NIHR | National Institute for Health Research

Journals Library

Check for updates

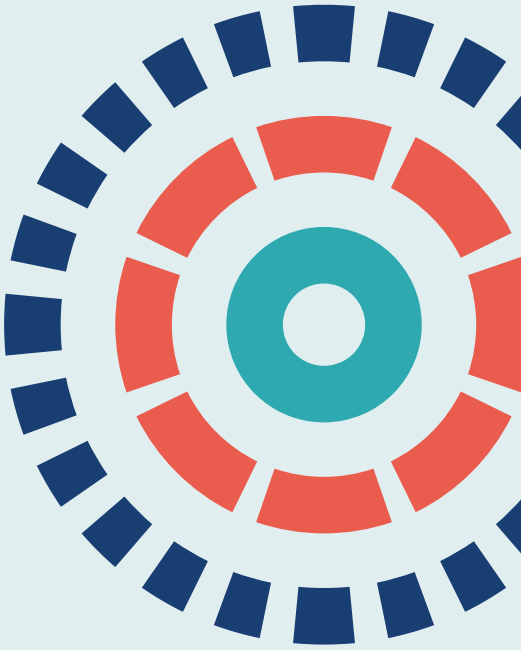
Programme Grants for Applied Research

Volume 7 • Issue 6 • September 2019

ISSN 2050-4322

Developing routinely recorded clinical data from electronic patient records as a national resource to improve neonatal health care: the Medicines for Neonates research programme

Neena Modi, Deborah Ashby, Cheryl Battersby, Peter Brocklehurst, Zoe Chivers, Kate Costeloe, Elizabeth S Draper, Victoria Foster, Jacquie Kemp, Azeem Majeed, Joanna Murray, Stavros Petrou, Katherine Rogers, Shalini Santhakumaran, Sonia Saxena, Yevgeniy Statnikov, Hilary Wong and Alys Young



DOI 10.3310/pgfar07060

Developing routinely recorded clinical data from electronic patient records as a national resource to improve neonatal health care: the Medicines for Neonates research programme

Neena Modi,[1]* Deborah Ashby,[2] Cheryl Battersby,[1] Peter Brocklehurst,[3] Zoe Chivers,[4] Kate Costeloe,[5] Elizabeth S Draper,[6] Victoria Foster,[7] Jacquie Kemp,[8] Azeem Majeed,[9] Joanna Murray,[10] Stavros Petrou,[11] Katherine Rogers,[12] Shalini Santhakumaran,[2] Sonia Saxena,[9] Yevgeniy Statnikov,[13] Hilary Wong[14] and Alys Young[12]

[1]Department of Medicine, Imperial College London, London, UK
[2]Imperial Clinical Trials Unit, Imperial College London, London, UK
[3]Birmingham Clinical Trials Unit, Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[4]Bliss, London, UK
[5]Centre for Genomics and Child Health, Queen Mary University of London, London, UK
[6]Department of Health Sciences, University of Leicester, Leicester, UK
[7]Department of Social Sciences, Edge Hill University, Ormskirk, UK
[8]National Programme of Care, NHS England, London, UK
[9]School of Public Health, Imperial College London, London, UK
[10]Development Media International, London, UK
[11]Division of Health Sciences, University of Warwick, Coventry, UK
[12]School of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK
[13]Society and College of Radiographers, London, UK
[14]Department of Paediatrics, University of Cambridge, Cambridge, UK

*Corresponding author

Declared competing interests of authors: Sonia Saxena was funded by a National Institute for Health Research (NIHR) Career Development Fellowship (CDF) (NIHR-CDF-2011-04-048). CNAM is funded by the Higher Education Funding Council for England and the NIHR Collaborations for Leadership in Applied Health Research and Care (CLAHRC) programme. Kate Costeloe reports that she was a member of the Neonatal Data Analysis Unit Board throughout this research programme. No authors have any financial or business relationships with Clevermed Ltd.

Published September 2019
DOI: 10.3310/pgfar07060

This report should be referenced as follows:

Modi N, Ashby D, Battersby C, Brocklehurst P, Chivers Z, Costeloe K, *et al*. Developing routinely recorded clinical data from electronic patient records as a national resource to improve neonatal health care: the Medicines for Neonates research programme. *Programme Grants Appl Res* 2019;**7**(6).

Programme Grants for Applied Research

ISSN 2050-4322 (Print)

ISSN 2050-4330 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full PGfAR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/pgfar. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Programme Grants for Applied Research* journal
Reports are published in *Programme Grants for Applied Research* (PGfAR) if (1) they have resulted from work for the PGfAR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Programme Grants for Applied Research programme

The Programme Grants for Applied Research (PGfAR) programme, part of the National Institute for Health Research (NIHR), was set up in 2006 to produce independent research findings that will have practical application for the benefit of patients and the NHS in the relatively near future. The Programme is managed by the NIHR Central Commissioning Facility (CCF) with strategic input from the Programme Director.

The programme is a national response mode funding scheme that aims to provide evidence to improve health outcomes in England through promotion of health, prevention of ill health, and optimal disease management (including safety and quality), with particular emphasis on conditions causing significant disease burden.

For more information about the PGfAR programme please visit the website: http://www.nihr.ac.uk/funding/programme-grants-for-applied-research.htm

This report

The research reported in this issue of the journal was funded by PGfAR as project number RP-PG-0707-10010. The contractual start date was in January 2009. The final report began editorial review in February 2017 and was accepted for publication in December 2018. As the funder, the PGfAR programme agreed the research questions and study designs in advance with the investigators. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PGfAR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

Developing routinely recorded clinical data from electronic patient records as a national resource to improve neonatal health care: the Medicines for Neonates research programme

Neena Modi,[1]* Deborah Ashby,[2] Cheryl Battersby,[1] Peter Brocklehurst,[3] Zoe Chivers,[4] Kate Costeloe,[5] Elizabeth S Draper,[6] Victoria Foster,[7] Jacquie Kemp,[8] Azeem Majeed,[9] Joanna Murray,[10] Stavros Petrou,[11] Katherine Rogers,[12] Shalini Santhakumaran,[2] Sonia Saxena,[9] Yevgeniy Statnikov,[13] Hilary Wong[14] and Alys Young[12]

[1]Department of Medicine, Imperial College London, London, UK
[2]Imperial Clinical Trials Unit, Imperial College London, London, UK
[3]Birmingham Clinical Trials Unit, Institute of Applied Health Research, University of Birmingham, Birmingham, UK
[4]Bliss, London, UK
[5]Centre for Genomics and Child Health, Queen Mary University of London, London, UK
[6]Department of Health Sciences, University of Leicester, Leicester, UK
[7]Department of Social Sciences, Edge Hill University, Ormskirk, UK
[8]National Programme of Care, NHS England, London, UK
[9]School of Public Health, Imperial College London, London, UK
[10]Development Media International, London, UK
[11]Division of Health Sciences, University of Warwick, Coventry, UK
[12]School of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK
[13]Society and College of Radiographers, London, UK
[14]Department of Paediatrics, University of Cambridge, Cambridge, UK

*Corresponding author n.modi@imperial.ac.uk

Background: Clinical data offer the potential to advance patient care. Neonatal specialised care is a high-cost NHS service received by approximately 80,000 newborn infants each year.

Objectives: (1) To develop the use of routinely recorded operational clinical data from electronic patient records (EPRs), secure national coverage, evaluate and improve the quality of clinical data, and develop their use as a national resource to improve neonatal health care and outcomes. To test the hypotheses that (2) clinical and research data are of comparable quality, (3) routine NHS clinical assessment at the age of 2 years reliably identifies children with neurodevelopmental impairment and (4) trial-based economic evaluations of neonatal interventions can be reliably conducted using clinical data. (5) To test methods to link NHS data sets and (6) to evaluate parent views of personal data in research.

Design: Six inter-related workstreams; quarterly extractions of predefined data from neonatal EPRs; and approvals from the National Research Ethics Service, Health Research Authority Confidentiality Advisory Group, Caldicott Guardians and lead neonatal clinicians of participating NHS trusts.

Setting: NHS neonatal units.

Participants: Neonatal clinical teams; parents of babies admitted to NHS neonatal units.

Interventions: In workstream 3, we employed the Bayley-III scales to evaluate neurodevelopmental status and the Quantitative Checklist of Autism in Toddlers (Q-CHAT) to evaluate social communication skills. In workstream 6, we recruited parents with previous experience of a child in neonatal care to assist in the design of a questionnaire directed at the parents of infants admitted to neonatal units.

Data sources: Data were extracted from the EPR of admissions to NHS neonatal units.

Main outcome measures: We created a National Neonatal Research Database (NNRD) containing a defined extract from real-time, point-of-care, clinician-entered EPRs from all NHS neonatal units in England, Wales and Scotland ($n = 200$), established a UK Neonatal Collaborative of all NHS trusts providing neonatal specialised care, and created a new NHS information standard: the Neonatal Data Set (ISB 1595) (see http://webarchive.nationalarchives.gov.uk/±/http://www.isb.nhs.uk/documents/isb-1595/amd-32–2012/index_html; accessed 25 June 2018).

Results: We found low discordance between clinical (NNRD) and research data for most important infant and maternal characteristics, and higher prevalence of clinical outcomes. Compared with research assessments, NHS clinical assessment at the age of 2 years has lower sensitivity but higher specificity for identifying children with neurodevelopmental impairment. Completeness and quality are higher for clinical than for administrative NHS data; linkage is feasible and substantially enhances data quality and scope. The majority of hospital resource inputs for economic evaluations of neonatal interventions can be extracted reliably from the NNRD. In general, there is strong parent support for sharing routine clinical data for research purposes.

Limitations: We were only able to include data from all English neonatal units from 2012 onwards and conduct only limited cross validation of NNRD data directly against data in paper case notes. We were unable to conduct qualitative analyses of parent perspectives. We were also only able to assess the utility of trial-based economic evaluations of neonatal interventions using a single trial. We suggest that results should be validated against other trials.

Conclusions: We show that it is possible to obtain research-standard data from neonatal EPRs, and achieve complete population coverage, but we highlight the importance of implementing systematic examination of NHS data quality and completeness and testing methods to improve these measures. Currently available EPR data do not enable ascertainment of neurodevelopmental outcomes reliably in very preterm infants. Measures to maintain high quality and completeness of clinical and administrative data are important health service goals. As parent support for sharing clinical data for research is underpinned by strong altruistic motivation, improving wider public understanding of benefits may enhance informed decision-making.

Future work: We aim to implement a new paradigm for newborn health care in which continuous incremental improvement is achieved efficiently and cost-effectively by close integration of evidence generation with clinical care through the use of high-quality EPR data. In future work, we aim to automate completeness and quality checks and make recording processes more 'user friendly' and constructed in ways that minimise the likelihood of missing or erroneous entries. The development of criteria that provide assurance that data conform to prespecified completeness and quality criteria would be an important development. The benefits of EPR data might be extended by testing their use in large pragmatic clinical trials. It would also be of value to develop methods to quality assure EPR data including involving parents, and link the NNRD to other health, social care and educational data sets to facilitate the acquisition of lifelong outcomes across multiple domains.

Study registration: This study is registered as PROSPERO CRD42015017439 (workstream 1) and PROSPERO CRD42012002168 (workstream 3).

Funding: The National Institute for Health Research Programme Grants for Applied Research programme (£1,641,471). Unrestricted donations were supplied by Abbott Laboratories (Maidenhead, UK: £35,000), Nutricia Research Foundation (Schiphol, the Netherlands: £15,000), GE Healthcare (Amersham, UK: £1000). A grant to support the use of routinely collected, standardised, electronic clinical data for audit, management and multidisciplinary feedback in neonatal medicine was received from the Department of Health and Social Care (£135,494).

Contents

List of tables

List of figures

List of boxes

List of abbreviations

ADHD	attention deficit hyperactivity disorder
ASD	autism spectrum disorder
AUC	area under the receiver operating characteristic curve
BAPM	British Association of Perinatal Medicine
Bayley-III	Bayley Scales of Infant and Toddler Development, third edition
BSID	Bayley Scales of Infant Development
BSID-II	Bayley Scales of Infant Development, second edition
CI	confidence interval
CQUIN	Commissioning for Quality and Innovation
CRF	case report form
DOR	diagnostic odds ratio
DQ	developmental quotients
EPR	electronic patient record
ESS	effective sample size
GBP	Great British pounds
GMDS	Griffiths Mental Development Scales
GMFCS	Gross Motor Function Classification System
HES	Hospital Episode Statistics
HINE	Hammersmith Infant Neurological Examination
HSROC	hierarchical summary receiver operator characteristic curve
ICD-9	*International Classification of Diseases*, Ninth Edition
ICD-10	*International Classification of Diseases*, Tenth Edition
IMD	Index of Multiple Deprivation
IQ	intelligence quotient
IQR	interquartile range
LSOA	lower-layer super output area
M-CHAT	Modified Checklist for Autism in Toddlers
MDI	Mental Development Index
NDAU	Neonatal Data Analysis Unit
NEC	necrotising enterocolitis
NICHD	National Institute of Child Health and Human Development
NNRD	National Neonatal Research Database
NPEU	National Perinatal Epidemiology Unit
ODN	Operation Delivery Network
ONS	Office for National Statistics
OR	odds ratio
PARCA-R	Parent Report of Children's Abilities-Revised
PiPS	Probiotic in Preterm infants Study
PPI	patient and public involvement
PPV	positive predictive values
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
Q-CHAT	Quantitative Checklist for Autism in Toddlers
QUADAS-2	Quality of Diagnostic Accuracy Studies version 2
ROC	receiver operating characteristic
ROP	retinopathy of prematurity
RSV	respiratory syncytial virus
SD	standard deviation
SDS	standard deviation score
TN	true negative
TP	true positive
VLBW	very low birthweight
VP	ventriculoperitoneal
WHO	World Health Organization

Plain English summary

Increasingly, health-care professionals record data in electronic patient records (EPRs) rather than traditional paper case notes. EPRs are a rich source of data with great potential to improve patient care, services and outcomes. We aimed to develop the use of EPRs to support neonatal specialised care, a high-cost NHS service for approximately 80,000 newborn infants each year.

We carried out six inter-related workstreams. We pooled data from newborn EPRs across all 200 NHS neonatal units and developed their use as a national resource. We tested the use of EPR data in research and health service evaluations. We assessed the reliability of EPR data for evaluating development in preterm babies at the age of 2 years. We compared EPR data against the same data recorded as part of a clinical research trial, and determined if we could link EPR data successfully with NHS administrative data. In a specific workstream, we obtained parent views on using routine clinical EPR data in research.

We show that it is possible for a clearly defined extract of EPR data to be stored in a National Neonatal Research Database as a resource for multiple purposes. We found that data from EPRs do not provide a reliable assessment of development at the age of 2 years in children who were born very preterm. Routine EPR clinical data show reasonable agreement with the same data recorded as part of a clinical research trial, and the data are higher in quality than similar data recorded for administrative purposes. We were able to link around two-thirds of EPR data with NHS administrative data. We found that in general there is strong parent support for sharing routine health data for research purposes.

Scientific summary

Background

Data obtained from electronic patient records (EPRs) have potential to advance patient care and to improve health services. Although this is an acknowledged national goal, problems in realising this aspiration have involved difficulties in data extraction, population coverage, regulations around holding identifying information, uncertain data quality and patient trust.

Approximately 80,000 newborn infants are admitted annually for neonatal specialised care, a high-cost NHS service. Three circumstances placed neonatal specialised services in a favourable position to realise the potential of clinical data: (1) a strong professional desire to develop a standardised Neonatal Data Set, evidenced by a series of working groups of the British Association of Perinatal Medicine dating from the 1990s to the present; (2) a specialist commercial EPR supplier working closely with clinicians; and (3) a national reorganisation of NHS neonatal services over 2003–12 into managed clinical networks and the consequent frequent transfer of infants in accordance with their clinical needs to neonatal units providing different levels of care, which provided impetus to share clinical data.

Members of the Medicines for Neonates research group were closely involved in these initiatives and developed this proposal with the aim of utilising point-of-care, clinician-entered EPR-derived clinical data to improve newborn care and services. The Medicines for Neonates applied research programme is based on the principle that information should be recorded once and not repeatedly, recorded to a high standard, and made available to support multiple outputs.

Objectives

We conducted six inter-related workstreams to:

1. secure agreement for the use of EPR data as a national resource, evaluate and improve neonatal EPR data quality, and develop and test their utility to support multiple outputs
2. test the hypothesis that EPR data are of comparable quality to research data
3. test the hypothesis that neurodevelopmental assessment at the age of 2 years, conducted during routine NHS follow-up and recorded in EPRs, can reliably identify children with neurodevelopmental impairments
4. test the hypothesis that trial-based and other economic evaluations of perinatal interventions can be reliably conducted using EPR data
5. develop and test methods to link EPRS with other NHS data sets
6. involve parents in evaluating parent views of the use of EPR data in research.

We extended our original proposal in two workstreams. In workstream 1, we additionally conducted a systematic review of databases holding data on infants admitted to neonatal units, utilised EPR data to conduct national surveillance of severe necrotising enterocolitis (NEC) (a feared gastrointestinal inflammatory disease predominantly affecting preterm neonates), and tested the use of EPR data in supporting clinical services by evaluating mortality. In workstream 3, we additionally assessed the social communication skills at the age of 2 years of very preterm children using a parent-completed questionnaire, and conducted a systematic review and meta-analysis to determine the sensitivity and specificity of early developmental assessment in identifying school-age cognitive deficits. *Figure a* illustrates the relationships between workstreams and chapters.

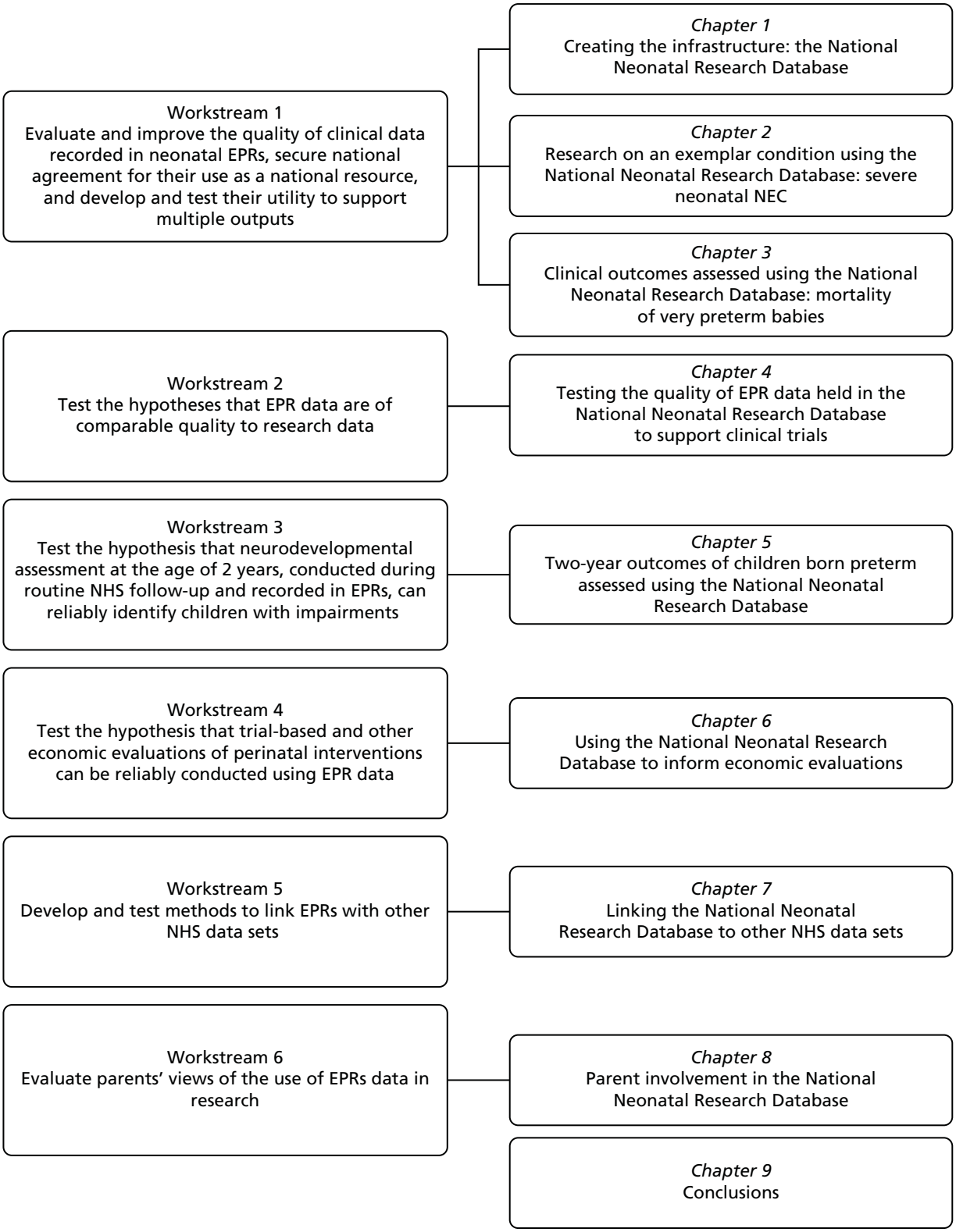


FIGURE a Relationships between medicines for neonates workstreams and chapters.

Methods

The Medicines for Neonates research programme ran for the period 1 July 2009 to 31 March 2015. In workstream 1, we obtained regulatory approvals to receive quarterly extracts of predefined data from neonatal EPRS through collaborative arrangements with the commercial supplier, the NHS trust hosting the Medicines for Neonates programme and the University research sponsor. We completed multiple application stages and public consultations leading to the submission of the defined data set for approval by the Health and Social Care Information Centre (now known as NHS Digital) as an NHS data standard. We developed algorithms and standard operating procedures for data management. We conducted a series of evaluations and addressed our research hypotheses using EPR data. We carried out a literature search of existing neonatal databases covering the period 1 January 2000 to 15 March 2015. We identified all cases of NEC requiring surgery or resulting in death over the complete 2-year period 2012–13 and we assessed variation in incidence across neonatal networks in England.

In workstream 2, we assessed the quality (completeness and accuracy) of EPR data in comparison with demographic, process and outcome variables obtained as part of a Health Technology Assessment-funded multicentre randomised clinical trial [i.e. Probiotic in Preterm infants Study (PiPS)]; Costeloe KL, Bowler U, Brocklehurst P, Hardy P, Heal P, Juszczak E, *et al.* A randomised controlled trial of the probiotic Bifidobacterium breve BBG-001 in preterm babies to prevent sepsis, necrotising enterocolitis and death: the Probiotics in Preterm infantS (PiPS) trial. *Health Technol Assess* 2016;**20**(66)].

In workstream 3, we employed a standard assessment tool [Bayley Scales of Infant and Toddler Development, third edition (Bayley-III scales)] to evaluate the neurodevelopmental status at the age of 2 years of children who were born before 32 weeks' gestation, and compared this with categorisation derived from EPR data recorded in the course of routine NHS follow-up. We evaluated the children's social communication skills as measured on a parent-completed questionnaire [Quantitative Checklist of Autism in Toddlers (Q-CHAT)] against normative data. We conducted a search on MEDLINE through the PubMed interface covering English-language literature published between 1 January 1990 and 31 March 2012 to determine the predictive validity of early developmental assessment in identifying cognitive deficit at school age.

In workstream 4, we compared health-care resource utilisation for infants recruited to the PiPS trial using three data sources: the PiPS trial case report forms, EPR-derived data and a combination of information from these two sources. Resource inputs captured by each data source were primarily valued using national tariffs and expressed in GBP (2012/13 prices). We estimated the level of agreement between the data sources and the level of precision of incremental cost-effectiveness for the probiotic evaluated in PiPS. For comparisons within trial by data source, differences in resource use and costs were tested using the independent sample *t*-test for continuous variables, the chi-squared test for categorical variables and the Mann–Whitney *U*-test for medians. For comparisons between the data sources, the levels of agreement in resource use and cost estimates for alternative combinations of data sources were estimated using the Lin concordance correlation coefficient.

In workstream 5, we obtained approval to receive NHS numbers and infant identifiers from the Confidentiality Advisory Group of the Health Research Authority and requested permission from all NHS trusts in England to receive these as extracts from their EPR data. We obtained Hospital Episode Statistics (HES) data from the Health and Social Care Information Centre (now NHS Digital) (https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics). We conducted two studies, utilising HES data covering the financial years 2005/6–2009/10, and HES and EPR-derived data for the calendar year 2010. We identified all individual birth episodes in HES, examined the completeness of HES recording and compared the total number of births with Office for National Statistics birth registrations. We used a deterministic approach to link the NNRD and HES records using the NHS number as a common unique identifier. We created a birth cohort of all infants born in English NHS hospitals and discharged during the period 1 April 2007 to 31 March 2008.

In workstream 6, we undertook a review of literature concerning public understanding of health data use for research purposes and contemporary e-health policy, identified relevant parameters and involved parents with previous experience of a child in neonatal care to assist in the design of a questionnaire directed at the parents of infants admitted to neonatal units. Materials were made available in eight languages in addition to English. These parents informed the research team of key questions regarding the routine use of babies' clinical data for research purposes and, thus, contributed to the content of the questionnaire. We recruited 29 NHS hospitals in England with neonatal care units as research sites. Research nurses approached parents to explain the study, provide written information and obtain consent.

Patient and public involvement (PPI) was a component of workstreams 1, 5 and 6.

Results

In workstream 1, we created a National Neonatal Research Database (NNRD) containing a defined extract from real-time, point-of-care, clinician-entered EPR NHS neonatal units and made this available for a variety of outputs. We achieved incremental coverage with data from 90% of English NHS neonatal units in 2010, 100% of English NHS neonatal units from 2012 onwards, and neonatal units in Wales from 2012 onwards. Scottish neonatal units joined in 2016.

We established a UK Neonatal Collaborative comprising all NHS Trusts providing neonatal specialised care, each of which provided Caldicott Guardian and Lead Neonatal Clinician approval for their data to be held in the NNRD. We created a new NHS Information standard, the Neonatal Data Set (ISB1595), comprising the predefined data held in the NNRD. We showed incidence of severe NEC to range from 7.55 per 1000 admissions [95% confidence interval (CI) 4.94 to 11.55] to 1.70 (95% CI 0.80 to 3.61), with no strong evidence of variation by network from the national average.

In workstream 2, we set out to test the hypothesis that EPR data are of comparable quality to research data; our results demonstrate that following data cleaning and merging, most key data items derived from EPR systems are of comparable quality to research data. We found completeness of data in the NNRD to be generally good. We assessed 2257 episodes of care from 1258 infants. Major discordance rates were low for 14 out of 15 patient characteristics, 9 out of 12 process measures and 10 out of 11 outcomes. The prevalence of adverse outcomes was < 6% with the exception of bronchopulmonary dysplasia (49.0%) and medical treatment for patent ductus arteriosus (20.3%). Specificity was > 85% for all outcomes, with the majority being > 90%. Specificity was high (> 85%) for all outcomes, sensitivity ranged from 50% to 100%, positive predictive values (PPV) ranged from 58.8 (95% CI 40.7 to 75.4) for a report of a porencephalic cyst to 99.7 (95% CI 99.2 to 99.9) for survival to discharge. Patient characteristics and the majority of NNRD items tested compare well against case report form (CRF) data. A small number of important outcomes are not currently reliably recorded in the EPRs.

In workstream 3, we recruited 190 children. We set out to test the hypothesis that neurodevelopmental assessment at the age of 2 years, conducted during routine NHS follow-up and the results of which are recorded in EPRs, can reliably identify children with neurodevelopmental impairments. The results demonstrate that neurodevelopmental assessment conducted during NHS follow-up has low sensitivity but high specificity for identifying children with neurodevelopmental impairments. Clinical neurodevelopmental data underestimated population prevalence of impairments following preterm birth by between 30% and 50%. We assessed the social communication skills of 141 very preterm children and found that they displayed greater social communication difficulties and autistic spectrum behaviours at 2 years than the general population. The systematic literature review revealed that neurodevelopmental assessment at approximately 2 years has low sensitivity but high specificity for identifying later school-age cognitive deficits.

In workstream 4, we set out to test the hypothesis that trial-based interventions and other economic evaluations of perinatal interventions can be reliably conducted using EPR data. The results revealed no statistically significant differences between NNRD data and data collected as part of randomised clinical trials for any resource input or cost category (i.e. that trial-based economic evaluations of neonatal interventions can be reliably conducted using the NNRD). When clinical trial data and NNRD data were compared, the agreement was relatively high for utilisation or cost of hospital stay by level of neonatal care, hospital transfers, retinopathy of prematurity screening and treatment, and surgery. However, for post-mortem examinations and cranial ultrasound scans, agreement fell below an acceptable threshold. The bulk of hospital resource inputs incorporated into a rigorously designed economic evaluation of a neonatal intervention in a UK context can be successfully and accurately extracted from the NNRD. We suggest that these results should be validated against other trials. Comparisons of cost-effectiveness outcomes between the NNRD and clinical trial data sources revealed low probability levels of miscoverage of incremental net monetary benefit when the NNRD acted as the sole source of resource use information. However, separate sensitivity analyses revealed that probability estimates of miscoverage for incremental net monetary benefit increased for both the death and the sepsis outcomes when the NNRD acted as the sole source of resource use information and clinical outcomes.

In workstream 5, we showed that completeness and quality of NNRD data are higher than NHS administrative (HES) data. The completeness of HES birth data varies substantially between hospitals. Approximately one-fifth of babies in HES have missing gestational age data and around 1.5% have a biologically implausible birthweight. We found that 1 in 10 neonates identified in HES is represented in the NNRD. There is > 95% agreement between HES and the NNRD for key items. We achieved linkage between HES and the NNRD for 61.3% of records. Linkage enhances the quality and scope of records substantially.

In workstream 6, we showed that there is a very high level of parent support for the routine use of health data for research purposes. Overall, 70% of the 1291 respondents were in agreement that their infant's clinical data be used for research, which rose to 77% if permission was asked and nearly 80% if identifying information was removed. Attitudes are moderated by level of education, previous children who had required neonatal care, and the degree of intensity of care received by their baby.

Conclusions

We have shown that it is possible to obtain high-quality data extracts from EPRS, achieve total population coverage and make data available as a national resource to support a wide range of outputs, researchers and organisations. Parent support for sharing clinical data for research is strong and underpinned by altruistic motivation.

Implications for health care

The Medicines or Neonates programme has established proof of concept for the use of EPR-derived clinical data in a wide range of research and health service evaluations. This opens the possibility of adapting the road map that we have established for other specialty areas with potential to bring about substantial NHS savings.

This study highlights the potential limitations of clinical data, in particular the necessity for high-quality recording. Clinical data are important to patient care and safety, and utilising routine clinical data for research is a secondary purpose. Completeness and quality checks can be automated for electronic data, and recording processes can be made 'user friendly' and constructed in ways to minimise the likelihood of missing or erroneous entries, which represents a major potential advance over traditional hand-written medical case notes. The development of criteria that provide assurance that data conform to prespecified

completeness and quality criteria would be an important development. This would enhance both patient care and research, each of which contributes to improving patient outcomes. This development would be especially beneficial in relation to research involving Investigational Medicinal Products as these are subject to the most stringent regulatory processes.

We developed and currently maintain the NNRD through academic endeavour, but processes to secure the stability of EPR-derived databases as national resources and their ongoing management are uncertain. A systematic approach to delivering neurodevelopment and neurocognitive screening of very preterm children by appropriately trained health-care personnel at ages that have optimum sensitivity and specificity for the identification of impairment requires consideration and evaluation. Finally, measures to extend public understanding and improve trust in the wider uses of clinical data are likely to be required if the full potential of clinical data are to be realised.

Research recommendations

Our principal recommendations are aimed at extending the benefits of EPR data, the outcomes of the Medicines for Neonates programme and the NNRD. Unlike the EPRs, data is not received in the NNRD in real time. As they are real time, EPR data are not appropriate for service evaluations or research because they change from moment to moment and have not undergone quality assurance processes. In contrast, the NNRD contains data that have been cleaned, merged and locked down in a permanent repository. We suggest testing the use of the NNRD to facilitate the delivery of a large-scale pragmatic national clinical trial and developing and testing methods to quality assure EPR data. The latter include, but are not limited to, involving parents, directing incentives at provider organisations, assigning local lead responsibility and automating certain procedures using machine-learning approaches. We also suggest linking the NNRD to other health, social care and educational data sets to facilitate the acquisition of lifelong outcomes across multiple domains.

Study registration

This study is registered as PROSPERO CRD42015017439 (workstream 1) and PROSPERO CRD42012002168 (workstream 3).

Funding

Funding for this study was provided by Programme Grants for Applied Research programme of the National Institute for Health Research (£1,641,471). Unrestricted donations were supplied by Abbott Laboratories (Maidenhead, UK: £35,000), Nutricia Research Foundation (Schiphol, The Netherlands: £15,000), GE Healthcare (Amersham, UK: £1000). A grant to support the use of routinely collected, standardised, electronic clinical data for audit, management and multidisciplinary feedback in neonatal medicine was received from the Department of Health and Social Care (£135,494).

Chapter 1 Creating the infrastructure: the National Neonatal Research Database

Abstract

Background: Successive UK governments have highlighted the potential of clinical data to advance patient care. Difficulties experienced by high-profile projects exemplify the challenges, including limited population coverage and clinical engagement, unknown data quality and public disquiet.

Aims: To develop the use of electronic patient record (EPR) data to improve neonatal specialised care, a high-cost NHS service.

Methods: We secured approvals from Caldicott Guardians, Lead Clinicians, the National Research Ethics Service and the Health Research Authority Confidentiality Advisory Group. We established a UK Neonatal Collaborative of provider NHS trusts. We collaborated with the Royal College of Paediatrics and Child Health and the national charity Bliss to develop a parent information leaflet. We conducted a systematic review to identify neonatal databases globally. We improved data completeness and quality through close interaction with clinical teams.

Results: We achieved 100% coverage of NHS neonatal units in England, Wales and Scotland (n = 185). We created a new NHS information standard, the Neonatal Data Set (ISB1595) and a National Neonatal Research Database (NNRD) containing a defined extract from real-time, point-of-care, clinician-entered EPRs. The NNRD is now used for a wide range and growing range of purposes including clinical and health services research, quality improvement programmes, national audit, commissioning support and national and regional benchmarking.

Conclusions: We have established proof of principle that EPR data may be employed to support patient care and clinical services through research and evaluation, and reduce the burden placed on busy clinical teams by providing a single national data source to service multiple outputs.

Background

Electronic patient records

Electronic patient records have been used increasingly over the last two to three decades and represent a rich data resource. Successive UK governments have recognised the potential of NHS clinical data to improve patient care and outcomes.[1–3] However, these aspirations have been slow to be adequately realised.

The challenges faced in harnessing the power of clinical data in health care are perhaps exemplified by the lessons of care.data and other high-profile UK projects. These include limited population coverage, weak clinical engagement, unknown data quality, regulatory uncertainty and public disquiet. Confidence in the concept of clinical data as a resource to improve patient care, despite an ambition to use these to improve standards, quality of care, accountability and patient choice, has been further damaged by escalating costs, critical media reports, breaches of data security and loss of public confidence by reports that personal data would be 'sold' to commercial organisations.

Neonatal specialised services

In the UK, neonatal specialised services (i.e. services for newborn infants requiring care over and above normal care) are currently provided by neonatal units operating in a series of mature clinical networks, each comprising around six to eight neonatal units. Neonatal networks were introduced as part of the

restructuring of neonatal services in response to a report by the Department of Health and Social Care in 2002. Each neonatal network was to be largely self-sufficient in providing care across the complete range of intensities (the traditional intensive, high dependency and special care levels). As a result, large numbers of infants were transferred between neonatal units providing different levels of care in accordance with their care requirements, with ultimate 'repatriation' to a neonatal unit closest to home in preparation for discharge. The concurrent need for clinical information to be readily transferable between NHS provider trusts was a cardinal driver for the introduction of EPR technologies into neonatal units.

Prior to the restructuring of neonatal services into networks, the British Association of Perinatal Medicine (BAPM) had commenced developing a 'minimum' Neonatal Data Set.[4] Neonatologists had long recognised the benefits of a uniform approach to recording clinical information, including the the ability to evaluate outcomes consistently at a national level. Over the period of the restructuring and subsequently, successive BAPM working parties made refinements to the 'minimum' Neonatal Data Set.[4]

Neonatal electronic patient records

Over the last decades, a UK-based commercial firm had developed a technical platform for neonatal data in close consultation with neonatal clinicians. This platform has evolved, with successive versions introduced into use over the years. Electronic systems were introduced across all NHS provider trusts from 2005 onwards. The EPR system in most widespread use includes fixed-choice and free-text items, with the NHS number as the principal identifier. Data are recorded daily throughout the neonatal inpatient stay. Clinician-entered diagnoses are converted to the *International Classification of Diseases*, Tenth Edition (ICD-10) codes.[5]

Thus, the reorganisation of neonatal services into managed clinical networks, an established, commercially available technical platform with a user front-end developed in close collaboration with neonatal clinicians, and a 'minimum' Neonatal Data Set established by a professional organisation, provided the three prime underpinning requirements on which to develop electronic clinical data for secondary purpose, including research and evaluation. Members of the Medicines for Neonates investigator group had been involved in several of the developments in relation to neonatal data described above, and electronic records more widely (e.g. as members of successive BAPM data working parties) and, hence, brought a wealth of experiential knowledge to the programme.

Aims

Our aim was to develop the use of EPR data for secondary purposes to support neonatal services and facilitate research to improve newborn care and outcomes. We also aimed to secure strong clinician engagement and parental support, implement measures to assess data quality systematically, and establish a new national resource.

Methods

We established a Programme Steering Committee comprising the Medicines for Neonates investigator group, an independent chairperson and independent members, including a patient and public involvement (PPI) representative. We conducted a systematic review to identify and describe existing neonatal databases. We investigated the regulatory processes required, and considered and tested ways in which to establish close clinical engagement, evaluate and improve data completeness and quality, and provide information to parents nationally.

Systematic review methods

We carried out an electronic search on MEDLINE (via Ovid), EMBASE (via Ovid), and CINAHL (Cumulative Index to Nursing and Allied Health Literature; via Athena), of publications covering the period 1 January 2000 to 15 March 2015. We applied language restrictions, including only English, French, German, Italian,

Russian and Spanish articles. We employed the following search terms: 'intensive care units, neonatal/' OR 'intensive care, neonatal/' OR 'neonatal intensive care units' OR 'NNU' OR 'NICU' OR 'neonatal ICU' AND 'infant/' OR 'neonat$' AND 'database$' or 'registry' OR 'registries' OR 'dataset$' OR 'data set$' OR 'vital statistics'. The literature search strategy is summarised in *Figure 1*.

We carried out grey literature searches on the Web of Science and the Ovid Maternity and Infant Care Databases using the free-text terms 'neonatal intensive care unit' AND 'infant' AND 'database'.

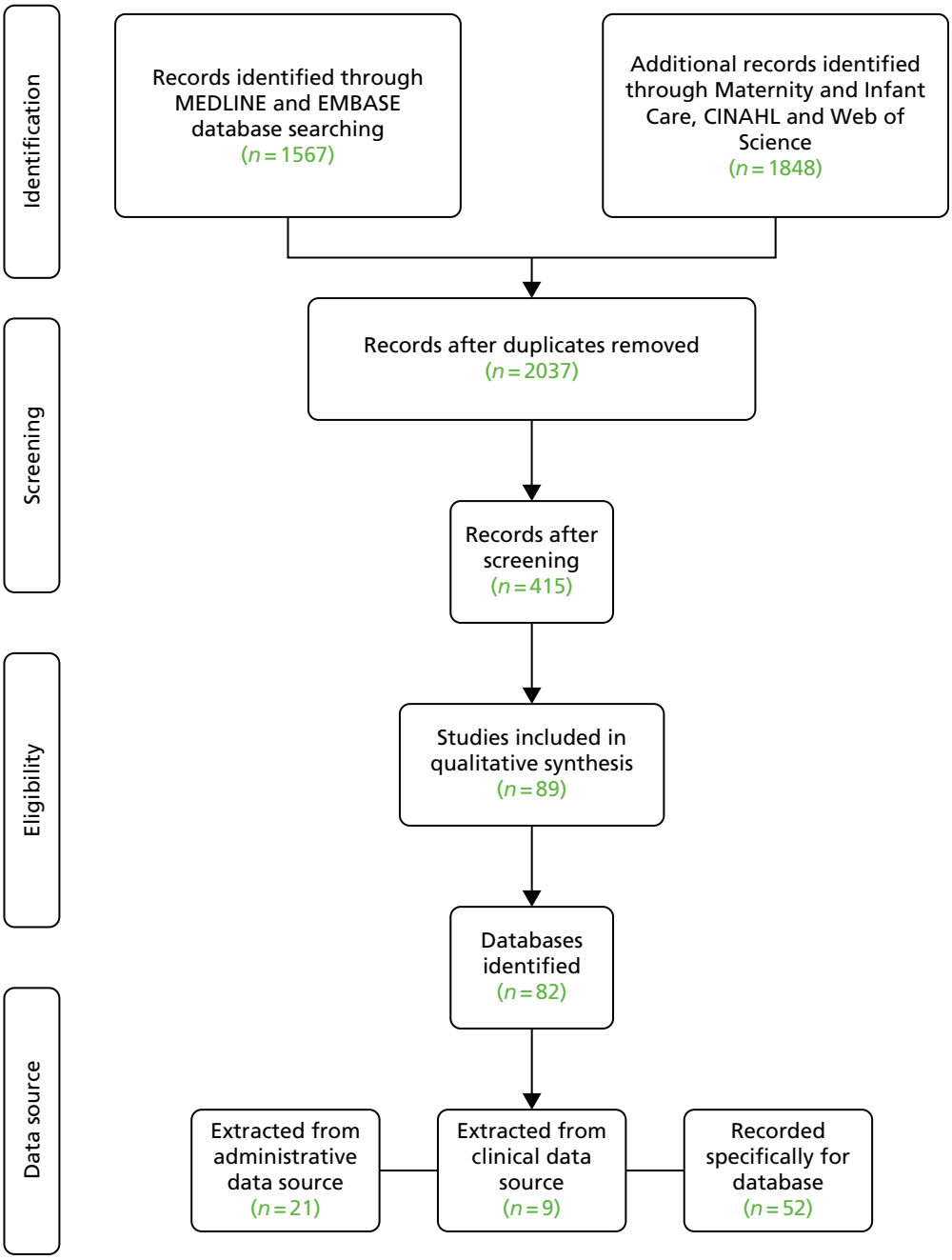


FIGURE 1 Literature search strategy.

We exported results, including abstracts, into EndNote X7 [Clarivate Analytics (formerly Thomson Reuters), Philadelphia, PA, USA]. Two researchers reviewed titles and abstracts to identify relevant publications and remove duplicate results. We retained publications that mentioned databases of patient-level information (administrative or clinical) and specified that data covered populations of infants from more than one neonatal unit. We reviewed full-text articles, references and websites. We entered extracted predefined information into Microsoft Excel® 2013 (Microsoft Corporation, Redmond, WA, USA) (*Table 1*).

