M., Truitt,E., Zupancic,J.A., Gafni,A., Kirpalani,H., Development and pretesting of a decision-aid to use when counseling parents facing imminent extreme premature delivery, Journal of Pediatrics, 160, 382-387, 2012

Gurka 2010

Gurka, M. J., LoCasale-Crouch, J., Blackman, J. A., Long-term cognition, achievement, socioemotional, and behavioral development of healthy late-preterm infants, Archives of pediatrics & adolescent medicine, 164, 525-32, 2010

Guy 2015

Guy, A., Seaton, S. E., Boyle, E. M., Draper, E. S., Field, D. J., Manktelow, B. N., Marlow, N., Smith, L. K., Johnson, S., Infants born late/moderately preterm are at increased risk for a positive autism screen at 2 years of age, Journal of Pediatrics, 166, 269-75.e3, 2015

Halbwachs 2013

Halbwachs, M., Muller, J. B., Nguyen The Tich, S., de La Rochebrochard, E., Gascoin, G., Branger, B., Rouger, V., Roze, J. C., Flamant, C., Usefulness of parent-completed ASQ for neurodevelopmental screening of preterm children at five years of age, PLoS ONE, 8, e71925, 2013

Hansen 2004

Hansen, B. M., Hoff, B., Uldall, P., Greisen, G., Kamper, J., Djernes, B., Hertel, J., Christensen, M. F., Andersen, E., Lillquist, K., Verder, H., Peitersen, B., Grytter, C., Agertoft, L., Andersen, E. A., Berg, A., Krag-Olsen, B., Sardeman, H., Jonsbo, F., Jorgensen, N. F., Christensen, N. C., Nielsen, F., Ebbesen, F., Pryds, O., Lange, A., Danish, Etfol Group, Perinatal risk factors of adverse outcome in very preterm children: a role of initial treatment of respiratory insufficiency?, Acta Paediatrica, 93, 185-9, 2004

Harrison & Neufeld 1997

Harrison, M. J., Neufeld, A., Women's experiences of barriers to support while caregiving, Health Care for Women International, 18, 591-602, 1997

Harvey 2013

Harvey, M. E., Nongena, P., Gonzalez-Cinca, N., Edwards, A. D., Redshaw, M. E., Parents' experiences of information and communication in the neonatal unit about brain imaging and neurological prognosis: A qualitative study, Acta Paediatrica, International Journal of Paediatrics, 102, 360-365, 2013

Hayden 2013

Hayden, J.A., van der Windt, D.A., Cartwright, J.L., Cote, P., Bombardier, C., Assessing Bias in Studies of Prognostic Factors, Annals of Internal Medicine, 158(4), 280-286, 2013.

Helderman 2012

Helderman, J. B., O'Shea, T. M., Kuban, K. C., Allred, E. N., Hecht, J. L., Dammann, O., Paneth, N., McElrath, T. F., Onderdonk, A., Leviton, A., Elgan study Investigators, Antenatal antecedents of cognitive impairment at 24 months in extremely low gestational age newborns, Pediatrics, 129, 494-502, 2012

Hellgren 2016

Hellgren, K. M., Tornqvist, K., Jakobsson, P. G., Lundgren, P., Carlsson, B., Kallen, K., Serenius, F., Hellstrom, A., Holmstrom, G., Ophthalmologic outcome of extremely preterm infants at 6.5 years of age: Extremely preterm infants in Sweden study (EXPRESS), JAMA Ophthalmology, 134, 555-562, 2016

Herber-Jonat 2014

Herber-Jonat, S., Streiftau, S., Knauss, E., Voigt, F., Flemmer, A. W., Hummler, H. D., Schulze, A., Bode, H., Long-term outcome at age 7-10 years after extreme prematurity - a prospective, two centre cohort study of children born before 25 completed weeks of gestation (1999-2003), Journal of Maternal-Fetal & Neonatal Medicine, 27, 1620-6, 2014

Higa Diez 2016

Higa Diez, M., Yorifuji, T., Kado, Y., Sanada, S., Doi, H., Preterm birth and behavioural outcomes at 8 years of age: a nationwide survey in Japan, Archives of Disease in Childhood, 101, 338-43, 2016

Hillemeier 2011

Hillemeier, M. M., Morgan, P. L., Farkas, G., Maczuga, S. A., Perinatal and socioeconomic risk factors for variable and persistent cognitive delay at 24 and 48 months of age in a national sample, Maternal and Child Health Journal, 15, 1001-1010, 2011

Himmelmann & Uvebrant 2014

Himmelmann, K., Uvebrant, P., The panorama of cerebral palsy in Sweden. XI. Changing patterns in the birth-year period 2003-2006, Acta Paediatrica, 103, 618-24, 2014

Hintz 2005

Hintz,S.R., Kendrick,D.E., Stoll,B.J., Vohr,B.R., Fanaroff,A.A., Donovan,E.F., Poole,W.K., Blakely,M.L., Wright,L., Higgins,R., Neurodevelopmental and growth outcomes of extremely low birth weight infants after necrotizing enterocolitis, Pediatrics, 115, 696-703, 2005

Hirvonen 2014

Hirvonen, M., Ojala, R., Korhonen, P., Haataja, P., Eriksson, K., Gissler, M., Luukkaala, T., Tammela, O., Cerebral palsy among children born moderately and late preterm, Pediatrics, 134, e1584-93, 2014

Hoffman 2015

Hoffman, L., Bann, C., Higgins, R., Vohr, B., Eunice Kennedy Shriver National Institute of Child, Health, Human Development Neonatal Research, Network, Developmental outcomes of extremely preterm infants born to adolescent mothers, Pediatrics, 135, 1082-92, 2015

Holmstrom 2014

Holmstrom, G. E., Kallen, K., Hellstrom, A., Jakobsson, P. G., Serenius, F., Stjernqvist, K., Tornqvist, K., Ophthalmologic outcome at 30 months' corrected age of a prospective Swedish cohort of children born before 27 weeks of gestation: the extremely preterm infants in sweden study, JAMA Ophthalmology, 132, 182-9, 2014

Hornman 2016

Hornman, J, de Winter, AF, Kerstjens, JM, Bos, AF, Reijneveld, SA, Emotional and Behavioral Problems of Preterm and Full-Term Children at School Entry, Pediatrics, 137, 2016

Hreinsdottir 2013

Hreinsdottir, J., Ewald, U., Strand Brodd, K., Ornkloo, H., von Hofsten, C., Holmstrom, G., Ophthalmological outcome and visuospatial ability in very preterm children measured at 2.5 years corrected age, Acta Paediatrica, 102, 1144-9, 2013

Hussey-Gardner 2002

Hussey-Gardner, B., McNinch, A., Anastasi, J. M., Miller, M., Early intervention best practice: collaboration among an NICU, an early intervention program, and an NICU follow-up program, Neonatal Network - Journal of Neonatal Nursing, 21, 15-22, 2002

Hutchinson 2013

Hutchinson, E. A., De Luca, C. R., Doyle, L. W., Roberts, G., Anderson, P. J., Victorian Infant Collaborative Study, Group, School-age outcomes of extremely preterm or extremely low birth weight children.[Erratum appears in Pediatrics. 2013 Oct;132(4):780], Pediatrics, 131, e1053-61, 2013

Hwang 2013

Hwang, Y. S., Weng, S. F., Cho, C. Y., Tsai, W. H., Higher prevalence of autism in Taiwanese children born prematurely: A nationwide population-based study, Research in Developmental Disabilities, 34, 2462-2468, 2013

Ignell Mode 2014

Ignell Mode, R., Mard, E., Nyqvist, K. H., Blomqvist, Y. T., Fathers' perception of information received during their infants' stay at a neonatal intensive care unit, Sexual & reproductive healthcare : official journal of the Swedish Association of Midwives, 5, 131-6, 2014

Indredavik 2005

Indredavik, M. S., Vik, T., Heyerdahl, S., Kulseng, S., Brubakk, A. M., Psychiatric symptoms in low birth weight adolescents, assessed by screening questionnaires, European Child & Adolescent Psychiatry, 14, 226-36, 2005

Joanna Briggs Institute 2014

Joanna Briggs Institute. Joanna Briggs Institute Reviewers’ Manual. Joanna Briggs Institute. Melbourne, 2014

Johnson 2008

Johnson, S., Wolke, D., Marlow, N., Developmental assessment of preterm infants at 2 years: Validity of parent reports, Developmental medicine and child neurology, 50, 58-62, 2008

Johnson 2009

Johnson,S., Fawke,J., Hennessy,E., Rowell,V., Thomas,S., Wolke,D., Marlow,N., Neurodevelopmental disability through 11 years of age in children born before 26 weeks of gestation, Pediatrics, 124, e249-e257, 2009

Johnson 2010

Johnson, S., Hollis, C., Kochhar, P., Hennessy, E., Wolke, D., Marlow, N., Psychiatric Disorders in Extremely Preterm Children: Longitudinal Finding at Age 11 Years in the EPICure Study, Journal of the American Academy of Child and Adolescent Psychiatry, 49, 453-463.e1, 2010

Johnson 2010a

Johnson,S., Hollis,C., Kochhar,P., Hennessy,E., Wolke,D., Marlow,N., Autism spectrum disorders in extremely preterm children, Journal of Pediatrics, 156, 525-531, 2010

Johnson 2011

Johnson, S., Wolke, D., Hennessy, E., Marlow, N., Educational outcomes in extremely preterm children: neuropsychological correlates and predictors of attainment, Developmental Neuropsychology, 36, 74-95, 2011

Johnson 2014

Johnson, S., Hollis, C., Marlow, N., Simms, V., Wolke, D., Screening for childhood mental health disorders using the Strengths and Difficulties Questionnaire: The validity of multi-informant reports, Developmental Medicine and Child Neurology, 56, 453-459, 2014

Johnson 2015

Johnson, S., Evans, T. A., Draper, E. S., Field, D. J., Manktelow, B. N., Marlow, N., Matthews, R., Petrou, S., Seaton, S. E., Smith, L. K., Boyle, E. M., Neurodevelopmental outcomes following late and moderate prematurity: A population-based cohort study, Archives of Disease in Childhood: Fetal and Neonatal Edition, 100, F301-F308, 2015

Johnson 2015a

Johnson, S., Matthews, R., Draper, E. S., Field, D. J., Manktelow, B. N., Marlow, N., Smith, L. K., Boyle, E. M., Early Emergence of Delayed Social Competence in Infants Born Late and Moderately Preterm, Journal of Developmental & Behavioral Pediatrics, 36, 690-9, 2015

Johnson 2015b

Johnson, S., Gilmore, C., Gallimore, I., Jaekel, J., Wolke, D., The long-term consequences of preterm birth: what do teachers know?, Developmental Medicine & Child Neurology, 57, 571-7, 2015

Johnson 2016

Johnson, S., Matthews, R., Draper, E. S., Field, D. J., Manktelow, B. N., Marlow, N., Smith, L. K., Boyle, E. M., Eating difficulties in children born late and moderately preterm at 2 y of age: a prospective population-based cohort study, American Journal of Clinical Nutrition, 103, 406-14, 2016

Joseph 2016a

Joseph, R. M., O'Shea, T. M., Allred, E. N., Heeren, T., Hirtz, D., Paneth, N., Leviton, A., Kuban, K. C. K., Prevalence and associated features of autism spectrum disorder in extremely low gestational age newborns at age 10 years, Autism Research., 2016

Joseph 2016b

Joseph, R. M., O'Shea, T. M., Allred, E. N., Heeren, T., Hirtz, D., Jara, H., Leviton, A., Kuban, K. C., Elgan Study Investigators, Neurocognitive and Academic Outcomes at Age 10 Years of Extremely Preterm Newborns, Pediatrics, 137, 2016

Kallen 2015

Kallen, K., Serenius, F., Westgren, M., Marsal, K., Express Group, Impact of obstetric factors on outcome of extremely preterm births in Sweden: prospective population-based observational study (EXPRESS), Acta Obstetricia et Gynecologica Scandinavica, 94, 1203-14, 2015

Kan 2008

Kan, E., Roberts, G., Anderson, P. J., Doyle, L. W., Victorian Infant Collaborative Study, Group, The association of growth impairment with neurodevelopmental outcome at eight years of age in very preterm children, Early Human Development, 84, 409-16, 2008

Keenan 2005

Keenan,H.T., Doron,M.W., Seyda,B.A., Comparison of mothers' and counselors' perceptions of predelivery counseling for extremely premature infants, Pediatrics, 116, 104-111, 2005

Kent 2012

Kent, A. L., Wright, I. M., Abdel-Latif, M. E., New South, Wales, Australian Capital Territory Neonatal Intensive Care Units Audit, Group, Mortality and adverse neurologic outcomes are greater in preterm male infants, Pediatrics, 129, 124-31, 2012

Kerstjens 2011

Kerstjens, J. M., de Winter, A. F., Bocca-Tjeertes, I. F., ten Vergert, E. M., Reijneveld, S. A., Bos, A. F., Developmental delay in moderately preterm-born children at school entry, Journal of Pediatrics, 159, 92-8, 2011

Kerstjens 2012

Kerstjens,J.M., Bocca-Tjeertes,I.F., de Winter,A.F., Reijneveld,S.A., Bos,A.F., Neonatal morbidities and developmental delay in moderately preterm-born children, Pediatrics, 130, e265-e272, 2012

Kerstjens 2013

Kerstjens,J.M., de Winter,A.F., Sollie,K.M., Bocca-Tjeertes,I.F., Potijk,M.R., Reijneveld,S.A., Bos,A.F., Maternal and pregnancy-related factors associated with developmental delay in moderately preterm-born children, Obstetrics and Gynecology, 121, 727-733, 2013

Kiechl-Kohlendorfer 2013

Kiechl-Kohlendorfer, U., Ralser, E., Pupp Peglow, U., Pehboeck-Walser, N., Fussenegger, B., Early risk predictors for impaired numerical skills in 5-year-old children born before 32 weeks of gestation, Acta Paediatrica, 102, 66-71, 2013

Kuzniewicz 2014

Kuzniewicz, M. W., Wi, S., Qian, Y., Walsh, E. M., Armstrong, M. A., Croen, L. A., Prevalence and neonatal factors associated with autism spectrum disorders in preterm infants, Journal of Pediatrics, 164, 20-25, 2014

Larroque 2008

Larroque, B., Ancel, P. Y., Marret, S., Marchand, L., Andre, M., Arnaud, C., Pierrat, V., Roze, J. C., Messer, J., Thiriez, G., Burguet, A., Picaud, J. C., Breart, G., Kaminski, M., Epipage Study group, Neurodevelopmental disabilities and special care of 5-year-old children born before 33 weeks of gestation (the EPIPAGE study): a longitudinal cohort study, Lancet, 371, 813-20, 2008

Larroque 2011

Larroque, B., Ancel, P. Y., Marchand-Martin, L., Cambonie, G., Fresson, J., Pierrat, V., Roze, J. C., Marpeau, L., Thiriez, G., Alberge, C., Breart, G., Kaminski, M., Marret, S., Epipage Study, group, Special care and school difficulties in 8-year-old very preterm children: the Epipage cohort study, PLoS ONE [Electronic Resource], 6, e21361, 2011

Lasby 2004

Lasby, K., Newton, S., von Platen, A., Neonatal transitional care, Canadian Nurse, 100, 18-23, 2004

Laughon 2009

Laughon, M., O'Shea, M. T., Allred, E. N., Bose, C., Kuban, K., Van Marter, L. J., Ehrenkranz, R. A., Leviton, A., Chronic lung disease and developmental delay at 2 years of age in children born before 28 weeks' gestation, Pediatrics, 124, 637-648, 2009

Lee & Lin 2013

Lee, T. Y., Lin, F. Y., Taiwanese parents' perceptions of their very low-birth-weight infant with developmental disabilities, Journal of Perinatal & Neonatal Nursing, 27, 345-52, 2013

Lee 2009

Lee, T. Y., Lee, T. T., Kuo, S. C., The experiences of mothers in breastfeeding their very low birth weight infants, Journal of Advanced Nursing, 65, 2523-2531, 2009

Leversen 2010

Leversen,K.T., Sommerfelt,K., Ronnestad,A., Kaaresen,P.I., Farstad,T., Skranes,J., Stoen,R., Elgen,I.B., Rettedal,S., Eide,G.E., Irgens,L.M., Markestad,T., Predicting neurosensory disabilities at two years of age in a national cohort of extremely premature infants, Early Human Development, 86, 581-586, 2010

Leversen 2011

Leversen, K. T., Sommerfelt, K., Ronnestad, A., Kaaresen, P. I., Farstad, T., Skranes, J., Stoen, R., Bircow Elgen, I., Rettedal, S., Egil Eide, G., Irgens, L. M., Markestad, T., Prediction of neurodevelopmental and sensory outcome at 5 years in Norwegian children born extremely preterm, Pediatrics, 127, e630-8, 2011

Leversen 2012

Leversen,K.T., Sommerfelt,K., Elgen,I.B., Eide,G.E., Irgens,L.M., Juliusson,P.B., Markestad,T., Prediction of outcome at 5 years from assessments at 2 years among extremely preterm children: a Norwegian national cohort study, Acta Paediatrica, 101, 264-270, 2012

Lewin 2015.

Lewin, S., Glenton, C., Munthe-Kaas, H., Carlsen, B., Colvin, C., Gülmezoglu, M., Noyes, J., Booth, A., Garside, R., Rashidian, A., Using Qualitative Evidence in Decision Making for Health and Social Interventions: An Approach to Assess Confidence in Findings from Qualitative Evidence Syntheses (GRADE-CERQual), PLoS Med 12(10), e1001895, 2015.

Little 2015

Little, A. A., Kamholz, K., Corwin, B. K., Barrero-Castillero, A., Wang, C. J., Understanding Barriers to Early Intervention Services for Preterm Infants: Lessons From Two States, Academic Pediatrics, 15, 430-438, 2015

MacKay 2010

MacKay, D. F., Smith, G. C., Dobbie, R., Pell, J. P., Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren, PLoS Medicine / Public Library of Science, 7, e1000289, 2010

MacKay 2013

MacKay, D. F., Smith, G. C. S., Dobbie, R., Cooper, S. A., Pell, J. P., Obstetric factors and different causes of special educational need: Retrospective cohort study of 407 503 schoolchildren, BJOG: An International Journal of Obstetrics and Gynaecology, 120, 297-307, 2013

Mansson & Stjernqvist 2014

Mansson, J., Stjernqvist, K., Children born extremely preterm show significant lower cognitive, language and motor function levels compared with children born at term, as measured by the Bayley-III at 2.5 years, Acta Paediatrica, 103, 504-11, 2014

Marlow 2005

Marlow, N., Wolke, D., Bracewell, M. A., Samara, M., E. PICure Study Group, Neurologic and developmental disability at six years of age after extremely preterm birth, New England Journal of Medicine, 352, 9-19, 2005

Marret 2007

Marret, S., Ancel, P. Y., Marpeau, L., Marchand, L., Pierrat, V., Larroque, B., Foix-L'Helias, L., Thiriez, G., Fresson, J., Alberge, C., Roze, J. C., Matis, J., Breart, G., Kaminski, M., Epipage Study, Group, Neonatal and 5-year outcomes after birth at 30-34 weeks of gestation, Obstetrics & Gynecology, 110, 72-80, 2007

Marshall 2003

Marshall, D. D., Primary care follow-up of the neonatal intensive care unit graduate, Clinics in Family Practice, 5, 243-263, 2003

Martin 2010

Martin,C.R., Dammann,O., Allred,E.N., Patel,S., O'Shea,T.M., Kuban,K.C., Leviton,A., Neurodevelopment of extremely preterm infants who had necrotizing enterocolitis with or without late bacteremia, Journal of Pediatrics, 157, 751-756, 2010

Martin 2013

Martin, A. J., Darlow, B. A., Salt, A., Hague, W., Sebastian, L., McNeill, N., Tarnow-Mordi, W., Performance of the Parent Report of Children's Abilities-Revised (PARCA-R) versus the Bayley Scales of Infant Development III, Archives of Disease in Childhood, 98, 955-8, 2013

May 1997

May, K. M., Searching for normalcy: mothers' caregiving for low birth weight infants, Pediatric Nursing, 23, 17-20, 1997

Merhar 2012

Merhar,S.L., Tabangin,M.E., Meinzen-Derr,J., Schibler,K.R., Grade and laterality of intraventricular haemorrhage to predict 18-22 month neurodevelopmental outcomes in extremely low birthweight infants, Acta Paediatrica, International Journal of Paediatrics, 101, 414-418, 2012

Migraine 2013

Migraine, A., Nicklaus, S., Parnet, P., Lange, C., Monnery-Patris, S., Des Robert, C., Darmaun, D., Flamant, C., Amarger, V., Roze, J. C., Effect of preterm birth and birth weight on eating behavior at 2 y of age, American Journal of Clinical Nutrition, 97, 1270-7, 2013

Mikkola 2005

Mikkola,K., Ritari,N., Tommiska,V., Salokorpi,T., Lehtonen,L., Tammela,O., Paakkonen,L., Olsen,P., Korkman,M., Fellman,V., Neurodevelopmental outcome at 5 years of age of a national cohort of extremely low birth weight infants who were born in 1996-1997, Pediatrics, 116, 1391-1400, 2005

Miyazaki 2016

Miyazaki, K., Furuhashi, M., Ishikawa, K., Tamakoshi, K., Hayashi, K., Kai, A., Ishikawa, H., Murabayashi, N., Ikeda, T., Kono, Y., Kusuda, S., Fujimura, M., Impact of chorioamnionitis on short- and long-term outcomes in very low birth weight preterm infants: the Neonatal Research Network Japan, Journal of Maternal-Fetal & Neonatal Medicine, 29, 331-7, 2016

Moore 2012

Moore, G. S., Kneitel, A. W., Walker, C. K., Gilbert, W. M., Xing, G., Autism risk in small- and large-for-gestational-age infants, American Journal of Obstetrics and Gynecology, 206, 314.e1-314.e9, 2012

Moore 2012a

Moore,T., Hennessy,E.M., Myles,J., Johnson,S.J., Draper,E.S., Costeloe,K.L., Marlow,N., Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies, BMJ, 345, e7961-, 2012

Moore 2012b

Moore, T., Johnson, S., Hennessy, E., Marlow, N., Screening for autism in extremely preterm infants: problems in interpretation, Developmental Medicine & Child Neurology, 54, 514-20, 2012

Munn 2014

Munn, Z., Moola, S., Riitano, D., Lisy, K., The development of a critical appraisal tool for use in systematic reviews addressing questions of prevalence, Int J Health Policy Manag, 3 (3), 123–128, 2014.