TABLE 1 Data extraction for systematic review

Name	Original definition from PROSPERO submission	Updated definition for final systematic review
Study identification	To include main author and year of publication (e.g. Smith, John *et al.* 2015)	Same as original but also includes websites for databases
Database name	The name of the database	No change
Primary purpose	Administrative, clinical, research, audit, other	No change
Country	Free text for country where database is based	No change
Scope	Regional, national or international	No change
Scope name	Free text to specify the region of country or countries	No change
Population limit	Admissions in hospital, births in hospital	Admission to neonatal units, all infants included in admission to neonatal units, gestational age and/or birthweight cut-off point, admissions or births in enrolled hospitals, health insurance enrolment, no limitations, entire region included
Data source	Recorded specifically for database or secondary-use database	Secondary-use database broken down as data extracted from clinical source (electronic health records) or data extracted from administrative source
Number of infants reported	Number	No change
Time period for number of infants reported	The range of years that the database spans from earliest time period that could be identified to the present (e.g. 2000–15)	Includes if database is still enrolling patients
Maternal characteristics	Mother's ethnicity (yes/no), mother's age (yes/no), mother's education (yes/no), mode of delivery (yes/no)	No change
Infant characteristics	Gestational age in weeks (yes/no), gestational age in days (yes/no), gestational age definition (free text), birthweight (yes/no), sex (yes/no), multiplicity (yes/no), infant identification (yes/no), infant identification type (free text), intervention (yes/no), intervention type (free text), diagnoses (yes/no), diagnoses coded (yes/no), laboratory samples (yes/no), abdominal X-rays (yes/no), retinopathy of prematurity (yes/no), cranial ultrasound (yes/no), post-discharge information (free text)	Same as before except for the following changes: gestational age definition (yes/no), post-discharge information (yes/no) and blood cultures (yes/no) instead of laboratory samples (yes/no)
Funding	Not collected	Hospital subscription, insurance, mixed funding including support from public body. No current funding support identified. Support from public body

Creating the Neonatal Data Set

We built on and extended the BAPM 'minimum' Neonatal Data Set (data items used to derive daily level of care, the currency underpinning the commissioning of neonatal specialised care services) and the mandated National Critical Care Minimum Data Set (NCC-MDS) (used for deriving neonatal Healthcare Resource Groups) to build a national 'Neonatal Data Set'. This comprised basic demographic details (e.g. date of birth, birthweight), clinical interventions captured daily (e.g. respiratory support, type of feeds, surgical procedures, high-cost drugs), clinical outcomes and diagnoses. Each data item is clearly defined in an accompanying metadata set, and mapped to existing national standards as well as ICD codes. There was a preliminary assessment of the compatibility of Neonatal Data Set items for conversion to Snomed computed tomography (CT) terminology (international medical nomenclature); the conclusion was that the Neonatal Data Set is compatible, but conversion would require clinical and technical resourcing.

With the support of the NHS Information Standards Board (now NHS Digital), we submitted the Neonatal Data Set for approval as a national NHS standard. Following initial submission, the Information Standards Board issued an 'advance notice' of the Neonatal Data Set standard. In the process to becoming a national standard, the Neonatal Data Set evolved through changes that came about following public consultation, review by terminology experts at the NHS data dictionary, and alignment to other national data sets. As a result, 25 data items were added to the revised Neonatal Data Set and an existing 28 items were recoded to reflect data dictionary terminology or other national criteria. Full approval of the standard was obtained in December 2013. The stages leading to approval are shown in *Table 2*. The current approved Neonatal Data Set (SCCI1595) for standard items and age 2 years items are provided as *Appendix 3*.

TABLE 2 Approval pathway for the Neonatal Data Set

Submission stage	Document reference	Document title	Version	Date
Needs	Needs stage	National Neonatal Data Set Needs Stage Submission	1.7	7 September 2012
Requirements gathering	Requirements stage	National Neonatal Data Set ISB 1595 Requirements Stage Submissions	0.4	6 December 2012
Draft and full approval	Review of central returns approval	Review of Central Returns Approval notification OR 2027 FT6 0001PMAND	1	13 August 2013
Draft and full approval	Submission	Neonatal Data Set ISB 1595 Submission	1.4	16 October 2013
Draft and full approval	Specification	Neonatal Data Set ISB 1595 Specification	0.5	16 October 2013
Draft and full approval	Data set	Neonatal Data Set ISB 1595 Release 1	2.1	16 October 2013
Draft and full approval	Evidence of consultation	Neonatal Data Set ISB 1595 Evidence of consultation	0.7	16 October 2013
Draft and full approval	Data discovery	Neonatal Data Set ISB 1595 Data Discovery	2.5	16 October 2013
Draft and full approval	Implementation and maintenance plan	Neonatal Data Set ISB 1595 Implementation and Maintenance Plan	0.5	16 October 2013
Draft and full approval	Issues and risks	Neonatal Data Set ISB 1595 Issues Log & Risk Register	10	16 October 2013

Results

Systematic review

Search results

The search yielded 2037 unique papers. Following a review of the titles and abstracts, 415 papers met our prespecified criteria. From these, we identified 82 databases and, for 52, data were recorded specifically for the database. In 21 papers, data were obtained from a primary administrative source and in nine papers data were obtained from a clinical source (*Figure 2*). Five countries accounted for the location of more than half (47/82) of all identified databases: the USA ($n = 24$), Canada ($n = 11$), the UK ($n = 7$) and Australia/New Zealand ($n = 5$). We provide details of the databases identified in *Table 3*.

Primary purpose of databases

Of the 38 national databases, the primary purpose was clinical in 18, administrative in 13 and research in seven. Of the 40 regional databases 15 were clinical, 23 were administrative and two were research orientated. We identified four international databases (two were clinical, one was research and one was surveillance).

Data sources

Specific data collection is required for 28 out of 38 national databases. Data are extracted from a primary administrative source for seven databases, and from a primary clinical source for three databases (UK: NNRD; USA: Consortium of Safe Labor Database and Pediatrix BabySteps Clinical Data Warehouse) (see *Table 3*). Twenty-one of the 40 regional databases require specific data collection and, for 14, the source is an administrative database and, for 5, the source is clinical (see *Table 3*). All four international databases require specific data recording.

Population coverage

Twenty-seven databases hold data on all admissions to neonatal units, and the remaining databases restrict data by gestational and/or birthweight cut-off points, and/or enrolment or insurance cover.

Funding sources

Of the 82 databases identified, 71 receive funding from public sources, eight are funded through hospital subscriptions and one is funded through private insurance. We were unable to identify the funding source for the French AUDIPOG (Association des Utilisateurs de Dossiers Informatisés en Pédiatrie, Obstétrique et Gynécologie) Network;[11] the NNRD was developed in part through public research funding, but has no ongoing funding support.

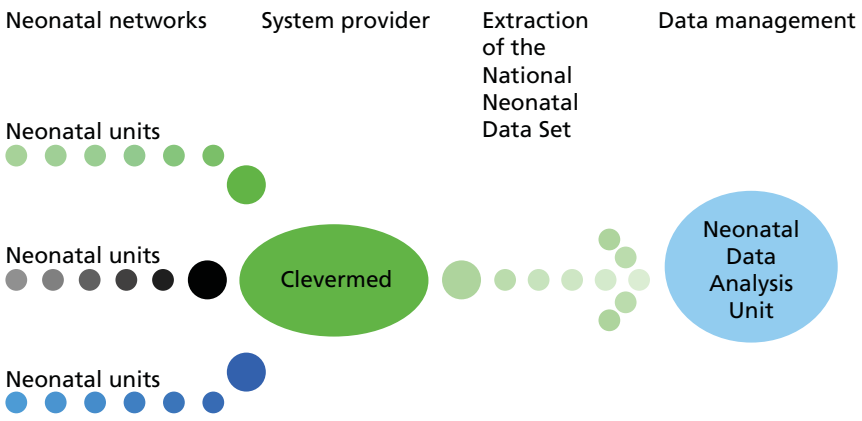


FIGURE 2 Data flows into the Neonatal Data Analysis Unit.

TABLE 3 Details of databases identified

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
1	Alberta Perinatal Health Program Database[6]	A	Canada	Regional (Alberta)	No limitations, entire region included	Extracted from administrative data source	2002–4 (yes)	Mother's age, mode of delivery	Birthweight, sex, multiplicity	Support from public body
2	Alere or Matria Health care/ Paradigm[7,8]	C	USA	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2003–7 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Hospital subscription
3	Arizona Newborn Intensive Care Program[9]	A	USA	Regional (Arizona)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Extracted from administrative data source	1994–8 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, identifier, intervention, ROP, post-discharge information	Support from public body
4	Asian Network on Maternal and Newborn Health[10]	R	Asia (Malaysia, Japan, Hong Kong and Singapore)	International	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2003–6 (yes)	Mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, diagnoses, abdominal X-rays, cranial ultrasound	Support from public body
5	AUDIOPOG Sentinel Network[11]	C	France	National	Admissions or births in enrolled hospitals	Recorded specifically for database	1994–2008 (yes)	Mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	No current funding support identified

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
6	Australia and New Zealand Neonatal Network[12]	C	Australia/ New Zealand	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1994–2012 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, blood cultures, ROP, post-discharge information	Mixed funding, including support from public body
7	Better Outcomes Registry and Network[13]	A	Canada	Regional (Ontario)	No limitations, entire region included	Recorded specifically for database	2006–10 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, ROP, post-discharge information	Support from public body
8	California Patient Discharge Linked Birth Cohort Database[14,15]	A	USA	Regional (California)	Admissions or births in enrolled hospitals	Extracted from administrative data source	1999–2004 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity	Support from public body
9	California Perinatal Quality Care Collaborative[16]	C	USA	Regional (California)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2005–11 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
10	Canadian Institute for Health Information Discharge Abstract Database[17]	A	Canada	National	No limitations, entire region included	Extracted from administrative data source	2002–10 (yes)	Mother's ethnicity, mother's age	Gestational age in weeks, birthweight, sex, identifier, intervention, ROP, post-discharge information	Support from public body
11	Canadian Neonatal Follow-Up Network[18]	C	Canada	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2010–11 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, diagnoses, blood cultures, ROP, post-discharge information	Support from public body
12	Canadian Neonatal Network[19]	C	Canada	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2013–14 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, diagnoses, blood cultures	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
13	Canadian Paediatric Surgery Network[20]	C	Canada	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2013–14 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, identifier, intervention, diagnoses, abdominal X-rays, cranial ultrasound	Support from public body
14	Children's Hospital Neonatal Database[21]	C	USA	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2010–11 (yes)	Mother's ethnicity	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Hospital subscription
15	Colorado Birth Certificate Database[22]	A	USA	Regional (Colorado)	No limitations, entire region included	Recorded specifically for database	2007–12 (yes)	Mother's ethnicity	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
16	Consortium of Safe Labor Database[23]	R	USA	National	Admissions or births in enrolled hospitals	Extracted from clinical data source (electronic health records)	2002–8 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body
17	Croatian Intensive Care network[24]	C	Croatia	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2004–5 (yes)	No variables identified from those sought	Identifier	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
18	Danish Medical Birth Registry[25]	A	Denmark	National	No limitations, entire region included	Recorded specifically for database	1997–2008 (yes)	Mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
19	Danish Neonatal Clinical Database (NeoBase)[26]	C	Denmark	Regional (North And South Jutland)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Extracted from clinical data source (electronic health records)	2005–6 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight, sex, identifier, intervention	Support from public body
20	Emilia-Romagna Health Agency[27]	A	Italy	Regional (Emilia-Romagna)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2002–9 (yes)	No variables identified from those sought	Gestational age in weeks	Support from public body
21	EPICure[28]	R	UK	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2006–7 (no)	Mother's ethnicity, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, intervention, ROP, post-discharge information	Support from public body
22	Erie County Register[29]	A	USA	Regional (Erie County, New York)	No limitations, entire region included	Recorded specifically for database	2006–8 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, identifier, intervention	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
23	EuroNeoNet[30]	C	European	International (Austria, Belgium, Croatia, Finland, France, Germany, Greece, Italy, Poland, Portugal, Russia, Slovenia, Spain, Switzerland, the Netherlands, Turkey and the UK)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2006–11 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, ROP, post-discharge information	Support from public body
24	Florida birth registry[31]	A	USA	Regional (Florida)	No limitations, entire region included	Recorded specifically for database	2009–10 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
25	Intermountain Health care[32]	C	USA	Regional (Utah)	Admissions or births in enrolled hospitals	Extracted from clinical data source (electronic health records)	2003–5 (yes)	No variables identified from those sought	Birthweight	Hospital subscription
26	Israel National VLBW Infant Database[33]	C	Israel	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1995–2003 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
27	Japanese Vital Statistics[34]	A	Japan	National	No limitations, entire region included	Recorded specifically for database	1999–2008 (yes)	Mother's age	Gestational age in weeks, birthweight, sex, multiplicity	Support from public body
28	Kaiser Permanente Medical Care Program[35,36]	C	USA	Regional (Northern California and Boston, Massachusetts)	Admissions or births in enrolled hospitals	Extracted from clinical data source (electronic health records)	1995–6 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Hospital subscription
29	Kids' Inpatient Databases[37]	R	USA	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2003–12 (yes)	Mother's ethnicity	Birthweight, sex	Support from public body
30	Linked Emergency Management and Research Institute – Department of Health and Family Welfare, Government of Gujarat[38]	A	India	Regional (Gujarat)	Admissions or births in enrolled hospitals	Extracted from administrative data source	2008–9 (yes)	Mother's age, mode of delivery	Gestational age in weeks	Support from public body
31	London Neonatal Transfer Service[39]	C	UK	Regional (London)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2005–11 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
32	Massachusetts Community Health Information Profile (MassCHIP) and PELL[40,41]	A	USA	Regional (Massachusetts)	No limitations, entire region included	Extracted from administrative data source	2002–10 (yes)	Mother's ethnicity, mother's age	Gestational age in weeks, multiplicity, intervention	Support from public body
33	Malaysian National Neonatal Registry[42]	C	Malaysia	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2006–7 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, blood cultures, abdominal X-rays, cranial ultrasound	Support from public body
34	Medicaid Analytic eXtract[43]	A	USA	National	Health insurance enrolment	Extracted from administrative data source	2006–8 (yes)	Mode of delivery	Identifier	Support from public body
35	Memorial Care Medical Centres: Perinatal database, Quality Improvement Database[44]	C	USA	Regional (California)	Admissions or births in enrolled hospitals	Recorded specifically for database	2002–3 (yes)	Mother's ethnicity, mode of delivery	Gestational age in weeks, identifier, intervention	Hospital subscription

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
36	Michigan Linked Records[45]	A	USA	Regional (Michigan)	No limitations, entire region included	Extracted from administrative data source	2003–4 (yes)	Mother's ethnicity	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
37	National Centre for Health Statistics linked live birth and infant death cohort file[46]	A	USA	National	No limitations, entire region included	Extracted from administrative data source	1998–9 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity	Support from public body
38	National Collaborative Perinatal Neonatal Network[47]	C	Lebanon	Regional (Greater Beirut)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2009–10 (not identified)	Mother's age, mother's education	Gestational age in weeks, birthweight, sex, multiplicity	Support from public body
39	National Institute for Health and Welfare: Medical Birth Register[48]	A	Finland	National	No limitations, entire region included	Recorded specifically for database	2012–13 (yes)	Mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
40	National Neonatal Database SEN1500[49]	C	Spain	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2002–5 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, diagnoses, blood cultures, abdominal X-rays, cranial ultrasound	Support from public body
41	National Neonatal Perinatal Database[50]	C	India	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2002–3 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, intervention	Support from public body
42	NNRD[51]	C	UK	National	Admission to neonatal units, all infants included	Extracted from clinical data source (electronic health records)	2009–11 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, diagnoses, blood cultures, abdominal X-rays, ROP, cranial ultrasound, post-discharge information	No current funding support identified

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
43	National Perinatal Information Centre[52]	A	USA	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2004–8 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, sex, identifier	Support from public body
44	National Perinatal Registry of Slovenia[53]	A	Slovenia	National	No limitations, entire region included	Recorded specifically for database	2012–13 (yes)	No variables identified from those sought	Gestational age in weeks	Support from public body
45	National Perinatal Registry, the Netherlands[54]	C	The Netherlands	National	No limitations, entire region included	Recorded specifically for database	2003–7 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
46	National Perinatal Data Collection[55]	A	Australia/ New Zealand	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Extracted from administrative data source	2001–5 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
47	National Registry of Respiratory Distress Syndrome in Romania[56]	C	Romania	National	Admissions or births in enrolled hospitals	Recorded specifically for database	2011–12 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight, sex, identifier, intervention	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
48	Neonatal Intensive Care Outcomes and Research Evaluation[57]	C	UK	Regional (Northern Ireland)	Admission to neonatal units, all infants included	Extracted from clinical data source (electronic health records)	1999–2000 (yes)	Mother's ethnicity, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, blood cultures, abdominal X-rays, cranial ultrasound	Support from public body
49	Neonatal Research Network of Japan[58]	R	Japan	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2003–8 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, intervention, blood cultures	Support from public body
50	NEOSANO's Perinatal Network in Mexico[59]	A	Mexico	Regional (Mexico City, Tlaxcala City and Oaxaca City)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2006–9 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, multiplicity	Support from public body
51	New Jersey Perinatal Linked Data-Set[60]	A	USA	Regional (New Jersey)	No limitations, entire region included	Extracted from administrative data source	1997–2005 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, identifier	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
52	New South Wales Newborn and Paediatric Emergency Transport Service[61]	C	Australia/ New Zealand	Regional (New South Wales)	Admissions or births in enrolled hospitals	Recorded specifically for database	1992–2001 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight, sex, intervention	Support from public body
53	New York State-wide Perinatal Data System[62]	A	USA	Regional (New York)	Admissions or births in enrolled hospitals	Extracted from administrative data source	1996–2003 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body
54	Newfoundland and Labrador Provincial Perinatal Program Database[63]	C	Canada	Regional (Newfoundland and Labrador)	Admissions or births in enrolled hospitals	Extracted from administrative data source	2001–9 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity	Support from public body
55	NICHD Neonatal Research Network Generic Database[64]	R	USA	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1998–2009 (yes)	Mother's ethnicity, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, ROP, post-discharge information	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
56	Nova Scotia Atlee Perinatal Database[65]	A	Canada	Regional (Nova Scotia)	No limitations, entire region included	Extracted from administrative data source	2002–11 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body
57	NSW Pregnancy and Newborn Services Network[66]	C	Australia/ New Zealand	Regional (New South Wales And Australian Centralised Territory)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1997–2006 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, intervention	Support from public body
58	Pediatrix BabySteps Clinical Data Warehouse[67]	C	USA	National	Admission to neonatal units, all infants included	Extracted from clinical data source (electronic health records)	1996–2010 (yes)	Mother's ethnicity, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, blood cultures	Hospital subscription
59	Perinatal and Neonatal Surveys in Saxony[68]	C	Germany	Regional (Saxony)	No limitations, entire region included	Recorded specifically for database	2001–5 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
60	Perinatal database of Middlesex Country, Canada[69]	C	Canada	Regional (Middlesex County, Ontario)	Admissions or births in enrolled hospitals	Recorded specifically for database	2002–11 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex	Support from public body
61	Perinatal Revision South[70]	C	Sweden	Regional (Southern Sweden)	Admissions or births in enrolled hospitals	Recorded specifically for database	1995–6 (yes)	Mode of delivery	Gestational age in weeks, birthweight, identifier	Support from public body
62	Perinatal Services British Columbia[71]	A	Canada	Regional (British Columbia)	Admissions or births in enrolled hospitals	Extracted from administrative data source	2004–14 (yes)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body
63	Population Health Research Data Repository[72]	A	Canada	Regional (Manitoba)	No limitations, entire region included	Extracted from administrative data source	2004–9 (yes)	Mother's age	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
64	Scottish Administrative Linked Data[73]	A	UK	National	No limitations, entire region included	Extracted from administrative data source	1981–2007 (yes)	Mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body
65	Seguro Medico para una Nueve Generacio[74]	A	Mexico	National	Health insurance enrolment	Recorded specifically for database	2008–9 (yes)	No variables identified from those sought	Gestational age in weeks, sex, identifier	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
66	Swedish Neonatal Quality Register[75,76]	C	Sweden	National	Admission to neonatal units, all infants included	Extracted from administrative data source	2001–2 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, blood cultures, ROP, post-discharge information	Support from public body
67	Swiss Neonatal Network[77]	C	Switzerland	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1996–2008 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, intervention, blood cultures, ROP, post-discharge information	Hospital subscription
68	Taiwan's National Health Insurance Research Database[78]	A	Taiwan	National	Health insurance enrolment	Extracted from administrative data source	1998–2001 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight, sex, intervention, ROP, post-discharge information	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
69	Tennessee Hospital Discharge Data System[79]	A	USA	Regional (Tennessee)	Admissions or births in enrolled hospitals	Extracted from administrative data source	2003–5 (yes)	Mother's ethnicity	Birthweight, sex, identifier, intervention	Support from public body
70	The National Neonatology Database[80]	C	The Netherlands	National	Admission to neonatal units, all infants included	Recorded specifically for database	2003–5 (yes)	No variables identified from those sought	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body
71	The Neonatal Survey[81]	C	UK	Regional (East Midlands and Yorkshire)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2008–10 (yes)	Mother's ethnicity, mother's age, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, infant identification, intervention, diagnoses, laboratory samples, ROP, post-discharge information	Support from public body

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
72	The WHO's Global Survey for Maternal and Perinatal Health[82,83]	S	International [Africa (Angola, Democratic Republic of Congo, Algeria, Kenya, Niger, Nigeria and Uganda), Latin America (Argentina, Brazil, Cuba, Ecuador, Mexico, Nicaragua, Paraguay and Peru) and Asia (Cambodia, China, India, Japan, Nepal, Philippines, Sri Lanka, Thailand and Vietnam)]	International	No limitations, entire region included	Recorded specifically for database	2004–8 (no)	Mother's age, mother's education, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
73	Vermont Oxford Network[84]	C	International (Australia, Brazil, Canada, China, Columbia, Finland, Germany, Hungary, Ireland, Italy, Kuwait, Malaysia, Namibia, Poland, Portugal, Qatar, Romania, Saudi Arabia, Singapore, Slovenia, South Africa, Spain, Switzerland, Taiwan, Turkey, United Arab Emirates, the UK and the USA)	International	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1990–2012 (yes)	Mother's ethnicity, mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention, blood cultures, abdominal X-rays, cranial ultrasound	Hospital subscription
74	Victorian Perinatal Data Collection Unit[85]	R	Australia/ New Zealand	Regional (Victoria)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1979–97 (no)	No variables identified from those sought	Birthweight, sex, identifier, ROP, post-discharge information	Support from public body
75	West Midlands Perinatal Institute[86]	C	UK	Regional (West Midlands)	No limitations, entire region included	Recorded specifically for database	2008–9 (yes)	No variables identified from those sought	No variables identified from those sought	No current funding support identified

continued

TABLE 3 Details of databases identified (*continued*)

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
76	Wisconsin Linked Birth Record File[87]	A	USA	Regional (Wisconsin)	Health insurance enrolment	Extracted from administrative data source	2001–2 (yes)	Mother's ethnicity, mother's age, mother's education, mode of delivery	Birthweight, sex, multiplicity, identifier	Support from public body
77	AOK National Insurance Entries[88]	A	Germany	National	Health insurance enrolment	Recorded specifically for database	2002–6 (yes)	No variables identified from those sought	No variables identified from those sought	Insurance
78	Regional Census Data[89]	A	Germany	Regional (Westfalen Lippe)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	1990–6 (yes)	No variables identified from those sought	Blood cultures	Support from public body
79	Neonatal Quality Assurance System[90]	A	Germany	Regional (Baden Wuertemberg)	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2003–4 (yes)	No variables identified from those sought	No variables identified from those sought	Support from public body
80	Bourgogne database[91]	A	France	Regional (Bourgogne)	Admissions or births in enrolled hospitals	Extracted from clinical data source (electronic health records)	2000–1 (yes)	Mode of delivery	Gestational age in weeks, birthweight, sex, multiplicity, identifier, intervention	Support from public body

Number	Name	Primary purpose	Country	Scope	Population limits	Data source	Time period covered by publication (database still enrolling patients)	Maternal variables	Infant variables	Funding source
81	Multicentre national database[92]	R	France	National	Admission to neonatal units, gestational age and/or birthweight cut-off point	Recorded specifically for database	2005–6 (yes)	No variables identified from those sought	Gestational age in weeks, sex, intervention	Support from public body
82	Hessen Neonatal Register[93]	A	Germany	Regional (Hessen)	No limitations, entire region included	Recorded specifically for database	1989–2012 (yes)	No variables identified from those sought	No variables identified from those sought	Support from public body

a, administrative; c, clinical; NICHD, National Institute of Child Health and Human Development; o, other; r, research; ROP, retinopathy of prematurity; s, surveillance; VLBW, very low birthweight; WHO, World Health Organization.

Summary
The NNRD is one of six national neonatal databases with ongoing data acquisition, primarily developed to support research. Uniquely, and in contrast to each of the other five (databases 16, 29, 49, 55 and 81 in *Table 3*), data in the NNRD are extracted from EPRs rather than being recorded specifically. There is complete national coverage of all admissions to neonatal units and no gestational age, birthweight, insurance cover or other restrictions.

Regulatory approvals

We obtained National Research Ethics Service approval in 2010 to establish a NNRD from extracts from EPRs, undertake projects within the Medicines for Neonates Programme, and employ the NNRD for NHS service evaluations and other research studies (REC reference number 10/H0803/151; provided as *Appendices 5* and *6*). We obtained approval in 2010 from the Confidentiality Advisory Group of the Health Research Authority [formerly Ethics and Confidentiality Committee of the National Information Governance Board; reference number ECC 8–05(f)/2010; provided as *Appendix 4*] to receive specific patient identifiers for the purpose of linking to Hospital Episode Statistics (HES) data.

Under standard operating procedures for Research Ethics Committees, site-specific approval is not required for studies conducted using research databases. There is no requirement for specific ethics approval for data collection centres that provide data because, under the Research Governance Framework, data collection centres are not regarded as research sites. NHS trust approval ('R&D' approval) is only required from the NNRD host institution (i.e. not from each data collection centre). However, we were informed that 'local collaborators at Data Collection Centres within the NHS will require internal permission from their NHS care organisation to collect and supply data relating to NHS patients'.[94] We addressed this requirement by seeking Caldicott Guardian and Lead Neonatal Clinician approval from every NHS trust providing neonatal specialised care, to receive a defined extract from their neonatal EPRs, to hold these in the NNRD and to use these in NHS service evaluations and Research Ethics Committee-approved research studies. We obtained approvals incrementally and all NHS trusts in England, Wales and Scotland have now granted approval.

The National Neonatal Research Database

The data items constituting the Neonatal Data Set (NND) are extracted from EPRs created by clinical staff on all admissions to neonatal units in England, Wales and Scotland. Neonatal units in Northern Ireland utilise the same EPR platform but, to date, the regulatory approvals governing data transfer into the NNRD have not been sought. Following receipt of the necessary approvals, retrospective data extraction was undertaken so that the NNRD contains data from 2007 to the present. The NNRD is updated quarterly, and to date it contains data on approximately half a million infants and > 5 million care days. All neonatal units across England, Wales and Scotland have approved the release of Neonatal Data Set data items for inclusion in the NNRD (the total number of neonatal units is approximately 200, which has fluctuated over the course of the Medicines for Neonates programme as neonatal units have merged or reorganised).

An NHS approved supplier, Clevermed Ltd (Edinburgh, UK), provides a web-based data capture platform known as Neonatal.Net or BadgerNet. Data held by Clevermed Ltd are stored on a secure N3 server and transmitted to the Neonatal Data Analysis Unit (NDAU) where they are used to create the NNRD after merging and cleaning of files. The NNRD is held on the NHS servers of Chelsea and Westminster NHS Foundation Trust. Data flows are shown in *Figures 2* and *3*.

Data management

At the NDAU, all data extracted from the neonatal EPRS are interrogated to identify duplicate, missing and potentially erroneous entries. Items are considered potentially erroneous if they fail a series of out of range, internal logic, and internal inconsistency checks.

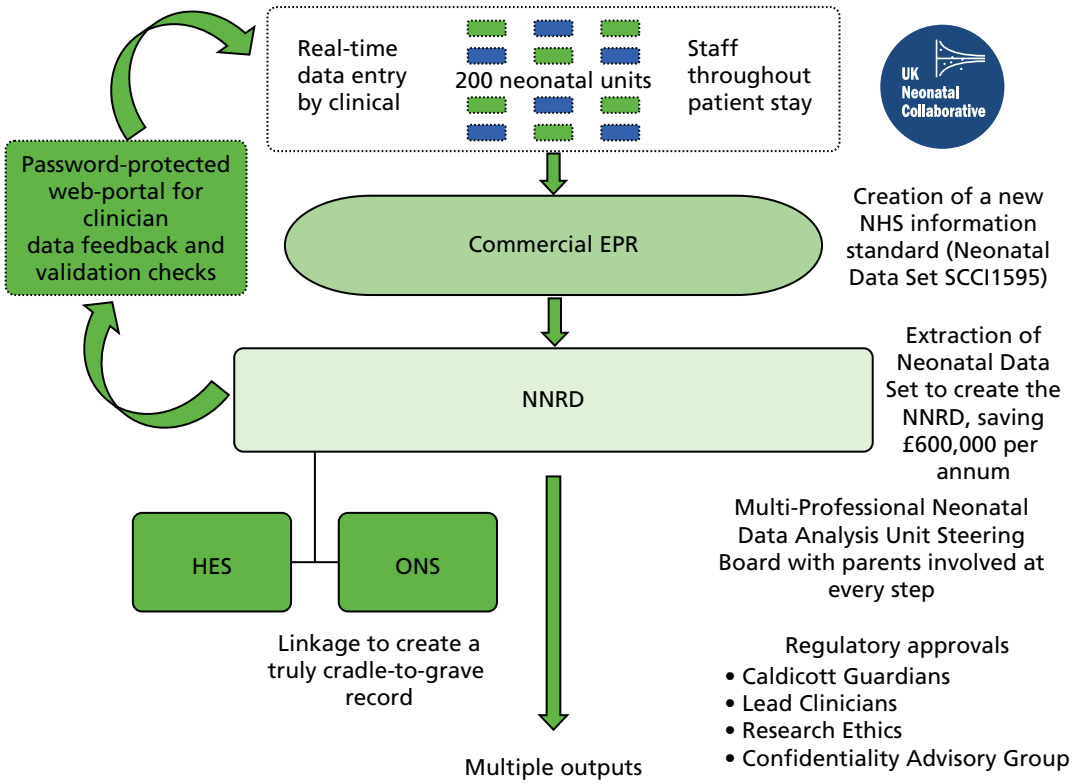


FIGURE 3 Data flows to create the NNRD. ONS, Office for National Statistics.

A web-based portal was created to notify neonatal unit lead clinicians of missing or potentially erroneous entries. If clinical teams amend errors or complete missing fields, this is done in the baby's EPR and this is sent to the NDAU at the next download. Initially, this process was confined to the data items (approximately 60 items) used for analyses for the National Neonatal Audit Programme. In addition, as the NNRD has become used for research studies, if key data items are required for specific projects then these are also subjected to the feedback loop process.

At the NDAU, patient episodes across multiple neonatal units are also merged to create a single file for each infant.

Clinician engagement

We termed NHS neonatal units contributing data to the NNRD the 'UK Neonatal Collaborative'. Of note was that, although only site-specific approval and NHS approvals are required for the NNRD host institution, we would adopt a policy of seeking the approval of each NHS trust's lead neonatal clinician for their data to be included in research studies. We adopted this practice in order to grow clinician engagement with the concept of the NNRD as a national resource and in recognition of their crucial contribution to acquiring the data.

Parent information leaflet

We collaborated with the Royal College of Paediatrics and Child Health, the national charity Bliss and the parents of newborn babies receiving specialised neonatal care to develop a parent information leaflet ('A Guide for Parents and Carers') that explains the multiple uses of the NNRD. This was approved by the National Research Ethics Service and the Ethics and Confidentiality Committee of the National Information Governance Board.

If the parent or carer of the infant does not wish for EPR data on their infant to be extracted for the NNRD, then they can notify the neonatal unit staff who will then notify the data entry system supplier to prevent the flow of the data. To date, no parent or carer has asked that their infant's data not be extracted.

Uses and outputs of the National Neonatal Research Database to date

The use of the NNRD to support a wide range of outputs grew rapidly over the course of the Medicines for Neonates Programme, and continues to expand. Examples of the multiple outputs from the NNRD are shown in *Figure 4*.

Conclusions

We have shown that it is possible to create a national data resource, the NNRD, from extractions from EPRs, which brings multiple benefits. This eliminates the need for multiple individual collections, with repetitive capture of many commonly required data items. This in turn reduces the burden of data capture all too often imposed on busy clinical teams, and reduces the risk of transcription errors and other errors.

Our literature search identified 82 databases worldwide that hold neonatal information. The NNRD is one of only six national neonatal databases primarily developed to support research, with ongoing data acquisition. Uniquely, data in the NNRD are extracted from EPRs rather than being recorded specifically and there is complete national coverage of all admissions to neonatal units with no gestational age, birthweight, insurance cover or other restrictions.

The Neonatal Data Set incorporates data required to fulfil all currently mandated UK requirements. In addition, the Neonatal Data Set is sufficiently comprehensive to make the need for subsequent addition of new national data items unlikely in the immediate future. However, should this be required, the process for incorporation of new items into the EPR is straightforward (see *Chapter 2, Research on an exemplar condition*).

We have demonstrated that the approach we have adopted has been successful. The NNRD is now used for a growing number of purposes by a growing number of research groups, professional organisations and government bodies. In effect, our approach has gone a considerable way towards fulfilling the vision set out by Florence Nightingale more than 100 years ago and, more recently, the principles set out in successive national information strategies. These include the Council for Science and Technology report *Better Use of Personal Information: Opportunities and Risks*,[95] the UK Clinical Research Collaboration Research and Development Advisory Group to Connecting for Health,[96] the Academy of Medical Sciences report entitled *Personal Data for Public Good: Using Health Information in Medical Research*,[97] the Department of Health and Social Care's entitled *Toolkit for High Quality Neonatal Services*,[98] and the aspiration articulated by the then Prime Minister, in numerous references to 'big data'. We believe that the NNRD may reasonably

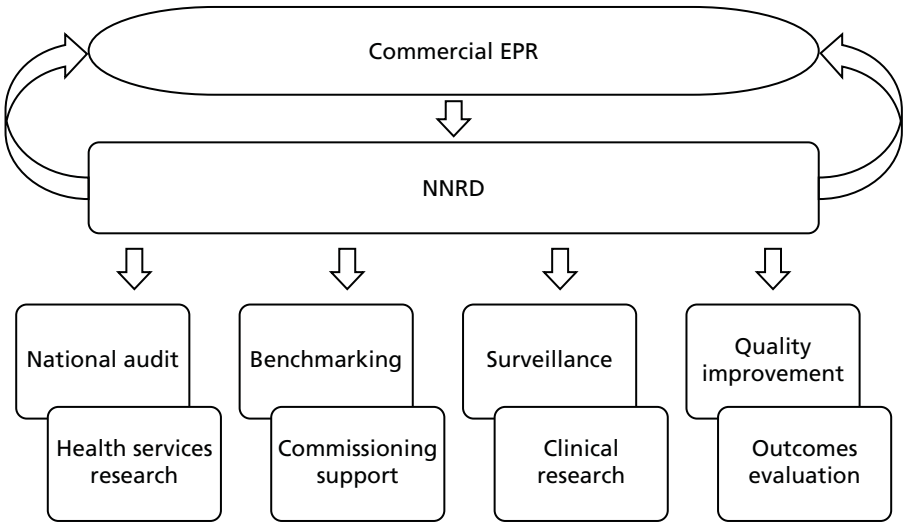


FIGURE 4 Examples of multiple outputs from the NNRD.

be termed an example of 'big data'. Although the term has been defined in various ways, Wang and Krishnan[99] state that 'A popular definition of big data is the "3V" model proposed by Gartner, which attributes three fundamental features to big data: high volume of data mass, high velocity of data flow, and high variety of data types'. The data in the NNRD does encompass each of these elements to varying extents, in contrast with many other clinical data sets that are much simpler.

In conclusion, we have established proof of the principle that EPR data may be employed successfully to support patient care and clinical services through research and a range of evaluations. We have shown that it is possible to reduce the burden placed on busy clinical teams by providing a single national data source to service multiple outputs.

Other Medicines for Neonates workstreams deal with issues of data quality, utility and patient (parent) involvement.

Implications for health care

The Medicines or Neonates programme has also established proof of concept for the use of EPR-derived clinical data in a wide range of research and health service evaluations. This opens up the possibility of adapting the road map that we have established for other specialty areas with potential to bring about NHS savings.

The NNRD has been developed and is currently maintained through academic endeavour, but processes to secure the stability of EPR-derived databases as national resources and their ongoing management are uncertain.

Research recommendations

A next step towards seizing the full potential of our approach for the benefit of the NHS and patient care would be to formally test the creation of another specialty database from EPRs using the road map that we have developed.

Chapter 2 Research on an exemplar condition: the use of the National Neonatal Research Database to study neonatal necrotising enterocolitis

Abstract

Background: Necrotising enterocolitis is a feared gastrointestinal inflammatory disease that predominantly affects preterm infants. The aetiology is uncertain and population incidence data are scant. Treatment is supportive and it includes surgery, but research is constrained by the relative rarity of the disease.

Aims and objectives: We utilised EPR-derived data to conduct population surveillance of severe NEC in England. Additional objectives were to inform the development of future clinical trials by identifying factors associated with severe NEC.

Methods: We secured the participation of every NHS neonatal unit in England in a prospective study. We extracted relevant data from the NNRD. We also obtained outcome data for infants who received NEC surgery or died from NEC at stand-alone paediatric surgical centres that do not use the neonatal EPRS.

Results: We identified 531 infants (462 who were born at < 32 weeks' gestation) with severe NEC (resulting in surgery and/or death) over the complete 2-year period 2013–14. Among the infants born before 32 weeks' gestation, neonatal network incidence ranged from 19.8 [95% confidence interval (CI) 9.1 to 30.4] to 47.4 (95% CI 32.5 to 62.4) per 1000 babies. We identified no strong evidence of variation between networks following adjustment for gestational age and birthweight standard deviation score (SDS), which were the only factors found to be independently associated with the disease.

Conclusions: The NNRD provides opportunity for rapid population surveillance of neonatal conditions and a source of baseline information to inform clinical trials but it requires strong clinician engagement.

Background

Necrotising enterocolitis is a feared gastrointestinal inflammatory disease that predominantly, but not exclusively, affects preterm infants. NEC is a principal cause of mortality and morbidity in very preterm infants.[100,101] The aetiology of NEC is uncertain and is likely to be multifactorial. Some studies have suggested that the most significant factor in determining NEC incidence is the neonatal unit in which an infant receives care, with the implication that variations in care affect risk.[102,103] In particular, there is a widespread view that enteral feeding regimens, including type of milk, affect the risk of NEC. However, lack of good evidence for specific feed-related interventions that affect the risk of NEC has resulted in variation in neonatal practice, entrenched clinical opinion and bewilderment among parents. Evidence is conflicting regarding whether or not antenatal steroid exposure, a strong predictor of neonatal survival, is associated with NEC.[102,104–110] A further difficulty is that the diagnosis of NEC can be difficult as signs are often non-specific and presentation is variable. No internationally agreed case definition exists, which makes comparisons between studies unreliable. The most frequently applied definitions include modified Bell's criteria,[111,112] which, although developed as criteria for staging after the diagnosis is made, have been widely adopted as a definition worldwide. There are also definitions from the *International Classification of Diseases,* Ninth Edition (ICD-9) or IC-10 codes, the Vermont Oxford Network (VON), the US Centers for Disease Control and Prevention (CDC) and definitions from individual study authors. None has been developed through evidence-based methodology or has undergone validation.

In the absence of evidence from randomised trials, an approach that is widely believed to be of benefit is to identify variation in NEC incidence between neonatal units or networks, in the hope that this might help highlight potentially beneficial clinical practices that can then be tested in future randomised controlled trials.

Aims

We aimed to build on the establishment of the UK Neonatal Collaborative and the existence of the neonatal EPRS and NNRD to examine aspects of this serious disease. Our objectives were to:

- conduct population surveillance of severe NEC
- identify factors associated with severe NEC
- evaluate variation in incidence across neonatal networks in England.

We also aimed to build engagement with local clinical staff responsible for recording neonatal data, through demonstration of the utility of EPR data in national research in an area considered a priority by parents and clinicians alike.

Methods

Approvals and agreements

We sought and obtained research ethics approval from the National Research Ethics Service (Dulwich Research Ethics Committee; reference number 11/LO/1430) and inclusion into the UK Clinical Research Network Portfolio (ID 11853). We invited the participation of all neonatal units in England. We sought agreement from the local UK Neonatal Collaborative lead clinicians to utilise data from their neonatal unit held in the NNRD on all live-born infants admitted over the complete 2-year period 2012–13.

In order to promote maximal engagement with the neonatal community and optimise data quality and completeness, we asked local clinical staff to ensure that the following data items were recorded in each infant's EPR: birthweight, gestational age, sex, mother's race, antenatal steroids, and clinical and abdominal X-ray (AXR) findings for infants in whom abdominal signs were being investigated. We excluded infants of mothers that were unbooked, booked in non-English networks, or for whom network of booking was unknown.

Identifying babies with severe necrotising enterocolitis in the National Neonatal Research Database

We defined 'severe NEC' as NEC confirmed at surgery or post-mortem or resulting in death (death certification and/or verified by neonatal team). We initially intended to capture outcomes on a specific section of the neonatal EPRs, the screen used to record details of abdominal X-rays taken to investigate clinical signs consistent with gastrointestinal pathology. However, despite regular quarterly feedback of data completeness to neonatal units, we found that only 25% of infants who proceeded to NEC surgery had completed this screen; hence, sole use of these data would underestimate the incidence of severe NEC. Therefore, we used data from the NNRD to identify infants who may have received surgery for NEC or died from NEC, and verified these outcomes with clinicians at neonatal units. In addition, outcomes were sought for infants who received NEC surgery or who died at the four stand-alone paediatric surgical centres which do not use the BadgerNet neonatal EPRs (Great Ormond Street, Sheffield, Alder Hey and Birmingham Children's Hospitals). Here, we describe the steps taken to identify and verify infants with severe NEC.

Step 1: data extraction

The variables extracted from the NNRD comprised static data (discharge/died status, cause of death, whether or not the post-mortem-confirmed NEC); daily data (NEC treatment: medical or surgical); episodic data (gastrointestinal diagnoses, discharge diagnoses, procedures during stay); AXR screen (whether or not

surgery was required, whether surgery was required but the patient was too sick, whether or not the surgery-confirmed NEC, whether or not histology-confirmed NEC).

Step 2: data verification

We identified infants from the following EPR locations using the predefined field listed:

Discharge diagnoses

- 'Necrotising enterocolitis – perforated'
- 'Necrotising enterocolitis – proven (on X-ray or at surgery)'
- 'Necrotising enterocolitis – confirmed'
- 'Cause of death includes necrotising enterocolitis'.

Abdominal X-ray screen

- 'Laparotomy-confirmed NEC'
- 'Histology-confirmed NEC'
- 'Post-mortem-confirmed NEC'
- 'Procedures screen'
- 'Laparotomy approach NEC'
- 'Colectomy and ileostomy NEC'
- 'NEC surgery performed'.

Combinations

- 'Necrotising enterocolitis' in 'Discharge diagnoses' and 'Laparotomy' in 'Procedures'
- 'Necrotising enterocolitis' and 'died' in discharge status field.

Step 3

The study lead at each neonatal unit or paediatric stand-alone hospital where the surgery was performed or where the infant died was contacted to verify data. The following data were verified:

- Gestation weeks and days.
- Birthweight.
- Did infant die in neonatal unit? (Yes/no.)
- Age of infant at surgery for NEC (if applicable).
- Was laparotomy performed? (Yes/no/required but too sick.)
- Did visualisation confirm NEC? (Yes/no.)
- Did histology confirm NEC? (Yes/no.)
- Was a peritoneal drain inserted? (Yes/no.)
- Was a post-mortem done? (Yes/no.)
- If yes, did the post-mortem confirm NEC? (Yes/no.)
- Was NEC a cause of death? (Yes/no).