Natarajan 2012

Natarajan,G., Pappas,A., Shankaran,S., Kendrick,D.E., Das,A., Higgins,R.D., Laptook,A.R., Bell,E.F., Stoll,B.J., Newman,N., Hale,E.C., Bara,R., Walsh,M.C., Outcomes of extremely low birth weight infants with bronchopulmonary dysplasia: Impact of the physiologic definition, Early Human Development, 88, 509-515, 2012

Neu & Robinson 2008

Neu,M., Robinson,J., Early weeks after premature birth as experienced by Latina adolescent mothers, MCN, American Journal of Maternal Child Nursing, 33, 166-172, 2008

Nicolaou 2009

Nicolaou,M., Rosewell,R., Marlow,N., Glazebrook,C., Mothers' experiences of interacting with their premature infants, Journal of Reproductive and Infant Psychology, 27, 182-194, 2009

Niela-Vilen 2015

Niela-Vilen, H., Axelin, A., Melender, H. L., Salantera, S., Aiming to be a breastfeeding mother in a neonatal intensive care unit and at home: A thematic analysis of peer-support group discussion in social media, Maternal and Child Nutrition, 11, 712-726, 2015

Nordmark 2001

Nordmark,E., Hagglund,G., Lagergren,J., Cerebral palsy in southern Sweden I. Prevalence and clinical features, Acta Paediatrica, 90, 1271-1276, 2001

Odd 2012

Odd,D.E., Emond,A., Whitelaw,A., Long-term cognitive outcomes of infants born moderately and late preterm, Developmental Medicine and Child Neurology, 54, 704-709, 2012

Odd 2013

Odd,D.E., Lingam,R., Emond,A., Whitelaw,A., Movement outcomes of infants born moderate and late preterm, Acta Paediatrica, 102, 876-882, 2013

Odd 2013a

Odd,D., Evans,D., Emond,A., Preterm Birth, Age at School Entry and Educational Performance, PLoS ONE, 8, -, 2013

Odd 2016

Odd, D., Evans, D., Emond, A., Preterm Birth, Age at School Entry and Long Term Educational Achievement, PLoS ONE, 11, e0155157, 2016

ONS 2015

Office for National Statistics. Families and Households: 2014. Office for National Statistics. UK, 2015.

ONS 2016

Office for National Statistics. Birth characteristics in England and Wales: 2015. Live births and stillbirths by gestational age at birth, 2015 birth cohort. Office for National Statistics. UK, 2016.

ONS 2016

Office for National Statistics. National Life tables, United Kingdom: 2013-2015. Office for National Statistics. UK, 2016.

O'Shea 2008

O'Shea, T. M., Kuban, K. C., Allred, E. N., Paneth, N., Pagano, M., Dammann, O., Bostic, L., Brooklier, K., Butler, S., Goldstein, D. J., Hounshell, G., Keller, C., McQuiston, S., Miller, A., Pasternak, S., Plesha-Troyke, S., Price, J., Romano, E., Solomon, K. M., Jacobson, A., Westra, S., Leviton, A., Neonatal cranial ultrasound lesions and developmental delays at 2 years of age among extremely low gestational age children, Pediatrics, 122, e662-e669, 2008

Padden & Glenn 1997

Padden, T., Glenn, S., Maternal experiences of preterm birth and neonatal intensive care, Journal of Reproductive and Infant Psychology, 15, 121-139, 1997

Pappas 2014

Pappas, A., Kendrick, D. E., Shankaran, S., Stoll, B. J., Bell, E. F., Laptook, A. R., Walsh, M. C., Das, A., Hale, E. C., Newman, N. S., Higgins, R. D., Chorioamnionitis and early childhood outcomes among extremely low-gestational-age neonates, JAMA Pediatrics, 168, 137-147, 2014

Payne 2013

Payne, A. H., Hintz, S. R., Hibbs, A. M., Walsh, M. C., Vohr, B. R., Bann, C. M., Wilson-Costello, D. E., Neurodevelopmental outcomes of extremely low-gestational-age neonates with low-grade periventricular-intraventricular hemorrhage, JAMA Pediatrics, 167, 451-459, 2013

Peacock 2012

Peacock, P. J., Henderson, J., Odd, D., Emond, A., Early school attainment in late-preterm infants, Archives of Disease in Childhood, 97, 118-20, 2012

Perrott 2003

Perrott,S., Dodds,L., Vincer,M., A population-based study of prognostic factors related to major disability in very preterm survivors, Journal of Perinatology, 23, 111-116, 2003

Petrini 2009

Petrini,J.R., Dias,T., McCormick,M.C., Massolo,M.L., Green,N.S., Escobar,G.J., Increased risk of adverse neurological development for late preterm infants, Journal of Pediatrics, 154, 169-176, 2009

Phillips-Pula 2013

Phillips-Pula, L., Pickler, R., McGrath, J. M., Brown, L. F., Dusing, S. C., Caring for a preterm infant at home: a mother's perspective, The Journal of perinatal & neonatal nursing, 27, 335-344, 2013

Plomgaard 2006

Plomgaard, A. M., Hansen, B. M., Greisen, G., Measuring developmental deficit in children born at gestational age less than 26 weeks using a parent-completed developmental questionnaire, Acta Paediatrica, 95, 1488-94, 2006

Potijk 2012

Potijk, M. R., de Winter, A. F., Bos, A. F., Kerstjens, J. M., Reijneveld, S. A., Higher rates of behavioural and emotional problems at preschool age in children born moderately preterm, Archives of Disease in Childhood, 97, 112-7, 2012

Potijk 2013

Potijk, M. R., Kerstjens, J. M., Bos, A. F., Reijneveld, S. A., de Winter, A. F., Developmental delay in moderately preterm-born children with low socioeconomic status: risks multiply, Journal of Pediatrics, 163, 1289-95, 2013

Potijk 2015

Potijk, M. R., de Winter, A. F., Bos, A. F., Kerstjens, J. M., Reijneveld, S. A., Behavioural and emotional problems in moderately preterm children with low socioeconomic status: a population-based study, European Child & Adolescent Psychiatry, 24, 787-95, 2015

Quigley 2012

Quigley, M. A., Poulsen, G., Boyle, E., Wolke, D., Field, D., Alfirevic, Z., Kurinczuk, J. J., Early term and late preterm birth are associated with poorer school performance at age 5 years: a cohort study, Archives of Disease in Childhood Fetal & Neonatal Edition, 97, F167-73, 2012

Rabie 2015

Rabie, N. Z., Bird, T. M., Magann, E. F., Hall, R. W., McKelvey, S. S., ADHD and developmental speech/language disorders in late preterm, early term and term infants, Journal of Perinatology, 35, 660-664, 2015

Rautava 2010

Rautava, L., Andersson, S., Gissler, M., Hallman, M., Hakkinen, U., Korvenranta, E., Korvenranta, H., Leipala, J., Tammela, O., Lehtonen, L., Development and behaviour of 5-year-old very low birthweight infants, European Child & Adolescent Psychiatry, 19, 669-77, 2010

Raynes-Greenow 2012

Raynes-Greenow, C. H., Hadfield, R. M., Cistulli, P. A., Bowen, J., Allen, H., Roberts, C. L., Sleep apnea in early childhood associated with preterm birth but not small for gestational age: a population-based record linkage study, Sleep, 35, 1475-80, 2012

Reijneveld 2006

Reijneveld, S. A., de Kleine, M. J., van Baar, A. L., Kollee, L. A., Verhaak, C. M., Verhulst, F. C., Verloove-Vanhorick, S. P., Behavioural and emotional problems in very preterm and very low birthweight infants at age 5 years, Archives of Disease in Childhood Fetal & Neonatal Edition, 91, F423-8, 2006

Reyna 2006

Reyna, B. A., Pickler, R. H., Thompson, A., A descriptive study of mothers' experiences feeding their preterm infants after discharge, Advances in Neonatal Care, 6, 333-40, 2006

Rieger-Fackeldey 2010

Rieger-Fackeldey, E., Blank, C., Dinger, J., Steinmacher, J., Bode, H., Schulze, A., Growth, neurological and cognitive development in infants with a birthweight <501 g at age 5 years, Acta Paediatrica, 99, 1350-5, 2010

Roberts et al., 2010

Roberts, G., Anderson, P. J., De Luca, C., Doyle, L. W., Changes in neurodevelopmental outcome at age eight in geographic cohorts of children born at 22-27 weeks' gestational age during the 1990s, Archives of Disease in Childhood: Fetal and Neonatal Edition, 95, F90-F94, 2010

Roberts 2011

Roberts, G., Anderson, P. J., Davis, N., De Luca, C., Cheong, J., Doyle, L. W., Developmental coordination disorder in geographic cohorts of 8-year-old children born extremely preterm or extremely low birthweight in the 1990s, Developmental medicine and child neurology, 53, 55-60, 2011

Robertson 2007

Robertson, C. M., Watt, M. J., Yasui, Y., Changes in the prevalence of cerebral palsy for children born very prematurely within a population-based program over 30 years, JAMA, 297, 2733-40, 2007

Rogers 2013

Rogers, C. E., Lenze, S. N., Luby, J. L., Late preterm birth, maternal depression, and risk of preschool psychiatric disorders, Journal of the American Academy of Child and Adolescent Psychiatry, 52, 309-318, 2013

Russell 2014

Russell, G., Sawyer, A., Rabe, H., Abbott, J., Gyte, G., Duley, L., Ayers, S., Parents' views on care of their very premature babies in neonatal intensive care units: a qualitative study, BMC Pediatrics, 14, 230, 2014

Salokorpi 2001

Salokorpi, T., Rautio, T., Sajaniemi, N., Serenius-Sirve, S., Tuomi, H., von Wendt, L., Neurological development up to the age of four years of extremely low birthweight infants born in Southern Finland in 1991-94, Acta Paediatrica, 90, 218-21, 2001