Other data extraction from the National Neonatal Research Database: data management

We extracted the following data for all babies from the NNRD: booking network, gestational age in completed weeks and days, birthweight, fetus number, antenatal steroids, maternal pyrexia in labour, whether or not mother received antibiotics, mode of delivery, maternal chorioamnionitis and maternal infection. We calculated birthweight SDS, standardised for sex and gestational age from UK World Health Organization (WHO) reference data.[113] We considered a birthweight SDS of < -4 or > 4 to be erroneous, and we treated these as missing values.

Analyses

Incidence and absolute numbers of cases of severe necrotising enterocolitis

We expected all very preterm infants (born before 32 weeks' gestation) to be admitted to a neonatal unit and, hence, derived the population incidence of severe NEC for this group. To ensure that this is a valid assumption, we compared the numbers of infants for whom data were present in the NNRD against the equivalent number obtained from the most recently available data from the Office for National Statistics (ONS) as this contains complete birth registrations. For preterm infants born before 32 weeks' gestation, we report the incidence of severe NEC (95% CIs) per 1000 infants admitted to neonatal care. In contrast, not all infants born at a gestational age of $\geq$ 32 weeks will necessarily be admitted to a neonatal unit and so, for this group, we only present the absolute numbers of cases of severe NEC. For infants who received surgery for NEC, we will report the median and interquartile ranges (IQRs) for postnatal age and postmenstrual age for the day of surgery by gestational bands.

Factors associated with severe necrotising enterocolitis in infants born at a gestational age of < 32 weeks

For preterm infants, we compared baseline characteristics for gestational age in completed weeks, birthweight SDS, fetus number, antenatal steroids, maternal pyrexia in labour, whether or not mother received antibiotics, mode of delivery, maternal chorioamnionitis and maternal infection between those with severe NEC and those who did not develop the condition. We used the chi-squared test, the application of Yate's correction and the *t*-test, as appropriate. We performed a stepwise multiple logistic regression; variables found to be significantly associated with severe NEC in the univariate logistic regression analysis ($p < 0.15$) were considered candidate variables for the multivariable logistic regression model. For the final multivariable model, we retained only variables that were significant independent predictors of NEC. We further investigated the effect of retaining and excluding antenatal steroid exposure in the final multivariable model. The level of statistical significance for all analyses was set at $p < 0.05$ using two-tailed comparisons.

Variation in the incidence of severe necrotising enterocolitis among infants born before 32 weeks' gestation

For preterm infants, we report the incidence of severe NEC (95% CIs) per 1000 infants admitted to neonatal care by network of booking. We assessed whether or not there was variation in the incidence of severe NEC at network level in two ways. First, we compared the rate in each neonatal network against the average incidence across England. Second, we compared each individual network against a reference network. We selected the reference as the network contributing the largest number of infants, to minimise the standard errors.

In the first approach, we used methods analogous to those used to calculate standardised mortality ratios (SMRs), assigning infants to the neonatal network of booking. We calculated the standardised severe NEC ratio (SNR) by dividing the observed number of severe NEC cases by the expected number of severe NEC cases. For the unadjusted SNR, the expected number of severe NEC cases was calculated as the total number of infants in the booking network multiplied by the overall severe NEC rate across England. For the adjusted SNR, the expected number of severe NEC cases was calculated by first estimating the probability of severe NEC for each infant using logistic regression, and adding up the probabilities to obtain the expected number of severe NEC cases in each network. The 95% CIs for the SNR were calculated using Byar's approximation[114] with correction for multiple testing, controlling the false discovery rate at 5%.[115] Variables included in the logistic regression to estimate the probability of severe NEC for each infant were gestational age (in completed weeks), birthweight SDS and antenatal steroids.

Funnel plots were used to illustrate the SNR at network level. The prediction limits were drawn corresponding to 95% (SD 2) 99.8% (SD 3) from the target SNR of 1, assuming the observed severe NEC rates follow a Poisson distribution. The limits were adjusted for multiple testing controlling the false discovery rate of 5%.

We expect 5% of networks to lie outside the 95% prediction limits, and 0.2% to lie outside the 99.8% prediction limits. For the second approach, we used multivariable logistic regression adjusted for antenatal steroid exposure and variables independently associated with severe NEC to derive the odds ratio (OR) for severe NEC in each network relative to the reference network. We corrected for multiple testing using the Bonferroni method. All *p*-values reported are two sided. Statistical analyses were performed in SAS® software, version 9.3 (SAS Institute Inc., Cary, NC, USA).

Data validation

We compared data from the NNRD with data from paper medical notes as part of a project evaluating a quality improvement project conducted by the East of England Neonatal Networks.[116] In brief, between 2011 and 2013, we assessed and fed back completeness and accuracy of NNRD data to participating neonatal teams involving 17 neonatal units. The study lead at each neonatal unit extracted a selection of data items from medical notes for two randomly selected infants discharged in the previous month. These data were sent to the NDAU and compared with NNRD data.

Results

Incidence and numbers of cases of severe necrotising enterocolitis

We extracted data on a total population cohort of 118,073 infants (*Figure 5*). We identified 531 infants with severe NEC. *Table 4* shows the proportion of infants with severe NEC who had surgery and survived to discharge from neonatal care, who had surgery and died, and who died without surgery, by gestational age bands. Of the total number of infants with severe NEC 79.7% (423/531) had surgery; of those who had surgery 32.9% (139/423) died; 20.3% (108/531) of infants with severe NEC had died without surgery (*Figure 6*).

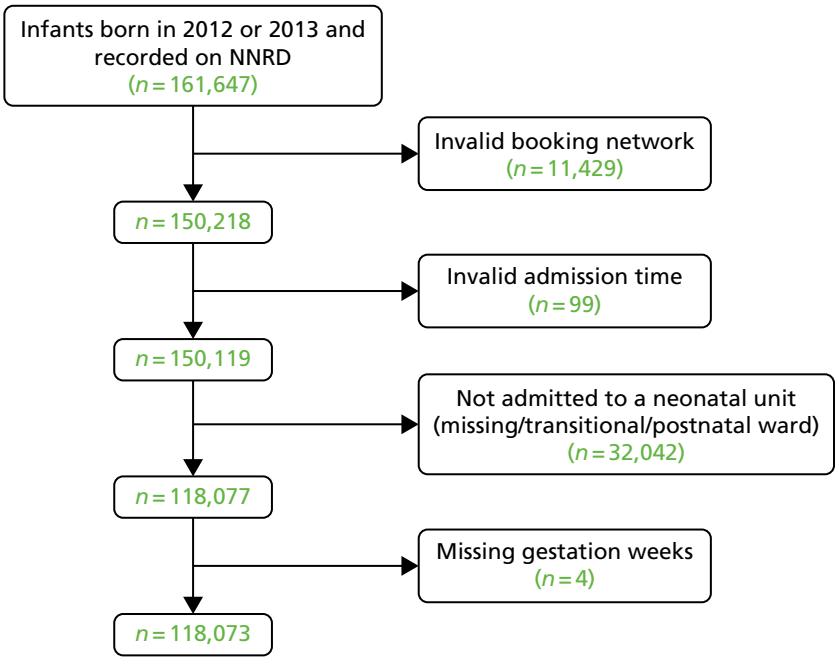


FIGURE 5 Flow chart showing derivation of the study population.

TABLE 4 Severe NEC, surgery and survival, by gestational age bands

	Infants with severe NEC ($N = 531$), *n*			
Gestation (completed weeks)	Surgery and survived	Surgery and died	Died without surgery	Total, *n*
22^{+0} to 25^{+6} ($n = 2035$)	101	58	41	200
25^{+0} to 28^{+6} ($n = 4331$)	97	56	42	195
29^{+0} to 31^{+6} ($n = 8312$)	42	12	13	67
32^{+0} to 36^{+6} ($n = 42,169$)	29	11	12	52
$\geq 37^{+0}$ ($n = 61,226$)	15	2	0	17
All gestations ($n = 118,073$)	284	139	108	531

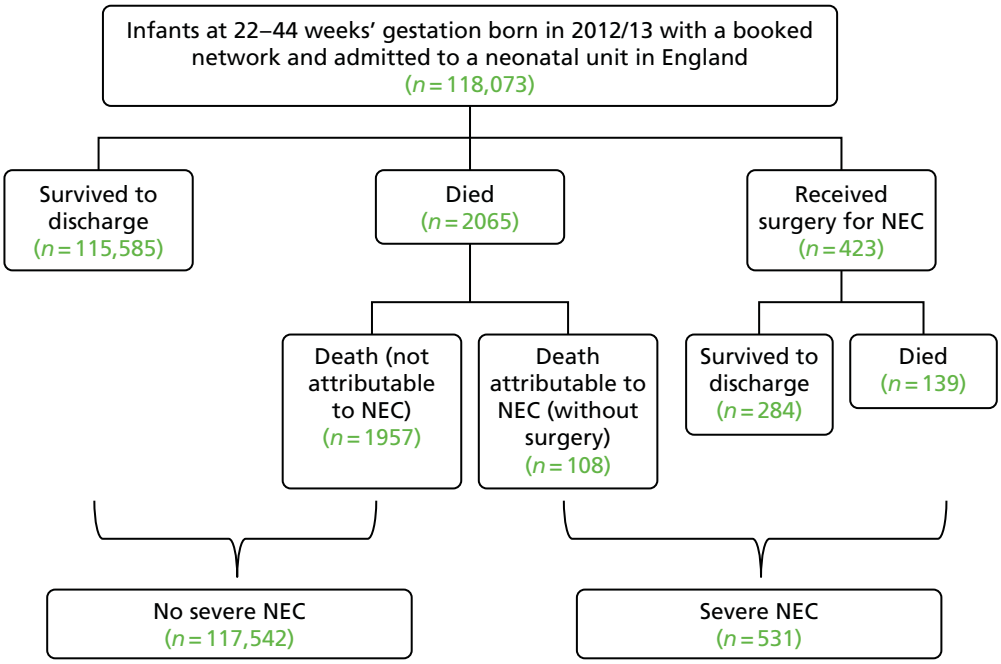


FIGURE 6 Flow chart showing the population with severe NEC.

Comparing ONS data with NNRD data from 2012 shows that of the infants who were born alive in England between the gestational ages of 25 and 31^{+6} weeks, 96–99% were admitted to a neonatal unit (*Figure 7*); the corresponding figures are 92% and 70% of infants born at 24 and at 23 weeks' gestation, respectively. The percentage of infants who were admitted to a neonatal unit starts to fall after 32 weeks' gestation. Therefore, as we have population data for infants born before 32 weeks' gestation, we present the incidence for these infants; but for infants who were born after 32 weeks' gestation, we present only the raw numbers.

The incidence of severe NEC was inversely related with gestational age ($p < 0.001$, test for trend). The highest incidence of severe NEC occurred in infants born at 24 weeks' gestation. The incidence per 1000 infants for all infants born before 32 weeks' gestation is 31.5 per 1000 (95% CI 28.7 to 34.3 per 1000), and by gestational age is as follows: 23 weeks, 96.4 (95% CI 66.8 to 125.9); 24 weeks, 112.4 (95% CI 90.0 to 134.8); 25 weeks, 86.9 (95% CI 68.4 to 105.5); 26 weeks, 58.9 (95% CI 45.1 to 72.7); 27 weeks, 41.0 (95% CI 30.6 to 51.3); 28 weeks, 39.0 (95% CI 30.1 to 48.0); 29 weeks, 10.9 (95% CI 6.5 to 15.3); 30 weeks, 9.5 (95% CI 5.8 to 13.2); and 31 weeks, 5.3 (95% CI 2.9 to 7.7). *Figure 8* shows that there is a sharp decline in the incidence of severe NEC at 29 weeks' gestation.

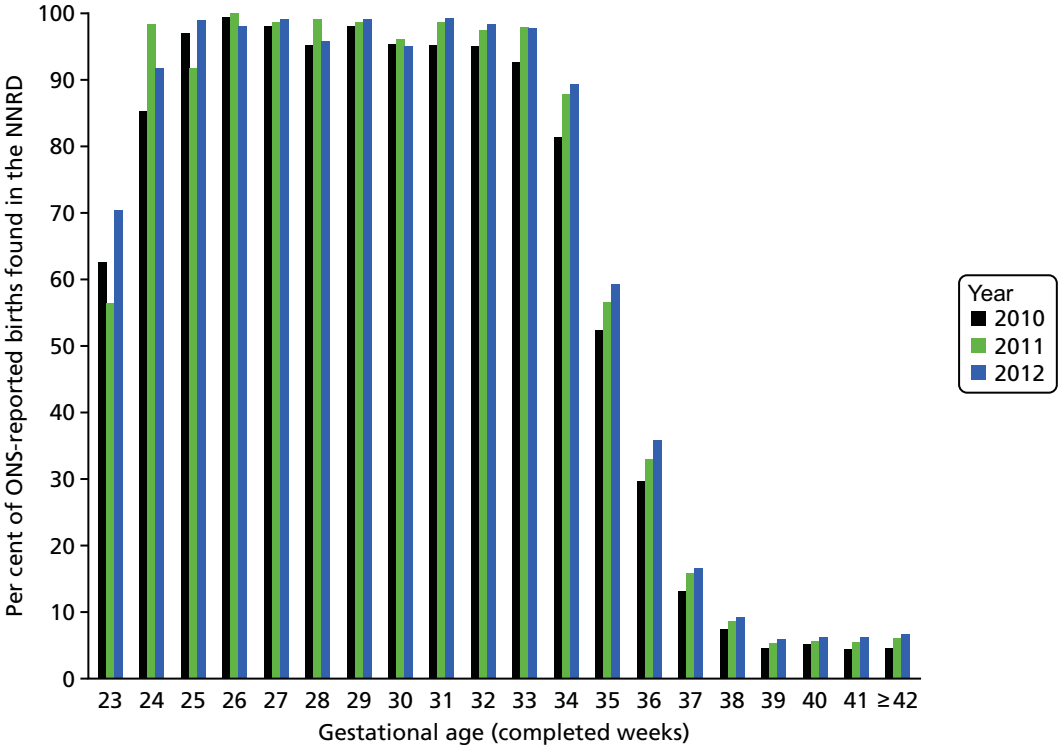


FIGURE 7 Percentage of ONS-reported live births in the NNRD.

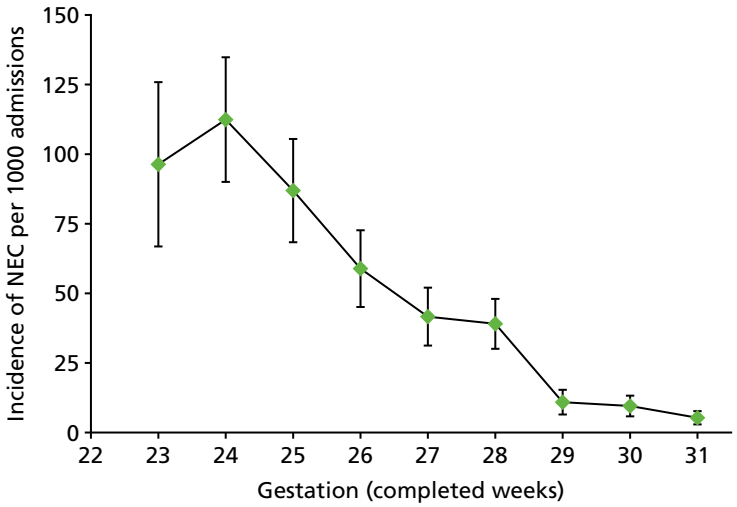


FIGURE 8 Incidence of severe NEC, infants born before 32 weeks' gestation. Incidence derived from raw data; bars are 95% CI; gestation category labelled 23 weeks represents 22 and 23 weeks (14 infants born at 22 weeks, none with severe NEC).

Of the 531 infants with severe NEC, 462 (87%) were born before 32 weeks' gestation, of whom 366 received surgery. *Table 5* shows the postnatal and postmenstrual age at NEC surgery for infants < 32 weeks' gestation. There is an inverse relationship between gestational age at birth and postnatal age at surgery (log-rank test < 0.001). The most immature infants born, before 26 weeks' gestation, receive surgery for NEC around the third to fourth week of life; in contrast infants who are born at 30–31 weeks' gestation undergo surgery in the second week of life.

TABLE 5 Postnatal and postmenstrual age at NEC surgery (infants born before 32 weeks' gestation)

Gestational age (weeks)	Number of infants	Age (days) at NEC surgery (median, IQR)	Postmenstrual age (completed weeks)
23	29	25 (12–37)	27 (25–29)
24	66	20 (12–38)	27 (26–30)
25	64	31 (12–53)	29 (26–32)
26	55	29 (15–39)	30 (28–31)
27	41	13 (9–31)	29 (28–31)
28	57	24 (14–36)	31 (30–33)
29	19	18 (9–32)	32 (30–33)
30	19	11 (7–25)	32 (31–33)
31	16	10 (8–17)	32 (32–33)
Total	366	22 (11–37)	30 (27–32)

Factors associated with severe necrotising enterocolitis for infants born before 32 weeks' gestation

We compared the patient characteristics of the 462 preterm infants who developed severe NEC against the 14,216 preterm infants without severe NEC. Infants who developed severe NEC were more immature, with a mean gestational age of 26.2 weeks compared with 28.5 weeks ($p < 0.001$). Gestational age, birthweight, birthweight SDS, fetus number, whether or not the mother received antibiotics in labour, and mode of delivery were significantly different between the two groups (see *Appendix 1, Table 53*).

Using univariate logistic regression for infants born before 32 weeks' gestation ($n = 14,678$), we identified fetus number, whether or not the mother had received antibiotics, mode of delivery, gestational age (completed weeks) and birthweight SDS to be significantly associated with severe NEC. After multivariable logistic regression analysis, only gestational age and birthweight SDS were significant independent predictors of severe NEC. We investigated the effect of including antenatal steroid exposure in the multivariable model and found that the conclusions were unchanged. As antenatal steroids are an important determinant of survival, we chose nonetheless to include this is our model. *Appendix 1, Table 54*, shows parameters for the final multivariable model that include gestational age, birthweight SDS and antenatal steroid exposure.

Incidence of severe necrotising enterocolitis by neonatal network

The incidence of severe NEC per 1000 infants born before 32 weeks' gestation, by network of booking, is shown in *Table 6*. The highest incidence was 47.4 cases per 1000 babies (95% CI 32.5 to 62.4); the lowest incidence was 19.8 (95% CI 9.1 to 30.4). The rate of severe NEC in England is 3.15% (462/14,678). We calculated adjusted NEC rates (adjusted for gestational age, birthweight SDS, antenatal steroid exposure) for infants born before 32 weeks' gestation. The adjusted funnel plot in *Figure 9* illustrates that the network-level variation in severe NEC rates is consistent with the pattern we would expect given that the population rate of NEC is 3.15% (i.e. we identified no strong evidence that any neonatal network differed from the mean national incidence of severe NEC).

For comparison between networks, we selected a reference network with the largest number of infants born before 32 weeks' gestation (London North East with 1014 infants) to minimise the standard errors. We identified a statistically significant difference in the incidence of severe NEC between the reference network and all the other 22 networks combined (overall p-value of < 0.001). We therefore proceeded to perform pairwise comparisons between each network and the reference network (see *Appendix 1, Table 55*). Following correction for multiple testing, we found no statistically significant difference in the incidence of severe NEC between any network and the reference network.

TABLE 6 Incidence of severe NEC among infants born before 32 weeks' gestation by network of booking

Network of booking	Total number of infants ($n = 14,678$)	Severe NEC cases, *n* (%)	Incidence of severe NEC per 1000 infants	95% CI
Bedfordshire and Hertfordshire	391	8 (2.05)	20.5	6.4 to 34.5
Cheshire and Merseyside	635	24 (3.8)	37.8	23.0 to 52.6
Eastern	800	37 (4.63)	46.3	31.7 to 60.8
Greater Manchester	871	35 (4.02)	40.2	27.1 to 53.2
Kent	470	15 (3.19)	31.9	16.0 to 47.8
Lancashire and South Cumbria	409	11 (2.69)	26.9	11.2 to 42.6
London North Central	421	9 (2.14)	21.4	7.6 to 35.2
London North East	1014	30 (2.96)	29.6	19.2 to 40.0
London North West	755	31 (4.11)	41.1	26.9 to 55.2
London South East	600	23 (3.83)	38.3	23.0 to 53.7
London South West	418	13 (3.11)	31.1	14.5 to 47.7
Midlands Central	780	37 (4.74)	47.4	32.5 to 62.4
Midlands South West	797	17 (2.13)	21.3	11.3 to 31.4
Midlands North	649	18 (2.77)	27.7	15.1 to 40.4
North Trent	623	20 (3.21)	32.1	18.3 to 45.9
Northern	773	19 (2.46)	24.6	13.7 to 35.5
Peninsula South West	344	7 (2.03)	20.3	5.4 to 35.3
South Central (North)	589	18 (3.06)	30.6	16.7 to 44.5
South Central (South)	639	14 (2.19)	21.9	10.6 to 33.3
Surrey and Sussex	591	26 (4.4)	44.0	27.5 to 60.5
Trent	520	15 (2.88)	28.8	14.5 to 43.2
Western	658	13 (1.98)	19.8	9.1 to 30.4
Yorkshire	931	22 (2.36)	23.6	13.9 to 33.4

Comparison of National Neonatal Research Database against East of England medical notes

There was high agreement (> 95%) for sex, gestational age, birthweight and discharge destination (see *Results*). Agreement for the length of parenteral nutrition improved from 80–90% to > 90% over the 26-month study duration ($p < 0.029$). The agreement for central line days was consistently > 80%, with no significant change over time. Agreement on type of discharge feed improved over time, from 50–60% to 70–80% ($p < 0.009$). Agreement for numbers of AXR and blood cultures was consistently low, at around 50%. Further results of this work are available in the published manuscript.[116]

Conclusions

We identified that over a complete 2-year period, 531 babies admitted to neonatal units in England, required surgery for NEC and/or died as a result of the condition. Of these, one-fifth died before they could receive surgery. Of note, we identify no strong evidence of statistically significant variation between neonatal networks.

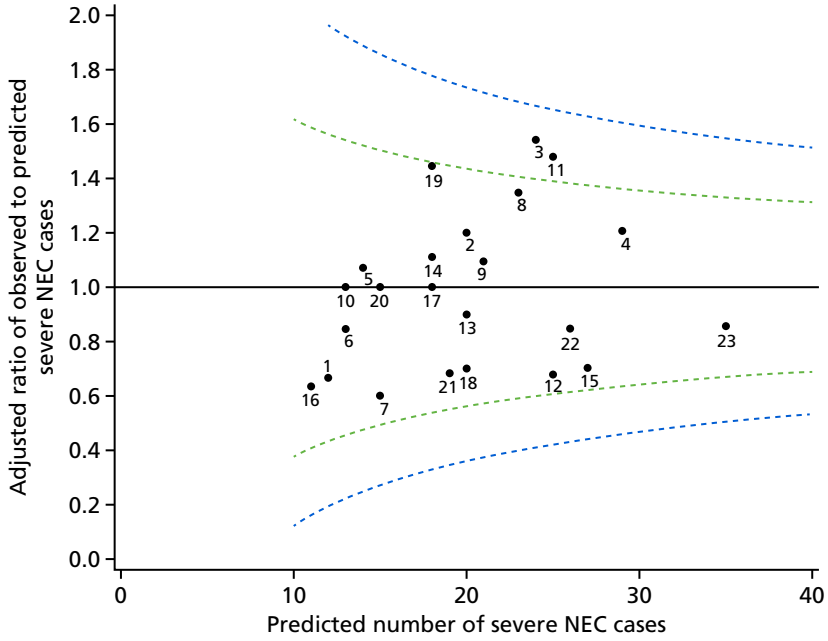


FIGURE 9 Funnel plot for infants born before 32 weeks' gestation, showing percentage of observed to predicted NEC cases estimated from the multivariable logistic regression model (adjusted for gestation, birthweight SDS and antenatal steroids) relative to average percentage of NEC cases in England. 1 = Bedfordshire and Hertfordshire; 2 = Cheshire and Merseyside; 3 = Eastern; 4 = Greater Manchester; 5 = Kent; 6 = Lancashire and South Cumbria; 7 = London North Central; 8 = London North West; 9 = London South East; 10 = London South West; 11 = Midlands Central; 12 = Midlands South West; 13 = Midlands North; 14 = North Trent; 15 = Northern; 16 = Peninsula South West; 17 = South Central (North); 18 = South Central (South); 19 = Surrey and Sussex; 20 = Trent; 21 = Western; 22 = Yorkshire; and 23 = London North East.

This study has several strengths. To our best knowledge, it is the largest complete population-based study of NEC. We obtained data from the NNRD in turn derived from the neonatal EPRs and, thus, the data recorded as part of this study remain a permanent part of the infant's clinical record. Through comparison with ONS data, we show that the NNRD contains data from the majority of infants born alive below 32 weeks' gestation. Uniquely, unlike previous studies in England and elsewhere,[117,118] we engaged the participation of every neonatal unit in the country. The use of the NNRD and the complete population coverage despite frequent transfers between neonatal units has meant that we were able to ascertain the final outcome for every baby. Incomplete ascertainment have been a common limitation of other studies. This has also allowed us to report the incidence of NEC by neonatal network. This is an important consideration; in a networked-based model of care, the delivery of care is a collaborative responsibility and clinical outcomes are in large part attributable to the network rather than to individual neonatal units. Previous studies that have only included tertiary centres are severely limited by selection bias, first, because these neonatal units are likely to care for only the sickest and most complex infants and, second, because of the omission of infants that die before transfer to a tertiary centre. Our study also provides complete ascertainment of all cases of severe NEC, regardless of birthweight or gestational age. We applied a stringent, consistent definition for NEC and confirmed each case individually with clinical teams to minimise contamination from diagnoses such as spontaneous intestinal perforation, dysmotility, feed intolerance, septic ileus or other ambiguous abdominal pathologies.

A limitation of our study is that we did not attempt to determine the incidence of less severe NEC. This is because the value of attempting to do so is open to question given the absence of an agreed case definition, the high degree of diagnostic subjectivity, and the poor sensitivity and specificity of many indicative clinical signs. NEC remains a clinical diagnosis using radiographic and clinical findings and there is increasing recognition of the difficulties caused by inconsistencies in the application of these criteria. This problem requires the identification of reliable biomarkers of NEC or the consistent application of an agreed case-definition purely for surveillance purposes.[118,119]

Consistent with other studies, we found the incidence of NEC to be inversely related to gestational age[120,121] other than for a lower incidence at 22 and 23 weeks' gestation. As the median time to onset of NEC in babies born before 32 weeks' gestation is 22 days, this is likely to be a consequence of the higher proportion of the most immature infants dying in the early neonatal period. We confirm a strong independent association between severe NEC and both gestational age and birthweight SDS. Previous studies are inconsistent regarding other risk factors for NEC including infant sex,[117,122,123] race,[120,124] mode of delivery,[102,104,125] antenatal steroids[102,104–110] and prolonged rupture of membranes.[102,126]

The lack of strong evidence of variation in the incidence of NEC at neonatal network level in England contrasts with widely held beliefs. Although we identified two neonatal networks with an incidence of severe NEC falling outside the upper 95% control limits of the funnel plot, this is not incompatible with what would be expected by chance alone. Adjusting for gestational age, birthweight and antenatal steroids did not alter our conclusions. Other population-based incidence data are limited. The EPICure 2 study, a population-based study, reported that 8% (95% CI 6% to 9%) of infants born at 22–26 weeks' gestation in England in 2006 that survived to discharge received a laparotomy for NEC,[127] a figure comparable to the equivalent figure of 6% (138/2275) that we found. Our data are also broadly similar to other published studies from Canada,[128] Australia,[129] and the USA,[130] but these largely employed birthweight- or neonatal unit-based selection criteria.

Implications for health care

A clear implication for health care is that the lack of significant variation in NEC between networks despite differences in many clinical practices, such as the use of probiotics and pasteurised human donor milk, the time to commence enteral feeds and the rate of enteral feed advancement, that are widely believed to affect risk, justifies caution in the imposition of inadequately evidence-based guidelines or quality-improvement approaches. The imposition of poorly evidence-based guidance is often justified on the basis that this provides consistency in care; an alternative conclusion is that this places patient safety at risk by exposing all patients, rather than just some, to a potentially less beneficial treatment approach.[131] A corollary is the paramount importance of assisting health-care staff to support the delivery of randomised controlled trials that seek to reduce uncertainties in everyday care practices.

Gestational age-specific population denominators are required to derive population incidences. In the UK, the ONS is the gold-standard source of population denominators based on birth registrations. The data available in the NNRD are timely, as these are downloaded from EPRs at quarterly intervals. In contrast, we were able to obtain ONS data only up to 2012. Improving the timeliness of ONS data would go some way towards improving the scope and utility of health-care evaluations.

Research recommendations

Necrotising enterocolitis is a cardinal cause of morbidity and mortality in the most immature infants. The low incidence rates require national and international collaboration to test preventative strategies in adequately powered randomised controlled trials. An important national advance would be the development of efficient approaches based on routinely recorded data. Implementation of EPR or database trials internationally requires agreement of core data sets that include predefined outcome measures and ancillary variables. Clinical engagement and contribution to reliable data entry into the neonatal EPRs was paramount to this study.

Chapter 3 Clinical outcomes assessed using the National Neonatal Research Database: mortality of very preterm babies admitted to NHS neonatal units

Abstract

Background: Although preterm survival analyses are widely used a health indicator, they have generally been based on historical data, which limits their relevance.

Aims: To evaluate (1) survival trends in relation to geographical region and socioeconomic status for infants born 22^{+0}–31^{+6} weeks' gestation and admitted to neonatal units in England over 2008–14, and (2) variations between neonatal networks in mortality to discharge over 2013–14.

Methods: We used logistic regression to model survival probability, joinpoint regression for trend analyses and multiple imputation for missing outcome and covariate data. We calculated unadjusted, risk-adjusted and gestation-specific survival.

Findings: (1) The cohort comprised 50,112 infants. There was an increase in survival to discharge from 88% in 2008 to 91.3% in 2014 [adjusted annual percentage change (APC) 0.46% (95% CI 0.3% to 0.62%); $p < 0.001$] and 28 days [2008: 91.4%; 2014: 93.5%; APC 0.27% (95% CI 0.11% to 0.44%); $p = 0.002$] with the greatest improvement for infants born at 22^{+0}–23^{+6} weeks' gestation (6.03%, 95% CI 2.47 to 3.53%; $p = 0.002$). Crude survival was lower for infants from the most deprived quintile than from the least deprived [89.5% (95% CI 88.9% to 90.1%) vs. 91.1% (95% CI 90.2% to 92.1%)] and it was lower in the Midlands and the East of England than in London [89.3% (95% CI 88.6% to 89.9%) vs. 91.1% (95% CI 90.3% to 91.8%)]. Regional variation remained after adjustment for socioeconomic status. (2) We analysed data on 15,255 infants. We identified no strong evidence that any neonatal network differed significantly from the national mean.

Interpretation: Analysis of population data over time is required to identify variation unlikely to be due to chance. Continued national improvement in very preterm survival masks significant north–south variation that is not explained by population characteristics.

Background

Preterm birth is now the primary cause of neonatal death and is associated with risks to health and well-being into childhood and beyond. The preterm birth rate is rising worldwide, and this growing population presents a substantial public health issue. Advances in obstetric and neonatal care have resulted in improved survival of preterm infants over the last few decades. In England, the EPICure study found that survival to discharge from hospital among admitted babies born between 22^{+0} and 25^{+6} weeks' gestation increased from 40% in 1995 to 53% in 2006.[124] Data such as these have been invaluable, but are now several years out of date. Evaluation of survival is a widely used health indicator. Keeping pace with changes in preterm survival is important for counselling parents and planning clinical services, but undertaking population-based studies is challenging.

The NNRD holds extracts from point-of-care EPRs for infants admitted to neonatal units in the UK, providing an opportunity to obtain precise, up-to-date estimates of neonatal survival.

Aims

We aimed to use data from the NNRD to conduct two analyses relating to infants born at 22^{+0} to 31^{+6} weeks' gestation and admitted to neonatal units in England.

First, we describe trends in survival between 2008 and 2014, evaluate regional variation and relationship to socioeconomic deprivation, and compare survival rates with those from the EPICure studies. Secondary aims were to examine changes in the time of death, develop a statistical model to predict the probability of survival to discharge for a given set of infant characteristics, and cross-validate NNRD data with data from the ONS.

In England, neonatal specialised care is delivered through Operation Delivery Networks (ODNs). Our second aim was to present unadjusted and adjusted SMRs for each ODN for admissions to neonatal specialised care over 2013–14.

Methods

We obtained data for infants born between 22^{+0} and 31^{+6} weeks' gestation between January 2008 and December 2014, admitted to a neonatal unit in England. We excluded infants with a birthweight SDS that was > 4 SDs from the gestation and sex-specific mean (UK-WHO preterm standards), as we considered these likely to be erroneous.[132]

We described the following population characteristics: gestational age (based on ultrasound or best obstetric estimate), birthweight, birthweight SDS (UK-WHO preterm growth charts), small for gestational age (SGA), singleton/multiple pregnancy, administration of any antenatal steroids (complete or incomplete course), vaginal/caesarean delivery, maternal age, maternal ethnicity, any smoking during pregnancy, the Index of Multiple Deprivation (IMD) 2010[133] quintile [based on rank of lower-layer super output area (LSOA)] by year of birth and for the cohort overall.

The primary outcome was survival to discharge from neonatal care. Secondary outcomes were survival to 28 days to facilitate comparison with other neonatal survival data, and time of death in days. The outcomes were determined by the discharge record for the last episode of care. If the last discharge was a transfer to another location for further clinical care and no subsequent data were available, the outcome was coded as missing.

To reduce the number of missing data, we attempted to link infants with missing outcomes to the ONS–HES mortality data set.[134] The NDAU has requested permission from all neonatal units to receive infants' NHS numbers for the purposes of data linkage following approval by the Confidentiality Advisory Group of the Health Research Authority; at the time of this study, permission had been received from 159 neonatal units (93%). Deterministic linkage was carried out on the basis of NHS number and HES ID, when available. If the NNRD record could be linked to a death in the ONS–HES mortality data set, this information was used for the 28-day survival outcome but not for the survival to discharge outcome as we were unable to determine whether or not the infant was still in neonatal care at the time of death.

All data extraction and linkage was carried out using SAS.

Statistical analysis

Direct standardisation was used to control for population differences, as this permits comparison of rates over time, unlike indirect standardisation. Survival rates were directly standardised for risk. Infants were grouped into 10 categories based on the probability of death predicted by the regression model. The thresholds for the 10 categories were calculated such that each group had an equal number of predicted deaths. The directly standardised rate for each period is the weighted sum of the survival rates in each risk

group, with the weights determined by the proportion of infants in each risk group in the whole study cohort. Sensitivity analyses were performed using 5 and 15 risk categories.

Prediction model

We used multivariable logistic regression to model the probability of death before discharge from neonatal care. Variables included in the regression model were gestational age, birthweight, sex, multiplicity of pregnancy (singleton/multiple) and administration of any antenatal steroids (no/yes). Previous research has shown these variables to be significant predictors of mortality. We used spline terms to model gestational age and birthweight and their interaction, using simpler functions if the fit was comparable. We included interactions between multiple birth, gestational age and birthweight, as the influence of these variables on survival is different for singleton and multiple pregnancies. As outcomes for babies from the same pregnancy are likely to be correlated, we used generalised estimating equations (GEEs) to account for the lack of independence.

We excluded observations if > 1% were missing. Otherwise, missing outcome and covariate data were imputed 25 times using multiple imputation with chained equations based on all other variables in the prediction model. Sensitivity analysis using complete cases was performed.

We carried out modelling using Stata® 12 (StataCorp LP, College Station, TX, USA). We present results as regression coefficients with standard errors (SEs). We produced isosurv graphs[135] using the ggplot2 package in R 2.13.2 (The R Foundation for Statistical Computing, Vienna, Austria) to show contours of survival probability by gestation and birthweight. We added birth year to the model used to generate the graphs so that predictions would be calibrated to the most recent year.

Model performance

We checked discrimination (ability to differentiate between babies that survived and those who died) by calculating the area under the receiver operating characteristic curve (AUC). We calculated the Brier score as a measure of overall model fit (range 0–1; 0 means better fit). We checked calibration (comparability of actual and predicted survival) using Cox's calibration in gestational age subgroups ($< 28^{+0}$ weeks and $\geq 28^{+0}$ weeks). If the model predicts perfectly across all survival probabilities, the intercept α will equal 0 and the slope β will equal 1. As model performance was assessed on the same data set used to build the model, these measures were corrected for optimism using 200 bootstrap samples.

Comparison with existing models

We compared model performance with three previously published models to predict death before discharge in preterm babies admitted to neonatal units: (1) the Clinical Risk Index for Babies II (CRIB II), a frequently used model based on data from infants born in 1998–9 and admitted to 35 UK neonatal units; (2) a more recent UK model using data from neonatal units in the East Midlands and Yorkshire region (The Neonatal Survey); and (3) the National Institute of Child Health and Development Neonatal Research Network model based on infants born at 22^{+0}–25^{+6} weeks' gestation in 1998–2003 who were admitted to 19 hospitals in the USA. The relevant subset of the NNRD cohort was used to match the population characteristics of the comparator model. The predicted survival rate, AUC, Brier score and the intercept and slope from Cox's calibration were compared between the NNRD and comparator models.

Time trend analysis

For survival to discharge and to 28 days, trends over time were analysed using joinpoint regression applied to quarterly periods using Joinpoint 4.2.0 software (National Cancer Institute, Bethesda, MD, USA). This method allows detection of changes in trend when the number and location of the changes are unknown. Rates were log-transformed so trends are presented as annual percentage change, which is the annual rate of change of the survival rate. Heteroskedastic errors were allowed using weighted least squares, with weights inversely proportional to the variance. As the number of contributing neonatal units increased over time, we repeated all time trend analyses using data from complete neonatal networks only as a sensitivity analysis. Differences in the time of death across years were tested using quantile regression.

Variation by region and Index of Multiple Deprivation quintile

We restricted this analysis to data from 2011 onwards as lower population coverage in earlier years may bias regional estimates. Infants were assigned to one of the four NHS commissioning regions (London, Midlands and East of England, North of England and South of England) based on LSOA of mother's residence. Crude and directly standardised rates of survival to discharge and associated 95% CIs were calculated for each region. Trends in crude survival were estimated and compared for each region using joinpoint regression; standardised trends by region were not calculated because of the low quarterly numbers in each risk group.

To examine whether or not survival experiences differ by socioeconomic deprivation, crude and directly standardised rates of survival to discharge were calculated for the highest and lowest IMD quintile and compared using RR. NHS commissioning region (categorical) and IMD decile (continuous) were added in the risk adjustment model to test whether or not there was evidence of residual variation across regions.

Validation with Office for National Statistics data

For validation purposes, the number of deaths before 28 days of infants born in England and Wales at 22 to 31 weeks' gestation in 2012 was compared in the NNRD data (denominator is neonatal unit admissions) and published ONS data (denominator is live births).[136] Data were compared for 2012 only, owing to comparability of sources.

Comparison with previous national data

The EPICure studies examined survival and morbidity outcomes for all infants born 22^{+0} to 25^{+6} weeks' gestation during 10 months of 1995 in the UK (EPICure) and all infants born at 22^{+0} to 26^{+6} weeks' gestation in 2006 (EPICure 2) in England.[124,137] Survival outcomes were reported separately for infants admitted to neonatal units, giving a population comparable to the NNRD cohort. We compared data on survival to discharge of admitted infants born at 22^{+0} to 25^{+6} weeks' gestation in England in EPICure, EPICure 2 and this study cohort using joinpoint regression to see whether or not the rate of improvement has changed.

Mortality by Operational Delivery Network

We obtained data from the NNRD on infants born in 2013 and 2014 at $\leq 31^{+6}$ weeks' gestation and admitted to neonatal care for whom the neonatal network of booking was known. Death was defined as death before discharge from neonatal care. We used multiple imputation (applying the mi routine in Stata, version 12) to impute missing outcome and covariate data; analysis of complete-case data was also performed. We present standardised mortality ratios (SMRs) for each ODN, assigning infants to the network of booking. Crude and adjusted SMRs are presented for 2013 and 2014 combined.

The SMR was calculated as the observed number of deaths divided by the expected number of deaths. The observed number of deaths was averaged over the imputed data sets so that infants with missing outcomes were included. For the unadjusted SMR, the expected number of deaths was calculated as the total number of infants multiplied by the overall mortality rate across all networks. For the adjusted SMR, the expected number of deaths was calculated by estimating the probability of death for each infant using logistic regression, and adding up the probabilities to obtain the expected number of deaths. The 95% CIs for the SMRs were calculated using Byar's approximation[138] with correction for multiple testing, controlling the false discovery rate at 5%.[115]

The logistic model used to estimate the probability of death was derived using data from babies born at $\leq 31^{+6}$ weeks' gestation in England in 2008–14. Multivariable logistic regression was used with survival to discharge from neonatal care as the outcome. Predictor variables were gestational age (typically the best obstetric estimate from antenatal ultrasound), birthweight, sex, multiplicity of pregnancy (singleton/multiple), administration of any antenatal steroids (no/yes). Spline terms were used to model gestational age and birthweight and their interaction. The association between gestation and mortality is known to be different among singletons and multiples;[139] a similar interaction effect has been shown for birthweight.[140] Interactions between multiple birth and gestational age/birthweight terms were therefore included.

As outcomes for infants from the same pregnancy are likely to be correlated, we used GEEs to account for the lack of independence. We used funnel plots to illustrate the variation in SMR. Funnel plot limits were drawn corresponding to 2 and 3 standard deviations (SDs) from the target SMR of 1, assuming the observed deaths follow a Poisson distribution. The limits were adjusted for multiple testing[141] controlling the false discovery rate at 5%.

Results

Population

Data were available for 71% of neonatal units from the beginning of 2008, 80% in 2009, 86% in 2010, 97% in 2011, 99% in 2012, and 100% in 2013 and 2014. There were 50,467 infants born between January 2008 and December 2014 at 22^{+0} to 31^{+6} weeks' gestation whose mothers were resident in England and who were admitted to a contributing neonatal unit. Thirty-eight infants were excluded owing to implausible birthweight for gestation. A further 317 observations (0.6%) were excluded because birthweight, sex or multiple birth status was missing, leaving 50,112 infants included in the study. Population characteristics were fairly similar across all 5 years (*Table 7*), although some differences were statistically significant. There was a slight increase in the proportion of infants born 30^{+0} to 31^{+6} weeks' gestation, from 39.6% in 2008 to 42.9% in 2014. The proportion of babies delivered by caesarean section increased every year, from 55% in 2008 to 59% in 2014, but this outcome was missing for around 9% of infants. Infants admitted to neonatal units tend to be from more deprived areas than the general population, and this became more marked over the study period: the 20% most deprived LSOAs contribute > 30% of the study population (increasing from 29% to 33% over the period), whereas the 20% least deprived LSOAs contribute only 13% (decreasing from 14% to 12%).

Predictive model

Parameter estimates from the logistic regression model are shown in *Appendix 1, Table 56*. Gestational age was modelled with a five-knot spline and birthweight was modelled as birthweight and birthweight^2 with interactions between linear birthweight with all gestational age terms, and birthweight^2 with linear gestational age included. *Appendix 2, Figures 28–35*, show isosurv plots for survival prediction. After correcting for optimism, the AUC was 0.84 and the Brier score was 0.07. The model was well calibrated for both gestational subgroups, with optimism-corrected slopes of 1.01 and 0.97, and intercepts of 0.01 and –0.1 for infants born at $< 28^{+0}$ weeks' and $\geq 28^{+0}$ weeks' gestation, respectively. *Table 8* shows the results comparing performance with comparator models. The NNRD model was better calibrated than the other models, with calibration intercepts and slopes closer to the target values of 0 and 1, respectively. The AUC was higher than the National Institute of Child Health and Human Development (NICHD) model, but still fairly low at 0.70. There were no other differences in the AUC or the Brier score.

Survival to discharge from 2008 to 2014

Of the 48,422 admitted infants for whom outcomes were known, 43,444 (89.7%) survived to discharge. There was no evidence of autocorrelation in any analyses (no change when altering the autocorrelation parameter), so results are presented without autocorrelation. There was an increase in the percentage of admitted infants who survived to discharge from 88% in 2008 to 91.3% in 2014. Survival increased with gestational age, from 35% for 22^{+0} to 23^{+6} weeks' gestation to 98% for 30^{+0} to 31^{+6} weeks' gestation. *Appendix 1, Table 57,* shows the associations between survival and infant characteristics for the whole cohort, based on complete data only. Crude survival rates were lower for boys, infants whose mothers did not receive antenatal steroids and infants born by vaginal delivery. Infants born to younger mothers, mothers who smoked and mothers from more deprived areas had lower crude survival rates.

The annual percentage change (APC) for crude survival was 0.51% (95% CI 0.35% to 0.67%; $p < 0.001$), and 0.46% (95% CI 0.30% to 0.62%; $p < 0.001$) after direct standardisation for risk of death. Results were similar when the only neonatal networks where all hospitals contributed data for the whole period were

TABLE 7 Population characteristics 2008 to 2014; percentages are of the total non-missing values; *p*-value from non-parametric trend test

Characteristics	Year, *n* (%)							Total (*N* = 50,112), *n* (%)	*p*-value for trend
	2008 (*N* = 6103)	2009 (*N* = 6487)	2010 (*N* = 7386)	2011 (*N* = 7733)	2012 (*N* = 7667)	2013 (*N* = 7367)	2014 (*N* = 7369)		
Gestational age (complete weeks)									
22^{+0} to 23^{+6}	195 (3.2)	160 (2.5)	198 (2.7)	165 (2.1)	205 (2.7)	198 (2.7)	228 (3.1)	1349 (2.7)	$p < 0.01$
24^{+0} to 25^{+6}	760 (12.5)	694 (10.7)	759 (10.3)	890 (11.5)	872 (11.4)	842 (11.4)	820 (11.1)	5637 (11.2)	
26^{+0} to 27^{+6}	1121 (18.4)	1219 (18.8)	1306 (17.7)	1401 (18.1)	1373 (17.9)	1238 (16.8)	1232 (16.7)	8890 (17.7)	
28^{+0} to 29^{+6}	1610 (26.4)	1734 (26.7)	2029 (27.5)	2064 (26.7)	1992 (26)	1997 (27.1)	1925 (26.1)	13,351 (26.6)	
30^{+0} to 31^{+6}	2417 (39.6)	2680 (41.3)	3094 (41.9)	3213 (41.5)	3225 (42.1)	3092 (42)	3164 (42.9)	20,885 (41.7)	
Birthweight (g)									
< 500	53 (0.9)	47 (0.7)	45 (0.6)	40 (0.5)	52 (0.7)	74 (1)	71 (1)	382 (0.8)	$p = 0.74$
500 to 999	2053 (33.6)	2061 (31.8)	2286 (31)	2523 (32.6)	2446 (31.9)	2332 (31.7)	2360 (32)	16,061 (32.1)	
1000 to 1499	2519 (41.3)	2811 (43.3)	3209 (43.4)	3310 (42.8)	3297 (43)	3148 (42.7)	3160 (42.9)	21,454 (42.8)	
1500 to 1999	1358 (22.3)	1472 (22.7)	1716 (23.2)	1737 (22.5)	1757 (22.9)	1700 (23.1)	1667 (22.6)	11,407 (22.8)	
≥ 2000	120 (2)	96 (1.5)	130 (1.8)	123 (1.6)	115 (1.5)	113 (1.5)	111 (1.5)	808 (1.6)	
SGA									
No	5211 (85.4)	5530 (85.2)	6305 (85.4)	6540 (84.6)	6569 (85.7)	6271 (85.1)	6261 (85)	42,687 (85.2)	$p = 0.62$
Yes	892 (14.6)	957 (14.8)	1081 (14.6)	1193 (15.4)	1098 (14.3)	1096 (14.9)	1108 (15)	7425 (14.8)	
Sex									
Female	2831 (46.4)	3099 (47.8)	3367 (45.6)	3547 (45.9)	3513 (45.8)	3278 (44.5)	3376 (45.8)	23,011 (45.9)	$p = 0.01$
Male	3272 (53.6)	3388 (52.2)	4019 (54.4)	4186 (54.1)	4154 (54.2)	4089 (55.5)	3993 (54.2)	27,101 (54.1)	
Multiplicity of pregnancy									
Singleton	4456 (73)	4714 (72.7)	5364 (72.6)	5628 (72.8)	5609 (73.2)	5522 (75)	5416 (73.5)	36,709 (73.3)	$p = 0.02$
Twins	1514 (24.8)	1626 (25.1)	1828 (24.7)	1889 (24.4)	1852 (24.2)	1675 (22.7)	1777 (24.1)	12,161 (24.3)	
Triplets or more	133 (2.2)	147 (2.3)	194 (2.6)	216 (2.8)	206 (2.7)	170 (2.3)	176 (2.4)	1242 (2.5)	

Characteristics	Year, *n* (%)							Total (*N* = 50,112), *n* (%)	*p*-value for trend
	2008 (*N* = 6103)	2009 (*N* = 6487)	2010 (*N* = 7386)	2011 (*N* = 7733)	2012 (*N* = 7667)	2013 (*N* = 7367)	2014 (*N* = 7369)		
Any antenatal steroids given									
No	738 (12.6)	728 (11.5)	868 (12.1)	864 (11.4)	879 (11.6)	773 (10.6)	766 (10.4)	5616 (11.4)	$p < 0.01$
Yes	5137 (87.4)	5585 (88.5)	6312 (87.9)	6724 (88.6)	6704 (88.4)	6552 (89.4)	6579 (89.6)	43,593 (88.6)	
Missing	228	174	206	145	84	42	24	903	
Mode of delivery									
Vaginal	2344 (45.2)	2557 (44.1)	2949 (43.6)	3080 (43.1)	3001 (42.6)	2848 (42.2)	2793 (41.1)	19,572 (43)	$p < 0.01$
Caesarean	2843 (54.8)	3246 (55.9)	3817 (56.4)	4070 (56.9)	4044 (57.4)	3896 (57.8)	3996 (58.9)	25,912 (57)	
Missing	916	684	620	583	622	623	580	4626	
Maternal age (years)									
< 20	531 (8.9)	520 (8.1)	630 (8.6)	581 (7.5)	527 (6.9)	469 (6.4)	450 (6.2)	3708 (7.5)	$p < 0.01$
20 to 24	1088 (18.3)	1201 (18.6)	1342 (18.2)	1498 (19.4)	1390 (18.2)	1248 (17)	1175 (16.1)	8942 (18)	
25 to 29	1526 (25.7)	1658 (25.7)	1900 (25.8)	1984 (25.7)	1986 (26)	1892 (25.8)	1934 (26.5)	12,880 (25.9)	
30 to 34	1499 (25.2)	1721 (26.7)	1962 (26.7)	2072 (26.9)	2123 (27.8)	2085 (28.5)	2165 (29.6)	13,627 (27.4)	
35 to 40	1023 (17.2)	1063 (16.5)	1206 (16.4)	1235 (16)	1216 (15.9)	1245 (17)	1192 (16.3)	8180 (16.5)	
> 40	270 (4.5)	290 (4.5)	321 (4.4)	335 (4.3)	396 (5.2)	389 (5.3)	386 (5.3)	2387 (4.8)	
Missing	166	34	25	28	29	39	67	388	

TABLE 8 Performance statistics; NNRD statistics reported separately for each model as the applicable populations differ

	Model					
Statistic	CRIB II	NNRD	Draper	NNRD	NICHD	NNRD
Number	16,652		16,445		6986	
Observed survival (%)	91.7		90.5		60.5	
Predicted survival (%)	92.6	91.5	89.3	90.5	52.2	60.2
AUC (95% CI)	0.82 (0.80 to 0.83)	0.81 (0.80 to 0.83)	0.81 (0.80 to 0.82)	0.82 (0.81 to 0.83)	0.59 (0.57 to 0.60)	0.71 (0.69 to 0.72)
Brier score	0.064	0.064	0.071	0.070	0.250	0.208
Cox						
α (95% CI)	–0.27 (–0.37 to –0.17)	–0.08 (–0.19 to 0.02)	–0.36 (–0.45 to –0.27)	0.03 (–0.07 to 0.13)	–0.40 (–0.45 to –0.35)	0.02 (–0.04 to 0.08)
β (95% CI)	0.77 (0.73 to 0.81)	0.97 (0.92 to 1.02)	0.88 (0.84 to 0.92)	1.01 (0.97 to 1.06)	0.42 (0.36 to 0.49)	1.09 (1.01 to 1.17)

examined [crude APC 0.56% (95% CI 0.35% to 0.77%); adjusted APC 0.53% (95% CI 0.33% to 0.73%)]. Sensitivity analysis of complete-case data and standardising for 5 and 15 categories gave very similar results.

Survival to 28 days

Fifty deaths were established by linkage with ONS, of which 20 were within 28 days of birth. There was an increase in the percentage of infants who survived to 28 days, from 91.4% in 2008 to 93.5% in 2014. Survival increased with gestational age from 48.4% for 22^{+0} to 23^{+6} weeks' gestation to 98.2% for 30^{+0} to 31^{+6} weeks' gestation. The APC for crude 28-day survival was 0.3% (95% CI 0.15% to 0.45%; $p < 0.001$), and 0.27% (95% CI 0.11% to 0.44%; $p = 0.002$) after direct standardisation for risk of death. Results were similar when only the neonatal networks in which all hospitals contributed data for the whole period were examined (crude APC 0.35%, 95% CI 0.19% to 0.52%; adjusted APC 0.3%; 95% CI 0.14% to 0.47%).

Time of death

A total of 24% of deaths occurred within 24 hours, 28% between 25 hours and 7 days, 26% between 8 days and 28 days and 23% beyond 28 days. There was no evidence of a change in the median and 25th percentile time of death, whereas the 75th percentile reduced from 2008 (27.2 days) to 2013 (20.8 days), but rose to 24.3 days in 2014 (estimated annual decrease 2008–14: 0.92 days, 95% CI 0.2 to 1.7 days; $p = 0.02$).

Trends in survival to discharge by gestational age

Figure 10 shows the joinpoint regression analysis for survival to discharge by gestational age group. Improvements were less marked with increasing gestation, ranging from an APC of 6.03% (95% CI 2.47% to 3.53%; $p = 0.002$) in infants born 22^{+0} to 23^{+6} weeks' gestation to no change in infants born 30^{+0} to 31^{+6} weeks' gestation (APC 0.01%, 95% CI –0.08% to 0.09%; $p = 0.9$).

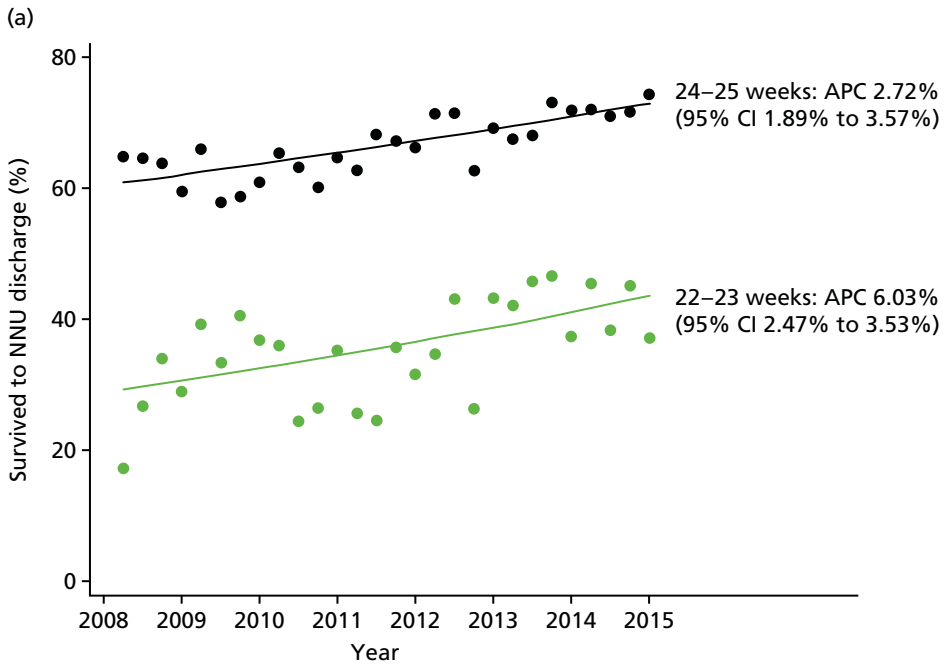


FIGURE 10 Joinpoint regression analysis for crude rates of survival to discharge for admitted infants born at (upper plot) 22^{+0} to 25^{+6} weeks' gestation and (lower plot) 26^{+0} to 31^{+6} weeks' gestation by birth year. APC, average percentage change. NNU, neonatal unit. (*continued*)

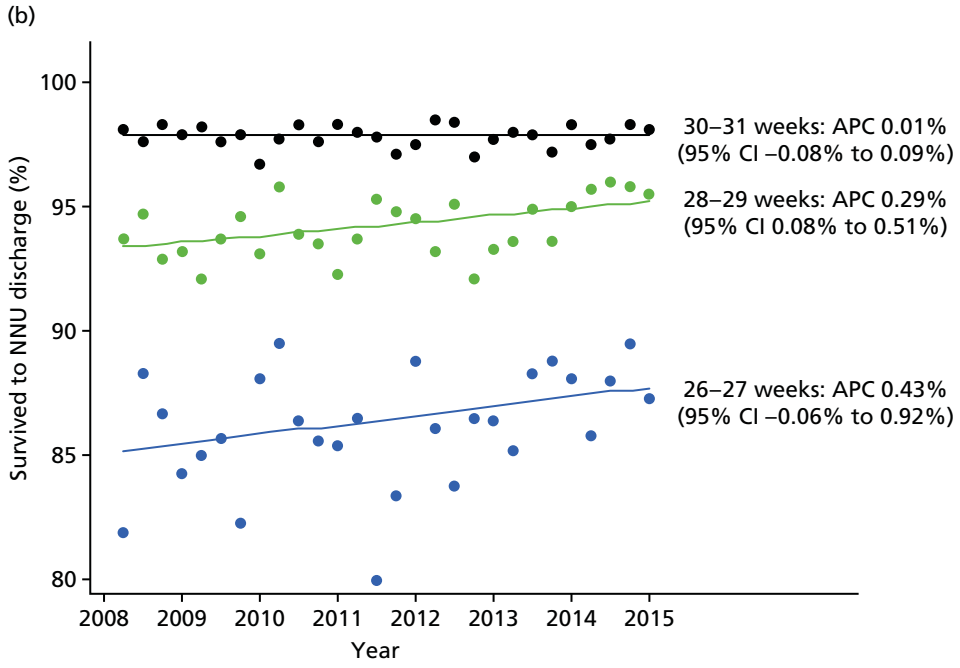


FIGURE 10 Joinpoint regression analysis for crude rates of survival to discharge for admitted infants born at (upper plot) 22^{+0} to 25^{+6} weeks' gestation and (lower plot) 26^{+0} to 31^{+6} weeks' gestation by birth year. APC, average percentage change. NNU, neonatal unit.

Variation by region and Index of Multiple Deprivation quintile using data from 2011 onwards

Crude survival varied from 89.3% (95% CI 88.6% to 89.9%) in the Midlands and the East of England to 91.1% (95% CI 90.3% to 91.8%) in London; after direct standardisation, the range was 89.2% (95% CI 87.3% to 91.1%) to 91.6% (95% CI 89.1% to 94.2%). Only London and the South of England showed improvements in crude survival (*Figure 11*).

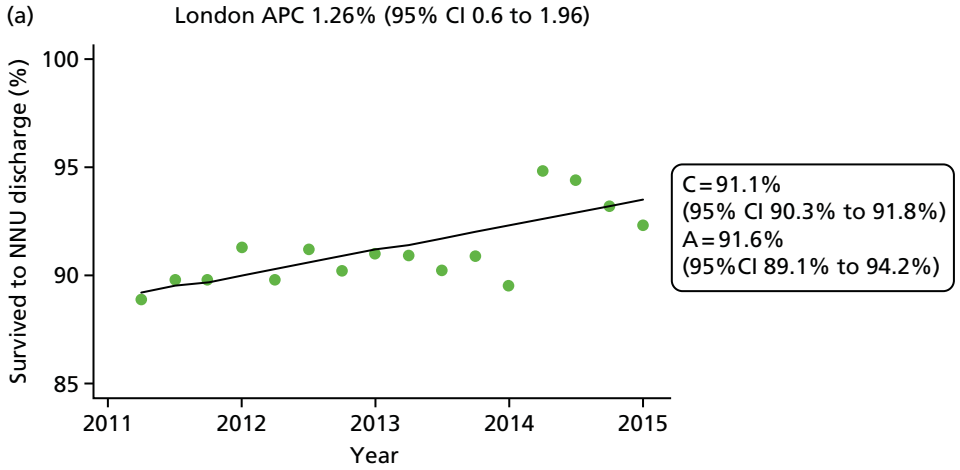


FIGURE 11 Joinpoint regression analysis for crude rates of survival to discharge for admitted infants born in 2011–14 at 22–31 weeks' gestation by NHS commissioning region. A, adjusted survival (%) over years 2011–14; C, crude survival (%) over years 2011–14. NNU, neonatal unit. (*continued*)

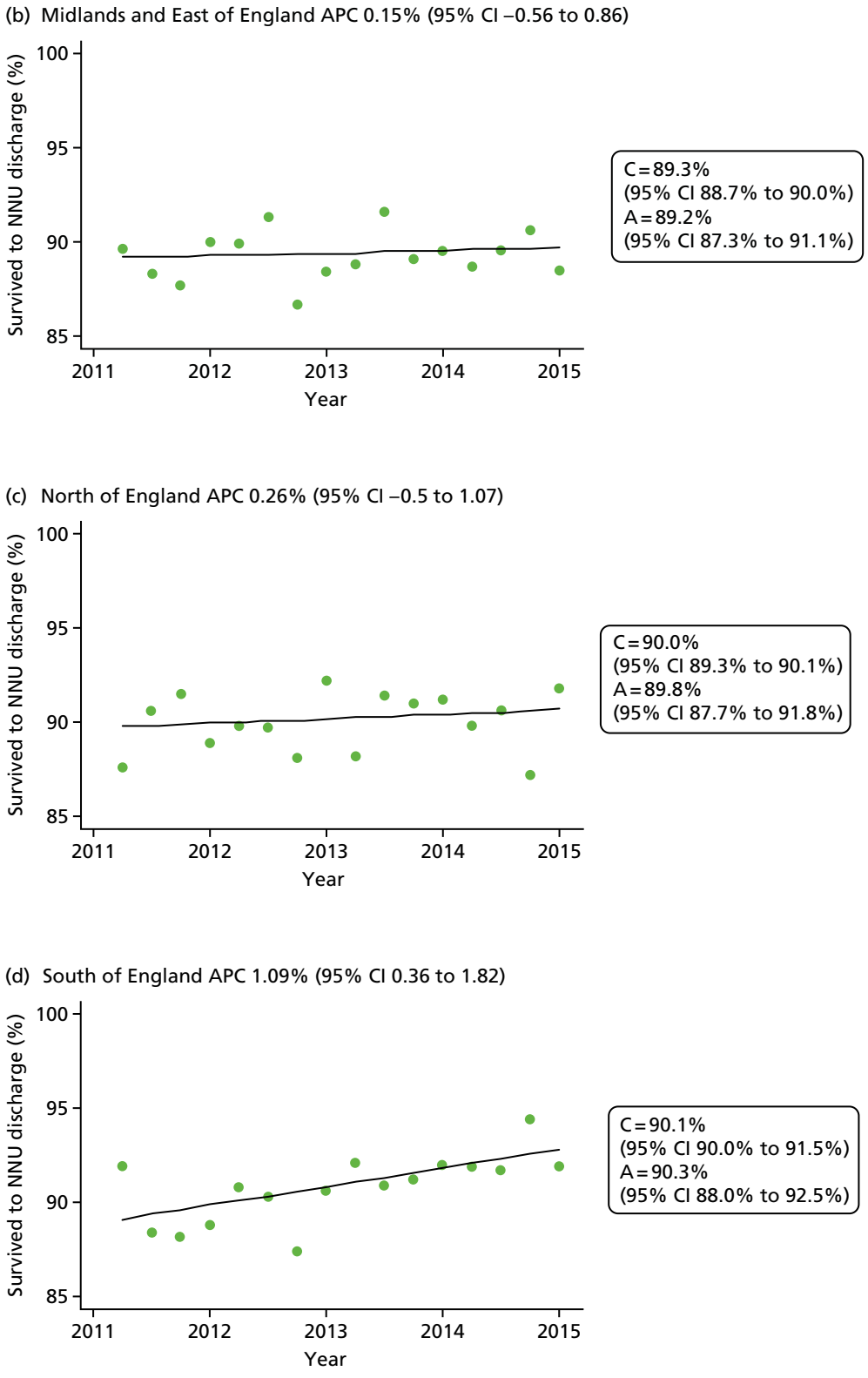


FIGURE 11 Joinpoint regression analysis for crude rates of survival to discharge for admitted infants born in 2011–14 at 22–31 weeks' gestation by NHS commissioning region. A, adjusted survival (%) over years 2011–14; C, crude survival (%) over years 2011–14. NNU, neonatal unit.

Infants from the most deprived quintile had lower survival rates than those from the least deprived quintile [89.5% (95% CI 88.9% to 90.1%) versus 91.1% (95% CI 90.2% to 92.1%)]; little difference remained after standardisation [89.8% (95% CI 87.9% to 91.5%) versus 90.1% (95% CI 87.1% to 93.2%)]. Inclusion of IMD decile in the risk adjustment model did not change results for each region, with evidence of residual variation across regions ($p < 0.001$).

Comparison with Office for National Statistics and EPICure data

The number of deaths before 28 days among admitted infants born 22^{+0} to 31^{+6} weeks' gestation recorded in the NNRD for England and Wales in 2012 was 801. This represents 81% (801/989) of the deaths recorded among live births for the same gestation range in England and Wales by the ONS. Most of the discrepancy occurred at earlier gestations: there were seven deaths among infants born 22 weeks' gestation in the NNRD, compared with 154 in the ONS.

There was no evidence of a change in the rate of improvement since the first EPICure study. Improvements in survival to discharge of infants born at 22^{+0} to 25^{+6} weeks' gestation and admitted to neonatal care in 1995 (EPICure),[124] 2006 (EPICure 2)[135] and 2008–14 (NNRD)[137] have continued at a similar rate (*Figure 12*).

Mortality by Operational Delivery Network

The number of infants were born in 2013–14 at $\leq 31^{+6}$ weeks' gestation and admitted to neonatal care for whom the neonatal network of booking was known was 15,255; 8.9% of those for whom the outcome was known (1327/14,837) died before discharge. The outcome was missing for 2.7% of infants. Infants with a missing outcome tended to be more vulnerable based on other neonatal characteristics. Antenatal steroid entries were missing for 0.5% of infants, and sex and multiple birth status SDSs for < 0.01%. The prediction model fit the data well, giving an area under the receiver operating characteristic (ROC) curve of 0.83 (95% CI 0.82 to 0.84).

The SMRs are shown on funnel plots, both unadjusted (*Figure 13*) and adjusted (*Figure 14*), with neonatal networks numbered (*Table 9* contains the key). Analysis of complete cases gave very similar results (largest absolute difference in SMR of 0.04).

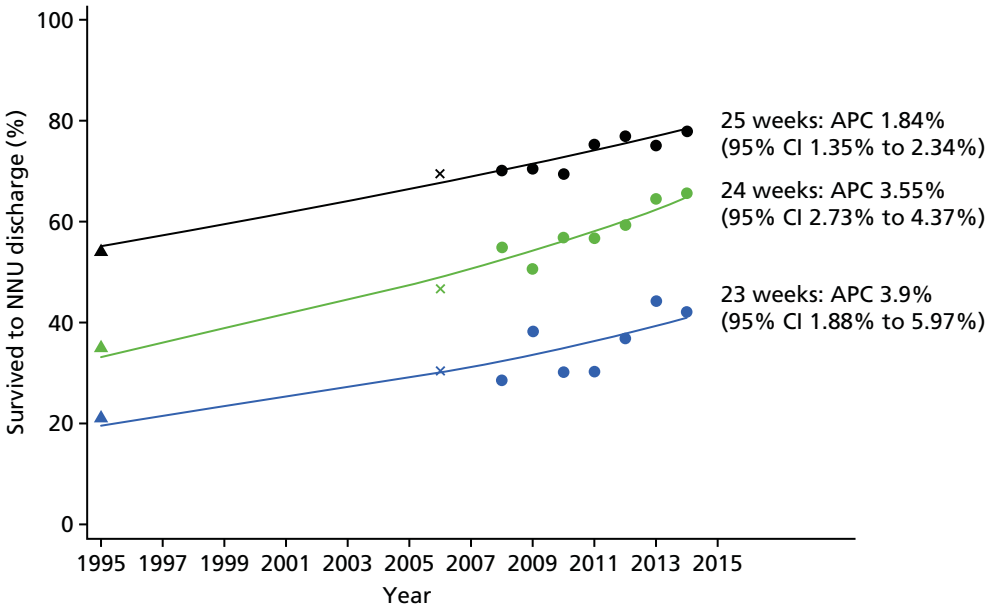


FIGURE 12 Survival to NNU discharge from 1995 to 2015 based on data from EPICure (triangle), EPICure 2 (cross) and NNRD (circles), for infants born at 23 (blue), 24 (green) and 25 (black) weeks' gestation. NNU, neonatal unit.

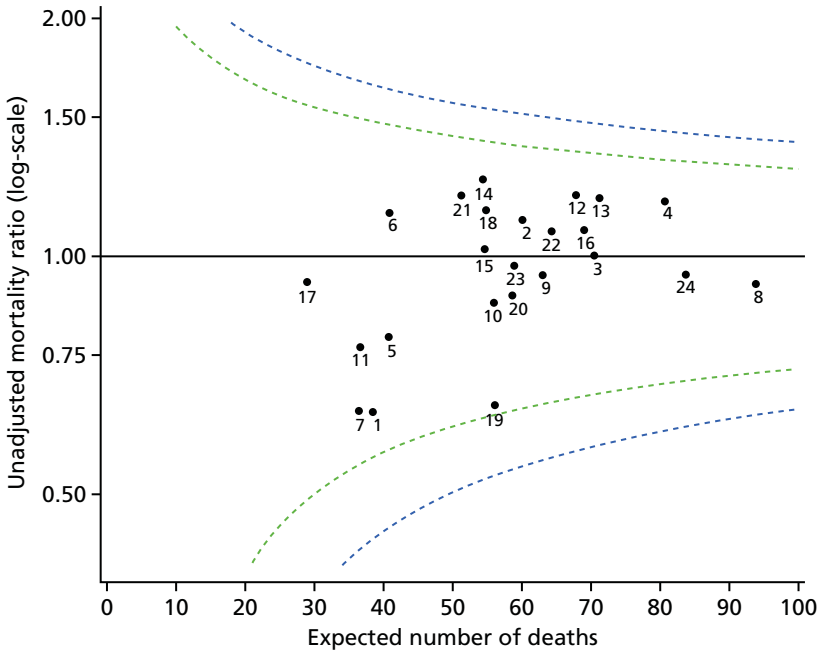


FIGURE 13 Funnel plot for unadjusted SMR for babies live-born ≤ 31^{+6} weeks' gestation in 2013–14 and admitted to neonatal care, by neonatal network of booking; control limits show 2 and 3 SDs from the mean after correction for multiple testing, assuming observed deaths follow a Poisson distribution; numbers correspond to neonatal networks in *Table 9*.

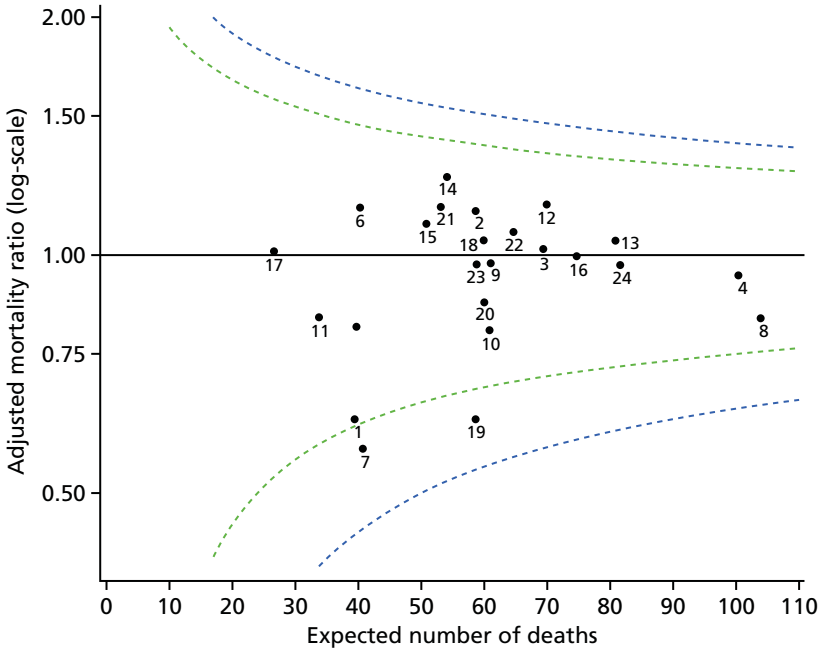


FIGURE 14 Funnel plot for adjusted SMR for babies live-born ≤ 31^{+6} weeks' gestation in 2013–14 and admitted to neonatal care, by neonatal network of booking; control limits show 2 and 3 SDs from the mean after correction for multiple testing, assuming observed deaths follow a Poisson distribution. Adjusted for gestation, birthweight SDS, sex, antenatal steroids accounting for correlation within multiple birth sets and missing data; numbers correspond to neonatal networks in *Table 9*.

TABLE 9 Unadjusted and adjusted SMR for babies live-born ≤ 31^{+6} weeks' gestation in 2013–14 and admitted to neonatal care, by neonatal network of booking

Code	Booked neonatal network	Total infants	SMR (95% CI)	
			Raw	Adjusted
1	Bedfordshire and Hertfordshire	422	0.64 (0.33 to 1.1)	0.62 (0.32 to 1.07)
2	Cheshire and Merseyside	659	1.11 (0.76 to 1.56)	1.14 (0.78 to 1.6)
3	Eastern	773	1 (0.69 to 1.4)	1.02 (0.71 to 1.42)
4	Greater Manchester	885	1.17 (0.86 to 1.56)	0.94 (0.69 to 1.26)
5	Kent	447	0.79 (0.45 to 1.28)	0.81 (0.46 to 1.31)
6	Lancashire and South Cumbria	448	1.13 (0.71 to 1.7)	1.15 (0.72 to 1.72)
7	London (North Central)	399	0.64 (0.32 to 1.12)	0.57 (0.29 to 1)
8	London (North East)	1029	0.92 (0.66 to 1.24)	0.83 (0.6 to 1.12)
9	London (North West)	691	0.95 (0.63 to 1.36)	0.98 (0.65 to 1.4)
10	London (South East)	614	0.87 (0.56 to 1.3)	0.8 (0.51 to 1.19)
11	London (South West)	402	0.77 (0.42 to 1.28)	0.83 (0.45 to 1.4)
12	Midlands (Central)	744	1.2 (0.85 to 1.63)	1.16 (0.83 to 1.58)
13	Midlands (South West)	781	1.18 (0.85 to 1.6)	1.04 (0.75 to 1.41)
14	Staffordshire, Shropshire and Black Country	596	1.25 (0.86 to 1.75)	1.26 (0.86 to 1.76)
15	North Trent	599	1.02 (0.67 to 1.48)	1.1 (0.72 to 1.59)
16	Northern	757	1.08 (0.76 to 1.49)	1 (0.7 to 1.38)
17	Peninsula (South West)	318	0.93 (0.5 to 1.57)	1.01 (0.54 to 1.71)
18	South Central (North)	602	1.14 (0.77 to 1.62)	1.05 (0.71 to 1.48)
19	South Central (South)	615	0.65 (0.38 to 1.02)	0.62 (0.37 to 0.98)
20	Surrey and Sussex	643	0.89 (0.58 to 1.31)	0.87 (0.57 to 1.28)
21	Trent	562	1.19 (0.8 to 1.7)	1.15 (0.77 to 1.64)
22	Wales	705	1.08 (0.74 to 1.5)	1.07 (0.74 to 1.5)
23	Western	646	0.97 (0.65 to 1.4)	0.98 (0.65 to 1.41)
24	Yorkshire	918	0.95 (0.67 to 1.3)	0.97 (0.69 to 1.33)

Conclusions

Survival between 2008 and 2014

We demonstrate that survival of very preterm infants admitted to neonatal units in England improved between 2008 and 2014, with the greatest improvement seen among infants born at the lowest number of weeks of gestation. However, survival did not improve consistently across the four NHS commissioning regions, with London and the South of England performing better than the Midlands, the East of England, and the North. Survival was lower for infants from more deprived areas, but regional differences in survival persisted after adjustment for socioeconomic differences.

A key strength of the study is the data. Over 50,000 very preterm infants were included, representing the vast majority of neonatal unit admissions in the country during the study period. Assurance on the quality and completeness of the data was provided through comparison with ONS data. Furthermore, neonatal units had the opportunity to validate survival outcomes and clinical characteristics for infants in the study population born after 2012 as part of work by the NDAU. The risk adjustment variables used in the study were limited to key, unambiguous clinical characteristics to minimise the risk of incomplete or inaccurate

data; < 3% of records had any missing data, and only 38 records had implausible birthweight for gestation. Several steps were taken to limit or investigate potential bias in the analysis. The improvements in survival remained when we examined only neonatal networks contributing data throughout the period, changed the number of categories used in risk adjustment and looked at infants with complete data only.

A limitation is that the population comprises infants admitted to a neonatal unit, thus excluding live-born infants who died before admission; this is because data capture in the EPRs is triggered by neonatal unit admission. However, this is the relevant population for neonatal services, and comparison with ONS data showed that > 80% of known deaths in this gestational age range were captured in the NNRD, with the shortfall at the earliest gestations likely to represent deaths before admission. Furthermore, analysis of live births does not guarantee a consistent population as there is variation across England in whether or not infants born at < 24 weeks' gestation are registered as live births. We have imputed missing outcomes on the assumption that the missingness presents no additional information beyond the other neonatal characteristics. This may not be so, as some missing outcomes might reflect infants who were transferred to specialist surgical providers that are not standard neonatal units and do not contribute data to the NNRD. However, no patterns were seen (e.g. with gestational age) for infants according to the reason for transfer, and the proportion of infants with missing outcomes was small.

Improvements in survival of very preterm infants have been demonstrated in other countries, but most examine survival of live-born infants rather than neonatal unit admissions. A population study from New South Wales and the Australian Capital Territory showed improved survival of admitted babies at 24 (50–60%), 25 (60–75%) and 26 (80–85%) weeks' gestation between 2000–1 and 2007–11, which is similar to results seen here.[142]

Mortality by Operational Delivery Network

We combined 2 years' data to provide improved power to detect significant deviation from average national performance. The average number of expected deaths in a network is 60. If a neonatal network with a patient case-mix leading to 60 expected deaths had a true underlying SMR of 1.3, the probability of the network's observed data falling above the 2 SD upper control limit is around 60% (i.e. there would be 60% power to alert the network as having potentially unusual performance). This is before widening the limits to allow for multiple testing, which reduces power further.

Note that if the SMR for a neonatal network lies outside the funnel, it will not necessarily have a CI that excludes 1. This is because they have different interpretations: the CI reflects uncertainty about the true SMR for that particular network, whereas the funnel plot limits reflect the variability we would expect to see in the SMR for similar neonatal networks. More specifically, the CI is the range in which we are 95% confident that the true SMR for the neonatal network lies, whereas the funnel plot limits represent the range in which we expect 95% of neonatal networks to lie.

Implications for health care

We have established the feasibility of monitoring neonatal outcomes at the national level using near-contemporaneous, routinely recorded EPR data. This has been achieved with wide professional support and our methods provide a template for future evaluations. We hope the opportunity to monitor survival will be adopted by health-care managers, clinicians and commissioners.

Many clinical outcomes, including mortality, are relatively rare. Hence, achieving adequate power to detect changes over time and in relation to factors, such as geographical region and patient or neonatal unit characteristics, requires a large sample. The use of EPR data facilitates the capture of large samples and, hence, offers opportunity for rigorous evaluation of clinical outcomes. The use of EPR data has also enabled up-to-date, rapid assessment ensuring that the information available to clinicians and health service mangers is near-contemporaneous, with minimal burden to clinical staff.

As data capture in the EPRs is triggered by neonatal unit admission, a limitation is that the total population denominator for all live births is not available in the NNRD. Through comparison with ONS data, we showed that > 80% of known deaths in the preterm gestational age range studied were captured in the NNRD, with the shortfall occurring among the earliest gestations and thus likely to represent deaths before admission rather than incomplete NNRD data. We and others have suggested that the neonatal EPR is modified to enable capture of live-born infants who die before admission, but this will not guarantee a consistent or complete population until the variation across England regarding whether or not infants born live at less than 24 weeks' gestation are registered as live births is addressed.

Research recommendations

Improved short-term survival over time has been previously recognised but the ensuing trends in later life outcomes are not well defined. A potential extension to our work is to evaluate national trends in developmental outcomes at 2 years, as these data are captured in the NNRD as part of the Royal College of Paediatrics and Child Health National Neonatal Audit Programme.

Our finding that preterm survival has not improved consistently across the four NHS commissioning regions and that this is not accounted for by socioeconomic differences is important. Factors, such as staffing levels and care practices, that might explain geographical variation in survival require consideration in order to reduce potential inequities in health-care delivery.

Chapter 4 Testing the quality of Electronic Patient Record data held in the National Neonatal Research Database to support clinical trials

Abstract

Background: Because data recorded in a trial case report form (CRF), widely considered 'gold standard', may already exist within an EPR, repeated collection is wasteful.

Aim: We tested the null hypothesis that EPR data from the NNRD are not of comparable quality to research data.

Methods: We compared NNRD data with data recorded independently in a NIHR trial CRF. We selected a broad range of patient characteristics, processes and outcomes. For each variable, we calculated major and minor discordance rates using predefined criteria, and the sensitivity, specificity and positive predictive values (PPV) of NNRD outcome variables in comparison with the gold standard CRF source.

Results: We assessed 2257 episodes of care in 1258 infants. Major discordance rates were low for 14 out of 15 patient characteristics, 9 out of 12 process measures and 10 out of 11 outcomes. The prevalence of adverse outcomes was < 6% with the exception of bronchopulmonary dysplasia (49.0%) and medical treatment for patent ductus arteriosus (PDA) (20.3%). Specificity was high (> 85%) for all outcomes, sensitivity ranged from 50% to 100%, and PPV ranged from 58.8% (95% CI 40.7% to 75.4%) for a report of a porencephalic cyst to 99.7% (95% CI 99.2% to 99.9%) for survival to discharge.

Conclusions: Patient characteristics and the majority of NNRD items tested compare well against CRF data. A small number of important outcomes are not currently reliably recorded in the EPRs. We recommend minor changes to EPR entry screens to improve outcome data, and testing of NNRD data use in a clinical trial.

Background

Randomised clinical trials are the gold standard method for evaluating therapeutic interventions. The data set recorded for trials involving hospital inpatients usually overlaps with, and may exist completely within, an EPR. Despite this, data collection for clinical trials is usually conducted independently of routine care. This results in duplication of effort for clinical staff and may increase the risk of transcription errors and missing data. The additional workload contributes to the high cost of trials and may act as a disincentive to busy clinicians to participate.

One of the most important principles of the Medicines for Neonates programme is that recording data only once and using them for a multitude of purposes will lead to higher-quality NHS data. This workstream addresses the question of whether or not the data recorded in neonatal EPRs as part of clinical care are of sufficient quality to support a clinical trial. Our objective was to compare routinely recorded EPR data with data recorded specifically for a NIHR-funded multicentre trial of an investigational medicinal product conducted in accordance with the principles of good clinical practice. The trial selected for comparison was the 'Probiotic in Preterm infants Study (PiPS)', funded by the Health Technology Assessment (HTA) programme.[143]

The PiPS trial[143] was a multicentre, double-blind, placebo-controlled, randomised trial of probiotic administration in preterm infants, designed to study the possible benefits of early administration of the probiotic *Bifidobacterium breve* BBG-001 (hereafter referred to in brief as 'the probiotic') to infants born before 31 weeks' gestation and recruited within 48 hours of birth. There were three primary outcomes: (1) any episode of NEC to Bell stage II or III, (2) any positive blood culture of an organism that is not recognised as a skin commensal on a sample drawn > 72 hours after birth and before 46 weeks' postmenstrual age or discharge if sooner (hereafter, sepsis for brevity) and (3) death before discharge from hospital.

Infants (n = 1315) from 24 neonatal units in the south-east of England were recruited to PiPS between July 2010 and July 2013. All of the recruiting hospitals and those to which the babies were likely to be transferred before their initial discharge home, with the exception of the Great Ormond Street and the Royal Brompton Hospitals (to which infants are occasionally referred for specialist care), use the neonatal EPRs and submit data to the NDAU.

Professor Kate Costeloe, who is a co-applicant for the Medicines for Neonates Programme, is also the chief investigator for the PiPS trial and, thus, was able to provide a facilitated opportunity to compare data held in the trial CRFs and those derived from the neonatal EPRs.

Aims

1. To assess the agreement between EPR-derived demographic, process and outcome variables held in the NNRD and equivalent CRF-derived variables held in the PiPS trial database.
2. To evaluate whether or not there was any decrease in discordance rates of compared items over the course of recruitment to the PiPS trial.

Methods

Data

Neonatal EPR data obtained from the NNRD were compared with those recorded independently on trial CRFs and held in the PiPS database. Items for comparison were selected either because the definitions were identical or because data in the NNRD could be used to derive the item as defined for PiPS trial requirements. Variables were selected to represent a broad range of patient characteristics, processes and outcome measures.

The 15 baseline patient data items comprised (1) expected date of delivery, (2) gestational age (weeks and days), (3) month and year of birth, (4) birthweight (g), (5) sex, (6) Apgar score at 5 minutes, (7) whether inborn or transferred, (8) singleton or multiple, (9) birth order, (10) maternal year of birth, (11) maternal ethnicity by NHS category, (12) LSOAS, as derived from postcode, (13) any antenatal corticosteroid given, (14) mode of delivery (vaginal vs. caesarean) and (15) instrumental delivery (*Table 10*).