Salt & Redshaw 2006

Salt, A., Redshaw, M., Neurodevelopmental follow-up after preterm birth: Follow-up after two years, Early Human Development, 82, 185-197, 2006

Samara 2008

Samara, M., Marlow, N., Wolke, D., E. PICure Study Group, Pervasive behavior problems at 6 years of age in a total-population sample of children born at </= 25 weeks of gestation, Pediatrics, 122, 562-73, 2008

Samara 2010

Samara, M., Johnson, S., Lamberts, K., Marlow, N., Wolke, D., Eating problems at age 6 years in a whole population sample of extremely preterm children, Developmental Medicine & Child Neurology, 52, e16-22, 2010

Schendel 1997

Schendel, D. E., Stockbauer, J. W., Hoffman, H. J., Herman, A. A., Berg, C. J., Schramm, W. F., Relation between very low birth weight and developmental delay among preschool children without disabilities, American Journal of Epidemiology, 146, 740-9, 1997

Schonhaut 2013

Schonhaut, L., Armijo, I., Schonstedt, M., Alvarez, J., Cordero, M., Validity of the ages and stages questionnaires in term and preterm infants, Pediatrics, 131, e1468-74, 2013

Serenius 2013

Serenius, F., Kallen, K., Blennow, M., Ewald, U., Fellman, V., Holmstrom, G., Lindberg, E., Lundqvist, P., Marsal, K., Norman, M., Olhager, E., Stigson, L., Stjernqvist, K., Vollmer, B., Stromberg, B., Express Group, Neurodevelopmental outcome in extremely preterm infants at 2.5 years after active perinatal care in Sweden, JAMA, 309, 1810-20, 2013

Shah 2012

Shah, T. A., Meinzen-Derr, J., Gratton, T., Steichen, J., Donovan, E. F., Yolton, K., Alexander, B., Narendran, V., Schibler, K. R., Hospital and neurodevelopmental outcomes of extremely low-birth-weight infants with necrotizing enterocolitis and spontaneous intestinal perforation, Journal of Perinatology, 32, 552-8, 2012

Shankaran 2004

Shankaran, S., Johnson, Y., Langer, J. C., Vohr, B. R., Fanaroff, A. A., Wright, L. L., Poole, W. K., Outcome of extremely-low-birth-weight infants at highest risk: Gestational age <24 weeks, birth weight <750 g, and 1-minute Apgar <3, American Journal of Obstetrics and Gynecology, 191, 1084-1091, 2004

Simard 2012

Simard, M. N., Luu, T. M., Gosselin, J., Concurrent validity of ages and stages questionnaires in preterm infants, Pediatrics, 130, e108-14, 2012

Singer 2001

Singer, L. T., Hawkins, S., Huang, J., Davillier, M., Baley, J., Developmental outcomes and environmental correlates of very low birthweight, cocaine-exposed infants, Early Human Development, 64, 91-103, 2001

Singh 2013

Singh, G. K., Kenney, M. K., Ghandour, R. M., Kogan, M. D., Lu, M. C., Mental Health Outcomes in US Children and Adolescents Born Prematurely or with Low Birthweight, Depression Research and Treatment, 2013, 570743, 2013

Skellern 2001

Skellern, C. Y., Rogers, Y., O'Callaghan, M. J., A parent-completed developmental questionnaire: follow-up of ex-premature infants, Journal of Paediatrics & Child Health, 37, 125-9, 2001

Sommer & Cook 2015

Sommer, C. M., Cook, C. M., Disrupted bonds - parental perceptions of regionalised transfer of very preterm infants: a small-scale study, Contemporary nurse, 50, 256-266, 2015

Stahlmann 2009

Stahlmann,N., Rapp,M., Herting,E., Thyen,U., Outcome of extremely premature infants at early school age: health-related quality of life and neurosensory, cognitive, and behavioral outcomes in a population-based sample in northern Germany, Neuropediatrics, 40, 112-119, 2009

Stene-Larsen 2014

Stene-Larsen, K., Brandlistuen, R. E., Lang, A. M., Landolt, M. A., Latal, B., Vollrath, M. E., Communication impairments in early term and late preterm children: A prospective cohort study following children to age 36 months, Journal of Pediatrics, 165, 1123-1128, 2014

Stoelhorst 2003

Stoelhorst, G. M. S. J., Rijken, M., Martens, S. E., Van Zwieten, P. H. T., Feenstra, J., Zwinderman, A. H., Wit, J. M., Veen, S., Developmental outcome at 18 and 24 months of age in very preterm children: A cohort study from 1996 to 1997, Early Human Development, 72, 83-95, 2003

Stoelhorst 2003a

Stoelhorst, G. M. S. J., Martens, S. E., Rijken, M., Van Zwieten, P. H. T., Zwinderman, A. H., Wit, J. M., Veen, S., Behaviour at 2 years of age in very preterm infants (gestational age <32 weeks), Acta Paediatrica, International Journal of Paediatrics, 92, 595-601, 2003

Stoll 2004

Stoll, B. J., Hansen, N. I., Adams-Chapman, I., Fanaroff, A. A., Hintz, S. R., Vohr, B., Higgins, R. D., Neurodevelopmental and growth impairment among extremely low-birth-weight infants with neonatal infection, Journal of the American Medical Association, 292, 2357-2365, 2004

Sukhov 2012

Sukhov, A., Wu, Y., Xing, G., Smith, L. H., Gilbert, W. M., Risk factors associated with cerebral palsy in preterm infants, Journal of Maternal-Fetal & Neonatal Medicine, 25, 53-7, 2012

Sullivan 2015

Sullivan, S., Joinson, C., Heron, J., Factors Predicting Atypical Development of Nighttime Bladder Control, Journal of Developmental & Behavioral Pediatrics, 36, 724-33, 2015

Sutton & Bajuk 1999

Sutton, L., Bajuk, B., Population-based study of infants born at less than 28 weeks' gestation in New South Wales, Australia, in 1992-3, Paediatric and Perinatal Epidemiology, 13, 288-301, 1999

Synnes 2006

Synnes, A.R., Lefebvre, F., Cake, H.A., Current status of neonatal follow-up in Canada, Paediatr Child Health, 11(5), 271–274, 2006.

Thomas 2009

Thomas, J., Feeley, N., Grier, P., The perceived parenting self-efficacy of first-time fgathers caring for very-low-birth-weight infants self-efficacy in fathers of VLBW infants, Issues in Comprehensive Pediatric Nursing, 32, 180-199, 2009

Tommiska 2003

Tommiska,V., Heinonen,K., Kero,P., Pokela,M.L., Tammela,O., Jarvenpaa,A.L., Salokorpi,T., Virtanen,M., Fellman,V., A national two year follow-up study of extremely low birthweight infants born in 1996-1997, Archives of Disease in Childhood Fetal and Neonatal Edition, 88, F29-F35, 2003

Toome 2008

Toome, L., Varendi, H., Ilgina, O., Jaanson, E., Kaasik, B., Kolk, A., Kruustük, K., Lõivukene, R., Maas, H., Meriste, S., Mägi, M-L., Männamaa, M., Pakosta, T. , Salong, K., Sander, V., Stelmach, T., Tänavsuu, T., Utsal, U., Riskivastsündinute jälgimise juhend lapse esimesel ja teisel eluaastal, Eesti Arst, 87, 389−403, 2008

Toome 2013

Toome,L., Varendi,H., Mannamaa,M., Vals,M.A., Tanavsuu,T., Kolk,A., Follow-up study of 2-year-olds born at very low gestational age in Estonia, Acta Paediatrica, 102, 300-307, 2013

Turner 2013

Turner,M., Winefield,H., Chur-Hansen,A., The emotional experiences and supports for parents with babies in a neonatal nursery, Advances in Neonatal Care, 13, 438-446, 2013

Van Marter 2011

Van Marter, L. J., Kuban, K. C. K., Allred, E., Bose, C., Dammann, O., O'Shea, M., Laughon, M., Ehrenkranz, R. A., Schreiber, M. D., Karna, P., Leviton, A., Does bronchopulmonary dysplasia contribute to the occurrence of cerebral palsy among infants born before 28 weeks of gestation?, Archives of Disease in Childhood: Fetal and Neonatal Edition, 96, F20-F29, 2011

Vasquez 1995

Vasquez, E., Creating paths: living with a very-low-birth-weight infant, Journal of obstetric, gynecologic, and neonatal nursing: JOGNN / NAACOG, 24, 619-624, 1995

Victorian Infant Collaborative 2000

Victorian Infant Collaborative, Postnatal corticosteroids and sensorineural outcome at 5 years of age, Journal of Paediatrics & Child Health, 36, 256-61, 2000

Victorian Infant Collaborative Study Group 1997

Victorian Infant Collaborative Study Group, Outcome at 2 years of children 23-27 weeks' gestation born in Victoria in 1991-92. The Victorian Infant Collaborative Study Group, Journal of Paediatrics & Child Health, 33, 161-5, 1997

Vincer 2006

Vincer,M.J., Allen,A.C., Joseph,K.S., Stinson,D.A., Scott,H., Wood,E., Increasing prevalence of cerebral palsy among very preterm infants: a population-based study, Pediatrics, 118, e1621-e1626, 2006

Vincer 2014

Vincer,M.J., Allen,A.C., Allen,V.M., Baskett,T.F., O'Connell,C.M., Trends in the prevalence of cerebral palsy among very preterm infants (<31 weeks' gestational age), Paediatrics and Child Health, 19, 185-189, 2014

Vohr 2000

Vohr, B. R., Wright, L. L., Dusick, A. M., Mele, L., Verter, J., Steichen, J. J., Simon, N. P., Wilson, D. C., Broyles, S., Bauer, C. R., Delaney-Black, V., Yolton, K. A., Fleisher, B. E., Papile, L. A., Kaplan, M. D., Neurodevelopmental and functional outcomes of extremely low birth weight infants in the National Institute of Child Health and Human Development Neonatal Research Network, 1993-1994, Pediatrics, 105, 1216-1226, 2000

Vohr 2005

Vohr,B.R., Wright,L.L., Poole,W.K., McDonald,S.A., Neurodevelopmental outcomes of extremely low birth weight infants <32 weeks' gestation between 1993 and 1998, Pediatrics, 116, 635-643, 2005

Vollmer 2012

Vollmer, B., Baral, V., Hart, D., Drew, H., Hall, M. , Standardised neurodevelopmental follow-up of preterm infants and high risk newborns at University Hospital Southampton NHS Foundation Trust , 2012

Walsh 2005

Walsh, M. C., Morris, B. H., Wrage, L. A., Vohr, B. R., Poole, W. K., Tyson, J. E., Wright, L. L., Ehrenkranz, R. A., Stoll, B. J., Fanaroff, A. A., National Institutes of Child, Health, Human Development Neonatal Research, Network, Extremely low birthweight neonates with protracted ventilation: mortality and 18-month neurodevelopmental outcomes, Journal of Pediatrics, 146, 798-804, 2005

Whiting 2011

Whiting, P.F., Rutjes, A.W.S., Westwood, M.E., Mallett, S., Deeks, J.J., Reitsma, J.B., Leeflang, M.M., Sterne, J.A.C., Bossuyt, P.M.M., QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies, Ann Intern Med, 155 (8), 529-536, 2011.

Whittingham 2014

Whittingham, K., Boyd, R. N., Sanders, M. R., Colditz, P, Parenting and prematurity: Understanding parent experience and preferences for support, Journal of Child and Family Studies, 23, 1050-1061, 2014

Wilson-Ching 2013

Wilson-Ching, M., Molloy, C. S., Anderson, V. A., Burnett, A., Roberts, G., Cheong, J. L., Doyle, L. W., Anderson, P. J., Attention difficulties in a contemporary geographic cohort of adolescents born extremely preterm/extremely low birth weight, Journal of the International Neuropsychological Society, 19, 1097-108, 2013

Wolke 2008

Wolke, D., Samara, M., Bracewell, M., Marlow, N., E. PICure Study Group, Specific language difficulties and school achievement in children born at 25 weeks of gestation or less, Journal of Pediatrics, 152, 256-62, 2008

Wong 2014

Wong, D., Abdel-Latif, M., Kent, A., Nicus Network, Antenatal steroid exposure and outcomes of very premature infants: a regional cohort study, Archives of Disease in Childhood Fetal & Neonatal Edition, 99, F12-20, 2014

Wood 2000

Wood, N. S., Marlow, N., Costeloe, K., Gibson, A. T., Wilkinson, A. R., Neurologic and developmental disability after extremely preterm birth. EPICure Study Group, New England Journal of Medicine, 343, 378-84, 2000

Wood 2005

Wood, N. S., Costeloe, K., Gibson, A. T., Hennessy, E. M., Marlow, N., Wilkinson, A. R., The EPICure study: Associations and entecedents of neurological and developmental disability at the 30 months of age following extremely preterm birth, Archives of Disease in Childhood: Fetal and Neonatal Edition, 90, F134-F140, 2005

Woodward 2011

Woodward, B. J., Papile, L. A., Lowe, J. R., Laadt, V. L., Shaffer, M. L., Montman, R., Watterberg, K. L., Use of the Ages and Stages Questionnaire and Bayley Scales of Infant

Development-II in neurodevelopmental follow-up of extremely low birth weight infants, Journal of Perinatology, 31, 641-6, 2011

Woythaler 2011

Woythaler, M. A., McCormick, M. C., Smith, V. C., Late preterm infants have worse 24-month neurodevelopmental outcomes than term infants, Pediatrics, 127, e622-9, 2011

Zhu 2012

Zhu, J. L., Olsen, J., Olesen, A. W., Risk for developmental coordination disorder correlates with gestational age at birth, Paediatric and Perinatal Epidemiology, 26, 572-577, 2012

1 # Glossary of terms

Term	Definition
Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Antenatal risk factors	Maternal Risk Factors that might increase the risk of developmental disorders to the unborn child
Antenatal steroids	Administration of a corticosteroid preparation to a pregnant woman, at risk of preterm birth. These are currently associated with a significant reduction in neonatal mortality, respiratory distress syndrome and intraventricular haemorrhage
Area under the curve (AUC)	Area under the curve (AUC) is a measure of how well a parameter can distinguish between two diagnostic groups (with condition/without condition), often visualised in a SROC plot.
Arm (of a clinical study)	Subsection of individuals within a study who receive one particular intervention, for example placebo arm.
Association	Statistical relationship between 2 or more events, characteristics or other variables. The relationship may or may not be causal.
Attrition bias	Systematic differences between comparison groups for withdrawal or exclusion of participants from a study.
Attention deficit/hyperactivity disorder (ADHD)	Being hyperactive and impulsive or having difficulties with concentration and attention that are excessive for the child’s age as rated by parents and/or teachers using standardised questionnaires or rating scales.
Autism spectrum disorder (ASD)	Autism spectrum disorder (ASD) is characterized by persistent deficits in social communication and social interaction across multiple contexts combined with restricted repetitive patterns of behaviour, interests or activities
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable) with which subsequent results are compared.
Bayley scale	A standardised assessment tool of cognition.
Behaviour problems	Aggressive, disruptive, delinquent or defiant behaviours that are inappropriate or excessive for the child’s age as rated by parents or teachers using standardised questionnaires or rating scales.
Bias	Influences on a study that can make the results look better or worse than they really are. Bias can occur by chance, deliberately or as a result of systematic errors in the design and execution of a study. It can also occur at different stages in the research process, for example during the collection, analysis, interpretation, publication or review of research data. For examples see Confounding factor, Performance bias, Publication bias Selection bias.
Bronchopulmonary dysplasia (BPD)	Oxygen dependency at a corrected age (i.e post menstrual age) of 36 weeks. This term is often used interchangeably with ‘Chronic Lung Disease’.
Case-control study	A study to find out the cause(s) of a disease or condition. This is done by comparing a group of patients who have the disease or condition (cases) with a group of people who do not have it (controls) but who are otherwise as similar as possible (in characteristics thought to be unrelated to the causes of the disease or condition). This means the researcher can look for aspects of their lives that differ to see if they may cause the condition. Such studies are retrospective because they look back in time from the outcome to the possible causes of a disease or condition.
Cerebral palsy (CP)	Cerebral palsy is a disorder of the development of movement and posture due to permanent non-progressive abnormalities of the brain. Depending on the site of damage, the types of cerebral palsy may vary

Term	Definition
	and include spastic, ataxic, athetoid and mixed types of the disorder. Severity of physical disability can vary. In addition, there may be other problems such as speech and language delay, cognitive development, perceptual skills, difficulties with sensation, behaviour and feeding, eating, drinking and swallowing.
Clinician	A healthcare professional who provides patient care. For example a doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of RCTs prepared by the Cochrane Collaboration).
Cohort study	A study with 2 or more groups of people – cohorts – with similar characteristics. One group receives a treatment, or is exposed to a risk factor or has a particular symptom and the other group does not. The study follows their progress over time and records what happens.
Comorbidity	A disease or condition that someone has in addition to the health problem being studied or treated.
Composite outcome	An outcome that combines several components measured into a single measure.
Confidence interval (CI)	There is always some uncertainty in research. This is because a small group of patients is studied to predict the effects of a treatment on the wider population. The confidence interval is a way of expressing how certain we are about the findings from a study, using statistics. It gives a range of results that is likely to include the 'true' value for the population. The CI is usually stated as '95% CI', which means that the range of values has a 95 in 100 chance of including the 'true' value. For example, a study may state that "based on our sample findings, we are 95% certain that the 'true' population blood pressure is not higher than 150 and not lower than 110". In such a case the 95% CI would be 110 to 150. A wide confidence interval indicates a lack of certainty about the true effect of the test or treatment – often because a small group of patients has been studied. A narrow confidence interval indicates a more precise estimate (for example if a large number of patients have been studied).
Confounding factor	Something that influences a study and can result in misleading findings if it is not understood or appropriately dealt with. For example, a study of heart disease may look at a group of people who exercise regularly and a group who do not exercise. If the ages of the people in the 2 groups are different, then any difference in heart disease rates between the 2 groups could be because of age rather than exercise. Therefore age is a confounding factor.
Continuous outcome	Data with a potentially infinite number of possible values within a given range. Height, weight and blood pressure are examples of continuous variables.
Corrected age (CA)	In children born preterm, age used for the first 2 years when assessing their functional and developmental skills (such as walking and talking). Calculated from their original due date (and not the date they were born).
Cost–benefit analysis (CBA)	Cost-benefit analysis is one of the tools used to carry out an economic evaluation. The costs and benefits are measured using the same monetary units (for example UK pounds) to see whether the benefits exceed the costs.
Cost–consequence analysis (CCA)	Cost-consequence analysis is one of the tools used to carry out an economic evaluation. This compares the costs (such as treatment and hospital care) with the consequences (such as health outcomes) of a test or treatment with a suitable alternative. Unlike cost–benefit analysis

Term	Definition
	or cost-effectiveness analysis, it does not attempt to summarise outcomes in a single measure (such as the quality adjusted life year) or in financial terms. Instead, outcomes are shown in their natural units (some of which may be monetary) and it is left to decision-makers to determine whether, overall, the treatment is worth carrying out.
Cost-effectiveness analysis (CEA)	Cost-effectiveness analysis is one of the tools used to carry out an economic evaluation. The benefits are expressed in non-monetary terms related to health, such as symptom-free days, heart attacks avoided, deaths avoided or life years gained (that is, the number of years by which life is extended as a result of the intervention).
Cost-effectiveness model	An explicit mathematical framework which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost–utility analysis (CUA)	Cost–utility analysis is one of the tools used to carry out an economic evaluation. The benefits are assessed in terms of both quality and duration of life, and expressed as quality adjusted life years (QALYs). See also Utility.
Delayed motor milestones	This term refers to the age by which most children have acquired skills of, for example, sitting alone and walking.
Developmental problems and disorders	A group of problems that become apparent during child development and often occur together. They are characterised by impairments of personal, social, academic or occupational functioning, ranging from very specific limitations to global impairments of social skills or cognition, as measured by parent or teacher reports and surveillance tools. The term 'disorder' applies if the condition is severe, persistent and pervasive enough to meet the criteria for a disorder in the International Statistical classification of diseases and related health problems (ICD) or the Diagnostic and statistical manual of mental disorders (DSM).
Developmental coordination disorder (DCD)	Developmental coordination disorder (DCD) is characterized by difficulties in acquiring and executing coordination skills resulting in impairment of activities of daily living
Dichotomous outcomes	Outcome that can take one of 2 possible values, such as dead/alive, smoker/non-smoker, present/not present (also called binary data).
Disability	Impairment, activity limitations and participation restrictions affecting the individual.
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.
Early years foundation stage	The foundation stage of education begins when children reach the age of three years. Many children attend an early education setting soon after their third birthday. The foundation stage continues until the end of the reception year and is consistent with the National Curriculum. It prepares children for learning in Year 1, when programmes of study for Key Stage 1 are taught.
Early years provider	A provider of early education places for children under five years of age. This can include state-funded and private nurseries as well as child minders.
Economic evaluation	An economic evaluation is used to assess the cost effectiveness of healthcare interventions (that is, to compare the costs and benefits of a healthcare intervention to assess whether it is worth doing). The aim of an economic evaluation is to maximise the level of benefits – health effects – relative to the resources available. It should be used to inform