The 13 processes or interventions during admission were (1) surgery for PDA, (2) medical treatment of PDA, (3) retinopathy of prematurity (ROP) treatment by laser or cryotherapy, (4) central venous line days, (5) intensive care days, (6) high-dependency care days, (7) whether or not transferred to another hospital, (8) discharge month and year (*Table 11*) and, in the first 14 days, (9) day of first milk feed, (10) type or types of milk (maternal milk, donor milk or formula) given on the first day of feeding, (11) a summary of all types of milk received over all of the days on which feeds were reported on both databases, (12) duration of exposure to any antibiotic or (13) duration of exposure to any antacid (*Table 12*).

TABLE 10 Items selected for comparison: baseline characteristics, including details of the data held in each database, with preset definitions of limits of agreement, and minor and major discrepancies

	Data held		Definition		
Item to be compared	PiPS	NNRD	Limits of agreement	Minor disagreement	Major disagreement
EDD	EDD	EDD	Up to ± 2 days	3–6 days	± 1 week
Gestational age (weeks and days)	Gestational age is computed from EDD	Gestational age (recorded independently of EDD)	Up to ± 2 days	3–6 days	± 1 week
Date of birth (month and year)	Date and time of birth	Month and year of birth	No difference	N/A	N/A
Birthweight (g)	Infant's birthweight (g)	Infant's birthweight (g)	30 g	30–100 g	> 100 g
Sex	Infant's sex (male/female/ indeterminate)	Infant's sex (male/female/ indeterminate)	No difference	N/A	N/A
Apgar score at 5 minutes	Apgar score at 5 minutes	Apgar score at 5 minutes	± 1	± 2	± 3 or more
Born in this hospital	Whether or not infant was born in this hospital	Place of birth	No difference	N/A	N/A
Singleton or multiple birth	Whether infant is a singleton or multiple birth	Whether infant is a singleton or multiple birth	No difference	N/A	N/A
Birth order	Birth order of infant	Birth order of infant	No difference	N/A	N/A
Maternal year of birth	Maternal date of birth	Maternal birth year	No difference	N/A	N/A
Maternal ethnicity	Maternal ethnicity (NHS categories)	Maternal ethnicity (NHS categories)	No difference	N/A	N/A
Maternal LSOA at time of infant's birth	Maternal LSOA derived from postcode	Maternal LSOA derived from postcode	No difference	N/A	N/A
Whether or not any antenatal steroids were given	Antenatal steroids and exact timing	Any antenatal steroids given, no detail of timing	No difference	N/A	N/A
Mode of delivery: caesarean or vaginal	Mode of delivery	Mode of delivery	No difference	N/A	N/A
Whether or not instrumental delivery	Whether forceps or ventouse were used for delivery	Mode of delivery	No difference	N/A	N/A

EDD, expected date of delivery; N/A, not applicable.

TABLE 11 Items selected for comparison: processes of care and interventions, including details of the data held in each database, with preset definitions of limits of agreement, and minor and major discrepancies

	Data held		Definition		
Item to be compared	PiPS	NNRD	Limits of agreement	Minor disagreement	Major disagreement
Surgery for PDA	While in this hospital, did the infant receive surgical ligation for PDA?	Daily data: surgery for PDA today Discharge diagnoses Procedures during stay	No difference	N/A	Any difference
Medical treatment for PDA with indometacin or ibuprofen	While in this hospital, has the infant received medical treatment with indometacin and/or ibuprofen for PDA?	Daily data: treatment for PDA Daily drugs	No difference	N/A	Any difference
Treatment for ROP with laser or cryotherapy	While in this hospital, has infant had ROP treated with laser/cryotherapy?	Daily data: treatment Discharge diagnoses Procedures during stay?	No difference	N/A	Any difference
Central venous line days	While in this hospital, what was the total number of days for which the infant had a central venous line [UVC, peripheral long line, BROVIAC® (Bard Access Systems Inc., Salt Lake City, UT, USA), etc.]	Daily data: lines in situ	± 2 days	3–4 days	± 5 or more days
Intensive care days	While in this hospital, what was the total number of days of intensive care days?	Daily data	± 2 days	3–4 days	± 5 or more days
High-dependency care days	While in this hospital, what was the total number of high-dependency care days?	Daily data	± 2 days	3–4 days	± 5 or more days
Transfer to another hospital	Whether or not was transferred to another hospital	Discharge details	No difference	N/A	Any difference
Discharge month and year	Date of discharge or death	Discharge details	No difference	N/A	Any difference

N/A, not applicable.

TABLE 12 Items selected for comparison: processes of care and interventions in the first 14 days, including details of the data held in each database, with preset definitions of limits of agreement, and minor and major discrepancies

	Data held		Definition		
Item to be compared	PiPS	NNRD	Limits of agreement	Minor disagreement	Major disagreement
Day of first milk feed	What day of life was milk commenced?	Daily feeding data	± 1 day	± 2 days	> 2 days
Type(s) of first milk feed	Milk on first day of receiving milk feed	Daily feeding data	No difference	N/A	Any difference
Summary of all types of milk in first 14 days	Daily feeding data for first 14 postnatal days	Daily feeding data	No difference	N/A	Any difference
Total number of days of antibiotics received during first 14 postnatal days	Names and total days of antibiotics received during the first 14 postnatal days	Daily drugs	± 1 day	± 2 days	> 2 days
Total number of days of antacid received during first 14 postnatal days	Total days of antacid use during the first 14 postnatal days	Daily drugs	± 1 day	± 2 days	> 2 days

N/A, not applicable.

The nine outcome data items were (1) worst stage of ROP in either eye, (2) bronchopulmonary dysplasia (BPD), defined by whether or not the infant required supplementary oxygen at 36 weeks' postmenstrual age, (3) mechanical respiratory support at 36 weeks' postmenstrual age, (4) cranial ultrasound findings, (5) survival to discharge, (6) any diagnosis of perforated NEC, (7) any abdominal surgery for NEC, (8) any gastrointestinal perforation and (9) length of stay (*Table 13*).

Changes to the original protocol

When this study was designed it was anticipated that recruitment to the PiPS trial would begin in 2009. The analysis could be conducted only once PiPS data for the individual infant were complete and any queries had been addressed. Although it was always appreciated that the order in which the data for trial recruits were signed off would not be sequential, it was nonetheless expected that it would be possible to receive data in batches every few months. The original protocol involved comparison of PiPS and NNRD data from the first 200 infants recruited, as well as feedback of rates of minor and major discrepancies to neonatal units, with the final analysis undertaken on the next 200 infants. There were delays to the start of PiPS recruitment, which did not begin until July 2010 and which did not achieve its target rates for over 1 year. There were further delays to PiPS data being signed off as complete because of problems with introducing a new automated data query system. It became clear that we could not provide feedback to neonatal unit staff after 200 cases and that we were likely to receive any PiPS data at the NDAU only towards the end of PiPS trial recruitment. It was agreed within the MfN Steering Committee that we would request an initial download of all available complete data for piloting the database merger and would then perform the comparative analysis on the final PiPS data set, including all recruits, and this was received in October 2013. There is a constant process of data scrutiny and feedback aimed at improving data completeness and accuracy in the NNRD. It was therefore agreed that we should examine the whole data set for changes in discrepancy rates over time, in order to address the second aim listed above.

Preparation of data for comparison

Data acquisition differs between the PiPS trial and the NNRD, the former being recorded specifically for trial purposes and the latter containing data extracted from point-of-care EPRs, designed to provide a complete clinical record. In this section, we will describe how the data sets were prepared to ensure that the comparisons related to the same infant and the same episode of care.

Neonatal units in NHS hospitals in England function as clinical networks. Neonatal units vary in the care that they provide, in that some provide intensive care only in an emergency, whereas others that do provide ongoing intensive care might not do so for extremely preterm infants. A small number of neonatal units provide specialist services, such as surgery and cardiology. In so far as is possible, infants are looked after in the neonatal unit closest to the family home. Those infants needing specialist care may require transfer to a neonatal unit offering the required expertise; hence the entire course between birth and eventual discharge home may comprise a series of 'episodes of care' of varying durations in different neonatal units. The variables that were compared for each infant between the PiPS database and the NNRD comprise 'once only' data (e.g. baseline demographic items such as birthweight), 'episodic' data (e.g. processes and interventions during a defined episode of care) and 'infant-level' data (e.g. outcomes summarised from multiple episodes of care).

Infants born between 23^{+0} and 30^{+6} weeks' gestational age and who were < 48 hours old were eligible for recruitment to the PiPS trial. Infants with a lethal congenital anomaly or any known gastrointestinal malformation known at birth, or with no realistic chance of survival, were excluded.

The CRF data collection was paper based, with four main collection forms: form 1 (entry), form 2 (daily data), form 3 (transfer/discharge) and form 4 (abdominal pathology). Form 1 (entry) is completed within 7 days of recruitment and returned to the National Perinatal Epidemiology Unit (NPEU), and it contains 'once only' information regarding the infant and maternal history [e.g. infant sex, expected date of delivery (EDD), maternal and obstetric details and infant's condition at birth). Form 2 (daily data) requires daily recording for the first 14 postnatal days from the day of birth of the type of milk received, the total daily volume of milk (ml/kg/day),

TABLE 13 Items selected for comparison: outcomes, including details of the data held in each database, with preset definitions of limits of agreement, and minor and major discrepancies

Variable to be compared, definitions	Data held		Definition		
	PiPS	NNRD	Limits of agreement	Minor disagreement	Major disagreement
Worst stage of ROP in any eye	Worst stage of ROP in ANY eye? (Stage 1–5)	Discharge diagnoses Ad hoc forms for each ROP examination	No difference	N/A	Any difference
BPD requiring oxygen at 36 weeks' postmenstrual age	If still in hospital at 36 weeks' postmenstrual age: date reached 36 weeks' postmenstrual age and whether receiving supplementary oxygen?	Daily data for oxygen use	No difference	N/A	Any difference
Requirement of mechanical respiratory support at 36 weeks' postmenstrual age	If still in hospital at 36 weeks' postmenstrual age was the infant receiving mechanical respiratory support	Daily data for respiratory support received	No difference	N/A	Any difference
Cranial ultrasound findings	While in this hospital, did the infant have any of the following abnormalities in their cranial ultrasound scan? • Haemorrhagic parenchymal infarct • Hydrocephalus (ventricular index > 4 mm above 97th centile[144]) • Porencephalic cyst • Periventricular leucomalacia	Discharge diagnoses Ad hoc forms for each cranial ultrasound examination	No difference	N/A	Any difference
Survival to discharge from neonatal care	Survival to discharge	Discharge details	No difference	N/A	Any difference
Gastrointestinal diagnoses	• Perforated NEC • Any abdominal surgery for NEC • Any gastrointestinal perforation	Discharge diagnoses Ad hoc form reporting abdominal radiography Daily surgery/NEC data	No difference	N/A	Any difference
Length of stay	What was the total length of stay in neonatal care?	Daily data Discharge details	± 1 day	± 2 days	± 3 or more days

N/A, not applicable.

the antibiotics by type, the antifungals and the antacids. Form 3 (transfer/discharge) is completed for each episode of care, terminating at discharge (whether to another hospital or home or at death). It contains only events occurring during that episode (e.g. diagnoses, procedures and treatments received, including, if the admission covers 36 weeks' postmenstrual age, whether or not the baby is still receiving supplementary oxygen or mechanical ventilatory support). Form 4 (abdominal pathology) is completed for any episode of proven or suspected abdominal pathology including NEC, for which the severity is staged using modified Bell criteria.

On receipt at the PiPS trial office at the NPEU, validation included a series of range, logic and missing data checks to identify inconsistencies within and across forms for the same baby. Some queries were resolved in-house according to predefined protocols; those that could not be resolved were reconciled between the staff in the trial office, the PiPS trial research nurses, the chief investigator and principal investigators with reference to the clinical notes, and documented accordingly. Data were double-entered onto a dedicated trial database.

The NNRD is organised into different files, including static 'once only' data, 'episodic' data for each admission to a different neonatal unit, 'daily data' recorded on a daily basis, and 'if and only' data recorded only if applicable (e.g. ad hoc abdominal X-ray forms, blood cultures).

Episode numbering and matching

Episode number 1 on the PiPS database is the admission to the neonatal unit where the infant was recruited to the trial [Form 1 (entry)] and may not be where the infant was born. In contrast, the first episode on the NNRD is always at the hospital of birth (i.e. episode 1 for PiPS may be episode 2 for the NNRD if a baby was transferred from the hospital of birth to a second hospital where PiPS recruitment took place).

It emerged that during processes of addressing missing data and queries, some PiPS episodes had been renumbered and did not appear on the database sequentially. An additional problem arose because data for PiPS were recorded for all episodes, including those spent on paediatric wards and in specialist surgical or cardiac centres that do not use the neonatal EPRs and, thus, these episodes were not present in the NNRD. Consequently, it was necessary to check the dates, anonymised patient ID and hospital name of each episode on each database and renumber, when necessary, to ensure that comparisons were indeed for the same episode. This was further complicated because of inconsistencies of the names of hospitals and NHS trusts entered as free text on the PiPS database. By contrast, the names of hospitals and NHS trusts are standardised in the EPRs through the use of drop-down menus.

Sources of items within the databases

Although many of the data points recorded for the PiPS trial are also present in the NNRD, there are differences in how some data were obtained. In general, data for PiPS were obtained by asking a direct question (e.g. 'during this episode of care did the infant have surgical ligation for a PDA?'); by contrast, information on whether or not an infant has surgery for a PDA can be entered into the NNRD by a range of routes including variables in the 'daily data', 'discharge diagnoses' and 'procedures' EPR fields.

Preparation of data sets for linkage

We compared data items only for episodes present in both the PiPS and the NNRD databases. We excluded episodes that could not be linked.

Step 1: linking infants to National Neonatal Research Database by matching electronic patient record ID ('Badger ID')

We carried out linkage using the unique Badger ID as identifier. This is held as an identifier within the NNRD and is generated at EPR level. The NPEU provided to Clevermed Ltd the NHS numbers and date of birth of all PiPS recruits, requesting the Badger ID. The PiPS data were then anonymised and provided to the NDAU, identified by the Badger ID only.

Step 2: linking episodic data

We linked individual PiPS episodes of care, after any necessary renumbering as described above, to episodes on the NNRD, using the Badger ID and episode number.

Step 3: linking daily data

The NNRD does not hold any dates. The time of events is indicated by a variable 'minutes from birth'. We used this variable to link data describing feeding, antibiotic and antacid use in the first 14 days.

Methods of comparison

Infant and maternal baseline characteristics: infant-level comparison

These data from the PiPS database were recorded from episode 1 when the infant was recruited into the PiPS study. Data from the corresponding episode at the same neonatal unit on the NNRD were extracted for comparison of all baseline characteristics.

Processes of care and interventions: episodic and infant level comparisons

The majority of processes of care and interventions were compared on an episodic level. Exceptions were as follows: first, the interventions specifically recorded for PDA and ROP that may be entered into multiple EPR fields and, hence, are available in multiple locations in the NNRD and are often carried over across episodes of care to provide the full medical history, and, second, the details of enteral feeds, antibiotics and antacids in the first 14 postnatal days, which are available across episodes on both databases regardless of transfer status, were all compared at an infant level.

To be confident of the accuracy of data for those infants included in the analysis of 'first day of milk feed' and 'type(s) of milk received on first day of feeding', we included only those that had complete daily data on the NNRD for all days prior to the first feed (i.e. this analysis differs from others in that the PiPS data are linked to selected eligible infants on the NNRD rather than linking NNRD to PiPS data). For the summary variable 'all types of milk received during first 14 postnatal days' we included all infants with linkable days on both databases on which detailed daily feeding data (maternal, donor breast or formula milk) were available.

Statistical methods for assessing agreement

We assessed items that were identical, and items with minor and major discordance, using predefined criteria based on clinical judgement and set a priori by KC and CB to mitigate bias.

We calculated for each variable of interest the proportion of babies for whom the NNRD and PiPS trial data differed and the 95% CIs for the proportion. For variables with discordance rates of $< 5\%$, we used the Poisson approximation to the binomial to calculate CIs; otherwise we used the Agresti and Coull method[145] for binomial CIs, as this method has better coverage properties.[146] As observations from the same hospital are not independent, we calculated the CIs for discordance using generalised linear models with variances estimated to allow for within-hospital correlation.[147] To test whether or not the discordance rate had changed over the course of recruitment, we calculated discordance rates for sequential quarters for five key variables (i.e. antenatal steroids, mode of delivery, day of first milk feed, type of first milk and central line days) and tested a time trend using linear regression with weights, to allow for a varying number of observations at each time point and adjusting for clustering by hospital. We assessed autocorrelation using residual plots and the Breusch–Godfrey test and accounted for this using the Prais–Winsten procedure, if necessary. To investigate whether or not discordance varied by recruitment site, we calculated discordance rates separately for each hospital for five key variables (i.e. antenatal steroids, mode of delivery, birthweight, EDD and central line days).

To check case ascertainment for binary clinical outcomes, we calculated sensitivity and specificity, treating PiPS data as the gold standard. For continuous variables, we calculated mean and median differences and 95% limits of agreement for the differences.

Sensitivity, specificity and positive predictive values

For binary outcome variables, we calculated the sensitivity, specificity and PPVs of NNRD data in comparison with PiPS data.

In the context of this data comparison, we have taken the prevalence of an outcome as the proportion of infants on the PiPS database with that outcome reported. The following definitions have been used:

- Sensitivity is the ability of the NNRD database to correctly classify an individual as 'diseased' as indicated by the gold standard PiPS database.
- Specificity is the ability of the NNRD database to correctly classify an individual as disease free as indicated on the PiPS database.
- Positive predictive value is the percentage of individuals who are identified on the NNRD as having the disease who actually do have the disease as indicated on PiPS database.

Sensitivity and specificity are characteristics of the test and are, in contrast to the PPV, unaffected by the prevalence of the outcome.

Regulatory issues

The establishment of the NNRD and the PiPS trial had each been approved by a Research Ethics Committee (REC) (10/H0803/151 and 09/H0604/30 respectively; patient information leaflet is provided in *Appendix 5*). Advice was sought from the REC chairperson regarding whether or not an additional application either as a stand-alone project or as an amendment to the existing approval for the PiPS trial was required before undertaking this analysis. We were advised that no such application was necessary.

A data-sharing agreement was then put in place between the NPEU, University of Oxford, where the PiPS trial data are held, and Imperial College London for the transfer of PiPS trial data. These data were stripped of identifiers, other than the Badger ID, and were sent to the NDAU. Database merging and all subsequent analyses took place at the NDAU.

Results

Linkage

A total of 1315 babies were recruited into the PiPS trial; the parents of five babies withdrew consent including consent for the use of any data and, therefore, data for 1310 babies were available for analysis. Clevermed Ltd was able to provide Badger ID for 1280 (98%) infants [no EPR data could be identified for 30 (2%) recruits into the PiPS trial]. This resulted in data for 2360 episodes of care being available on the PiPS database (*Figure 15*).

Of the second episodes on the NNRD, 81 were the first episode on PiPS because the baby was recruited at a different hospital from that of birth. There were 103 episodes on PiPS from 22 infants who could not be reliably matched on the NNRD, because they were in a ward or a hospital that was not entering data onto BadgerNet, because of duplicate reports of the same episode on the NNRD or, in the case of two infants, because they could not be found in the NNRD. All data for these 22 infants were excluded from the analyses, leaving 2257 episodes of care from 1258 infants eligible for comparison (see *Figure 15*).

Infant and maternal characteristics

We compared baseline infant and maternal baseline characteristics for 1258 infants. The numbers of infants with missing data for each variable in both databases are reported along with the minor and major discordances calculated using the predefined criteria (*Table 14*). Gestational age on the PiPS database is calculated from the EDD, whereas gestational age in weeks and days is recorded directly on the EPR and, hence, the NNRD.

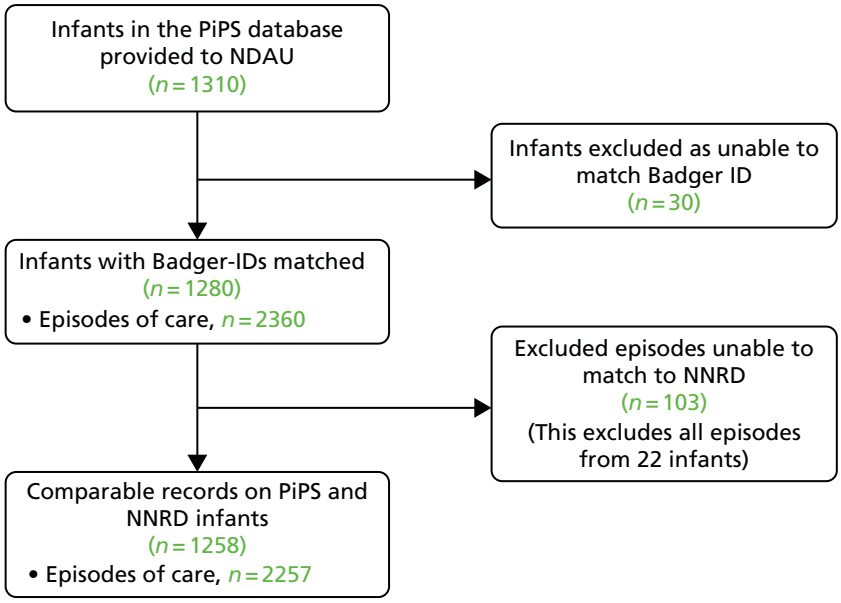


FIGURE 15 Records from PiPS CRF and the NNRD.

TABLE 14 Comparison of baseline infant and maternal characteristics

Baseline variable	Number of comparable infants	Missing data, *n* (%)		Discordance[a]			
				Any		Major	
		PiPS	NNRD	Rate (%)	95% CI (%)	Rate (%)	95% CI (%)
EDD	1142	0	116 (9.2)	9.9	8.3 to 11.8	4.1	3.0 to 5.5
Gestational age	1142	0	1	20.4	18.2 to 22.7	3.0	2.1 to 4.1
Month of birth	1257	0	1 (0.08)			0	0
Year of birth	1257	0	1 (0.08)			0	0
Birthweight	1257	0	1 (0.08)	1.7	1.0 to 2.6	0.9	0.4 to 1.6
Sex	1256	0	2 (0.16)			0.2	0.02 to 0.6
Apgar score at 5 minutes	1192	33 (2.62)	63 (5.01)	2.6	1.8 to 3.7	0.8	0.4 to 1.5
Born in this hospital	1257	1 (0.08)	0			1.5	0.9 to 2.4
Singleton or multiple	1257	0	1 (0.08)			1.1	0.6 to 1.9
Birth order	1257	0	1 (0.08)			0.6	0.3 to 1.3
Maternal year of birth	1255	0	3 (0.24)			1.4	0.9 to 2.3
Maternal ethnicity (NHS categories)	1185	10 (0.79)	64 (5.09)			10.2	8.6 to 12.1
Maternal LSOA	1090	24 (1.91)	151 (12.08)			16.5	14.4 to 18.8
Any antenatal steroids given	1243	9 (0.7)	6 (0.58)			2.4	1.76 to 3.4
Caesarean or vaginal delivery	1201	1 (0.08)	56 (4.5)			8.7	7.2 to 10.4
Instrumental delivery	1248	9 (0.72)	1 (0.08)			1.1	0.6 to 1.9

a Infants with missing data are excluded from calculations of discordance.
Shaded cells appear for variables for which any discordance is classified as major discordance.

Missing data on the PiPS database were few (< 3%): EDD was missing for 9.2% of infants in the NNRD, LSOA for 12.1% and maternal ethnicity for 5.1%. Rates of 'any discordance', defined as ±3 days, were 9.9% and 20.4% for EDD and gestational age, respectively. Major discordance rates were < 10% for all variables except for maternal ethnicity (10.2%) and maternal LSOA (16.5%).

Processes

We were able to compare 2257 linked episodes for 1258 infants (*Table 15*) using the predefined criteria. Major discordances, defined as ± 5 days, were 10.2% and 11.2% for the duration of high-dependency care and central venous lines, respectively. Discordance for medical treatment of a PDA was 6.0%. For all other variables, discordancies were < 5%.

Feeding data

Of the 1258 infants whose records could be matched, 29 on the PiPS database and 35 on the NNRD were reported as having no enteral feeding in the first 14 days. Of the 1223 for whom both databases contained any days with completed details of feeds given, 343 on the NNRD had missing days before the first reported feed and so they could not reliably be included in the analysis of first feeding.

The analysis of the summary of all milk feeds received over the first 14 days includes days when a report was completed confirming that no milk had been given. There were five infants for whom the NNRD contained no reports on any day on whether or not any feed was given; thus, this analysis includes data for 16,203 days from 1253 infants for whom feeding data were complete on both databases.

There was high agreement for day of first milk feed, with 2.8% major discordance (≥ 2 days difference) (*Table 16*). However, there was high disagreement for the type or types of milk given on first day of milk feed (22.3%) and for the summary of different milks given over the first 14 days (13.8%).

TABLE 15 Comparison of processes and interventions, excluding feeds and medicines in the first 14 days

Variable	Number of comparable records	Discordance			
		Any		Major	
		Rate (%)	95% CI (%)	Rate (%)	95% CI (%)
Comparison by episode					
Intensive care days	2257	8.5	7.4 to 9.7	3.9	3.1 to 4.8
High-dependency care days	2257	14.2	12.8 to 15.7	10.2	9.0 to 11.5
Central venous line	2257	20.5	18.9 to 22.2	11.2	10.0 to 12.6
Length of stay	2257	4.0	3.2 to 4.9	3.3	2.6 to 4.2
Transfer to another hospital	2257			2.2	1.6 to 2.9
Discharge month	2257			2.3	1.8 to 3.1
Discharge year	2257			0.5	0.3 to 0.9
Comparison at infant level					
Surgery for PDA	1258			1.7	1.1,2.6
Medical treatment of PDA with ibuprofen or indometacin	1258			6.0	4.8 to 7.5
ROP treatment by laser or cryotherapy	1258			1.6	1.0 to 2.5

Shaded cells appear for variables for which any discordance is classified as major discordance.

TABLE 16 Comparison of feeds and medicines in the first 14 postnatal days

Variable	Number of comparable records	Discordance			
		Any		Major	
		Rate (%)	95% CI (%)	Rate (%)	95% CI (%)
First 14 postnatal days					
Day of first milk feed	880	6.7	5.2 to 8.6	2.8	1.8 to 4.2
Type(s) of first milk feed	880			22.3	19.6 to 25.1
Summary of all types of milk in first 14 days	1253			13.8	12.0 to 15.8
Whether or not any antibiotic given in first 14 days	1258			0.6	0.2 to 1.1
Number of days that antibiotics were given	1258	21.4	19.2 to 23.7	9.0	7.6 to 10.8
Whether or not any antacid was given	1258			5.1	4.0 to 6.4
Number of days antacid given	1258	6.8	5.6 to 8.4	4.8	3.7 to 6.2

Shaded cells appear for variables for which any discordance is classified as major discordance.

Antacids and antibiotics

Antacid and antibiotic administration details are recorded on the EPRs in the 'daily medications' field by selecting from a drop-down menu, the completion of which is not essential. Absent data were included in the analysis as indicating that antibiotic and/or antacid was not given. All 1258 infants were considered eligible to be included in this comparison (see *Table 16*). Although whether or not any antibiotics were given in the first 14 days had high agreement and only 0.6% discordance, the number of days of antibiotic use had a major discordance (> 2 days) of 9.0% and a high 'any discordance' rate (±2 days) of 21.4%. Reporting of antacid indicated 5.1% discordance for any use and 9.0% for the number of days given.

Outcomes compared at infant level

Only 877 infants who were still inpatients at 36 weeks' postmenstrual age were eligible for comparison of oxygen supplementation and ventilatory support at that time; all other outcomes were summarised and compared for all 1258 infants from all linkable episodes of care. Any disagreement was prespecified as 'major'. Discordance is < 10% for all outcomes except the continued use of oxygen at 36 weeks' postmenstrual age, which had a discordance rate of 13.3% (*Table 17*).

Sensitivity and specificity

We report sensitivity, specificity and PPVs for NNRD data in *Table 18*. In the conventional context, the sensitivity would be the probability of a test that correctly identifies an individual with a disease as 'diseased'; in this context, it is the probability of being 'NNRD disease positive' when disease is present. Sensitivity for outcomes other than survival to discharge, which is 100%, ranges between 50% and 87%. There is a 50–87% chance that infants with the disease are 'NNRD disease positive'. Therefore, there is under-reporting of disease in the NNRD. Specificity was > 85% for all outcomes, with the majority being > 90%. Infants without the disease have a high probability of being 'NNRD disease negative', NNRD correctly identifying infants without disease. With the exception of BPD and medical treatment for PDA, which have a prevalence of 49.0% and 20.3% respectively, the prevalence of these adverse outcomes is low, at < 6%.

The PPV of all outcomes with the exception of treated ROP (71.4, 95% CI 56.7 to 83.4), perforated NEC (66.0, 95% CI 51.2 to 78.8) and a range of details of cerebral ultrasound scans was > 75.

TABLE 17 Comparison of outcomes by infant

Variable	Number of comparable records	Major discordance	
		Rate (%)	95% CI (%)
Outcomes			
Worse stage of ROP in any eye	1258	2.0	1.3 to 2.9
Whether or not infant was receiving supplementary oxygen at 36 weeks postmenstrual age	877	13.3	11.2 to 15.8
Whether or not infant was receiving mechanical respiratory support at 36 weeks postmenstrual age	877	9.2	7.4 to 11.3
Any diagnosis of perforated NEC	1258	2.1	1.4 to 3.1
Any gastrointestinal perforation	1258	1.7	1.1 to 2.6
Any abdominal surgery for NEC	1258	2.8	1.9 to 3.9
Haemorrhagic parenchymal infarct	1258	2.7	1.9 to 3.8
Hydrocephalus	1258	1.4	0.8 to 2.2
Periventricular leucomalacia	1258	1.7	1.1 to 2.6
Porencephalic cyst	1258	2.8	1.9 to 3.9
Survival to discharge from neonatal care	1258	0.2	0.02 to 0.6

By hospital analysis

We compared major discordance rates for five variables (i.e. antenatal steroids, mode of delivery, birthweight, EDD and days with a central line) across the 24 PiPS recruiting hospitals (*Table 19*). The discordance rates of mode of delivery and central line days, which in general have higher major discordance, show striking variation at different hospitals, with mode of delivery varying from 0.0% to 18.5% and central line days from 2.7% to 28.6%.

Trends over time

There were no significant changes in discordance rates over time for any of the five selected variables: any discordance in antenatal steroids (–0.5%, 95% CI –1.4% to 0.4%; $p = 0.27$) or mode of delivery (–0.5%, 95% CI –2.9% to 1.8%; $p = 0.64$), major discordance in day of first milk (–0.2%, 95% CI –2.0% to 1.5%; $p = 0.78$), major discordance in type of milk on first day (0.8%, 95% CI –5.2% to 6.8%; $p = 0.78$) and central line days (–1.2%, 95% CI –6.2% to 3.8%; $p = 0.64$).

Conclusions

Our study was designed to test whether or not data entered onto the neonatal EPRs, after the processes they undergo before being entered into the NNRD, are of sufficient completeness and accuracy to be used for a 'gold standard' assessment of a therapeutic intervention, that is, for a randomised controlled trial. For the majority of data items, discordance was low and was comparable to research data. For some of the rarer outcomes, sensitivity was low but PPV was relatively high, suggesting that infants diagnosed as having disease on the NNRD are correctly identified, but that a high proportion with disease are missed. It is not possible, with the configuration of the electronic data system, to distinguish between missing and discordant data. By comparing like items within NNRD and the database of the PiPS trial involving an investigational medical product performed to standards compliant with ICH-GCP, our study not only addresses the primary research question but also stands as a valuable audit of NNRD quality. For this analysis, and to preserve the integrity of the PiPS trial, the trial data had to be considered to be accurate and, indeed, the extent to which missing data were chased and inconsistencies were queried between trial and local clinical staff far exceeded what would be possible for population-based routine data. However, items such as types of different milk

TABLE 18 Sensitivity, specificity and positive predictive values of key processes and outcomes reported on the NNRD as determined by comparison with PiPS data

Variable	PiPS, *n*		Prevalence, % (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)
	Positives	Negatives				
Processes						
Surgery for PDA	60	1198	4.8 (3.7 to 6.1)	70.0 (56.8 to 81.2)	99.4 (98.8 to 99.8)	85.7 (72.8 to 94.1)
Medical treatment of PDA with ibuprofen or indometacin	256	1002	20.3 (18.2 to 22.7)	71.9 (65.9 to 77.3)	99.5 (98.8 to 99.8)	97.4 (93.9 to 99.1)
ROP treatment by laser or cryotherapy	41	1217	3.3 (2.3 to –4.4)	85.4 (70.8 to 94.4)	98.8 (98.1 to 99.4)	71.4 (56.7 to 83.4)
Outcomes						
Whether or not infant required supplementary oxygen at 36 weeks' postmenstrual age	430	447	49.0 (45.7 to 52.4)	86.7 (83.2 to 89.8)	86.6 (83.1 to 89.6)	86.1 (82.5 to 89.3)
Whether or not infant required mechanical respiratory support at 36 weeks' postmenstrual age	214	663	24.4 (21.6 to 27.4)	90.7 (85.9 to 942.2)	90.8 (88.3 to 92.9)	76.1 (70.4 to 81.2)
Any diagnosis of perforated NEC	43	1215	3.4 (2.5 to 4.6)	76.7 (61.4 to 88.2)	98.6 (97.8 to 99.2)	66.0 (51.2 to 78.8)
Any gastrointestinal perforation	55	1203	4.4 (3.3 to 5.7)	83.6 (71.2 to 92.2)	98.9 (98.2 to 99.4)	78.0 (65.3 to 87.7)
Any abdominal surgery for NEC	73	1185	5.8 (4.6 to 7.2)	67.1 (55.1 to 77.7)	99.1 (98.3 to 99.5)	81.7 (69.6 to 90.5)
Haemorrhagic parenchymal infarct	53	1205	4.2 (3.2 to 5.5)	69.8 (55.7 to 81.7)	98.8 (98.0 to 99.3)	71.2 (56.9 to 82.9)
Hydrocephalus	24	1234	1.9 (1.2 to 2.8)	50.0 (29.1 to 70.9)	99.5 (98.9 to 99.8)	66.7 (41.0 to 86.7)
Porencephalic cyst	39	1219	3.1 (2.2 to 4.2)	51.3 (34.8 to 67.6)	98.9 (98.1 to 99.4)	58.8 (40.7 to 75.4)
Periventricular leucomalacia	40	1218	3.2 (2.3 to 4.3)	62.5 (45.8 to 77.3)	99.5 (98.9 to 99.8)	80.6 (62.5 to 92.5)
Survival to discharge from neonatal care	1159	99	92.1 (90.5 to 93.6)	100.0 (99.7 to 100.0)	97.0 (91.4 to 99.4)	99.7 (99.2 to 99.9)

TABLE 19 Any and major discordance rates with 95% CIs for five key variables (i.e. antenatal steroids, mode of delivery, birthweight, EDD and central line days) by PiPS recruiting hospitals

Hospital	Key variable									
	Antenatal steroids		Mode of delivery		Birthweight		EDD		Central line days	
	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)
Barnet	29	3.45 (0.09 to 19.20)	27	7.4 (0.9 to 2.4)	29	0	27	0	65	9.2 (4.0 to 19.0)
Basildon	10	0	10	0	10	10.0 (0.0 to 42.6)	8	12.5 (0.1 to 49.2)	25	28 (14.1 to 47.8)
Croydon	10	0	11	0	11	0	8	0	17	5.9 (0.0 to 28.9)
Guy's & St Thomas'	93	0	90	10.0 (5.1 to 18.1)	94	0	91	1.1 (0.03 to 6.10)	146	13.7 (9.0 to 20.3)
Homerton	248	2.4 (0.9 to 5.3)	237	6.3 (3.8 to 10.3)	249	0.4 (0.01 to 2.20)	236	4.7 (2.3 to 8.3)	309	12.3 (9.1 to 16.5)
King's College Hospital	22	4.5 (0.1 to 25.3)	23	8.7 (1.2 to 28.0)	23	4.3 (0.1 to 24.2)	23	4.3 (0.1 to 24.2)	44	18.2 (9.2 to 32.2)
Luton and Dunstable	31	0	29	10.3 (2.8 to 27.2)	31	3.2 (0.08 to 18.00)	27	7.4 (1.0 to 24.5)	42	7.1 (1.8 to 19.7)
Medway Maritime	73	2.7 (0.3 to 9.9)	72	13.9 (7.5 to 23.9)	73	0	69	1.4 (0.04 to 8.10)	93	14.0 (8.2 to 22.6)
Newham General Hospital	59	6.8 (2.2 to 16.6)	57	17.5 (9.6 to 29.6)	59	0	53	9.4 (3.7 to 20.7)	110	7.3 (3.5 to 13.9)
North Middlesex	22	4.55 (0.1 to 25.3)	21	4.8 (0.1 to 26.5)	22	0	22	0	69	15.9 (9.0 to 26.5)
Oxford John Radcliffe	69	1.4 (0.04 to 8.10)	67	4.4 (0.9 to 13.1)	71	0	61	0	74	2.7 (0.3 to 9.8)
Queen's Hospital Romford	56	8.9 (3.5 to 19.7)	53	7.5 (2.5 to 18.4)	56	1.8 (0.05 to 10.00)	52	1.9 (0.05 to 10.70)	80	7.5 (3.2 to 15.7)
Royal Sussex County	26	0	21	9.5 (1.4 to 30.1)	27	7.4 (1.0 to 24.5)	23	4.3 (0.1 to 24.2)	42	28.6 (17.1 to 43.7)

Hospital	Key variable									
	Antenatal steroids		Mode of delivery		Birthweight		EDD		Central line days	
	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)	Total comparable records (episodes)	Major discordance, % (95% CI)
Southend Hospital	19	0	18	5.6 (0.0 to 27.6)	20	0	18	0	31	3.2 (0.08 to 18.0)
St George's Hospital	54	0	52	13.5 (6.4 to 25.6)	56	1.8 (0.05 to 1.0)	52	1.9 (0.05 to 10.7)	67	9.0 (3.8 to 18.5)
St Peter's Hospital	91	0	90	8.9 (4.4 to 16.8)	91	0	83	4.8 (1.3 to 12.3)	102	10.8 (6.0 to 18.4)
The Royal London Hospital	68	2.9 (0.4 to 10.6)	65	4.6 (1.0 to 13.5)	69	1.4 (0.04 to 8.1)	60	1.7 (0.04 to 9.3)	156	10.9 (6.8 to 16.8)
Tunbridge Wells Hospital	32	3.1 (0.08 to 17.40)	32	12.5 (4.4 to 28.7)	32	0	31	6.5 (0.8 to 21.7)	54	3.7 (0.4 to 13.4)
University College London	90	4.4 (1.2 to 11.4)	86	11.6 (6.3 to 20.3)	90	0	79	5.1 (1.6 to 12.7)	125	7.2 (3.7 to 13.3)
University Hospital Lewisham	21	0	21	4.7 (0.1 to 26.5)	21	0	20	0	33	6.1 (0.7 to 20.6)
Watford General Hospital	26	0	27	18.5 (7.7 to 37.2)	27	0	25	16 (5.8 to 35.3)	43	7.0 (1.7 to 19.3)
Whipps Cross University Hospital	28	0	27	0	28	0	25	0	83	4.8 (1.3 to 12.3)
Whittington Hospital	7	0	5	0	7	0	7	14.3 (0.5 to 53.3)	26	7.7 (1.0 to 25.3)
William Harvey Hospital	59	3.3 (0.4 to 12.2)	60	6.7 (2.2 to 16.4)	61	3.3 (0.4 to 11.8)	42	14.3 (6.3 to 28.2)	68	14.7 (8.0 to 25.2)
Total	1243		1201		1257		1142		1904	

feeds are often poorly recorded and inaccuracies were probably present in both data sources. For some items, such as days of intensive care and days with central lines in place, we had assumed that those completing trial CRF would be likely to refer to the EPRs and that there might be bias in favour of low discordance. In the event, the discordance of these items was relatively high, which possibly suggests that, rather than trust the routine data, researchers extracted the data from the clinical notes.

Much of the complexity of our analysis is not pertinent to the primary question but arose because of the preliminary work involved in linking episodes of care, which was essential in order to be confident that the process and outcome data being compared were for the same baby at the same time. It is important in considering the results of this study not to be side-tracked by these issues and lose sight of the main objective. In general, we found that simple objective baseline items compare well with those recorded for trial purposes.

Our study also identifies areas for improvement. We considered data completeness in the NNRD at three levels: first, whether or not an infant recruited into PiPS appeared on the NNRD; second, whether or not all of the episodes of care reported to PiPS were identified; and, third, whether or not individual clinical items, recorded once, daily or across episodes of care, were identified. For 2% of recruits into the PiPS trial, no EPR data could be identified. Whether this was because of errors in the date of birth and NHS number on either the PiPS database or the NNRD or, which however seems unlikely, because the infants were never entered onto the EPR is unclear, but this certainly needs to be better understood.

Data for a further 103 episodes, including all of the data for 22 babies, were lost because they could not be linked. Episodes were linked by dates of admission and discharge and by hospital name. Possible reasons for failure to link are inaccuracies in these dates, inconsistency in whether or not short stay episodes (e.g. transfer out of a neonatal unit for specialist ophthalmological assessment for a few hours) was considered an inpatient or an outpatient episode, and inconsistencies in the names of hospitals and NHS trusts. The first and second reasons above apply equally to EPR and PiPS entries and overcoming them requires clear rules to be applied, and the third reason could easily be addressed by ensuring that hospital names are standardised for research data.

We found the completeness of baseline data on the NNRD to generally be good, with the exception of maternal ethnicity and LSOA (derived from maternal postcode), 5-minute Apgar score and vaginal/caesarean birth. The last two items are particularly surprising as the variables are important clinically. In time, it is probable that real-time linkage between maternal and infant records will exist so that key items, such as these, feed directly into the infant record. In contrast to the CRF, where process and outcome data are recorded in answer to specific questions, a number of important NNRD data items are acquired by entries into EPR tick-box lists, or opportunistic entries in response to episodic events. As a consequence, if an item is ticked it is likely to be true, but in the absence of a tick it is impossible to know definitely if an intervention was not performed, if a condition was not present or if the item was simply overlooked and the data are genuinely missing. This difficulty could be readily addressed through reformatting of EPR entries.

We found low discordance for most baseline data. An exception, the apparent high discordance for the type of feed given on the first day, probably arises from an unrealistic expectation that we could capture the full extent of variability in patient care practices. With this exception, the other principal reason for the high discordance of some items describing process and outcomes appears to reside in the organisation of data entries within the EPR and the consequent impossibility of distinguishing negative from missing items.

We identified high specificity but with low sensitivity for some important outcomes. A probable explanation for this is that the computer screens completed for the EPR in general lack direct questions about presence and absence of outcomes. Instead, reporting is dependent on the outcome being recorded in one of a number of places, which may include an ‘ad hoc’ form that a busy junior doctor may overlook. The probable result of this is under-reporting of outcomes and uncertainty as to whether outcomes that are not recorded are true negatives or simply missing.

The PPV is influenced by the prevalence of a condition. If prevalence is low, a positive report on the NNRD is less likely to true. The fact that the PPV is generally high, despite low overall prevalence for key outcomes, highlights the potential utility of the NNRD as a large and growing population database. Smaller local or regional databases would be unlikely to have adequate statistical power to detect clinically important signals.