Term	Definition
	and support the decision-making process; it is not supposed to replace the judgement of healthcare professionals. There are several types of economic evaluation: cost–benefit analysis, cost–consequence analysis, cost-effectiveness analysis, cost-minimisation analysis and cost–utility analysis. They use similar methods to define and evaluate costs, but differ in the way they estimate the benefits of a particular drug, programme or intervention.
Education, Health and Care plan (EHC plan):	An EHC plan details the education, health and social care support that is to be provided to a child or young person who has special educational needs (SEN) or a disability. It is drawn up by the local authority after a needs assessment of the child or young person has determined that an EHC plan is necessary, and after consultation with relevant partner agencies.
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	A measure that shows the magnitude of the outcome in 1 group compared with that in a control group. For example, if the absolute risk reduction is shown to be 5% and it is the outcome of interest, the effect size is 5%. The effect size is usually tested, using statistics, to find out how likely it is that the effect is a result of the treatment and has not just happened by chance.
Emotional problems	Symptoms of anxiety, phobias or moodiness that are inappropriate or excessive for the child's age as rated by parents or teachers using standardised questionnaires or rating scales.
Enhanced developmental surveillance programme	A program in which developmental assessment is monitored actively at set times and using specific tools with the intention of detecting developmental problems and disorders.
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (for example infection, diet) and interventions.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including research studies and expert opinion (of clinical professionals or patients).
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Executive function	Executive functions are a set of inter-related cognitive processes that are used to organise and regulate thoughts and actions. These processes are important for guiding learning and behaviour, and comprise skills such as inhibition, impulse control, emotional control, working memory, cognitive flexibility and planning.
Extrapolation	An assumption that the results of studies of a specific population will also hold true for another population with similar characteristics.
False negative (FN)	A diagnostic test result that incorrectly indicates that an individual does not have the disease of interest, when they do actually have it.
False positive (FP)	A diagnostic test result that incorrectly indicates that an individual has the disease of interest, when they actually do not have it.
Feeding problems	A difficulty in physically managing to suck, use the tongue to manage semi-solids, chew solids, swallow, independently feed using mealtime utensils, difficulties emotionally tolerating certain food tastes and textures, difficulties engaging socially in the mealtime context and a reduction in mealtime communication with significant others.
Feeding support	Feeding support involves practical hands-on intervention with the families caring for the child with feeding problems. This may involve observing a mealtime, and offering practical strategies and emotional support to minimise stress, and optimise a safe and calm environment

Term	Definition
	by monitoring swallow safety, providing advice and support to manage oral intake through equipment and compensatory strategies, managing non – nutritive / oral care; ensuring the carer communication style supports the mealtime.
Fine motor skill	Fine motor skill (or dexterity) the coordination of small muscles, in movements—usually involving the synchronization of hands and fingers—with the eyes. Fine motor skills aid the growth of intelligence and develop continuously throughout the stages of human development
Follow-up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.
Forest plot	A graphical representation of the individual results of each study. The plot also allows readers to see the heterogeneity among the results of the studies. The results of individual studies are shown as squares or dots centred on each study's point estimate. A horizontal line runs through each square or dot to show each study's confidence interval.
Generalisability	The extent to which the results of a study hold true for groups that did not participate in the research.
Gestational age (GA)	Gestational age is the number of days and weeks since a mother's last menstrual period.
Global developmental delay	A general delay in development or learning, or failure to reach age-appropriate milestones in more than 2 domains of gross motor and fine motor, speech and language, cognition, personal and social development, or activities of daily living during the first four years of life, as rated by parents using validated questionnaires or checklists.
Gold standard	A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease.
GRADE, GRADE profile	A system developed by the GRADE Working Group to address the short-comings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Gross and fine motor delay	A delay in attaining the gross and fine motor performance typically associated with a particular age group. The extent of the delays can be determined by administering standardised assessment tools which will generate summary scores for gross and fine motor development in relation to a norm-referenced or typically developing sample of children. The standardisation process also generates cut-off scores which are used to determine the level: mild, moderate or severe delays; typical performance or accelerated performance. Repeated or serial assessment and the use of confidence intervals are recommended to reflect the wide variation in gross and fine motor development between children.
Gross motor skill	Gross motor skill is the co-ordination and movement of the arms, legs, and other large body parts and result in actions such as sitting, running, crawling, swimming, sports such as football. The majority of gross motor proficiency develops early in life in a predictable developmental sequence but can be further refined with training throughout life.
Gross motor function classification system (GMFCS)	A 5-level classification system that describes the gross motor function of children and youth with cerebral palsy on the basis of their self-initiated movement with particular emphasis on sitting, walking, and wheeled mobility.
Hazard ratio (HR)	A hazard is the rate at which events happen, so that the probability of an event happening in a short time interval is the length of time multiplied by the hazard. Although the hazard may vary with time, the assumption

Term	Definition
	in proportional hazard models for survival analysis is that the hazard in one group is a constant proportion of the hazard in the other group. This proportion is the hazard ratio.
Hearing impairment	Reduced ability of hearing both speech and everyday sounds. Hearing loss can be conductive (middle ear difficulties), sensorineural (inner ear difficulties), or a combination of both. Many children have pre-lingual deafness which can be mild, moderate, severe or profound, and may require the support of hearing aids.
Healthy Child Programme	The Healthy Child Programme covers pregnancy and the first five years of a child's life, focusing on a universal preventative service that provides families with a programme of screening, immunisation, health and development reviews, supplemented by advice around health, wellbeing and parenting.
Health economics	Study or analysis of the cost of using and distributing healthcare resources.
Herpetic meningitis	Inflammation of the membranes that surround the brain (the meninges) caused by a herpes virus.
Heterogeneity	The term is used in meta-analyses and systematic reviews to describe when the results of a test or treatment (or estimates of its effect) differ.
Hyperactivity impulsivity	Hyperactivity impulsivity refers to periods of limited attention with associated impulsive behaviours. The combination of both inattention and impulsivity can impact on learning and development as well as the ability to complete everyday functional activities.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Inattention	Inattention is the inability to maintain attention skills. It may be due to peripheral noise or activity preventing ability to sustain attention in either a structured learning, every day or play context. Inattention can impact on language processing, receptive language, memory and cognitive skills.
Incidence	The incidence of a disease is the rate at which new cases occur in a population during a specified period.
Inclusion criteria (clinical study)	Specific criteria that define who is eligible to participate in a clinical study.
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental cost	The extra cost linked to using one test or treatment rather than another. Or the additional cost of doing a test or providing a treatment more frequently.
Indirectness	The available evidence is different to the review question being addressed, in terms of population, intervention, comparison and outcome (PICO).
Intellectual disability	Intellectual disability (intellectual developmental disorder) is characterised by deficits in general cognitive abilities (such as reasoning and abstract thinking) and impairment of adaptive function that affects several aspects of daily life. In the ICD-10 this is defined as an IQ score more than 2 standard deviations below the mean.
Internalising behaviours	A combination of mood and emotions such as anxiety.
Intervention	In medical terms this could be a drug treatment, surgical procedure, diagnostic or psychological therapy. Examples of public health interventions could include action to help someone to be physically active or to eat a more healthy diet.
Intraventricular haemorrhage (IVH)	Intraventricular haemorrhage (IVH) refers to bleeding within the brain usually diagnosed on ultrasound. It has most commonly been graded

Term	Definition
	according to the Papile classification. Grade 1 - germinal matrix haemorrhage; Grade 2 - IVH without ventricular dilatation; Grade 3 - IVH with blood distending the ventricular; Grade 4 - IVH extending into adjacent brain parenchyma, this is more accurately referred to as Periventricular venous haemorrhagic infarction (PVHI).
Key stage 1 (KS1)	The national curriculum is organised into blocks of years called 'key stages' (KS). At the end of each key stage, the teacher will formally assess a child's performance. KS1 is the block at primary school from Year 1 to Year 2 (when a child will typically be between the ages of 5 and 7 years). At the end of KS1 (Year 2) the majority of children who are able are required to take the new National tests and teacher assessments in English, maths and science (introduced in 2016).
Key stage 2 (KS2)	The national curriculum is organised into blocks of years called 'key stages' (KS). At the end of each key stage, the teacher will formally assess a child's performance. KS2 is the block at primary school from Year 3 to Year 6 (when a child will typically be between the ages of 7 and 11 years). At the end of KS2 (Year 6) the majority of children who are able are required to take the new National tests and teacher assessments in English, maths and science (introduced in 2016).
Key stage 3 (KS3)	The national curriculum is organised into blocks of years called 'key stages' (KS). At the end of each key stage, the teacher will formally assess a child's performance. KS3 is the block at the first 3 years of secondary school from Year 7 to Year 9 (when a child will typically be between ages 11 and 14 years).
Key stage 4 (KS4)	The national curriculum is organised into blocks of years called 'key stages' (KS). At the end of each key stage, the teacher will formally assess a child's performance. KS4 is the block at the last two years of compulsory education, meaning Year 10 and Year 11 (when a child will typically be between ages 14 and 16 years).
Language delay	Development of speech and language skills, but at a slower rate compared to typical development. Delay may mean that early babbling is slow to emerge, and first words develop later than usually expected. Language delay can impact on a child's emotional and communicative confidence when with other children.
Licence	See Product licence.
Likelihood ratio (LR)	The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LR+) is sensitivity divided by (1 minus specificity).
Loss to follow-up	Individuals who have withdrawn from a study or otherwise were not participating in the studyat the point of follow-up.
Low educational attainment	A vague descriptor meaning a child's progress in learning in one or more areas is lower or below than expected when compared to national expectations of same age peers.
Major cerebral lesions	Significant structural brain abnormalities or areas of damage to brain tissue seen on cranial ultrasound or magnetic resonance imaging (MRI). In the references for this guideline, cerebral (brain) lesions were classified as 'major' if they included intraventricular haemorrhage with ventricular distension (blood filling the ventricles and extending them), grade 4 is intra-parenchymal (periventricular venous infarct).
Managing feeding	Managing feeding involves communication between the multi-disciplinary team when supporting parents and carers of infants and children who have feeding difficulties. The management may involve a paediatrician overseeing the case; a speech and language therapist monitoring swallow safety, providing advice and support to manage oral intake through equipment and compensatory strategies, managing non-

Term	Definition
	nutritive/oral care; ensuring the carer communication style supports the mealtime; an occupational therapist to maximise independent feeding skills where possible; a physiotherapist to manage postural stability during the mealtime; a dietitian to monitor calorific intake and weight gain; a clinical psychologist to reduce the risks of any behaviours that impact on the mealtime dynamic.
Mean	An average value, calculated by adding all the observations and dividing by the number of observations.
Median	The value of the observation that comes half-way when the observations are ranked in order.
Meta-analysis	A method often used in systematic reviews. Results from several studies of the same test or treatment are combined to estimate the overall effect.
Motor problem	Any motor difficulty with acquiring or executing tasks requiring motor coordination described by parents or carers using a questionnaire or checklist.
Multiple pregnancy/multiple birth	A pregnancy of two or more foetuses.
Multivariate model	A statistical model for analysis of the relationship between 2 or more predictors, (independent) variables and the outcome (dependent) variable.
National audit	A systematic process for setting and monitoring standards of clinical care. Whereas ‘guidelines’ define what the best clinical practice should be, ‘audit’ investigates whether best practice is being carried out. Clinical audit can be described as a cycle or spiral. Within the cycle there are stages that follow a systematic process of establishing best practice, measuring care against specific criteria, taking action to improve care and monitoring to sustain improvement. The spiral suggests that as the process continues, each cycle aspires to a higher level of quality.
National curriculum	This sets out a clear, full and statutory entitlement to learning for all pupils, determining what should be taught and setting attainment targets for learning. It also determines how performance will be assessed and reported.
Neonatal bacterial meningitis	Inflammation of the membranes surrounding the brain (the meninges) caused by a bacterial infection. Can occur in early (<7 days) or late onset (>7 days) forms and sometimes occurs as a complication of a more generalised septicaemia.
Neonatal encephalopathy	Abnormal neurological behaviour in the neonatal period, this has a wide range of aetiologies.
Necrotising enterocolitis (NEC)	A bowel condition of multifactorial aetiology that predominantly affects preterm babies. It is characterised by inflammation of the bowel, feed intolerance and physiological instability. It is usually treated by withholding milk feeds, antibiotics and if necessary surgical intervention. Often abbreviated to NEC.
Neonatal factors	Factors that impact the baby on the neonatal unit. The neonatal period is more strictly defined as the first 28 days if life.
Neonatal hearing screening	Hearing test done prior to discharge from hospital, to help identify babies who have permanent hearing loss as early as possible, it is universal in UK.
Neonatal sepsis	Blood culture-positive sepsis that is treated with antibiotics for more than 5 days
Neurodevelopmental disorders	A group of conditions with onset in the developmental period that frequently co-occur and are characterized by impairments of personal, social, academic or occupational functioning ranging from very specific limitations of, for example, aspects of learning to global impairments of

Term	Definition
	social skills or cognition. Behaviour and emotional problems commonly co-exist.
Observational study	Individuals or groups are observed or certain factors are measured. No attempt is made to affect the outcome. For example, an observational study of a disease or treatment would allow 'nature' or usual medical care to take its course. Changes or differences in one characteristic (for example whether or not people received a specific treatment or intervention) are studied without intervening. There is a greater risk of selection bias than in experimental studies.
Odds ratio (OR)	Odds are a way to represent how likely it is that something will happen (the probability). An odds ratio compares the probability of something in one group with the probability of the same thing in another. An odds ratio of 1 between 2 groups would show that the probability of the event (for example a person developing a disease, or a treatment working) is the same for both. An odds ratio greater than 1 means the event is more likely in the first group. An odds ratio less than 1 means that the event is less likely in the first group. Sometimes probability can be compared across more than 2 groups – in this case, one of the groups is chosen as the 'reference category' and the odds ratio is calculated for each group compared with the reference category. For example, to compare the risk of dying from lung cancer for non-smokers, occasional smokers and regular smokers, non-smokers could be used as the reference category. Odds ratios would be worked out for occasional smokers compared with non-smokers and for regular smokers compared with non-smokers. See also Confidence interval, Relative risk.
Oro motor feeding problems	Oro motor feeding problems involve difficulties with functional and consistent movements of the musculature of the lips, tongue, mouth and jaw. This can be due to neurological motor or sensory planning.
Orthoptic vision screening	The vision check at 4 years of age recommended by the national screening committee.
Parent Report of Children's Abilities-Revised (PARCA-R)	Parent Report of Children's Abilities-Revised (PARCA-R) is a questionnaire used as a screening tool for assessing global developmental delay, early intellectual disability or language problems between 22 and 26 months of age.
Parenchymal lesions	Areas of damaged brain tissue seen on cranial ultrasound or magnetic resonance imaging (MRI). Parenchymal lesions may be distinct from bleeding within the ventricles as in this case brain tissue may not be damaged.
Passivity	Withdrawn behaviour.
p value	The p value is a statistical measure that indicates whether or not an effect is statistically significant. For example, if a study comparing 2 treatments found that one seems more effective than the other, the p value is the probability of obtaining these results by chance. By convention, if the p value is below 0.05 (that is, there is less than a 5% probability that the results occurred by chance) it is considered that there probably is a real difference between treatments. If the p value is 0.001 or less (less than a 1% probability that the results occurred by chance), the result is seen as highly significant. If the p value shows that there is likely to be a difference between treatments, the confidence interval describes how big the difference in effect might be.
Perinatal risk factors	Factors in the period immediately surrounding birth which are measurable and may confer risk of later problems.
Periventricular leukomalacia (PVL)	Softening of the white brain matter around the ventricles of the brain leading to a cystic or honeycomb appearance to this area of brain.
Postnatal factors	Factors that impact the baby that occur after birth.

Term	Definition
Postnatal steroids	Administration of a corticosteroid preparation to a baby after birth. This is given to a select group of ill, ventilated preterm babies, in an attempt to facilitate their extubation from the ventilator and reduce the risk of chronic lung disease (bronchopulmonary dysplasia).
Postmenstrual age	Postmenstrual age, is an infant's age in weeks from the time of the last menstrual period. It is calculated by adding the gestational age of the infant at the time of birth with the chronological age (age of baby in days/weeks/months) after birth.
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Preterm	Born before 37 weeks' gestation.
Preterm baby	Also known as premature baby. Refers to a baby born at fewer than 37 weeks' gestational age.
Prevalence	The prevalence of a disease is the proportion of a population that are cases at a point in time.
Primary care	Healthcare delivered outside hospitals. It includes a range of services provided by GPs, nurses, health visitors, midwives and other healthcare professionals and allied health professionals such as dentists, pharmacists and opticians.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.
Prospective study	A research study in which the health or other characteristic of participants is monitored (or 'followed up') for a period of time, with events recorded as they happen. This contrasts with retrospective studies.
Protocol (review)	A document written prior to commencing a review that details exactly how evidence to answer a review question will be obtained and synthesised. It defines in detail the population of interest, the interventions, the comparators/controls and the outcomes of interest (PICO).
Publication bias	Publication bias occurs when researchers publish the results of studies showing that a treatment works well and don't publish those showing it did not have any effect. If this happens, analysis of the published results will not give an accurate idea of how well the treatment works. This type of bias can be assessed by a funnel plot.
Quality of life	See Health-related quality of life.
Quality adjusted life year (QALY)	A measure of the state of health of a person or group in which the benefits, in terms of length of life, are adjusted to reflect the quality-of-life. One QALY is equal to 1 year of life in perfect health. QALYS are calculated by estimating the years of life remaining for a patient following a particular treatment or intervention and weighting each year with a quality-of-life score (on a scale of 0 to 1). It is often measured in terms of the person's ability to perform the activities of daily life, and freedom from pain and mental disturbance.
Randomised controlled trial (RCT)	A study in which a number of similar people are randomly assigned to 2 (or more) groups to test a specific drug or treatment. One group (the experimental group) receives the treatment being tested, the other (the comparison or control group) receives an alternative treatment, a dummy treatment (placebo) or no treatment at all. The groups are followed up to see how effective the experimental treatment was.

Term	Definition
	Outcomes are measured at specific times and any difference in response between the groups is assessed statistically. This method is also used to reduce bias.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Relative risk (RR)	The ratio of the risk of disease or death among those exposed to certain conditions compared with the risk for those who are not exposed to the same conditions (for example the risk of people who smoke getting lung cancer compared with the risk for people who do not smoke). If both groups face the same level of risk, the relative risk is 1. If the first group had a relative risk of 2, subjects in that group would be twice as likely to have the event happen. A relative risk of less than 1 means the outcome is less likely in the first group. Relative risk is sometimes referred to as risk ratio.
Reporting bias	See Publication bias.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.
Retinopathy of prematurity (ROP)	A condition of the eye affecting mainly premature babies, usually those who have received oxygen therapy. It is thought to be caused by disorganized retinal blood vessel growth, which can result in scarring, and in severe cases, retinal detachment and blindness. All preterm babies at risk for this are screened for ROP in England.
Retrospective study	A research study that focuses on the past and present. The study examines past exposure to suspected risk factors for the disease or condition. Unlike prospective studies, it does not cover events that occur after the study group is selected.
Review question	The plan or set of steps to be followed in a study. A protocol for a systematic review describes the rationale for the review, the objectives and the methods that will be used to locate, select and critically appraise studies, and to collect and analyse data from the included studies.
Screening tool	The method used to screen for the presence of a condition/disease in a population.
Secondary care	Care provided in hospitals.
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	Selection bias occurs if: the characteristics of the people selected for a study differ from the wider population from which they have been drawn; or there are differences between groups of participants in a study in terms of how likely they are to get better.
Sensitivity	How well a test detects the thing it is testing for. If a diagnostic test for a disease has high sensitivity, it is likely to pick up all cases of the disease in people who have it (that is, give a 'true positive' result). But if a test is too sensitive it will sometimes also give a positive result in people who don't have the disease (that is, give a 'false positive'). For example, if a test were developed to detect if a woman is 6 months pregnant, a very sensitive test would detect everyone who was 6 months pregnant but would probably also include those who are 5 and 7 months pregnant. If the same test were more specific (sometimes referred to as having higher specificity), it would detect only those who are 6 months pregnant and someone who was 5 months pregnant would get a negative result (a 'true negative'). But it would probably also miss some people who were 6 months pregnant (that is, give a 'false negative'). Breast screening is a 'real-life' example. The number of women who are recalled for a second breast screening test is relatively high because the test is very sensitive. If it were made more specific, people who don't

Term	Definition
	have the disease would be less likely to be called back for a second test but more women who have the disease would be missed.
Shunt	A drainage mechanism for relieving blockage of and pressure from cerebro-spinal fluid in the ventricles.
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p<0.05$).
Sleep apnoea	Pausing and apparently not breathing during sleep usually due to obstruction in the nasal passages.
Social isolation	Social isolation is where a child may have very little contact with others through play or communication. Specifically, this may be more marked in a school context where a child may lack confidence or competence with the language skills necessary to be able to play, learn and engage with others.
Social problems	Having immature social skills, difficulties maintaining friendships or interacting with peers, or showing signs of social withdrawal that are inappropriate or excessive for the child's age as rated by parents or teachers using standardised questionnaires or rating scales.
Speech and language problems	Speech and language disorders are characterized by speech, language understanding or expression markedly below that expected for age resulting in limitations in communication, social participation or academic achievement.
Special educational needs and disability (SEND)	Often also called special educational needs (SEN). A child or a young person has SEND if they have a learning difficulty or disability which calls for special educational provision to be made for him or her. A child or young person has a learning difficulty or disability if he or she has a significantly greater difficulty in learning than the majority of others of the same age; has a disability which prevents or hinders him or her from making use of facilities of a kind generally provided for others of the same age in mainstream schools or mainstream post-16 institutions. For a child under two years of age, special educational provision means educational provision of any kind.
Specificity	The proportion of true negatives that are correctly identified as such. For example, in diagnostic testing the specificity is the proportion of non-cases correctly diagnosed as non-cases. In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers. See also Sensitivity.
Specific learning disorders	Specific learning disorders are characterized by impaired learning of academic skills, reading, writing, or maths.
Stakeholder	An organisation with an interest in a topic on which NICE is developing a clinical guideline or piece of public health guidance. Organisations that register as stakeholders can comment on the draft scope and the draft guidance. Stakeholders may be: manufacturers of drugs or equipment national patient and carer organisations NHS organisations organisations representing healthcare professionals.
Standard deviation (SD)	A measure of the spread or dispersion of a set of observations, calculated as the average difference from the mean value in the sample.
Strengths and Difficulties Questionnaire (SDQ)	The Strengths and Difficulties Questionnaire (SDQ) is a brief behavioural screening questionnaire assessing children aged 3 to 16 years. It exists in several versions to meet the needs of researchers, clinicians and educationalists.
Subgroup analysis	An analysis in which the intervention effect is evaluated in a defined subset of the participants in a trial, or in complementary subsets.