The accurate reporting of BPD in this comparison is problematic in that it is dependent on the correct identification of the date on which the infant reaches 36 weeks' postmenstrual age. In the PiPS trial, it was agreed a priori that gestational age and all subsequent assessments of age should be based on the EDD entered at birth with no later changes, whereas on the NNRD a baseline gestation is entered independently of EDD and there is the possibility that clinicians might subsequently revise their view around gestation. It was notable that one of the most frequent reasons why the PiPS staff had to query trial data was because the date the clinical staff had taken as 36 weeks' postmenstrual age was inconsistent with the EDD. As with other areas where differences exist, this could be improved with agreed adoption of standard rules for the determination of these data.

While we were unable to conduct the study as originally planned, providing specific feedback on preliminary data concordance before the final analysis, it was nonetheless disappointing that we were unable to identify any decrease in discordance over time. However, this is at variance with the experience of data required for the National Neonatal Audit Programme, which also utilises data from the NNRD, where a year-on-year improvement has been identified.[148] This may be a consequence of the introduction of regular feedback of missing and potentially erroneous entries, with opportunity for clinical teams to address these and make corrections to the EPR as part of a logged, auditable process.

It is beyond the scope of this study to explore the variation in discordance at different hospitals, and indeed the variation in recruitment rates and the generally low prevalence of adverse outcomes reduce the statistical power of these analyses. However, variation in outcome between hospitals and neonatal clinical networks is an important area of health services research and these data demonstrate the potential utility of NNRD data for this purpose. One area that would be helpful to explore is the possibility that the presence of dedicated staff for data entry and an identified lead for data collection are associated with increased completeness of data and lower discordance between data collected for routine and research purposes.

With the increased adoption of EPRs into clinical practice and the recognition of the importance of extracting the maximum value from the resultant databases, there is increasing interest in their use to support clinical trials. This has included the use of routine data to facilitate the identification of eligible participants[149] and the integration into routine systems of specific items needed for the trial data set.[150,151] In neonatal medicine, data repositories established primarily for observational research and/or benchmarking and audit purposes are increasingly used to support both the identification of recruits and trial conduct[152,153] and to obtain trial outcomes directly from the database.[154] We are unaware of any previous exploration of the possibility of extracting neonatal trial data from repositories of EPR data.

Implications for health care

Our study indicates that the use of NNRD data derived from the neonatal EPRs offers a good opportunity to facilitate clinical research and to reduce the burdens imposed on clinical teams and investigators by data recording requirements. However, our study also identifies areas that require attention before this potential can be exploited. A further important implication of our study is in revealing deficiencies in neonatal medical records. Not only are these used in day-to-day neonatal patient care, but these data are also used to inform the clinical summary and are the basis of hospital performance reports including quality indicators, benchmarking and national audit. Formal examination of the quality and completeness of NHS data is rarely if ever undertaken. This has potentially grave implications for the reliability of the inferences that can be drawn from interrogation of much NHS data. An important strength of our study is in bringing this issue to attention.

In order to provide complete national coverage, EPR coverage needs to be extended to those few inpatient sites not currently providing data to the BadgerNet platform, principally some neonatal surgical centres and independent (private) hospitals.

Research recommendations

The problems relating to data entry that we describe could be readily addressed, in theory, through redesign and reorganisation of EPR entry screens. The intention would be to ensure that, in so far as possible, entries are made in response to simple objective questions with options to provide unambiguous answers. Therefore, we intend to engage with the commercial supplier of the neonatal EPR data to request incorporation of certain relatively minor and straightforward adaptations that are necessary.

Our study highlights the necessity of implementing systematic examination of NHS data quality and completeness and testing methods to improve these measures. These include the involvement of parents (or patients) in quality assuring their data, formal 'sign off' by a senior manager, and incentives [e.g. Commissioning for Quality and Innovation (CQUIN) payments] for achieving predefined data quality standards.

Finally, our study highlights the importance of close clinical involvement in EPR data entry. This issue is considered further in subsequent chapters.

Chapter 5 Two-year neurodevelopmental outcomes of children who were born preterm, assessed using the National Neonatal Research Database

Abstract

Background: Information on the neurodevelopmental outcomes of children who were born very preterm is an important health metric that is required for multiple purposes.

Aims: To assess (1) the agreement between neurodevelopmental outcome information obtained from EPR data held in the NNRD and a gold-standard assessment, (2) the social communication skills of children using a parent-completed questionnaire and (3) the predictive value of early assessments for later cognitive deficits.

Methods: We assessed children at the age of 2 years to a research standard and obtained equivalent information from the NNRD. We invited parents to complete a questionnaire: the Quantitative Checklist of Autism in Toddlers (Q-CHAT). We conducted a systematic review and meta-analysis of early developmental assessment for identifying school-age cognitive deficits.

Results: We completed a formal neurodevelopmental assessment of 190 children; the parents of 141 children completed the Q-CHAT. The neurodevelopmental assessment conducted during NHS follow-up and recorded in the EPRs has low sensitivity but high specificity for identifying children with neurodevelopmental impairment. Very preterm children display greater early childhood social communication difficulties and autistic behaviour than the general population. Early neurodevelopmental assessment has low sensitivity but high specificity for identifying later school-age cognitive deficits.

Conclusions: Neurodevelopmental data in the EPRs underestimate population prevalence of impairment following preterm birth. Very preterm children may benefit from systematic approaches beyond the age of 2 years to identify autistic spectrum disorder (ASD) characteristics and cognitive deficits.

Background

Overview

Around 6000–7000 children who were born very preterm (< 32 weeks' gestation) are admitted to NHS neonatal units each year. They are at substantial risk of adverse neurodevelopmental outcomes. Severe disability rates of between 5% and 56% are reported[155] and long-term studies show that the adverse consequences of preterm birth are still apparent in adolescence and adulthood.[156,157]

Information on the later neurodevelopmental outcomes of preterm infants is necessary for several reasons. For an individual child, outcome assessment is needed to ensure that disability, when present, is identified and timely intervention is provided. Professionals require up-to-date outcome information to counsel, advise and support parents. Neonatal unit and population-based outcome data are essential for service planning, benchmarking and evaluation of the impact of neonatal specialised services and their cost. Neurodevelopment is also a common outcome measure in epidemiological, observational and clinical research.

Most neonatal services attempt to provide follow-up assessments up to around 2 years of age. Ceasing systematic assessment at an earlier age would risk confounding by transient neurological dystonia, which mimics cerebral palsy but which improves or resolves completely during the first year.[158] The literature suggests that a reliable early diagnosis of moderate to severe cerebral palsy can be made by 18 months corrected age, and a reliable early diagnosis of mild cerebral palsy can be made by 24 months corrected age.[159] Although assessment tools, such as the Bayley Scales of Infant Development (BSID), provide standardised mental (cognitive) scores from as early as 12 months of age, the correlation of the early mental scores with subsequent IQ at school age is unclear. Two recent cohort studies reported moderate to substantial agreement between BSID, second edition (BSID-II), Mental Development Index (MDI) at the age of 2 years and full-scale IQ at the age of 5 years among infants born before 30 weeks' gestation or with very low birthweight (VLBW) (i.e. birthweight of < 1500 g).[160,161] Conversely, Hack *et al.*[162] described a considerable reduction in the proportions of extremely low-birthweight infants (i.e. birthweight of < 1000 g) who were diagnosed with cognitive impairment (defined as standardised cognitive scores < 70), from 39% at 20 months to 16% at 8 years of age, when the children were tested sequentially. Applying the same diagnostic criteria, Roberts *et al.*[163] also found a reduction in the proportions of very preterm (< 27 weeks' gestation) and extremely low-birthweight infants with cognitive impairment, from 27.3% at the age of 2 years to 19.3% at the age of 8 years.

All follow-up programmes, whether for clinical or research purposes, incur significant costs related to the employment of trained staff, interim assessments, long-term tracking, data management and analysis, and the need for financial and logistic support to be sustained long term. This is often the main constraint on maintaining follow-up assessments. In the UK, there are currently no nationally agreed, implemented and funded policies for very preterm follow-up.

Types of neurodevelopmental outcome measures

Cerebral palsy is the most commonly quoted outcome in neonatal follow-up studies. It is an umbrella term used to describe a group of non-progressive permanent disorders of movement and posture that occur following damage to the developing fetal or infant brain. It is most commonly described based on the nature of the neurological abnormality (e.g. spastic, dyskinetic or dystonic) and the topography of limb involvement.

Even in the absence of cerebral palsy, preterm infants experience abnormal patterns of motor development and neuromotor dysfunction.[164] Several authors have described the presence of transient dystonia, which may mimic cerebral palsy, in the first year of life in almost one-third of VLBW cohorts.[158,165] A meta-analysis of studies of children who were born very preterm (≤ 32 weeks' gestation) reported motor scores of between 0.57 and 0.88 SDs behind their term-born peers.[166]

The most common disability among preterm children is developmental or cognitive delay.[167,168] Cognitive ability can be described using developmental quotients (DQs) or intelligence quotients (IQs) derived through standardised developmental or intelligence tests. Conventionally, a standardised DQ or IQ > 2 SDs below the population mean is used to define impairment or disability, as it represents the lowest functioning 2.3% of the population. The prevalence of developmental or cognitive impairment exists as a gradient that is inversely related to gestational age.[169] A population-based comparison of school-age children who were born before 28 weeks' gestation or with a birthweight of < 1000 g with term-born controls revealed a 0.7 SD reduction in IQ points in the preterm children, after adjusting for sociodemographic factors and exclusion of children with neurosensory impairment.[170] In the EPICure 2 study, which followed infants born before 27 weeks' gestation in 2006 in England, 35% of survivors assessed at the age of 3 years had cognitive scores (predicted MDI) that were > 1 SD below the normative mean.[28]

Preterm infants have delays in receptive language processing,[171] expressive language acquisition,[172,173] articulation and phonological short-term memory.[172,174,175] A meta-analysis of 12 studies published by Barre *et al.*[176] in 2011 reported that very preterm (< 32 weeks' gestation) infants perform between 0.38 and 0.77 SDs below their term-born counterparts in areas of expressive and receptive language. A metaregression

of six studies for the difference in language scores between very preterm infants and term-born controls against the age at assessment between 3 and 12 years suggested that the deficit in language function deteriorated with increasing age.[177]

Published data in the past 20 years estimate that hearing impairment affects between 1.5% and 9% of infants born very preterm, although < 1% had severe bilateral sensorineural hearing loss uncorrectable with hearing aids.[28,167,178–181]

Retinopathy of prematurity resulting from disordered retinal vascular development is a major threat for vision loss in preterm infants and high-risk groups receive regular screening ophthalmic screening examinations.[182] In the UK, ROP affects approximately 17% of infants born very preterm and/or VLBW[183] and it accounts for around 3% of all childhood vision loss.[184]

There is an increased risk of attention deficit hyperactivity disorder (ADHD) and emotional and social disorders, including ASD, among very preterm/VLBW children, compared with the general population.[185–188] Case–control studies have indicated a twofold to threefold increase in the risk of ADHD in very preterm/VLBW infants, compared with term-born controls.[188] The estimated prevalence of ASD has been reported to be 5% in children with a birthweight of < 2000 g[189] and 8% in children who were born at < 26 weeks' gestation.[190] This represents an approximate tenfold increase over the 2–9 per 1000 prevalence estimate in the general population.[191,192]

Nine out of 12 case–control studies published in 1980–2001 and included in a meta-analysis reported an increase in internalising behaviour among the very preterm/VLBW cases at ages 5–12 years; 9 out of 11 studies also reported an increase in externalising behaviour.[169] However, in a more recent meta-analysis based on parents' and teachers' ratings, the difference in internalising behaviour scores reported between preterm/VLBW cases and full-term controls was small (preterm cases' scores were < 0.28 SDs below the scores of the term-born controls), and for externalising behaviour the difference was negligible.[193] The EPICure 1 study reported that extremely preterm children were 3.5 times more likely to have anxiety disorders than their term-born classmate controls.[194]

Standardised developmental and neuropsychological tests

Overview

Standardised developmental tests are considered the 'gold-standard' method for assessing a child's development. The tests provide an inventory of key developmental milestones and are 'standardised' through administration to a large group of children (the normative sample).[195] Standardised scores are age adjusted with a normalised distribution and typically have a mean of 100 and SD of 15. Standardised tests are designed to be administered by qualified examiners who adhere to stringent administration and scoring protocols. From around 1930 to the present day, there has been a continuous and approximately linear increase in the standardised test score.[196,197] Therefore, tests need to be updated and standardised with a contemporary normative sample to remain valid.

There is a range of standardised assessment tools available. A review commissioned by the Department of Health Policy Research Programme to consider tools that can be used as part of the 2- to 2.5-year Healthy Child Programme[198] to monitor child development at population level was completed by the Policy Research Unit in the Health of Children, Young People and Families at the University College London Institute of Child Health. The report included a comprehensive analysis of the advantages and disadvantages of 13 different measures identified through a systematic literature search. In neonatal outcome studies, the BSID is the most commonly used.

Bayley Scales of Infant Development

The Bayley Scales of Infant Development, second edition (BSID-II), recognised to be highly reliable and valid, was the developmental test of choice among most major neonatal research studies, including the EPICure studies,[167,168] the Victorian Infant Collaborative Study[199] and the National Institute of Child Health and Human Development Neonatal Research Network.[200] Despite its popularity, the BSID-II has been criticised for the lack of separate assessments for language and non-verbal skills and for gross and fine motor performance.

The Bayley Scales of Infant and Toddler Development, third edition (Bayley-III),[201] standardised on a cohort of 1700 children in the USA in 2004, ameliorated these shortcomings by providing a more comprehensive assessment in separate cognitive, communication and motor domains, with subscale scores in receptive and expressive languages and fine and gross motor skills. However, several studies have raised concerns that, when compared with the BSID-II, the Bayley-III underestimates neurodevelopmental impairment.[202–205] Crucially, one of the key differences in the standardisation procedure between the two editions was the inclusion of 'clinical cases' (children with cognitive, physical and behavioural issues) to constitute approximately 10% of the Bayley-III standardisation sample. This was made on the basis that excluding these conditions with higher risk for developmental impairment that are normally present in the general population would falsely inflate the average test scores. However, the effect of these clinical cases in the normative sample appeared to be an increase in discrepancy between BSID-II and Bayley-III scores particularly in the lower functioning range,[203,204] leading to an overestimation of ability when the Bayley-III is used in children with suboptimal development. Some studies have developed conversion algorithms or suggested different cut-off scores to determine developmental delay, in order to allow comparison between cohorts.[203,204,206]

The Bayley-III Social-Emotional questionnaire was derived from the Greenspan Social-Emotional Growth Chart, which was reported to have a sensitivity of 67.2% and a specificity of 97.8% in identifying children with ASD.[207] The questionnaire is designed to be completed by parents and is structured according to the anticipated acquisition of functional emotional milestones between birth and 42 months of age. It was standardised on the same normative cohort as the Bayley-III and, therefore, produces a composite score with a mean of 100 and SD of 15.

Standard neurological examination

The use of a standard neurological examination in conjunction with a gross motor functional assessment increases the diagnostic accuracy for cerebral palsy.[159,208] The Hammersmith Infant Neurological Examination (HINE) is a simple, quantitative method for assessing children between the ages of 2 and 24 months to assess their cranial function, posture, movement, tone and reflexes, and it yields an optimality score.[209] The optimality score is valid for use in children who were born preterm.[210]

Assessment of autistic features

Several authors have studied the use of the Modified Checklist for Autism in Toddlers (M-CHAT) among preterm populations. The M-CHAT has promising test characteristics (sensitivity 87%, specificity 99%, PPV 80%, NPV 99%) when validated in a mixed population of unselected and high-risk children.[211] When applied to the preterm population, high positive screening rates of 25% in VLBW infants[212] and 21–41% in infants born before 28 weeks' gestation[213,214] were found. The M-CHAT is poor at differentiating autistic symptoms from neurosensory, cognitive and motor impairments and the specificity of screening for ASD in the preterm population is confounded by the high prevalence of these coexisting morbidities.[213–215] High positive screening rates were also found with other screening tools, such as the Communication and Symbolic Behaviour Scales Developmental Profile Infant-Toddler Checklist,[216] the Infant/Toddler Sensory Profile,[217] and the Pervasive Developmental Disorders Screening Test, 2nd edition.[218,219] A major revision of the M-CHAT, the Quantitative Checklist for Autism in Toddlers (Q-CHAT), has been published.[220]

The Q-CHAT is a parent-completed questionnaire that aims to identify children at risk for autism with a 5-point rating scale (0–4) instead of a binary scoring system for each item. In a preliminary report, Q-CHAT scores from an unselected group of 754 toddlers aged between 17 and 26 months (mean age 21.2 months),

living in Cambridgeshire in the UK, followed a near-normal distribution and were significantly lower (more normal) than the scores of children with ASD.[220] The Q-CHAT has not yet been validated as an ASD screening tool.

Classification of neurodevelopmental outcomes

National Perinatal Epidemiology Unit/Oxford classification of functional status at 2 years

In 1993, a working group of experts formed by NPEU and the former Oxford Regional Health Authority developed a standard minimum data set relevant to the measurement of health status in early childhood.[221] This consisted of patient identifiers (NHS numbers of mother and child, and child's date of birth), sociodemographic measures (postcode, mother's age at delivery, age last in full-time education and support status at birth), perinatal variables (birthweight, gestation, gender, plurality, hospital of birth, and presence of congenital anomaly) and information on the child's health and functional status in eight clinical domains at the age of 2 years, based on responses to 11 key questions. This set of 11 key questions became known as the 'Health Status Questionnaire' or the 'NPEU/Oxford criteria for disability'.

In 2007, a working group of the British Association of Perinatal Medicine (BAPM) and the National Neonatal Audit Project based in the Royal College of Paediatrics and Child Health[222] specified a data set based on the model of the NPEU/Oxford criteria to allow standardised classification of preterm children at 2 years corrected age into one of three outcome groups: (1) normal, (2) impairment without severe disability (or mild–moderate disability) or (3) severe disability. Moderate agreement between the NPEU/Oxford criteria and other methods of assessing disability had been reported.[223] The NPEU/Oxford classification had been used by several studies in the UK to report 2-year outcomes of preterm children,[179,224] most notably the EPICure 1 study.[167,168] Comparing the disability profile of the EPICure 1 cohort at 30 months and at 6 years, the use of the NPEU/Oxford classification at 30 months corrected age had 50% sensitivity and 93% specificity for moderate or severe disability at 6 years of age.[225]

Functional classification of cerebral palsy

The Gross Motor Function Classification System (GMFCS) is a method for categorising the gross motor functional abilities of children with cerebral palsy.[226] It is widely used internationally and it has proven to be reliable for classifying children with cerebral palsy to allow comparisons between different studies.[227–229] Classification using GMFCS at the age of 2 years has been found to be stable over time.[229] The Manual Ability Classification System is designed for children between the ages of 4 and 18 years and it provides a description of how children with cerebral palsy use their hands to handle objects in daily activities.[230]

Neonatal follow-up programmes in the UK

Neonatal follow-up programmes are not universal in the UK. Some neonatal networks have set up regional projects,[231] but the cost of setting up and running a follow-up programme, which includes training staff, is considerable.[232] Most neonatal units offer routine clinical follow-up for infants born very preterm, but the proportion that actually receive the assessment is unknown. In addition, the approach to the assessment of neurodevelopment during routine clinical follow-up varies widely. Children may be assessed by a neonatal or community paediatrics consultant, staff grade doctor, associate specialist, trainee doctor, advanced neonatal nurse practitioner or an occupational or developmental therapist.

In the UK, there is an established surveillance programme to monitor the health and development of all children.[198,199] In the 1990s, several studies investigated the extent to which data recorded during routine service delivery can be used to report the outcomes of survivors of neonatal intensive care.[224,233,234] The Trent Neonatal Follow-up Project reported that most of the data required to meet the NPEU/Oxford minimum data set could be extracted from routinely available information systems.[233] However, the quality of the data was variable and there was no standardisation in the interpretation or documentation of clinical assessments. An exercise on data linkage between the neonatal register and the community child health surveillance database produced 'error-free' linkage (using the identifiers date of birth, birthweight and gestation) in only 53.9% of children who had received neonatal intensive care. Modi and Carpenter[235] reported similar problems

when they reviewed the use of district and regional child health database in the North Thames Region to ascertain the 2-year health status of children who were born at < 29 weeks' gestation. They were able to retrieve child health surveillance records for only 2 out of 80 children surviving to 2 years. When Johnson and King[234] used the routine child health information system to compile a list of children with motor or sensory disability, they failed to identify 162 out of 446 (36.3%) children listed on the coexisting population register of cerebral palsy, sensorineural deafness and severe vision loss. Since 1992, several reports have highlighted the need for data collection on the later morbidity of survivors of neonatal intensive care.[236–239] The Audit Commission[237] proposed that all neonatal units collect data in a nationally agreed format. The Department of Health and Social Care (DHSC)'s report *Changing Childbirth* recommended the development of a system of data collection to enable meaningful comparison of perinatal statistics.[239]

The British Association of Perinatal Medicine (BAPM) first published standards for hospitals providing neonatal services in 2001, recommending that the later health status of survivors at particular risk of disability should be ascertained up to at least a corrected age of 2 years and standardised guidelines for the definition of disability should be used.[240]

Despite these recommendations, routine outcome reporting of health outcomes following preterm births remains largely unavailable. In 2007, the National Audit Office reported that evidence of neonatal outcomes, other than the traditional indicator of mortality rates, was still sparse.[241]

The National Institute for Health and Care Excellence (NICE), in 2010, published a list of statements that define high-quality specialist neonatal care.[242] This included evidence of processes to enable collection of health outcome data on babies who receive specialist neonatal care. The NICE guideline for developmental follow-up of preterm infants is also currently being developed and is anticipated to be published in full in 2017. Some neonatal networks have included a target for 2-year assessment of very preterm infants in the CQUIN payment framework to encourage follow-up and data collection.

Parent-completed questionnaires

Parent-completed questionnaires have been developed as a low-cost alternative to developmental tests to identify children with disabilities. The level of agreement between parental perceptions and paediatrician assessments is inconsistent[243–245] and may be influenced by parent sociodemographic factors. The validity of the revised Parent Report of Children's Abilities-Revised (PARCA-R),[246,247] the Parent's Evaluation of Developmental Status (PEDS),[248] the Functional Status II (FS-II) questionnaire,[249] the Ages and Stages Questionnaire (ASQ)[250] and a questionnaire adapted from the Griffiths Developmental Scales[251] had been evaluated in the preterm population. In particular, the PARCA-R was found to have good diagnostic utility for moderate to severe cognitive and language impairment when validated against the BSID-II (reported sensitivity 85%; specificity 87%)[247] and the Bayley-III (sensitivity 75–94%; specificity 79–89%),[252] and had been used for outcome reporting in neonatal studies.[253,254] Although the typical response rates to postal questionnaires were reported to be between 52% and 61%,[255] Field *et al.*,[256] when testing parent-completed questionnaires as a source of outcome data at 2 years following neonatal discharge, recorded a 90% response rate by maintaining contact with the families in the form of Christmas and birthday cards.

Electronic patient records

In the UK, most community child health services hold clinical information from child health surveillance programmes on electronic information systems, although these systems vary from one NHS trust to another and the data are not routinely passed back to neonatal units. In the past decade, all neonatal units in the UK have moved towards routinely recording clinical information in an EPR to facilitate shared care within neonatal networks. The BadgerNet platform is most widely used (www.clevermed.com/). In 2007, a standardised format for the recording of 2-year neurodevelopmental and health status, adapted from the NPEU/Oxford classification of disability, was developed by the Thames Regional Perinatal Group Outcomes Group. This was incorporated into the EPR in 2008. Since 2009, the National Neonatal Audit Programme, delivered by the Royal College of Paediatrics and Child Health, has been using data held in the NNRD for audit purposes, including 2-year health status of children who were born at < 30 weeks'

gestation. The programme has promoted outcome data recording, with an increase in the number of participating neonatal units documenting any 2-year outcome data on the eligible infants from 51 out of 170 units (30%) in 2009 to 158 out of 179 units (88%) in 2013.

Aims and objectives

The aim of this workstream was to evaluate the reliability and utility of neurodevelopmental outcome information on children who were born very preterm obtained in the course of NHS follow-up care. Specific objectives were to:

- compare the agreement between outcome data obtained during NHS follow-up assessments and recorded in the EPR, and outcomes obtained through a formal neurodevelopmental assessment conducted to a research standard
- characterise the early social communication skills and autistic-like traits in children at the age of 2 years who were born very preterm
- perform a systematic review of published literature and meta-analysis of the sensitivity and specificity of early developmental assessment in predicting school-age cognitive deficit in children who were born very preterm.

Methods

Approvals and registration

Approval was received from the National Research Ethics Committee (REC 10/H0720/35) and the NHS Research and Development Department of each study site. The study was adopted onto the UK Clinical Research Network Portfolio (ID 8626). The systematic review meta-analysis was registered on PROSPERO, an international prospective register of systematic reviews (CRD42012002168).

Study sites

Study sites were selected to provide representation of infants from a wide range of ethnic and socioeconomic backgrounds as well as include neonatal units where clinicians of different grades and specialties provide follow-up assessments of preterm infants. Study sites were restricted to within Greater London and Cambridge for practical reasons. The lead consultant responsible for post-discharge follow-up at each hospital was invited to be the local collaborator for the study. No hospital declined participation.

Participants

Eligible participants were children who were born at < 30 weeks' gestation and attended the routine NHS follow-up assessments at participating hospital sites between the corrected ages of 20 and 28 months (age adjusted for prematurity) during the recruitment period (June 2010 to July 2012).

To prevent 'practice effect' bias from repeated testing, children who had received Bayley-III assessment, either as part of their routine NHS assessment or owing to enrolment in other research studies, were excluded. Children from non-English-speaking families were also excluded because the Bayley-III neurodevelopmental assessment was developed to be administered in English and parents would not be able to complete the Q-CHAT and Bayley-III Social-Emotional questionnaires independently.

Recruitment

Local collaborators (i.e. clinical consultants) identified eligible participants and approached parents, who were sent the study information sheet and a letter of invitation to participate (see *Appendix 5*). If they were interested in participating or wished to discuss the study, they were asked to provide their contact details on a pre-printed response form and send it to the researcher in a pre-paid envelope. Alternatively, parents were given study information at the time of their child's NHS follow-up appointment.

Researcher training

The researcher received training on Bayley-III assessment techniques through attendance at a 2-day training workshop, followed by practice sessions that were supervised by Bayley-III expert trainers. Techniques were accredited through a pilot assessment that was independently scored. A score of 100% agreement on all items on the assessment scales was achieved. To ensure reliability and consistency during the study, validation sessions were attended by an observer, who scored assessments administered by the researcher on non-study participants who were born at < 30 weeks' gestation and were 20–28 months old (corrected age). The interobserver agreement between scores was evaluated and the researcher received feedback. The researcher also received training in the standardised neurological examination based on the HINE[209] from an expert trainer.

The research assessment

At the time of assessing the participant, the researcher was blinded to the results from the NHS assessment. The assessments took place in an outpatient clinic room at the same site as the routine NHS assessments. Each participant was accompanied by one or both parent(s). For the Bayley-III assessment, the participant was seated either at a children's table or on his/her parent's lap at the office desk. The test items were administered sitting across the table facing the participant. In the case of twins or triplets, one child was assessed at a time.

Timing of research assessment

The intention was to complete the research assessment within 1 month before or after the participant's NHS follow-up assessment. To minimise potential information bias caused by changes in development during the interval between the NHS and the research assessments, the intention was to administer the research assessment before the routine assessment in approximately half the cohort and after in the other half.

Assessment of cognition, language and neuromotor development

Participants' cognitive, language and motor development were assessed using the Bayley-III.[201] Each test item was scored as 1 (pass) or 0 (fail). If the participant refused to respond to any test item, it was scored as 'failed'. For each scale, the sum of the scores for all items tested between the basal and ceiling levels constitute the participant's raw score. Two types of norm-referenced scores were obtained: scaled scores, which are standardised to a mean of 10 and SD of 3; and composite scores, which have a mean of 100 and SD of 15. For the cognitive scale, both the scaled score and the composite score were derived from the raw score. The language composite score was derived from the sum of the receptive communication and expressive communication scaled scores. Similarly, the motor composite score was obtained from the sum of the fine motor and gross motor scaled scores. The algorithm developed by Moore *et al.*[204] was used to convert the Bayley-III cognitive and language scores into a predicted BSID-II MDI, for the purpose of comparing the classification of neurodevelopmental outcomes into categories of severity based on the two scores. The algorithm is:

$$\begin{aligned}\text{Predicted BSID-II MDI} = {} & 88.8 - [61.6 \times (\text{Bayley-III language composite score}/100)^{-1}] \\ & + (0.67 \times \text{Bayley-III cognitive composite score}).\end{aligned} \quad (1)$$

On the Bayley-III Social-Emotional questionnaire, parents were asked to rate how often their child demonstrated certain behaviours. Scores for each item were allocated according to behaviour frequency as follows: all of the time (5 points), most of the time (4 points), half of the time (3 points), some of the time (2 points), none of the time (1 point) and can't tell (0 point). A score of 0 (equivalent to 'can't tell') was given to questions with incomplete responses; if more than one response was given, the response with the highest score was used.

Assessment for neurological deficits and cerebral palsy

The tools utilised were the standardised neurological examination, based on the HINE,[209] and the 'extremely low gestational age newborn' algorithm, as a structured guide to diagnose and classify cerebral palsy into topography-based categories of quadriparesis (at least three-limb involvement), diparesis (involvement of one or both lower limbs) and hemiparesis (involvement of one side of the body). Functional severity of cerebral palsy was classified into five levels based on the GMFCS;[226,228,257] social communication abilities were judged using the parent-completed Q-CHAT[220] and Bayley-III Social-Emotional[201] questionnaires; parents were sent the questionnaires prior to the appointment; the Q-CHAT consisting of 25 items used a 5-point Likert scale (0–4 points) and scores were allocated according to the methods described by the research team that developed it.[220] Questionnaires with more than six incomplete responses were excluded. The scores from all items were summed to obtain a total Q-CHAT score within a possible range of 0–100.

Record of observed behaviour during the research assessment

The Behavioural Observation Inventory included in the standard Bayley-III was used to record behaviour observed during the assessment. Thirteen types of behaviour were noted: positive affect, enthusiasm, exploration, ease of engagement, co-operativeness, appropriate activity level, adaptability to change, alertness, distractibility, appropriate motor tone, tactile defensiveness, fear or anxiety, and negative affect. Numerical scores were assigned for each behaviour: a score of 2 was given if the behaviour was 'observed most of the time', a score of 1 was given if it was 'observed some of the time' and a score of 0 was given if it was 'never or rarely observed'. The presence of 'distractibility', 'tactile defensiveness', 'fear/anxiety' and 'negative affect' were reverse-scored. Using the same form, the parent(s) or caregiver accompanying the participant was asked to rate how much the child's behaviour during the assessment was representative of his/her usual conduct. A score of 2 points was given for 'very typical (child is like this most of the time)', a score of 1 was given for 'somewhat typical' and a score of 0 was given for 'not at all typical'. Hence, two behavioural rating scores (each with maximum score of 26) were obtained; an examiner rated the behavioural score for the frequency of positive behaviour and a parent rated the score for the typicality of behaviour.

Classification of impairment from the research assessment

Participants were classified into categories of neurodevelopmental status using two methods:

1. SD score groups 'higher than –1 SD', '–1 to –2 SDs' and 'lower than –2 SDs' based on their Bayley-III scores.
2. 'No', 'mild–moderate' and 'severe' impairment groups according to the modified NPEU/Oxford criteria.

Classification of impairment

The Bayley-III composite score was used to assign the SD score group in the cognitive domain. In the language and motor domains, composite scores were derived from combining scaled scores from the receptive and expressive communication subtests and the fine motor and gross motor subtests, respectively. Therefore, if a child had a specific impairment in only one subtest, it is possible for compensation from the other subtest to occur, resulting in a composite score within the normal range. Hence, the scaled score was used to identify specific impairment in the subdomains of receptive communication, expressive communication, fine motor skills and gross motor skills. In the combined language and motor domains, impairment was taken as the worst category of outcome assigned in the respective subdomains and based on the Bayley-III composite scores. The overall outcome of each participant was based on the worst category of impairment from the cognitive, language and motor domains. Participants who received Bayley-III scores of lower than –1 SD were considered to have at least a mild form of impairment and scores of lower than –2 SDs were considered to represent at least moderate to severe impairment.

Impairments were also classified according to the modified NPEU/Oxford criteria.

Outcome data from NHS follow-up assessments

Participants were assessed by their local clinicians as part of their routine NHS post-discharge follow-up. Assessors were blinded to the results of the research assessment. Results were entered into the electronic '2-year outcome' form on the BadgerNet EPR as required for the National Neonatal Audit Programme. The specific questions for the development (cognitive), communication and motor domains were as shown in *Box 1*.

BOX 1 Questions for the development (cognitive), communication and motor domains

Question reference	
Development (cognitive)	
D1	Is the child's development between 3 and 6 months behind corrected age?
D2	Is the child's development between 6 and 12 months behind corrected age?
D3	Is the child's development > 12 months behind corrected age?
Receptive communication	
RC1	Does this child have difficulty with understanding outside of familiar context?
RC2	Is this child unable to understand words or signs?
Expressive communication	
EC1	Does this child have any difficulty with communication?
EC2	Does this child have difficulty with speech (< 10 words/signs)?
EC3	Does the child have fewer than five meaningful words, vocalisation or signs?
Fine motor	
FM1	Does this child have any difficulty with the use of one hand?
FM2	Does this child have difficulty with the use of both hands?
FM3	Is this child unable to use hands (i.e. to feed)?
Gross motor	
GM1	Does this child have any difficulty walking?
GM2	Is this child's gait non-fluent or abnormal reducing mobility?
GM3	Is this child unable to walk without assistance?
GM4	Is this child unstable or needs to be supported when sitting?
GM5	Is this child unable to sit?

A positive response to any of the questions implied the presence of impairment. Questions D3, RC2, EC3, FM3, GM3 and GM5 denote the criteria for severe impairment. Additional information on whether or not the child was diagnosed with cerebral palsy, whether or not a standardised neurodevelopment test was used during the NHS assessment and whether or not the child was difficult to assess were also entered. The electronic form could be completed by the examining health professional or by administrators, such as secretaries or data entry clerks, based on the information given to them by the examiner.

With parental consent, the participants' unique identifier on the NNRD (the Badger ID) was obtained from the local collaborator at each study site. Neonatal and 2-year outcome data were then obtained from the NNRD with assistance from the NDAU data managers.

Classification of disability based on National Neonatal Research Database data

Participants were classified into categories of 'no', 'mild–moderate' and 'severe' impairment within each outcome domain (i.e. cognitive, receptive communication, expressive communication, fine motor and gross motor) using the electronic 2-year outcome data and according to the algorithm outlined in *Figure 16*. A missing response did not count as a 'no'; therefore, complete data entry is required to assign participants as having no impairment. An overall level of impairment was defined based on the worst outcome from the five domains.

In addition, for the purpose of assessing selection bias, the following data were extracted from the NNRD for all infants born between 1 January 2008 and 31 December 2010 at gestational ages of < 30 weeks and discharged from the participating study sites (the 'baseline population'): gestation at birth, birthweight, sex, ethnicity, singleton or multiple pregnancy, mode of delivery, days of mechanical ventilation, oxygen therapy at 36 weeks' corrected gestational age, maternal age and the IMD based on maternal residence at the time of birth. The IMD is a summary measure of relative area deprivation, calculated through a weighted combination of scores from 38 different indicators covering factors, such as income, employment, education, health, living environment and crime, for each area in England, using national census data. The IMD was obtained based on the postcode of the mother at the time of birth of her child and according to the English Indices of Deprivation 2010.[258]

Statistical tests

Data were coded for analysis using Microsoft Excel® 2007 (Microsoft Corporation, Redmond, WA, USA). Data were double-entered, examined and outliers were verified. All analyses were performed using Stata statistical package, version 11.0.

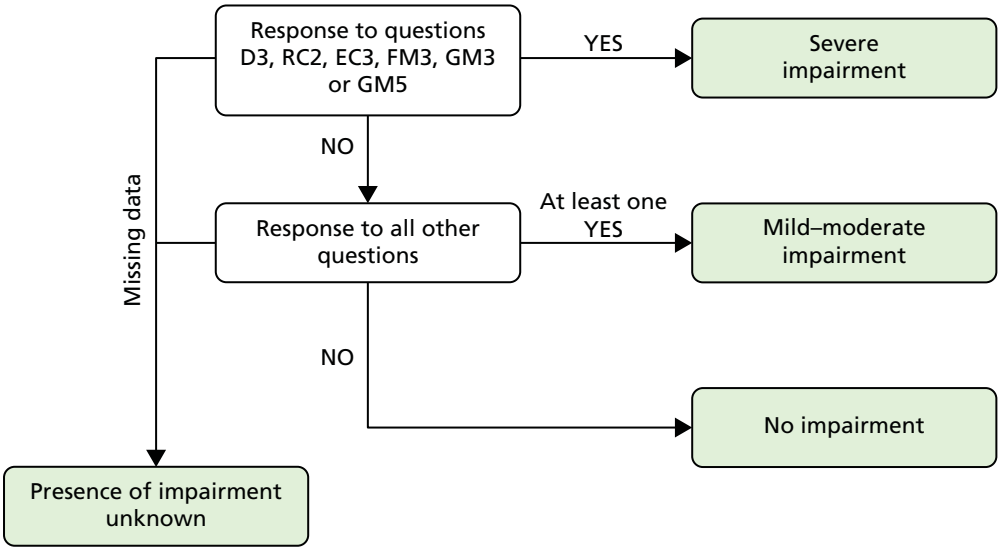


FIGURE 16 Algorithm for the classification of impairment using data from NHS assessments.

Quantitative variables are presented as means and SDs for normally distributed data, or medians and IQRs when the distribution was skewed. Qualitative variables are presented as numbers of subjects and percentages. Differences between categorical variables were analysed using Pearson's chi-squared test. For continuous variables, Student's *t*-test was used for parametric comparison and the Mann–Whitney *U*-test was used for non-parametric comparison. The *p*-values derived from statistical tests are presented and the conventional 5% level is used to define statistical significance. Several key statistical measures used in the analyses are described below.

The validity of an assessment, in the context of this research, refers to the ability of the assessment to differentiate accurately between children with and without neurodevelopmental impairment, as defined. It is described using sensitivity and specificity, which are derived through a 2 × 2 table (*Table 20*).

Sensitivity is the proportion of children with impairment who were accurately identified as having impairment from the assessment under evaluation. It is calculated as:

$$\text{Sensitivity} = {}^{TP}/_{(TP+FN)}. \quad (2)$$

Specificity is the proportion of children without impairment who were accurately identified by the assessment under evaluation and is calculated with the formula:

$$\text{Specificity} = {}^{TN}/_{(TN+FP)}. \quad (3)$$

Sensitivity and specificity calculations are expressed as either proportions or percentages with corresponding 95% CIs. Values of < 0.7 (or 70%) were interpreted as low, values of 0.7 to 0.85 (70% to 85%) were interpreted as moderate and values of > 0.85 (> 85%) were interpreted as high.[259]

Cohen's kappa statistic was used to compare the agreement in classifying neurodevelopmental outcomes into the three categories of 'no', 'mild–moderate' and 'severe' impairment (ordinal data). The κ coefficient is a measure of the proportion of agreement above that is due to chance alone and is calculated by:

$$\kappa = \frac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}}. \quad (4)$$

A κ value of 1 indicates perfect agreement and a value of 0 reflects agreement that is no better than by chance. Unweighted and weighted forms of κ coefficient were obtained. The purpose of the weighting was to derive a coefficient that provided a closer reflection of the clinical implications of disagreement between the ordinal categories. It is clinically more important to distinguish patients with impairments from those without impairment than to differentiate between the severity of 'mild–moderate' and 'severe' impairments. Hence, in the calculations for the weighted κ coefficient, discrepancy between 'mild–moderate' and 'severe' impairments was considered to be partial agreement.

TABLE 20 True and false positives and negatives

	Reference 'gold standard' assessment	
Assessment under evaluation	**Children with impairment**	**Children without impairment**
Tested positive for impairment	True positives	False positives
Tested negative for impairment	False negatives	True negatives

The weighting matrix used is shown in *Table 21*.

The κ values were interpreted according to the standards proposed by Landis and Koch: 0–0.4, slight to fair agreement; 0.4–0.6, moderate agreement; and 0.6–1, substantial to perfect agreement.[260]

Sample size

A precision analysis for the estimated sensitivity of the NHS assessment in identifying children with Bayley-III scores of < –2 SDs in study 1 was used to calculate the target sample size for recruitment. The desired sensitivity of a developmental test is conventionally between 70% and 80%. The precisions (widths of CI) of the observed sensitivity and specificity of a test vary depending on sample size and the observed estimates. The aim was to achieve a precision of 95% CI half-width within 10% for the estimated sensitivity of identifying children with Bayley-III scores of < –2 SDs by the NHS assessment.

Based on the London Perinatal Networks 2008 Annual Report,[261] it was estimated that approximately 500 children are born at < 30 weeks' gestation and survive to discharge from the participating hospitals per year. Assuming that 10% of these children have Bayley-III scores of < –2 SDs, with an unstratified random sample, 650 participants would be required to achieve a CI half-width within 10% for an estimated sensitivity of 80%. We attempted to recruit a stratified sample to include higher proportions of children with medium and high risk for impairment to improve the precision of the study while maintaining a practical sample size[262] (*Table 22*).

Of the infants born at < 30 weeks' gestation who survived to discharge in London in 2008, 20% were born at ≤ 25 weeks' gestation (higher-risk group), 30% were born at 26 to 27 weeks' gestation (medium-risk group) and 50% were born at 28 to 29 weeks' gestation (lower-risk group). Assuming that 25% of higher-risk, 15% of medium-risk and 0% of lower-risk children achieve Bayley-III scores of < –2 SDs, and the sensitivity of identifying different severity of impairment is the same for all risk groups, *Table 22* shows various sample size options and the resulting CI half-width for different sensitivity estimates. We aimed to recruit 500 children (i.e. 200 from the higher-risk group, 200 from the medium-risk group and 100 from the lower-risk group) over the 2-year recruitment period.

Representativeness of the study population

Neonatal and sociodemographic characteristics of study participants and non-participants were compared using data extracted from the NNRD.

Comparing classification of impairments

To estimate how comparable the three levels of impairment (i.e. none, mild–moderate and severe) based on the modified NPEU/Oxford criteria are to the three Bayley-III SD score groups of '> –1 SD', '–1 to –2 SDs' and '< –2 SDs', using only the data obtained from the research assessment, the two sets of criteria were cross-tabulated and the unweighted and weighted Cohen's κ coefficients were calculated. This was also performed for each neurodevelopmental domain.

TABLE 21 Weighting matrix

	Level of impairment based on assessment I[b]		
Level of impairment based on assessment I[a]	None	Mild–moderate	Severe
None	1	0	0
Mild–moderate	0	1	0.5
Severe	0	0.5	1

1, full agreement; 0.5, partial agreement; 0, no agreement.

a The Cohen's κ statistic was used for comparisons where: assessment I = classification using Bayley-III scores and assessment II = classification based on the modified NPEU/Oxford criteria assessment I = NHS assessment and assessment II = research assessment.

TABLE 22 Precision of estimated sensitivity for different sample sizes and sensitivity estimates

Size of strata		Sample size	Estimated proportion with Bayley-III scores of < –2 SDs	Estimated sensitivity (%)	95% CI half width (%)
Unstratified sample		650	10%	80	9.7
Unstratified sample		650	10%	50	12.2
Higher risk	200	500	25%	80	8.9
Medium risk	200		15%		
Lower risk	100		0%		
Higher risk	200	500	25%	50	11.2
Medium risk	200		15%		
Lower risk	100		0%		
Higher risk	100	300	25%	80	11.4
Medium risk	150		15%		
Lower risk	50		0%		
Higher risk	100	300	25%	50	14.2
Medium risk	150		15%		
Lower risk	50		0%		

Any participant with Bayley-III scores of < –1 SD was considered to have at least mild impairment. Taking the research assessment to be the reference 'gold standard', the sensitivity and specificity of the NHS data in identifying children with *any* impairment were calculated and the sensitivity and specificity of the *severe* impairment category in the NHS data for identifying children with Bayley-III scores of < –2 SDs were calculated.

To account for correlation clustering by study sites, robust standard errors were used to calculate the 95% CI for the estimated sensitivities and specificities. To examine the effect of correlated outcomes within multiple birth sets, analyses were repeated on all singleton births and one randomly selected child from each multiple birth set.

Weighted and unweighted κ coefficients were used to measure the concordance between the research and the NHS assessments, again matching the 'no impairment' category to Bayley-III scores of higher than –1 SD, 'mild–moderate' to Bayley-III scores of between –1 and –2 SDs and 'severe' to Bayley-III scores of lower than –2 SDs.

Defining question sets for identifying severe impairment

The NPEU/Oxford expert group suggested that a criterion of –3 SD scores be used to represent 'severe cognitive (developmental) disability' at the age of 2 years.[221] However, this cut-off point was not feasible because of the floor effect of the Bayley-III cognitive composite scores, which ranged between 55 and 145. A post hoc analysis was therefore performed to evaluate if applying broader criteria at the severe end of the impairment spectrum would improve the validity of NHS data in identifying children with Bayley-III scores of lower than –2 SDs. For this, impairment categories were re-defined as 'none', 'mild' and 'moderate–severe'. Referring back to the '2-year' questions on the EPR, participants who received a positive response to the following questions were recategorised into the 'moderate–severe' category (*Table 23*).

Participants who received a positive response to any of the other questions were classified as having mild impairment. The sensitivity and specificity of the 'moderate–severe' category in predicting children with Bayley-III scores of lower than –2 SDs was then calculated and the concordance of NHS data classified into

TABLE 23 Moderate–severe categorisation

Category	Question
D2	Is this child's development between 6 and 12 months behind corrected age?
RC1	Does this child have difficulty with understanding outside of familiar context?
EC2	Does this child have difficulty with speech (< 10 words/signs)?
FM2	Does this child have difficulty with the use of both hands?
GM2	Is this child's gait non-fluent or abnormal, reducing mobility?
GM4	Is this child unstable or needs to be supported when sitting?

these new categories with the Bayley-III SD score groups, matching 'no impairment' to Bayley-III scores of higher than –1 SD, 'mild impairment' to Bayley-III scores of between –1 to –2 SDs and 'moderate–severe' impairment to Bayley-III scores of lower than –2 SDs.

Variables associated with the validity of NHS neurodevelopmental data

The effect of the following factors on the validity of the NHS data were examined: gestation at birth, sex, supplemental oxygen requirement at 36 weeks corrected gestational age, IMD quintile at the time of assessment, English as the only language spoken at home, corrected age at NHS assessment, use of a standardised neurodevelopmental test or screening test during NHS follow-up, grade of NHS assessor, time interval between NHS and research appointments, behaviour during the research assessments as measured by the examiner-rated behavioural score, and whether or not the NHS assessor thought that the child was difficult to test during the NHS assessment. Cross-tabulations and the calculation of the sensitivities and specificities of NHS assessment, stratified by the factor under study, were performed for each domain of neurodevelopment.

Assessment of social communication and autistic traits in early childhood

For the purpose of assessing the applicability of the Q-CHAT for children who were born preterm, children with cerebral palsy and children with severe neurosensory impairments (defined as a hearing deficit not correctable with hearing aids or a visual deficit not correctable with glasses) were excluded from this analysis. Differences in characteristics between respondents and non-respondents, and between respondents and the 'baseline population', were compared to evaluate selection bias.

The overall and sex-specific Q-CHAT scores from the study population were compared with published scores from the general population [general population overall mean 26.7 (SD 7.8), mean for boys 27.5 (SD 7.8), mean for girls 25.8 (SD 7.7)][220] using the Student's *t*-test. Differences in the distributions of item-specific scores between the study cohort and the general population in each category of autistic-like behaviour were examined by chi-squared tests. To overcome the chi-squared test restriction for low numbers, the proportions in adjacent score categories were combined to ensure that all expected values were larger than 5.[263]

The correlation between the Q-CHAT scores and the Bayley-III cognitive, language and motor composite scores was explored using linear regression to determine if any observed differences in Q-CHAT scores between the study population and the general population were explained by delayed neurodevelopment in the preterm population. Post hoc analysis of the correlation between subcategorical Q-CHAT scores (total score from items within each category of autistic-like behaviour) and Bayley-III cognitive, language and motor composite scores was carried out with Bonferroni correction for multiple testing.

The following neonatal and sociodemographic factors were analysed for possible association with Q-CHAT scores: gestation at birth, birthweight *z*-score, sex, single versus multiple pregnancy, white versus non-white ethnicity, maternal age at birth, mode of delivery, length of mechanical ventilation, supplemental oxygen requirement at 36 weeks' corrected gestational age and IMD quintile at the time of completion of the Q-CHAT. The current IMD quintile for participant was chosen rather than the birth quintile. Comparing the

IMD quintiles at birth with those at the time of assessment, 177 (83.9%) participants continued to live within the same IMD quintile, 13 (9.2%) moved to a more deprived IMD quintile and 15 (10.6%) moved to a less deprived quintile. Linear regression models were created to determine the association between predictive variables and Q-CHAT scores. To account for correlated outcomes within multiple birth sets, cluster bootstrap analysis was used to estimate standard errors and the resultant 95% CI. Variables identified to be significant at a 5% level in univariable models were included in forward stepwise multivariable regression analyses to determine the independent effect of each factor on Q-CHAT scores. Post hoc analysis was conducted to explore possible interactions between ethnicity, Bayley-III language scores and IMD.

Using an arbitrary cut-off score of 2 SDs above the general population mean for Q-CHAT scores and 2 SDs below the standardised mean for Bayley-III Social-Emotional scores, participants were classified as 'at risk for ASD'. A scatterplot was used to examine the relationship in score distribution between the Q-CHAT and Bayley-III Social-Emotional questionnaires and the agreement between the questionnaires in identifying children 'at risk' was measured using Cohen's κ statistic.

Systematic literature review and meta-analysis

A systematic electronic literature search was conducted on MEDLINE for information on the early developmental outcomes and corresponding school-age cognitive outcomes of preterm children. The methods adopted in this review were based on recommendations outlined in the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*.[264,265] Results are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.[265]

Any cohort or matched-control studies published since 1 January 1990 on study populations of infants born $\leq$ 32 weeks' gestation and/or who had a birthweight of < 1500 g (VLBW), in which at least two serial assessments, consisting of a neurodevelopmental assessment conducted between 1 and 3 years of age and a cognitive assessment at $\geq$ 5 years of age, were conducted and reported using validated standardised psychometric assessments [e.g. BSID, Griffiths Mental Development Scales (GMDS), Wechsler Preschool and Primary Scale of Intelligence] were considered for inclusion in the review. Assessments conducted before 1 year of age were not included because impairment, particularly if mild, may not be evident at this stage. Studies with populations that did not meet the gestation or birthweight criteria or reported outcomes using non-standardised assessments including measures of academic attainment were excluded. Studies that reported only outcomes in language or executive function (e.g. memory) were excluded as they would not reflect the overall cognitive function of the study populations. Case reports, narrative reviews, editorials, letters and comments on published articles were excluded.

The electronic search was conducted on MEDLINE through the PubMed interface on 13 April 2012, covering English-language literature published between 1 January 1990 and 31 March 2012. Search terms were selected a priori through a preliminary review of the literature. The following search terms were used both as keywords and as subject headings: (combinations of 'preterm' or 'premature' with 'infant' or 'neonate' or 'children') or ('low birthweight' or 'extremely low birthweight') and ('cogniti*' or 'neurodevelopment*' or 'mental retardation' or 'disability' or 'intelligence' or 'IQ'). The 'explode' feature was used with subject headings to include articles categorised under more specific subheadings. The detailed search strategy was as follows:

(('preterm children'[tiab] OR 'premature children'[tiab]) OR ('premature infant'[tiab] OR ('preterm infant'[tiab]) OR ('preterm neonate'[tiab] OR 'premature neonate'[tiab]) OR ('Infant, Premature'[MeSH]) OR ('Infant, Very Low Birthweight'[MeSH]) OR ('very low birthweight'[tiab] OR 'very low birthweight'[tiab]) OR ('extremely low birthweight'[tiab]) OR ('extremely low birthweight'[tiab])) AND ((cogniti*[tiab]) OR (neurodevelopment*[tiab]) OR (mental retardation) OR ('Developmental Disabilities'[Mesh] OR disability[tiab]) OR (intelligence[tiab] OR IQ[tiab])).

The electronic search was supplemented by a manual search of the reference lists of studies that met the inclusion criteria.

The titles and abstracts of studies retrieved from the literature search were screened to identify studies that reported developmental and/or cognitive outcomes among preterm children who were born before 32 weeks' gestation and/or were VLBW. These were grouped into three categories: (1) studies that reported both early developmental outcomes between ages 1 and 3 years as well as school-age cognitive outcomes at ≥ 5 years, (2) studies that reported only early developmental outcomes and (3) studies that only reported school-age cognitive outcomes. The author lists for articles in groups (2) and (3) were matched to identify assessments and publications on the same population at different time points. Studies that satisfied the initial screening process were retrieved for full-text evaluation for final inclusion in the review.

The quality of included studies was assessed using a checklist adapted from the Quality of Diagnostic Accuracy Studies version 2 (QUADAS-2) appraisal tool.[266] The aim was to provide a qualitative judgement for the risk of bias and the applicability of each study to the review question. The QUADAS-2 tool uses 'signalling questions' to assess bias in four domains: patient selection, index test, reference standard, and flow of participants through the study and timing of the index test. The applicability of the study to the review question in the first three domains was also assessed. In the context of this review, the index tests referred to the early developmental assessments and the reference standards were the school-age cognitive assessments. An essential feature of QUADAS-2 was the tailoring of the signalling questions to enable review-specific appraisal. *Table 24* lists the signalling questions and the quality standards set for this review. By appraising against the set standards, each study was given a rating of 'low', 'high' or 'unclear' for risk of bias and concerns regarding applicability in each domain. No summary 'quality score' was generated as such scores lack statistical justification and are not comparable across different scoring systems.[267] It was decided not to exclude any study on the basis of its quality, to achieve a review on the topic that was as comprehensive as possible.

From each included study, the following data were extracted into a table (unpublished data were sought from study authors through e-mail requests): study characteristics (i.e. location, city, country); sampling method (i.e. single centre, multicentre, population based); inclusion and exclusion criteria; anticipation

TABLE 24 Review-specific signalling questions and standards for appraisal of study quality

Domain	Patient selection	Index test (early developmental assessment)	Reference standard (school-age cognitive assessment)	Flow and timing
Signalling questions	1. Was a consecutive or random sample of patients enrolled? 2. Did the study avoid inappropriate exclusion?	1. Was an age-appropriate validated standardised assessment tool used?	1. Was an age-appropriate validated assessment standardised assessment tool used? 2. Were the assessors blinded to the results of the early developmental test?	1. Were all eligible infants participants receive the same assessments? 2. Were all participants included in the analysis?
High risk of bias	Non-consecutive or random sampling methods; additional inclusion criterion not based on birthweight or gestational age	Inappropriate test used for population under study	Inappropriate test used for population under study or assessors were not blinded to results of early developmental test	Participants received different assessments or dropout rates were > 30%
High concerns regarding applicability	Subcohort of infants (e.g. only IUGR infants were included) recruited. Infants born before 1990, as they would differ from the target population in terms of neonatal care received and severity/pattern of diseases experienced	Non-universal tests (e.g. only standardised in a specific population). Outdated versions of assessments (e.g. published before 1990)	Non-universal tests (e.g. only standardised in a specific population). Outdated versions of assessments (e.g. published before 1990)	

IUGR, intrauterine growth restriction.

and/or follow-up rates (as percentage of eligible survivors); final sample size included in meta-analysis (i.e. number of participants who completed both early and school-age assessments); early developmental and school-age cognitive assessment tool used; and study population characteristics [i.e. year(s) of birth of participants, mean or median gestational age, mean or median birthweight, ages at assessment, mean test scores at assessment, data on the predictive validity of early developmental assessments].

For this review, mild–moderate deficit was defined as developmental or cognitive test scores of between 1 and 2 SDs below the means of the standardised or control groups. Severe deficit was defined as test scores of lower than 2 SDs below the means of the standardised or control groups. In studies for which a control group of children who were born at full term were recruited and assessed simultaneously, the mean and SD of the control group were used as the references for defining the presence of deficits. Data on the number of 'true-positive', 'false-positive', 'false-negative' and 'true-negative' cognitive deficits identified by early assessments were collated from each study. If serial assessments were performed at different time points, data obtained from participants at the oldest age were included in the meta-analysis. The estimated sensitivity and specificity with corresponding 95% CI for mild–moderate and severe deficits were calculated.

Meta-analysis

The goals of the meta-analysis were to evaluate the variation in the estimates of the diagnostic accuracy (sensitivity and specificity) of early developmental assessments between studies and to combine results from all studies to yield a more precise estimate than is possible from individual studies. Coupled forest plots were generated to depict the ranges of sensitivity and specificity derived from the studies. Homogeneity of the sensitivities and specificities from the studies were tested using chi-squared tests. It has been noted that, in meta-analyses of diagnostic tests, significant between-study heterogeneity often exists. One source of heterogeneity is attributable to variations in diagnostic threshold and the related 'trade off' between sensitivity and specificity.[264,268] This may occur even when the same diagnostic criterion was applied across the studies (as was in this review) because of, for example, inherent differences in the spectrum of impairments in the patient populations or interobserver interpretation of test performances. To examine this, a scatterplot of the true-positive rate (TPR) (or sensitivity) against the false-positive rate (or 1 – specificity) for each study was created and the Spearman correlation coefficient was computed. 'Threshold effect' was demonstrated when the points assume the shape of a ROC curve and the sensitivity and specificity were significantly correlated. In this circumstance, separate pooling of sensitivities and specificities that ignore the correlation between the two measures would lead to an underestimation of the diagnostic accuracy.[269] It is possible to combine estimates using the Moses–Littenberg method to generate a summary ROC curve.[268,270] However, this does not allow for between-study variation. Instead, the Rutter and Gatsonis approach was used to fit a hierarchical summary ROC (HSROC) curve of the data.[271] The HSROC model accounts for both sampling variation within study at a lower level and between-study heterogeneity at a higher level using random effects. It models the log-odds of a positive test result in each study and each impairment group as a function of the positivity threshold in each study and the true impairment status, with model parameters describing the accuracy and asymmetry of the ROC curves. The output includes a summary operating point (pooled values for sensitivity and specificity) with 95% confidence region and a 95% prediction region for a forecast of the true sensitivity and specificity in a future study. As this is a hierarchical model, the summary operating point represents an average of study effects rather than a common effect. Individual study effects may differ considerably because of heterogeneity, and this variation is represented by the 95% prediction region.

The possible association of the diagnostic validity with study-level variables that could account for the observed heterogeneity among studies was investigated using metaregression methods for continuous variables and subgroup analysis for categorical variables. The variables were gestational age, birthweight, age at early assessment, age at late assessment, time interval between assessments, year of birth of participants, prevalences of total and severe impairment, the developmental assessment tool used, and the inclusion/exclusion of neurosensory impaired participants. For categorical variables, couple forest plots stratified by the subgroups were generated to allow for visual assessment of the differences in diagnostic validity between subgroups.

For continuous variables, scatterplots of sensitivity and specificity against each study-level covariates were generated by taking the mean value for continuous variables within each study except for year of birth, when the earliest date was used, as the mean/median value was not available. Bivariate models[272] were used to formally test whether or not sensitivity and specificity were associated with study-level covariates. Bivariate models are equivalent to HSROC models when no covariate is included.[273] When including covariates, the bivariate model measures the association with sensitivity and specificity (on the logit scale), whereas the HSROC model measures the association with the accuracy and threshold parameters; therefore, the former was chosen for ease of interpretation. For each study-level covariate, associations with sensitivity and specificity were tested separately; likelihood ratio test was then used to test both associations jointly. Results are reported as estimated ORs with associated 95% CI and *p*-values.

For the studies that reported data from multiple assessments at different time points, scatterplots were created of sensitivity and specificity against mean age at assessment to explore the stability of sensitivity and specificity estimates over time. For reviews of interventional trials, the funnel plot, a graphical display of the estimates of study effects plotted against their sample size or precision (standard error), is the recommended method for examining publication bias. Statistical tests, such as Egger's regression test and Begg's rank correlation, are used to test for funnel plot asymmetry, which would indicate the presence of publication bias and other sample size-related effects. The appropriate method for investigating publication bias for studies of diagnostic test accuracy is unclear. Funnel plots of the estimates of log-diagnostic odds ratio (DOR) against corresponding precision were proposed.[274] The DOR is a single statistic measure of diagnostic performance that is defined as:

$$\mathrm{DOR} = \frac{\mathrm{TP} \times \mathrm{TN}}{\mathrm{FP} \times \mathrm{FN}}. \tag{5}$$

Therefore, the larger the DOR, the more accurate the test is. In the *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0*,[275] the application of tests for funnel plot asymmetry designed for use in randomised trials, including the Egger and Begg tests, is specifically discouraged as these are associated with inflated type I error rates. Instead, a regressions test for the association between the log-DOR and the 'effective sample size (ESS)', developed by Deeks *et al.*,[276] was suggested. The ESS is a function of the number of non-diseased (n_1) and diseased (n_2) participants, in which:

$$\mathrm{ESS} = \frac{(4n_1n_2)}{n_1 + n_2}. \tag{6}$$

Following the proposed methods outlined in the paper by Deeks *et al.*,[276] the possibility of publication and other sample size-related effects was investigated by developing funnel plots of log-DOR against $1/\mathrm{ESS}^{1/2}$ and tested for plot asymmetry using linear regression of the two variables, weighted by ESS.

For this review, forest plots were generated using RevMan, version 5.2 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark). All other analyses were performed using Stata statistical package, version 11.0, and SAS.

Results

Two-year neurodevelopmental outcomes

Two hundred and eight children were recruited from 13 hospitals (Addenbrooke's, Cambridge; Queen's, Romford; Chelsea and Westminster, London; Ealing Hospital, London; Hillingdon Hospital, London; Homerton University, London; Newham, London; North Middlesex University, London; Northwick Park, London; Royal London, London; St Thomas', London; West Middlesex, London; Whipps Cross Hospital, London).

Figure 17 shows the flow of children through recruitment to the completion of research and NHS assessments. Two hundred and four children completed all the subtests of the Bayley-III assessment. One child with ataxic cerebral palsy could not be assessed for the cognitive and language scales, two children did not co-operate for the receptive communication assessment and one child did not co-operate for the gross motor assessment. Of the children who completed the research assessments, three did not attend their routine NHS follow-up visit. Data from the NHS assessment were not entered onto the EPR by the examining clinician in 15 cases and the overall category of impairment could not be assigned for nine children because of missing EPR data. The 190 children for whom both research and NHS data were available in at least one outcome domain formed the study cohort. A complete set of data in all outcome domains was available for 177 children. Although the original plan was to stratify recruitment based on the gestational ages of the eligible children, it became clear during the recruitment phase that the final study cohort would be smaller than the initial projection. Therefore, all children whose parents agreed to participate were recruited and assessed.

The characteristics of the study population were compared with the 'baseline population' (all infants born between 1 January 2008 and 31 December 2010, at gestational ages below 30 weeks and discharged from the participating hospitals) (*Table 25*). Participants received a shorter duration of mechanical ventilation and were less likely to be receiving oxygen therapy at 36 weeks postmenstrual age than the baseline population ($p < 0.001$). The study population was comparable to the baseline population in terms of gestational age, birthweight, sex, proportions of singletons, mode of delivery and maternal age, and consisted of larger proportions of children of white ethnicity and those born to mothers living in the least deprived IMD quintile.

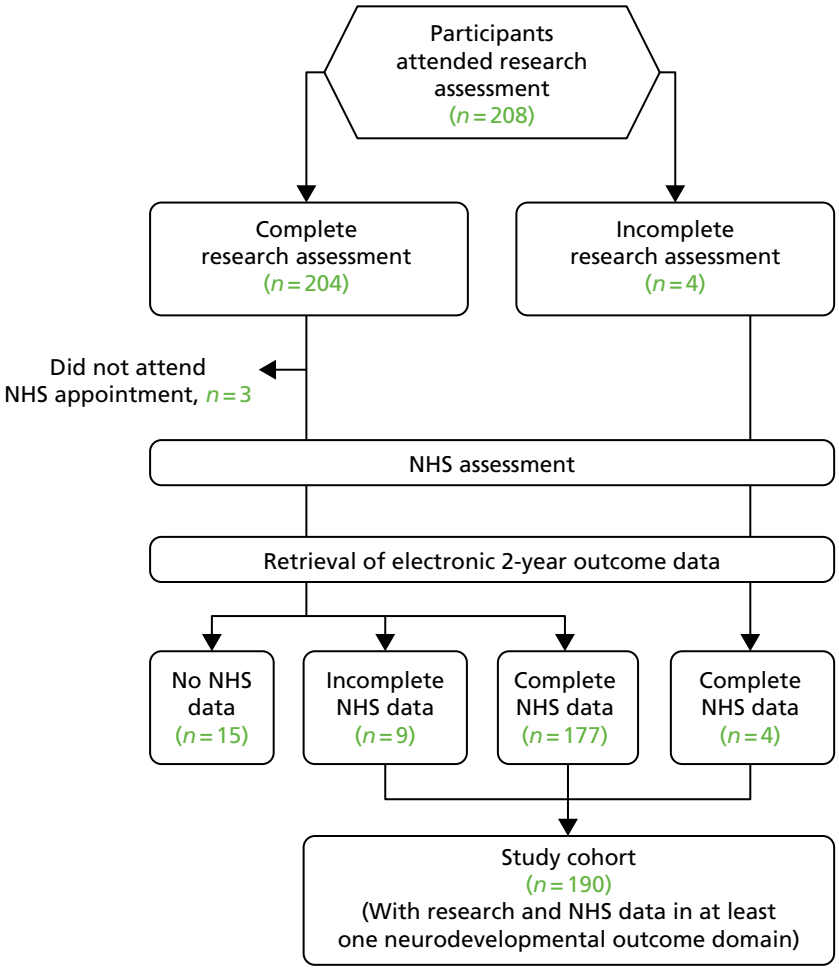


FIGURE 17 Flow chart of children through research and NHS assessments to form the study population.

TABLE 25 Demographic and neonatal characteristics of participants and non-participants (born before 30 weeks' gestation in 2008–10 and discharged from participating sites)

Characteristics	Population		*p*-value
	Study (*N* = 190)	Baseline (*N* = 1037)	
Gestation (completed weeks)			
Median (IQR), range	27 (26–29), 23–29	27 (26–29), 22–29	0.25
Birthweight (g)			
Median (IQR), range	965 (790–1140), 490–1720	1000 (812–1200), 455–1990	0.08
Sex, *n* (%)			
Girls	99 (52.1)	444 (42.8)	0.19
Boys	91 (47.9)	503 (48.5)	
Missing	0 (0.0)	90 (8.7)	
Ethnicity, *n* (%)			
White	88 (46.3)	364 (35.1)	0.03
Black	50 (26.3)	287 (27.7)	
Asian	41 (21.6)	239 (23.1)	
Mixed	0 (0.0)	33 (3.2)	
Other	11 (5.8)	52 (5.0)	
Missing	0 (0.0)	62 (6.0)	
Pregnancy, *n* (%)			
Singleton	147 (77.4)	690 (66.5)	0.26
Multiples	43 (22.6)	250 (24.1)	
Missing	0 (0.0)	97 (9.4)	
Mode of delivery, *n* (%)			
Vaginal	74 (39.0)	475 (45.8)	0.22
Caesarean	103 (54.2)	540 (52.1)	
Missing	13 (6.8)	22 (2.1)	
Maternal age (years)			
Mean (SD)	31.9 (6.7)	31.0 (6.4)	0.08
IMD quintile at birth, *n* (%)			
One (least deprived)	19 (10.0)	43 (4.2)	0.01
Two	20 (10.5)	81 (7.8)	
Three	26 (13.7)	144 (13.9)	
Four	52 (27.4)	268 (25.8)	
Five (most deprived)	73 (38.4)	477 (46.0)	
Missing	0 (0.0)	24 (2.3)	
Length of mechanical ventilation (days)			
Median (IQR), range	0 (0–3), 0–54	4 (0–18), 0–444	< 0.001
Oxygen therapy at 36 weeks corrected age, *n* (%)			
Yes	54 (28.4)	466 (44.9)	< 0.001
No	136 (71.6)	574 (55.1)	

Nevertheless, a wide range of ethnic groups was represented, reflecting the diversity of the population living in London. Consequentially, 92 (48.4%) children were raised in a bilingual or multilingual environment.

The mean (SD) corrected age of the children at assessment was 24.8 (2.2) months. The research assessment took place at a median (IQR) interval of 8 (0–27) days after the children received their NHS assessment, with a range of between 89 days before and 82 days after the NHS assessment.

Based on information given by the parents, 30 (15.8%) children had a visual defect including reduced visual acuity and/or squints, although only 11 (5.8%) required glasses. A total of 16 (8.4%) children had a hearing impairment, of whom three (1.6%) wore hearing aids.

The children performed significantly worse than the normative population, in which Bayley-III scores were standardised in all domains other than fine motor skills (*Table 26*).

Based on the worst score achieved in the cognitive, language and motor Bayley-III domains, 114 (61.3%) children were classified as having scores of higher than –1 SD from the standardised mean, 42 (22.6%) children had scores of between –1 and –2 SDs and 30 (16.1%) children had scores of lower than –2 SDs from the standardised mean.

In the cognitive domain, 156 (82.1%) children obtained Bayley-III scores of higher than –1 SD, 26 (13.7%) children had scores of between –1 and –2 SDs and seven (3.7%) children had scores of lower than –2 SDs from the standardised mean.

Nineteen (10.0%) children had specific expressive communication impairment (scaled score of < 7), with no impairment in receptive communication. Five of these children would have been classified as having no impairment based on the Bayley-III language composite score alone (language composite score of higher than –1 SD, i.e. ≥ 85) because of the compensation from the receptive communication subtest. Based on the worst SD score category from the receptive and expressive communication subtests and the language composite score, 120 (63.2%) children achieved scores of higher than –1 SD, 42 (22.1%) had scores of between –1 and –2 SDs and 25 (13.2%) had scores of lower than –2 SDs from the standardised mean.

Motor function was generally intact among the children, with only 11 children (5.8%) receiving scores of between –1 and –2 SDs and 11 (5.8%) children having scores of lower than –2 SDs from the standardised mean. The mean (SD) predicted BSID-II MDI for the study population was 77.9 (20.5) and was significantly lower than the mean Bayley-III cognitive and language composite scores ($p < 0.001$ for both). *Figure 18* shows the proportions of children with scores of higher than –1 SD (≥ 85), scores of between –1 and –2 SDs (70–84), scores of between –2 and –3 SDs (55–69) and scores of lower than –3 SDs (< 55) from

TABLE 26 Mean Bayley-III scores (scaled and composite scores) of study population

Domain	Score	Mean (SD) Bayley-III scores	*p*-value[a]
Cognitive	Cognitive composite ($n = 189$)	92.65 (12.8)	< 0.001
Language	Receptive communication scaled ($n = 187$)	8.0 (2.4)	< 0.001
	Expressive communication scaled ($n = 189$)	7.5 (2.5)	< 0.001
	Language composite ($n = 187$)	87.0 (13.6)	< 0.001
Motor	Fine motor scaled ($n = 190$)	10.2 (2.5)	0.23
	Gross motor scaled ($n = 189$)	8.6 (2.3)	< 0.001
	Motor composite ($n = 189$)	96.7 (12.7)	< 0.001

a *p*-value from Student's *t*-test comparing the study population mean scores to the Bayley-III standardised scaled score mean of 10 and standardised composite score mean of 100.

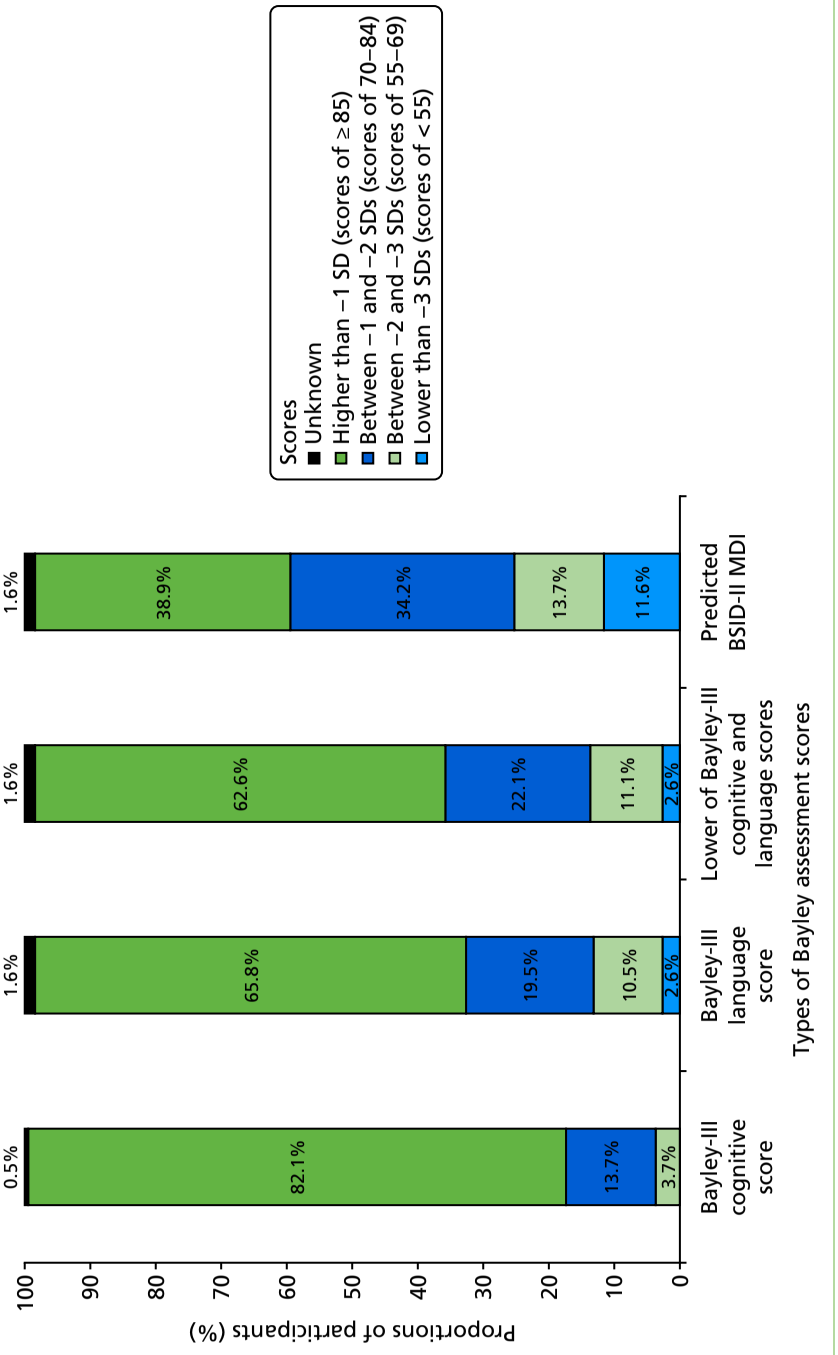


FIGURE 18 Neurodevelopmental status by Bayley-III scores and predicted BSID-II MDI.

the standardised mean for the Bayley-III cognitive and language composite scores individually, when the lower of the two Bayley-III scores was used and for the predicted BSID-II MDI score. A post hoc analysis using McNemar's test was performed to compare the proportions diagnosed with impairment using these cut-off scores. The proportions of children classified with impairment using a cut-off point of < 85 on the Bayley-III cognitive score (17.4%), language score (32.6%) and the lower of the cognitive and language scores (35.8%) were statistically dissimilar to the proportions with predicted BSID-II MDI < 70 (25.3%) ($p < 0.001$). However, the proportions of children with predicted BSID-II MDI of < 55 (11.6%) were similar to the proportions who scored < 70 (classified with severe impairment) on the Bayley-III language composite score (13.1%; $p = 0.26$) and the lower of the cognitive and language scores (13.7%; $p = 0.16$).

Classification using the modified NPEU/Oxford criteria showed that, for the communication domain, the assignment of outcome was heavily influenced by the presence of specific expressive communication impairment: 107 (56.3%) children were classified as having no language impairment, 55 (28.9%) children had mild–moderate language impairment and 27 (14.2%) children had severe language impairment. Thirteen (6.8%) children had isolated gross motor impairment with no fine motor difficulties, and only one (0.5%) child had specific fine motor impairment. The combined motor outcome was normal in 172 (90.5%) children. Nine (4.7%) children were classed as having mild–moderate motor impairment and nine (4.7%) children had severe motor impairment.

Evaluating the concordance in the classification of neurodevelopmental status by Bayley-III scores and NPEU/Oxford criteria, of the 187 children tested for their communication skills, 144 (77.0%) were classified in the same category. Of the other children, none differed by more than one category. The weighted kappa coefficient (κ) was 0.59 (95% CI 0.49 to 0.69), which indicated moderate agreement between the two criteria for the communication outcome. For the motor domain, 180 out of 189 (95.2%) children were classified in the same category. Classification differed by one category for eight children. One child who was assessed as having severe motor impairment by the Bayley-III was classified as having 'no impairment' on the NPEU/Oxford criteria. The weighted κ for concordance between the two methods in the motor domain was 0.76 (95% CI 0.62 to 0.93), which represented substantial agreement.

The per cent agreement across all assessed items was 97.2% (69/71 items in agreement) in the first session (midway) and 98.6% (69/70 items in agreement) in the second session end of study assessments).

Neurodevelopmental outcomes from NHS electronic patient record data

Children attended their NHS follow-up assessment at a mean (SD) corrected age of 24.4 (2.3) months. Data were entered on the EPR '2-year outcome' screen by clinical consultants in 111 (58.4%) cases [36 (19.0%) by consultant neonatologists, 42 (22.1%) by hospital paediatrics consultants and 33 (17.4%) by community paediatrics consultants], trainee doctors in 15 (7.9%) cases, staff-grade doctors in 58 (30.5%) cases and administrative staff in six (3.2%) cases. Sixty-seven (35.3%) children received standardised neurodevelopmental assessment or screening tests during their NHS appointment [19 (10.0%) using the Schedule of Growing Scales, 44 (23.2%) using the GMDS and 4 (2.1%) using the Alberta Infant Motor Scale]. *Table 27* shows the responses to each question on the electronic form and the classification of impairment for the developmental domains. The classification of overall neurodevelopmental outcome was possible in 181 children, of whom 124 (68.5%) had no impairments, 38 (21.0%) had mild–moderate impairments and 19 (10.5%) had severe impairments.

The proportions of children classified into each category of impairment by the research assessment (using Bayley-III scores and NPEU/Oxford criteria) and by the NHS assessment are displayed in *Appendix 2*, *Figures 36–39*.

TABLE 27 Responses to questions on the electronic '2-year outcome' form and classification of impairment based on NHS data

Domain	Question	Response, *n* (%)			Classification of impairment, *n* (%)
		No	Yes	Missing	
Cognitive	D1	154 (81.1)	35 (18.4)	1 (0.5)	None: 141 (74.2)
	D2	171 (90.0)	19 (10.0)	0 (0.0)	Mild–moderate: 42 (22.1)
	D3	183 (96.3)	6 (3.2)	1 (0.5)	Severe: 6 (3.2)
					Unknown: 1 (0.5)
Receptive communication	RC1	174 (91.6)	13 (6.8)	3 (1.6)	None: 174 (91.6)
	RC2	183 (96.3)	5 (2.6)	2 (1.1)	Mild–moderate: 8 (4.2)
					Severe: 5 (2.6)
					Unknown: 3 (1.6)
Expressive communication	EC1	149 (78.4)	40 (21.1)	1 (0.5)	None: 143 (75.3)
	EC2	153 (80.5)	36 (18.9)	1 (0.5)	Mild–moderate: 32 (16.8)
	EC3	176 (92.6)	13 (6.8)	1 (0.5)	Severe: 13 (6.8)
					Unknown: 2 (1.1)
Combined language[a]					None: 141 (74.2)
					Mild–moderate: 30 (15.8)
					Severe: 14 (7.4)
					Unknown: 5 (2.6)
Fine motor	FM1	188 (98.9)	2 (1.1)	0 (0.0)	None: 186 (97.9)
	FM2	189 (99.5)	1 (0.5)	0 (0.0)	Mild–moderate: 2 (1.1)
	FM3	188 (98.9)	2 (1.1)	0 (0.0)	Severe: 2 (1.1)
					Unknown: 0 (0.0)
Gross motor	GM1	178 (93.7)	12 (6.3)	0 (0.0)	None: 174 (91.6)
	GM2	177 (93.2)	9 (4.7)	4 (2.1)	Mild–moderate: 5 (2.6)
	GM3	182 (95.8)	8 (4.2)	0 (0.0)	Severe: 8 (4.2)
	GM4	186 (97.9)	4 (2.1)	0 (0.0)	Unknown: 3 (1.6)
	GM5	188 (98.9)	2 (1.1)	0 (0.0)	
Combined motor[a]					None: 173 (91.1)
					Mild–moderate: 6 (3.2)
					Severe: 8 (4.2)
					Unknown: 3 (1.6)

continued

TABLE 27 Responses to questions on the electronic '2-year outcome' form and classification of impairment based on NHS data (*continued*)

Domain	Question	Response, *n* (%)			Classification of impairment, *n* (%)
		No	Yes	Missing	
Overall[a]					None: 124 (65.3) Mild–moderate: 38 (20.0) Severe: 19 (10.0) Unknown: 9 (4.7)

a Combined language impairment was judged as the worst category of outcome from receptive and expressive communication, and combined motor impairment was judged as the worst category of outcome from fine and gross motor. Overall impairment was based on the worst category of outcome from cognitive, combined language and combined motor domains.

Cross-tabulations to compare the agreement between the research and the NHS assessments for the classification of 'any impairment' and 'severe impairment' are displayed in *Tables 28* and *29*. The estimated sensitivities and specificities for NHS assessments in each developmental domain are presented taking the research assessment as the 'gold standard'. The CIs for sensitivities and specificities were calculated using robust standard errors to account for clustering of data by study sites. Sensitivity analyses revealed that potential correlated outcomes from siblings did not affect the results. Therefore, the results presented included data from all participating children.

The validity of the NHS assessments in identifying children with no impairments was high, with estimated specificities ranging between 83.9% and 100.0% for 'any impairment', and between 96.6% and 100.0% for 'severe impairment'. However, the validity of the NHS and the research assessment in identifying and categorising children with impairments was variable. The sensitivities for identifying gross motor impairment were high, particularly when the impairment was severe. In the cognitive domain, the sensitivity for the identification of any impairment was 69.7% (95% CI 55.1% to 84.3%) but dropped to only 28.6% (95% CI 5.0% to 52.2%) for the identification of severe impairment. Of the seven children diagnosed with severe cognitive impairment through the research assessment, two were also classified as having impairment in the 'severe' category in the NHS data set, four were classified as having impairment 'mild–moderate' and one was classified as having 'no' impairment; hence, the disagreement in the classification occurred mainly between the 'mild–moderate' and 'severe' categories. Agreement between the NHS and research assessments was worst in the language domain, especially in receptive communication, where the sensitivity in identifying the presence of any impairment was only 23.1% (95% CI 6.7% to 39.5%). In the combined language domain, the 21 'false negatives' for severe impairment, based on Bayley-III classification, were evenly distributed among the impairment categories in the NHS data ('severe', $n = 9$; 'mild–moderate', $n = 6$; 'no' impairments, $n = 6$). The sensitivities were estimated with low precision (wide CIs), particularly in the motor domains and with severe impairments when the prevalence of impairment was low.

Although the sensitivities in the receptive communication and fine motor domains appeared considerably higher when impairment was assigned using the NPEU/Oxford criteria than using Bayley-III scores for the research assessment, this was driven by the small numbers in the 'false-negative' cells and the estimated sensitivities were associated with wide and overlapping CIs, which suggested that the differences in sensitivities may not be statistically significant.

The sensitivities and specificities of NHS assessment in identifying cognitive deficit were 69.7% (95% CI 55.1% to 84.3%) and 83.9% (95% CI 75.6% to 92.1%) for the presence of any impairment and 28.6% (95% CI 5.0% to 52.2%) and 97.8% (95% CI 95.1% to 100.0%) for severe impairments (see *Tables 28* and *29*). The analyses were repeated using the predicted BSID-II MDI scores as the 'gold standard'. Using the

TABLE 28 Results of cross-tabulations comparing the NHS and research categorisation of impairment and the sensitivities and specificities of the NHS assessment in identifying children with any impairment against the 'gold-standard' research assessment

Domain of development[a]	Method of classification of impairment for research assessment	Identification of impairment by NHS assessment against the 'gold standard' research assessment					
		True positives, *n* (%)	False negatives, *n* (%)	False positives, *n* (%)	True negatives, *n* (%)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Cognitive	Bayley-III scores (*n* = 188)	23 (12.2)	10 (5.3)	25 (13.3)	130 (69.1)	69.7 (55.1 to 84.3)	83.9 (75.6 to 92.1)
Receptive communication	Bayley-III scores (*n* = 184)	9 (4.9)	30 (16.3)	4 (2.2)	141 (76.6)	23.1 (6.7 to 39.5)	97.2 (94.6 to 99.9)
	NPEU/Oxford (*n* = 186)	8 (4.3)	13 (7.0)	5 (2.7)	160 (86.0)	38.1 (10.7 to 65.5)	97.0 (94.8 to 99.2)
Expressive communication	Bayley-III scores (*n* = 187)	32 (17.1)	22 (11.8)	13 (7.0)	120 (64.2)	59.3 (46.5 to 72.0)	90.2 (82.2 to 98.3)
	NPEU/Oxford (*n* = 187)	39 (20.9)	41 (21.9)	6 (3.2)	101 (54.0)	48.8 (33.9 to 63.6)	94.4 (88.9 to 99.9)
Combined language	Bayley-III scores (*n* = 182)	33 (18.1)	29 (15.9)	11 (6.0)	109 (59.9)	53.2 (42.0 to 64.5)	90.8 (83.5 to 98.2)
	NPEU/Oxford (*n* = 184)	38 (20.7)	39 (21.2)	6 (3.3)	101 (54.9)	49.4 (34.7 to 64.0)	94.4 (88.9 to 99.9)
Fine motor	Bayley-III scores (*n* = 190)	3 (1.6)	9 (4.7)	1 (0.5)	177 (93.2)	25.0 (0.0 to 59.7)	99.4 (98.3 to 100.0)
	NPEU/Oxford (*n* = 190)	4 (2.1)	1 (0.5)	0 (0.0)	185 (97.4)	80.0 (28.4 to 99.5)	100.0 (98.0 to 100.0)
Gross motor	Bayley-III scores (*n* = 186)	12 (6.5)	4 (2.2)	1 (0.5)	169 (90.9)	75.0 (49.9 to 100.0)	99.4 (98.1 to 100.0)
	NPEU/Oxford (*n* = 187)	11 (5.9)	5 (2.7)	2 (1.1)	169 (90.4)	68.8 (45.5 to 92.0)	98.8 (97.1 to 100.0)
Combined motor	Bayley-III scores (*n* = 186)	13 (7.0)	8 (4.3)	1 (0.5)	164 (88.2)	61.9 (32.9 to 90.9)	99.4 (98.1 to 100.0)
	NPEU/Oxford (*n* = 187)	12 (6.4)	5 (2.7)	2 (1.1)	168 (89.8)	70.6 (48.8 to 92.4)	98.8 (97.0 to 100.0)
Overall	Bayley-III scores (*n* = 177)	40 (22.6)	25 (14.1)	16 (9.0)	96 (54.2)	61.5 (52.5 to 70.6)	85.7 (77.4 to 94.0)

a Combined language impairment was judged as the worst category of outcome from receptive communication and expressive communication, and combined motor impairment was judged as the worst category of outcome from fine motor and gross motor. Overall impairment was based on the worst category of outcome from the cognitive, language and motor domains.