Term	Definition
Summary receiver operating characteristic (SROC)	Summary receiver operating characteristic (SROC) is a graphical plot that summarises the diagnostic accuracy of a given test by taking into consideration the trade-off between sensitivity and specificity.
Systematic review	A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. It may include a meta-analysis.
Toileting problems	Problems with toileting including problems with toilet training; awareness and refusal; irregular bowel habit; soiling; constipation; enuresis including bed wetting.
True negative	A diagnostic test result that correctly indicates that an individual does not have the disease of interest when they actually do not have it.
True positive	A diagnostic test result that correctly indicates that an individual has the disease of interest when they do actually have it.
Two year (corrected) age	Two years of age corrected is calculated from the Estimated Day Of Delivery based on 40 weeks pregnancy.
Univariate	Analysis which separately explores each variable in a data set.
Visual impairment	A visual impairment is a decreased ability to see to a degree that causes problems not correctable by usual means such as wearing glasses. It is classified in terms of severity based on best corrected distance acuity in the better eye into 3 main categories: Visual Impairment (VI), Severe Visual Impairment (SVI) and Blind (BL). Assessing precise visual acuity is difficult in babies and young children and a vision impairment may be diagnosed by a baby’s inability to visually fix and follow on objects.
Wechsler Preschool and Primary Scales of Intelligence Fourth Edition (WPPSI-IV)	A standardised assessment which measures cognitive development for pre-schoolers and young children (age range 2:6 – 7:7).

Acronyms and abbreviations

Table 65: Acronyms and abbreviations

AB	Antibiotic
AGA	Appropriate for Gestational Age
ADD	Attention Deficit Disorder
ADHD	Attention Deficit/Hyperactivity Disorder
ADI-R	Autism Diagnostic Interview-Revised
ADOS-2	Autism Diagnostic Observation Schedule 2
AGA	Appropriate for Gestational Age
AIMS	Abnormal Involuntary Movement Scale
ALSPAC	Avon Longitudinal Study of Parents and Children
ANS	Antenatal Steroids
AOR	Adjusted Odds Ratio
ASD	Attention Spectrum Disorder
ASQ	Ages and Stages Questionnaire
ASSQ	Autism Spectrum Screening Questionnaire
AUC	Area Under the Curve
BAI	Beck Anxiety Inventory
BASC	Behaviour Assessment System
Bayley	See BSID
BITSEA	Brief Infant Toddler Social Emotional Assessment
BLS	Basic Life Support
BMI	Body Mass Index
BPD	Bronchopulmonary Dysplasia
BRIEF	Behaviour Rating Inventory of Executive Function
BSF-R	Bayley Short Form Research Edition
BSID	Bayley Scales of Infant Development
BSID-I	Bayley Scales of Infant Development First Edition
BSID-II	Bayley Scales of Infant Development Second Edition
BSID-II-NL	Bayley Scales of Infant Development Second Edition Dutch version
BSID-III	Bayley Scales of Infant Development Third Edition
BSID-III-NL	Bayley Scales of Infant Development Third Edition Dutch version
BW	Birthweight
BWZ	Birthweight Z score
CA	Corrected Age
CADS-A	Conners' ADHD/DSM-IV Scale Self-Report Form
CADS-P	Conners' ADHD/DSM-IV Scale for Parents
CBCL	Child Behaviour Checklist
CCTR	Cochrane Central Register of Controlled Trials

CESD-R	Center for Epidemiologic Studies Depression Scale -Revised
CFT	Culture Fair Intelligence Tests
ChIPS	Children's Interview for Psychiatric Syndromes
CI	Confidence Interval
CLD	Chronic Lung Disease
CP	Cerebral Palsy
CRIB-score	Clinical Risk Index for Babies Scoring System
CSSA	Comprehensive Scales of Student Abilities
d	Days
DCD	Developmental Coordination Disorder
DCDQ	Developmental Coordination Disorder Questionnaire
DAS	Differential Ability Scale
DAWBA	Development and Well Being Assessment
D-KEFS	Delis-Kaplan Executive Function Scale
DQ	Developmental Quotient
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders
DSRS	Depression Self-Rating Scale
EFCS	Executive Function Composite Score
ELGAN	Extremely Low Gestational Age Newborns (study)
EPIBEL	Extremely Preterm Infants in Belgium Study
EPIPAGE	French Etude Epidemiologique Sur Les Petits Ages Gestationnels (study)
EPT	Extremely Preterm
ES	Estimate
EXPRESS	Extremely Preterm Infants Study in Sweden
FN	False Negative
FP	False Positive
FSIQ	Full Scale Intelligence Quotient
FSP	Foundation Stage Profile
FTF	Five to Fifteen Questionnaire
GA	Gestational Age
GCSE	General Certificates of Secondary Education
GMDS	Griffiths Scales of Mental Development
GMFCS	Gross Motor Function Classification System
GMFM	Gross Motor Function Measure
GRADE	Grading of Recommendations Assessment, Development and Evaluation
H	High (quality)
Hh	Household
HAWIK	Hamburg Wechsler Intelligence Test for Children
HDR	Hospital Discharge Register
HELLP	Haemolysis. Elevated Liver Enzymes. Low Platelet Count

HINE	Hammersmith Infant Neurological Examination
HTA	Health Technology Assessments
ICD	International Statistical Classification of Diseases and Related Health Problems
ICDS	Infant and child Development services
ICH	Intracerebral Haemorrage
IFSP	Individualised family service plan
IQ	Intelligence Quotient
IUGR	Intrauterine Growth Restriction
IVH	Intraventricular Haemorrhage
K-ABC	Kaufman Assessment Battery for Children
KS1-4	Key Stage 1 to 4
KSPD	Kyoto Scale of Psychological Development
L	Low (quality)
LAMBS	Late to Moderately Preterm Birth Study
LGA	Large for Gestational Age
LMPT	Late and Moderately Preterm
LR+	Positive Likelihood Ratio
LR-	Negative Likelihood Ratio
M	Moderate (quality)
MABC	Movement Assessment Battery for Children
MAP	Miller Assessment for Preschoolers
M-CHAT	Modified Checklist for Autism in Toddlers
MCS	Millennium Cohort Study
MDI	Mental Development Index
MDT	Multidiciplinary team
MGA	Mean Gestational Age
MLBW	Moderately Low Birth Weight
MND	Minor Neuromotor Dysfunction
mo	Months
MPC	Mental Processing Composite
MRI	Magnetic Resonance Imaging
NBW	Normal Birth Weight
NDI	Neurodevelopmental Impairment
NEC	Necrotising Enterocolitis
NECCPS	North of England Collaborative Cerebral Palsy Survey
NEPSY	Developmental Neuropsychological Assessment
NGA	National Guideline Alliance
NHS	National Health Service
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
NICHD	National Institute of Child Health and Human Development
NICU	Neonatal Intensive Care Unit
NR	Not Reported

NRN	Neonatal Research Network
NSC	National screening Committee
NSW	New South Wales
OCD	Obsessive Compulsive Disorder
ONS	Office for National Statistics
OR	Odds Ratio
OWLS	Oral and Written Language Scales
PAPA	Preschool Age Psychiatric Assessment
PARCA-R	Parent Report of Children's Abilities-Revised
PDI	Psychomotor Development Index
perc	Percentile
PIVH	Periventricular-Intraventricular Haemorrhage
PLASC	Pupil Level Annual School Census
PLS-3	Preschool Language Scale-3
PNS	Postnatal Steroids
PPVT-R	Peabody Picture Vocabulary Test-Revised
PRC	Parent Report Composite
PRIDE	Maryland's Premature Infant Developmental Enrichment
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines
PROM	Premature Rupture of Membranes
PTSD	Post-Traumatic Stress Disorder
PVL	Periventricular Leukomalacia
QALYs	Quality adjusted life years
RDS	Respiratory Distress Syndrome
ref	Reference
RF	Risk Factor
ROP	Retinopathy of Prematurity
RR	Relative Risk
SCARED	Screen for Child Anxiety Related Emotional Disorders
SCPE	Surveillance of Cerebral Palsy in Europe
SCID	Severe Combined Immune Deficiency
SD	Standard Deviation
SDQ	Strengths and Difficulties Questionnaire
SGA	Small for Gestational Age
SEN	Special Educational Needs
SEND	Special Educational Needs and Disabilities
sens	Sensitivity
SES	Socioeconomic Status
SNAP-II	Score of Neonatal Acute Physiology-II
SON-R	Snijders-Oomen Nonverbal Intelligence Test
spec	Specificity
SROC	Summary Receiver Operating Characteristic
TEA-Ch	Test of Everyday Attention for Children

TN	True Negative
TOVA	Test of Variables of Attention
TP	True Positive
TRF	Teacher Report Form
VL	Very low (Quality)
VLBW	Very Low Birth Weight
vs	Versus
WHO	World Health Organization
WIAT-II	Wechsler Individual Achievement Test Second Edition
WIAT-II	Wechsler Individual Achievement Test Third Edition
WISC	Wechsler Intelligence Scale for Children
WISC-III	Wechsler Intelligence Scale for Children Third Edition
WISC-IV	Wechsler Intelligence Scale for Children Fourth Edition
weeks’	Weeks
WPPSI	Wechsler Preschool and Primary Scale of Intelligence
WPPSI-R	Wechsler Preschool and Primary Scale of Intelligence Revised
WRAT-3	Wide-Range Achievement Test 3
y	Years

Appendices

Appendix A: Scope

Appendix B: Stakeholders

Appendix C: Declarations of interest

Appendix D: Review protocols

Appendix E: Search strategies

Appendix F: Prisma flow charts

Appendix G: Excluded studies

Appendix H: Health economic analysis on identification of problems and disorders

Appendix I: Resource impact analysis of delivery of enhanced support and surveillance

Appendix J: Forest plots and receiver operating curves

Appendix K: Evidence tables

Appendix L: Supplementary tables