TABLE 29 Results of cross-tabulations comparing the NHS and research categorisation of impairment and the sensitivities and specificities of the NHS assessment in identifying children with severe impairment against the 'gold-standard' research assessment

Domain of development[a]	Method of classification of impairment for research assessment	Identification of severe impairment by NHS assessment against the 'gold standard' research assessment					
		True positives, *n* (%)	False negatives, *n* (%)	False positives, *n* (%)	True negatives, *n* (%)	Sensitivity, % (95% CI)	Specificity, % (95% CI)
Cognitive	Bayley-III scores (*n* = 188)	2 (1.1)	5 (2.7)	4 (2.1)	177 (94.1)	28.6 (5.0 to 52.2)	97.8 (95.1 to 100.0)
Receptive communication	Bayley-III scores (*n* = 184)	3 (1.6)	8 (4.3)	2 (1.1)	171 (92.9)	27.3 (0.0 to 62.9)	98.8 (97.3 to 100.0)
	NPEU/Oxford (*n* = 186)	2 (1.1)	1 (0.5)	3 (1.6)	180 (96.8)	66.7 (4.9 to 100.0)	98.4 (95.3 to 99.7)
Expressive communication	Bayley-III scores (*n* = 187)	7 (3.7)	5 (2.7)	6 (3.2)	169 (90.4)	58.3 (36.6 to 80.0)	96.6 (92.7 to 98.7)
	NPEU/Oxford (*n* = 187)	9 (4.8)	18 (9.6)	4 (2.1)	156 (83.4)	33.3 (12.0 to 54.7)	97.5 (93.7 to 99.3)
Combined language	Bayley-III scores (*n* = 182)	9 (4.9)	12 (6.6)	5 (2.7)	156 (85.7)	42.9 (14.2 to 71.5)	96.9 (92.9 to 99.0)
	NPEU/Oxford (*n* = 184)	9 (4.9)	15 (8.2)	5 (2.7)	155 (84.2)	37.5 (15.4 to 59.6)	96.9 (92.9 to 99.0)
Fine motor	Bayley-III scores (*n* = 190)	1 (0.5)	1 (0.5)	1 (0.5)	187 (98.4)	50.0 (0.0 to 100.0)	99.5 (97.1 to 100.0)
	NPEU/Oxford (*n* = 190)	1 (0.5)	0 (0.0)	1 (0.5)	188 (98.9)	100.0 (2.5 to 100.0)	99.5 (97.1 to 100.0)
Gross motor	Bayley-III scores (*n* = 186)	8 (4.3)	1 (0.5)	0 (0.0)	177 (95.2)	88.9 (51.8 to 99.7)	100.0 (97.9 to 100.0)
	NPEU/Oxford (*n* = 187)	8 (4.3)	1 (0.5)	0 (0.0)	178 (95.2)	88.9 (51.8 to 99.7)	100.0 (97.9 to 100.0)
Combined motor	Bayley-III scores (*n* = 186)	8 (4.3)	2 (1.1)	0 (0.0)	176 (94.6)	80.0 (44.4 to 97.5)	100.0 (97.9 to 100.0)
	NPEU/Oxford (*n* = 187)	8 (4.3)	1 (0.5)	0 (0.0)	178 (95.2)	88.9 (51.8 to 99.7)	100.0 (97.9 to 100.0)
Overall	Bayley-III scores (*n* = 177)	13 (7.3)	12 (6.8)	5 (2.8)	147 (83.1)	52.0 (23.8 to 80.2)	96.7 (92.5 to 99.9)

a Combined language impairment was judged as the worst category of outcome from receptive communication and expressive communication, and combined motor impairment was judged as the worst category of outcome from fine motor and gross motor. Overall impairment was based on the worst category of outcome from the cognitive, language and motor domains.

cut-off point of MDI < 85 (–1 SD) to define mild–moderate impairment and < 70 (–2 SDs) to define severe impairment, there was a reduction in the sensitivities and an increment in the corresponding specificities [sensitivity 39.3% (95% CI 30.2% to 49.0%) and specificity 94.6% (95% CI 86.7% to 98.5%) for any cognitive impairment; sensitivity 12.5% (95% CI 4.7% to 25.2%) and specificity 100% (95% CI 97.4% to 100%) for severe cognitive impairment]. However, if thresholds of MDI < 70 (–2 SDs) for mild–moderate and < 55 (–3 SDs) for severe impairments were used, the results were similar to the reported findings using the Bayley-III [sensitivity 64.6% (95% CI 49.5% to 77.8%), specificity 87.7% (95% CI 81.0% to 92.7%) for any impairment; sensitivity 18.2% (95% CI 5.1% to 40.3%), specificity 98.8% (95% CI 95.7% to 99.9%) for severe impairment].

The concordance of the research and the NHS assessments as measured by κ was consistent with the findings from the estimated sensitivities and specificities. The agreement between NHS and research assessment was substantial in the motor domain with weighted $\kappa > 0.6$. In the cognitive and communication domains, agreement was moderate at best.

Post hoc analysis of the validity of NHS assessments using a different question set to identify 'moderate–severe' impairment

The purpose of this post hoc analysis was to assess if, by applying a broader criterion to define 'moderate–severe' impairment, the validity of the NHS data in identifying children with Bayley-III scores of lower than –2 SDs could be improved. Children were reclassified as having moderate–severe impairment if they met the broader criterion in the NHS data. The results are displayed in *Appendix 1, Table 58*. The use of a broader category of moderate–severe impairment improved the sensitivity of the NHS data, although this was at a cost of a small reduction in specificity. The biggest increase in sensitivity was observed in the cognitive and expressive communication domains.

Variables affecting the validity of the NHS assessments

As the diagnostic validity of the NHS assessment did not differ between the use of Bayley-III scores and NPEU/Oxford criteria for classifying impairment, subgroup analyses were performed using only the results from Bayley-III assessments. Lower prevalence of impairment with higher gestational age at birth across all domains was observed, with apparent reduction in sensitivity but increased specificity of NHS assessments in identifying overall impairment with increasing gestational age. Sensitivity in identifying cognitive impairment was higher if a standardised neurodevelopmental test was used during NHS assessment. Accuracy in identifying impairment also appeared higher across all domains with increasing postnatal age at assessment. However, as the CIs for the estimated sensitivities and specificities overlapped widely, the observed effect of these factors on the diagnostic validity of NHS assessment can be conservatively considered statistically insignificant.[277] Similarly, there was no clear effect of the exposure to English language, the grade of the NHS assessor, IMD and the time interval between NHS and research assessments on the validity of NHS assessment.

Behaviour during assessments and the effect on study findings

The prevalence of impairment was significantly higher among children who were difficult to assess during the NHS assessment (86.7% vs. 31.7% for impairment in any domain; $p < 0.001$) or who had received lower examiner-rated behaviour scores (less positive behaviour) (73.8% for impairment in any domain among children with behaviour score of ≤ 22 vs. 28.5% among children with behaviour score of > 22; $p < 0.001$). However, challenging behaviour demonstrated during assessments did not appear to affect the test validity of the NHS assessment against the research assessment. The prevalence of impairment, sensitivity and specificity of NHS assessment did not differ by parent-rated behaviour scores.

Hammersmith Infant Neurological Examination and diagnosis of cerebral palsy

Forty-seven (24.7%) children had a suboptimal global score (< 73/78) on the HINE. In general, in the preterm population, although scores below 73 are suboptimal, those with scores of > 64 will walk independently by 2 years, those with scores of < 64 but > 52 will sit independently by 2 years and those with scores of < 52 will not be able to do either. The proportions of participants who achieved suboptimal score in each subsection were as follows: 30 (15.8%) for cranial nerve function, 39 (20.5%) for posture, 16 (8.4%) for movement, 36 (18.9%) for tone and 8 (4.2%) for reflexes.

Nine (4.7%) children were found to have cerebral palsy during the research assessment. The HINE scores for these children (median 53, IQR 38.5–59.5) were significantly lower than those without cerebral palsy (median 78, IQR 74–78; $p < 0.001$) and consistent with published data.[210] Two children had spastic quadriplegia, five had spastic diplegia, one had three-limb involvement and one had dyskinetic cerebral palsy. The gross motor function varied from GMFCS level 1 (walks without limitations) for one child with spastic diplegia to GMFCS level 5 (transported in manual wheelchair) for the child with dyskinetic cerebral palsy and one of the children with spastic quadriplegia. Most children with spastic diplegic cerebral palsy functioned at GMFCS level 2 (walks with limitations).

Two children with spastic diplegia were not identified to have cerebral palsy in the NHS data. The topographic classifications entered in the NHS data for all other children identified to have cerebral palsy were in agreement with the research assessment.

Early childhood social communication difficulties

The Quantitative Checklist for Autism in Toddlers (Q-CHAT) questionnaire was sent to the parents of all 208 children who attended the research assessment. Ten children were assessed to have major functional impairments (nine with cerebral palsy and one with severe hearing impairment) and were ineligible for this study. The parents of three children who declined to participate in the research assessment agreed to complete the Q-CHAT and Bayley-III Social-Emotional questionnaires. A total of 150 questionnaires, including eight from children who were ineligible, were returned. One questionnaire with seven missing responses was treated as a non-respondent and excluded, leaving 141 participants (70.1% of eligible participants) for the analyses.

Non-respondents were more likely to be parents of girls (66.7%; $p = 0.02$). Nonetheless, both boys and girls were equally represented in the respondent group. Respondents showed over-representation of children of white ethnicity who were born to mothers living in less deprived IMD quintiles, with significantly shorter duration of mechanical ventilation and who were less likely to have required supplemental oxygen therapy at 36 weeks corrected age (see *Appendix 1, Table 59*).

The mean corrected age of the respondents was 24.7 (SD 2.6, range 18.5–35.6) months at the time of completion of the questionnaires. The mean Bayley-III composite score of the 138 respondents who completed the assessment was 94.6 (SD 13.0) for the cognitive scale, 87.7 (SD 13.0) for the language scale and 98.0 (SD 10.1) for the motor scale.

The Q-CHAT scores of the study population (mean 33.7, SD 8.3, range 15–55) were normally distributed and significantly higher (less favourable) than the published general population scores (difference in means 7.0, 95% CI 5.6 to 8.3; $p < 0.001$) (*Figure 19*). In contrast with the higher scores described in boys in the general population, no sex differences in Q-CHAT scores were observed in the preterm population ($p = 0.85$).

The distribution of scores between the preterm study cohort and the general population differed significantly in 17 items. In all of these items, there were greater proportions of preterm children receiving higher scores, indicating greater social communication difficulties and autistic behaviour characteristics. The differences were most prominent in the categories of restricted, repetitive, stereotyped behaviour (seven out of nine items differ significantly), communication (three out of four items) and sensory abnormalities (all three items). Only four out of the nine items exploring social relatedness were scored differently in the preterm population.

On multivariable testing, cognitive and motor function did not appear to affect Q-CHAT scores ($p = 0.18$ for cognitive scores and $p = 0.67$ for motor scores). Bayley-III language composite scores independently predicted Q-CHAT scores in a linear fashion (correlation coefficient –0.51; $p = 0.001$) and accounted for 24.5% of the variance in Q-CHAT scores. The relationship between language and Q-CHAT scores was attributable to expressive communication ability (regression coefficient Bayley-III expressive communication subscale scores and Q-CHAT scores: –1.35, 95% CI –1.96 to –0.74, correlation coefficient –0.43; $p < 0.001$). There was no association between receptive communication ability and Q-CHAT scores ($p = 0.22$).

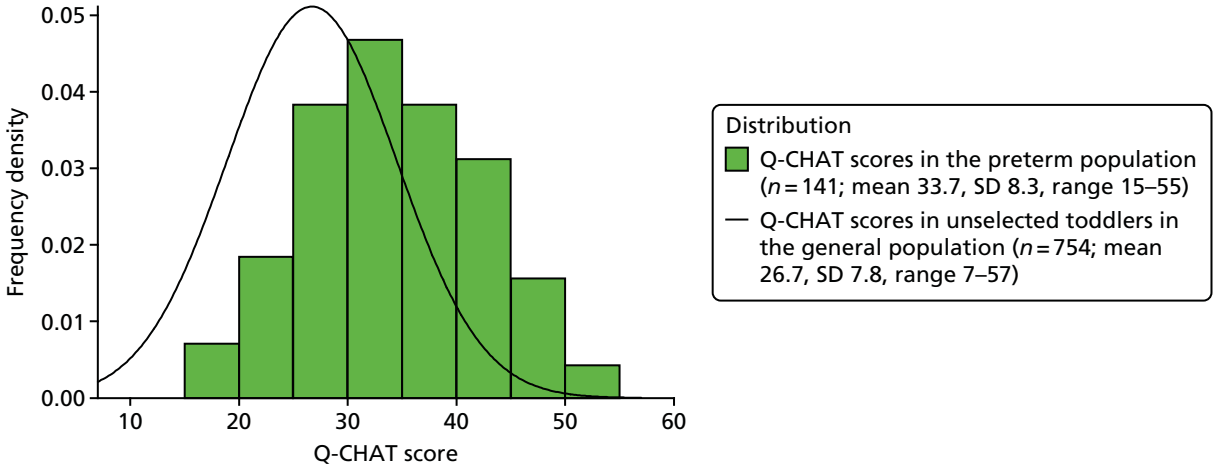


FIGURE 19 Histogram of Q-CHAT scores of the preterm study population with superimposed distribution of published Q-CHAT scores of unselected toddlers (general population).

Non-white ethnicity and living in deprived areas were associated with higher Q-CHAT scores in univariable analyses. Although non-white children were more likely to live in areas of higher deprivation (test for trend $p < 0.001$), there was no interaction between ethnicity and IMD in the association with Q-CHAT scores ($p = 0.72$). As lower Bayley-III language scores were observed among non-white children (mean difference 7.31, 95% CI 3.07 to 11.5; $p < 0.001$) and children living in more deprived areas (mean decrease of 1.89, 95% CI 0.24 to 3.55; $p = 0.03$) points per IMD quintile increase in deprivation, language ability was considered to be a potential confounder in the relationship between ethnicity, IMD and Q-CHAT scores. There was no interaction between Bayley-III language scores and IMD quintiles ($p = 0.88$) or ethnicity ($p = 0.51$). The final multivariable regression model included all variables found to be statistically significant during univariable analysis (Bayley-III language composite score, ethnicity and IMD) and is displayed in *Appendix 1, Table 60*.

The Bayley-III Social-Emotional questionnaire was completed in 140 out of the 141 eligible respondents to the Q-CHAT questionnaires. The Bayley-III Social-Emotional score distribution of the preterm population (mean 97.8, SD 17.2, range 55–145) did not differ significantly from the standardised norm of mean 100 and SD of 15 ($p = 0.12$). Twenty-three (16.5%) children had Q-CHAT scores of higher than 2 SDs above the general population mean (i.e. > 42.3). Only five (3.6%) children scored lower than 2 SDs below the standardised mean (i.e. > 70) for the Bayley-III Social-Emotional scale. There was poor concordance between the questionnaires, with only three children classed to be 'at risk' for ASD by both questionnaires and a resulting Cohen's κ coefficient of 0.17 (95% CI 0 to 0.36).

Systematic literature review and meta-analysis

The PRISMA flow diagram is shown in *Figure 20*. The electronic literature search yielded 3600 unique citations (one of which was a duplicate). Application of search limits excluded 343 non-English articles and 413 articles published before 1990. Two additional studies were identified through manual search and author correspondence. Sixty-eight studies were selected for full-text evaluation from the title/abstract screen and 44 met the eligibility criteria. By matching 375 articles that reported the conduct of early developmental assessments with 323 articles that reported school-age assessments, 10 additional studies (in 23 articles) were identified. Data required for the review and meta-analysis were extractable directly from six articles. The authors of 18 of the remaining 48 studies contributed unpublished data for this review. The list of included studies is in *Appendix 1, Table 61*. For simplicity of referencing, studies that are represented by more than one article are denoted by the first author and year of publication of the earliest article in all tables and figures.

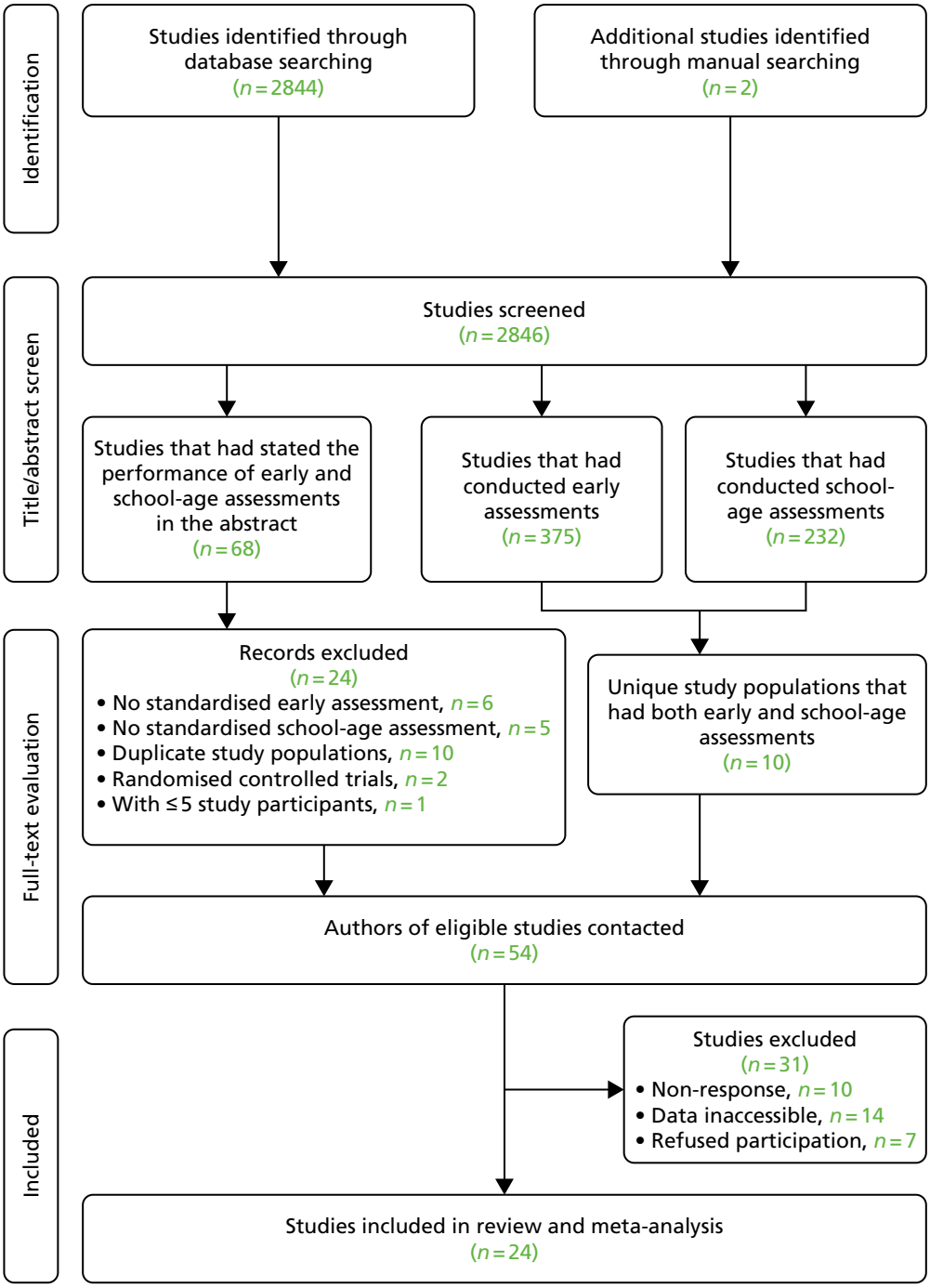


FIGURE 20 The PRISMA flow diagram depicting the literature search process.

Description of included studies

The studies were conducted in Europe (12 studies), the USA (seven studies), Australia (three studies), New Zealand (one study) and Israel (one study). Sample sizes ranged from 11 to 313 participants. Most studies restricted the recruitment of participants to a single institution (15 studies), three studies were multicentre and six studies adopted a geographical population-based sampling method. The inclusion criteria were wholly based on birthweight in nine studies, and on gestational age in five studies and based on both birthweight and gestational age in five studies. For the other studies, additional inclusion criteria applied, including intrauterine growth restriction,[278] spastic diplegia,[279] specific neonatal diagnoses[280,281] and low parental socioeconomic status.[282] The participants in six studies consisted of children who were born at > 32 weeks' gestation and with a birthweight of > 1500 g, but the authors were able to provide relevant

data limited to the subgroup that meet the criteria for this review. Tools used in each study and the ages of application are also listed in *Appendix 1, Table 62*. As the studies spanned a period of > 30 years, different editions of the same assessment tool were recorded.

Study populations

From these 24 studies, a total of 3133 children who were born at ≤ 32 weeks' gestation and/or with a birthweight of < 1500 g received both early and school-age assessments. The mean gestational ages at birth ranged from 25.0 to 33.1 weeks and the mean birthweights were between 675 g and 1298 g. A total of 37.0% (1159 children) of the included populations were born in the years 1972–90, 49.6% (1555 children) in 1991–2000 and 13.4% (419 children) in 2000–5. Children with known genetic syndromes and congenital anomalies were excluded from the studies. Children with severe neurosensory (including blindness and deafness) and motor impairment were likely to be under-represented in the cohort, as 13 studies (contributing 48% of the final sample) excluded children who were unable to complete the assessments as a result of their physical disabilities.[161,278,281–291] The actual number of children excluded from the analysis for this reason is unknown, as not all studies provided this information. In the studies by Claas *et al.*[292] and Fedrizzi *et al.*,[279] no child was unable to complete the assessment because of the presence of physical disability. In studies that included participants who were 'too physically disabled to be tested',[160,162,163,173,280,293,294] these children were assigned a nominal score that was equivalent to being more than 2–4 SDs below the population mean.

Developmental and cognitive assessments

Ten studies reported the results of developmental assessments conducted between 12 and 24 months corrected age and 11 studies reported the results at 24 months corrected age. In three of these studies,[161,283,293] a repeat assessment was conducted at age 3 years. Fedrizzi *et al.*[279] reported results at 3 years and Smith *et al.*[282] reported results at 3.5 years chronological age.

The results of the school-age cognitive assessment were available at the ages of 5–6 years in 16 studies, 7–10 years in 11 studies and > 10 years in three studies. Cohen,[286] Reuss *et al.*,[288] Marlow *et al.*,[225] Smith *et al.*,[282] and Wolke and Meyer[173] conducted multiple school-age assessments at different time points for their study populations.

The proportion of children diagnosed with developmental impairment (test scores of > 1 SD below standardised or control group mean) varied widely among studies, ranging from 6.0%[284] to 67.0%.[162] The reported prevalence of school-age cognitive deficit was between 5.0%[286] and 67.4%[225] for mild–moderate (1–2 SDs below the mean) and between 0.0%[279,286] and 37.8%[225] for severe impairment (> 2 SDs below the mean). In six studies,[163,280,282,287,294] the categorisation of outcomes was based on the mean and SD of the scores achieved by concurrently recruited term-born controls. Wolke and Meyer[173] used cohort-specific cut-off points derived from a normative sample representative of the total population of infants in the Bavarian region to categorise impairments. It should be noted that the study population in Smith *et al.*[282] was from middle to low socioeconomic groups and the mean test score achieved by the control group was about 0.5 SDs below the standardised mean. Using the results from the control group in this case could lead to an underestimation of the prevalence of impairment in this study. If the test standardised norm values were used, the prevalence of cognitive impairment diagnosed at 8 years of age would increase from 24.0% to 36.0% for mild–moderate and from 6.0% to 6.6% for severe impairment.

Quality of included studies: results of QUADAS-2 appraisal

Appendix 1, Table 62, shows the details of the quality of each included study based on the QUADAS-2 appraisal tool, and in *Appendix 2, Figure 40*, the proportions of studies that were considered at 'low', 'high' and 'unclear' risk for bias and applicability in each domain are displayed. The loss to follow-up of > 30% of the eligible birth cohort was a main source of selection bias in the included studies. Risk of information bias is low but may be introduced in three studies[279,283,292] because of the lack of blinding of assessors performing the school-age assessments to the results of the early developmental tests. It was unclear if blinding occurred in the studies by Roberts *et al.*,[163] Reuss *et al.*,[288] Smith *et al.*[282] and Tommiska *et al.*[290] Although the overall risk of bias was low, there is high concern for the applicability of the results

from the studies to our current population in > 50% of the studies. This is because many of the included studies were conducted more than 20 years ago and, therefore, the characteristics of the study populations would be different and the assessment tools have been superseded by newer versions.

Predictive validity of early developmental assessment

The results of the cross-tabulations and the estimated sensitivities and specificities of early assessments for identifying any cognitive deficit for each study, in the form of coupled forest plots ordered by the sample size of the study, are shown in *Figure 21* with the same information for the diagnosis of severe cognitive impairment. In studies for which participants were examined at different time points, only the results from the assessment performed at the oldest age are presented. This gives a final sample size of 3060 children for the meta-analysis. There was significant heterogeneity in the reported sensitivities and specificities among studies ($p < 0.001$ for both). The estimated sensitivities of diagnosing any impairment ranged from 17.0% to 90.5% and the corresponding estimated specificities ranged from 46.8% to 98.4%. For the diagnosis of severe impairment, the range of sensitivities was 0.0% to 100.0% and the range of specificities was 70.8% to 100.0%. The sensitivity of detecting severe impairment could not be estimated in the studies by Cohen[286] and Fedrizzi *et al.*[279] as no participant had severe impairment. There appears to be a wider range and poorer precision (wider CIs) in the estimated sensitivity than in the specificity across studies. This may reflect the presence of heterogeneity or more likely as a result of estimates of sensitivity being based on smaller samples than estimates of specificity. The estimated sensitivity of 0.0% for severe impairment was based on a denominator of 1[283,284,290] and 10[293] diagnosed cases at school-age assessments. In general, the larger the sample size, the more precise (the smaller the 95% CI) the sensitivity estimates. The precision of specificity estimates appears to be high with the CI half-widths in 10 studies being < 10.0%.[160,161,173,225,280–282,285,288,291]

Meta-analytic pooled estimates of sensitivity and specificity

There was significant correlation between estimated sensitivities and specificities (see *Appendix 2*, *Figure 41*; Spearman's rho –0.76; < 0.001). Therefore, the weighted averages of sensitivities and specificities were not computed separately. The pooled measures were estimated from the Rutter and Gatsonis HSROC curves that are presented in *Appendix 2*, *Figure 42*, for the presence of any impairment and severe impairments. The summary points and 95% CI regions are mapped out in the figures as well as the 95% prediction regions, which provide a forecast of the true sensitivity and specificity in a future study. The summary points corresponded to a pooled sensitivity of 55.0% (95% CI 45.7% to 63.9%) and pooled specificity of 84.1% (95% CI 77.5% to 89.1%) for the identification of any impairment. For the diagnosis of severe impairment, the pooled sensitivity was 39.2% (95% CI 26.8% to 53.3%) and pooled specificity was 95.1% (95% CI 92.3% to 97.0%).

Validity of early assessment assessed at different time points

In three studies,[161,283,293] participants were assessed at two different time points for the early developmental assessments. In the five studies by Cohen,[286] Reuss *et al.*,[288] Marlow *et al.*,[225] Smith *et al.*,[282] and Wolke and Meyer,[173] participants received school-age cognitive assessments more than once. In *Figure 22*, the sensitivity and specificity for the identification of any impairment are plotted over the age at developmental assessment for the three studies that examined early assessment at two different time points. *Figure 23* shows similar plots for the results obtained at serial school-age assessments in the five studies. It would appear, from these graphical displays, that the specificity of early assessment in excluding cognitive deficit remains relatively stable over time whereas no real correlation between sensitivity and age at assessment was apparent.

Metaregression: association of study-level variables with diagnostic validity

The ORs and 95% CIs, together with the corresponding *p*-values, for the association of study-level variables with sensitivity and specificity of identifying cognitive deficit by early developmental assessment are presented in *Table 30*. There was reduction in specificity with increased observed prevalence of impairment in the study population. For each 1% increase in cognitive impairment prevalence, the odds of identifying an additional 'true-negative' case among those with no cognitive impairment reduced by

(a)

Study	TP	FP	FN	TN	Sample size	Sensitivity (95% CI)	Specificity (95% CI)
Charkaluk 2011[285]	33	69	19	192	313	0.63 (0.49 to 0.76)	0.74 (0.68 to 0.79)
Wolke 1999[173]	81	43	35	95	254	0.70 (0.61 to 0.78)	0.69 (0.60 to 0.76)
Veelken 1991[291]	40	14	55	125	234	0.42 (0.32 to 0.53)	0.90 (0.84 to 0.94)
Hack 2005[162]	67	67	7	59	200	0.91 (0.81 to 0.96)	0.47 (0.38 to 0.56)
Marlow 2005[225]	83	20	47	43	193	0.64 (0.55 to 0.72)	0.68 (0.54 to 0.78)
Reuss 1996[288]	18	14	23	135	190	0.44 (0.28 to 0.60)	0.91 (0.85 to 0.95)
Roberts 2010[163]	51	30	43	62	186	0.54 (0.44 to 0.65)	0.67 (0.57 to 0.77)
Vermeulen 2001[295]	13	12	10	150	185	0.57 (0.34 to 0.77)	0.93 (0.87 to 0.96)
Smith 2006[282]	34	31	12	84	161	0.74 (0.59 to 0.86)	0.73 (0.64 to 0.81)
Orchinik 2011[294]	61	31	13	34	139	0.82 (0.72 to 0.90)	0.52 (0.40 to 0.65)
Gray 1995[281]	7	4	32	83	126	0.18 (0.08 to 0.34)	0.95 (0.89 to 0.99)
Munck 2012[160]	9	2	15	98	124	0.38 (0.19 to 0.59)	0.98 (0.93 to 1.00)
Kilbride 1990[293]	45	23	9	45	122	0.83 (0.71 to 0.92)	0.66 (0.54 to 0.77)
Claas 2011[292]	12	14	12	63	101	0.50 (0.29 to 0.71)	0.82 (0.71 to 0.90)
Gray 2006[287]	29	12	18	40	99	0.62 (0.46 to 0.75)	0.77 (0.63 to 0.87)
Potharst 2012[161]	10	4	13	69	96	0.43 (0.23 to 0.66)	0.95 (0.87 to 0.98)
McGrath 2000[280]	8	1	19	60	88	0.30 (0.14 to 0.50)	0.98 (0.91 to 1.00)
Tommiska 2003[290]	4	11	6	51	72	0.40 (0.12 to 0.74)	0.82 (0.70 to 0.91)
Bruggink 2010[284]	2	1	10	37	50	0.17 (0.02 to 0.48)	0.97 (0.86 to 1.00)
Bowen 1996[283]	7	4	2	32	45	0.78 (0.40 to 0.97)	0.89 (0.74 to 0.97)
Bassan 2011[278]	2	8	1	21	32	0.67 (0.09 to 0.99)	0.72 (0.53 to 0.87)
Skranes 1998[289]	1	3	3	14	21	0.25 (0.01 to 0.81)	0.82 (0.57 to 0.96)
Cohen 1995[286]	3	2	1	11	17	0.75 (0.19 to 0.99)	0.85 (0.55 to 0.98)
Fedrizzi 1993[279]	1	1	1	8	11	0.50 (0.01 to 0.99)	0.89 (0.52 to 1.00)

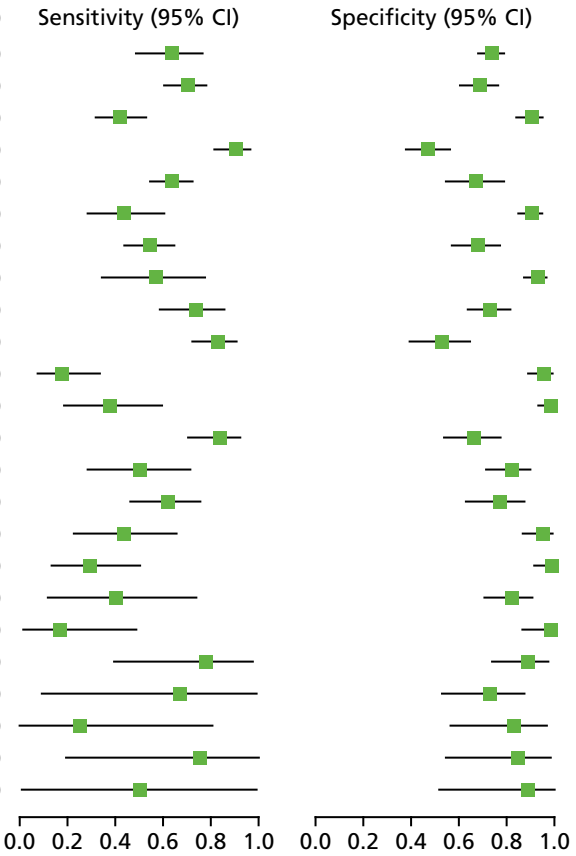


FIGURE 21 Results of cross-tabulations and coupled forest plots of the estimated sensitivities and specificities of early developmental assessments in identifying the presence of (a) coupled forest plots for the identification of any cognitive impairment; and (b) coupled forest plots for the identification of severe cognitive impairment. Sensitivities and specificities are expressed as proportions. (*continued*)

(b)

Study	TP	FP	FN	TN	Sample size	Sensitivity (95% CI)	Specificity (95% CI)
Charkaluk 2011[285]	5	23	3	282	313	0.63 (0.24 to 0.91)	0.92 (0.89 to 0.95)
Wolke 1999[173]	44	30	14	166	254	0.76 (0.63 to 0.86)	0.85 (0.79 to 0.89)
Veelken 1991[291]	11	13	20	190	234	0.35 (0.19 to 0.55)	0.94 (0.89 to 0.97)
Hack 2005[162]	29	49	3	119	200	0.91 (0.75 to 0.98)	0.71 (0.63 to 0.78)
Marlow 2005[225]	21	5	52	115	193	0.29 (0.19 to 0.41)	0.96 (0.90 to 0.98)
Reuss 1996[288]	3	5	9	173	190	0.25 (0.05 to 0.57)	0.97 (0.94 to 0.99)
Roberts 2010[163]	18	24	10	134	186	0.64 (0.44 to 0.81)	0.85 (0.78 to 0.90)
Vermeulen 2001[295]	1	9	1	174	185	0.50 (0.01 to 0.99)	0.95 (0.91 to 0.98)
Smith 2006[282]	3	11	5	142	161	0.38 (0.09 to 0.76)	0.93 (0.88 to 0.96)
Orchinik 2011[294]	26	26	7	80	139	0.79 (0.61 to 0.91)	0.75 (0.66 to 0.83)
Gray 1995[281]	1	0	6	119	126	0.14 (0.00 to 0.58)	1.00 (0.97 to 1.00)
Munck 2012[160]	5	1	4	114	124	0.56 (0.21 to 0.86)	0.99 (0.95 to 1.00)
Kilbride 1990[293]	0	6	10	106	122	0.00 (0.00 to 0.31)	0.95 (0.89 to 0.98)
Claas 2011[292]	1	4	6	90	101	0.14 (0.00 to 0.58)	0.96 (0.89 to 0.99)
Gray 2006[287]	9	7	6	77	99	0.60 (0.32 to 0.84)	0.92 (0.84 to 0.97)
Potharst 2012[161]	2	1	7	86	96	0.22 (0.03 to 0.60)	0.99 (0.94 to 1.00)
McGrath 2000[280]	3	0	5	80	88	0.38 (0.09 to 0.76)	1.00 (0.95 to 1.00)
Tommiska 2003[290]	0	0	1	71	72	0.00 (0.00 to 0.97)	1.00 (0.95 to 1.00)
Bruggink 2010[284]	0	0	1	49	50	0.00 (0.00 to 0.97)	1.00 (0.93 to 1.00)
Bowen 1996[283]	1	3	0	41	45	1.00 (0.03 to 1.00)	0.93 (0.81 to 0.99)
Bassan 2011[278]	1	1	0	30	32	1.00 (0.03 to 1.00)	0.97 (0.83 to 1.00)
Skranes 1998[289]	1	1	1	18	21	0.50 (0.01 to 0.99)	0.95 (0.74 to 1.00)
Cohen 1995[286]	0	0	0	17	17	Not estimable	1.00 (0.80 to 1.00)
Fedrizzi 1993[279]	0	0	0	11	11	Not estimable	1.00 (0.72 to 1.00)

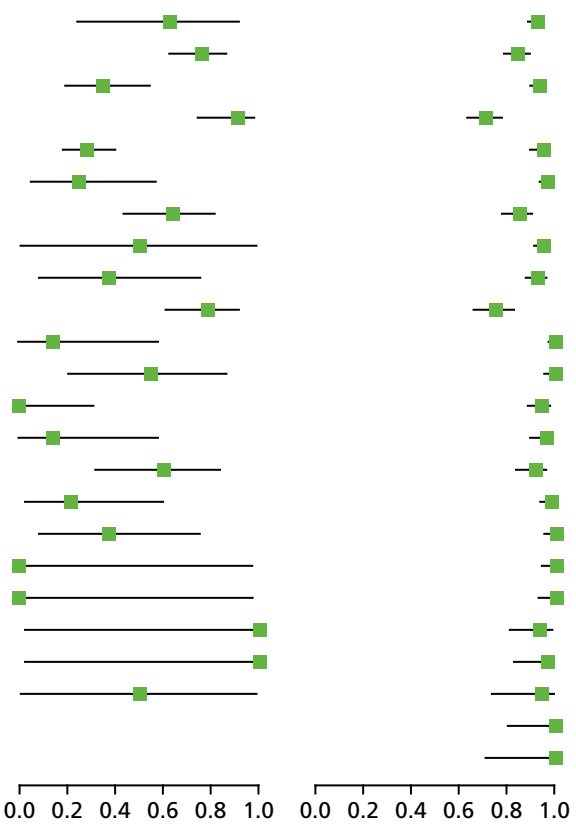


FIGURE 21 Results of cross-tabulations and coupled forest plots of the estimated sensitivities and specificities of early developmental assessments in identifying the presence of (a) coupled forest plots for the identification of any cognitive impairment; and (b) coupled forest plots for the identification of severe cognitive impairment. Sensitivities and specificities are expressed as proportions.

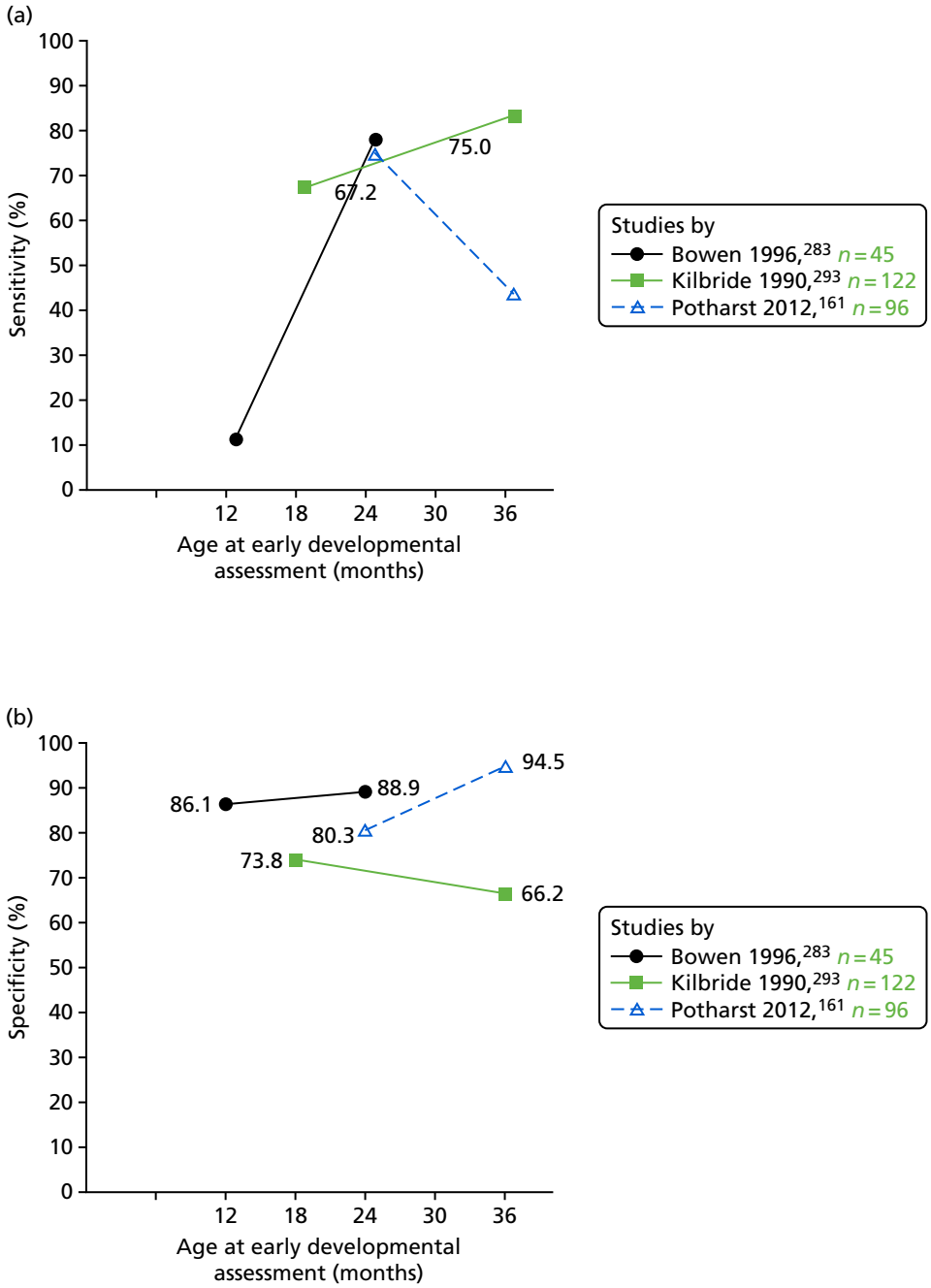


FIGURE 22 Line graphs demonstrating the change in (a) sensitivity and (b) in specificity when early developmental assessments were repeated at different ages in three studies.

3% ($p = 0.01$). The associations between mean gestational age and mean birthweight and specificity of identifying cognitive impairment reached borderline statistical significance (specificity increased with mean gestational age and mean birthweight of the study population). Post hoc analysis revealed no association between the prevalence of impairment reported in each study and the mean gestational age ($p = 0.55$) and mean birthweight ($p = 0.95$) of the study population; therefore, excluding the speculation that the observed association between specificity and mean gestational age and birthweight was mediated by the prevalence of impairment. The age at the assessments, the time interval between early and school-age assessments and the year of participant birth were not associated with sensitivity or specificity and, therefore, did not explain the heterogeneity present between studies.

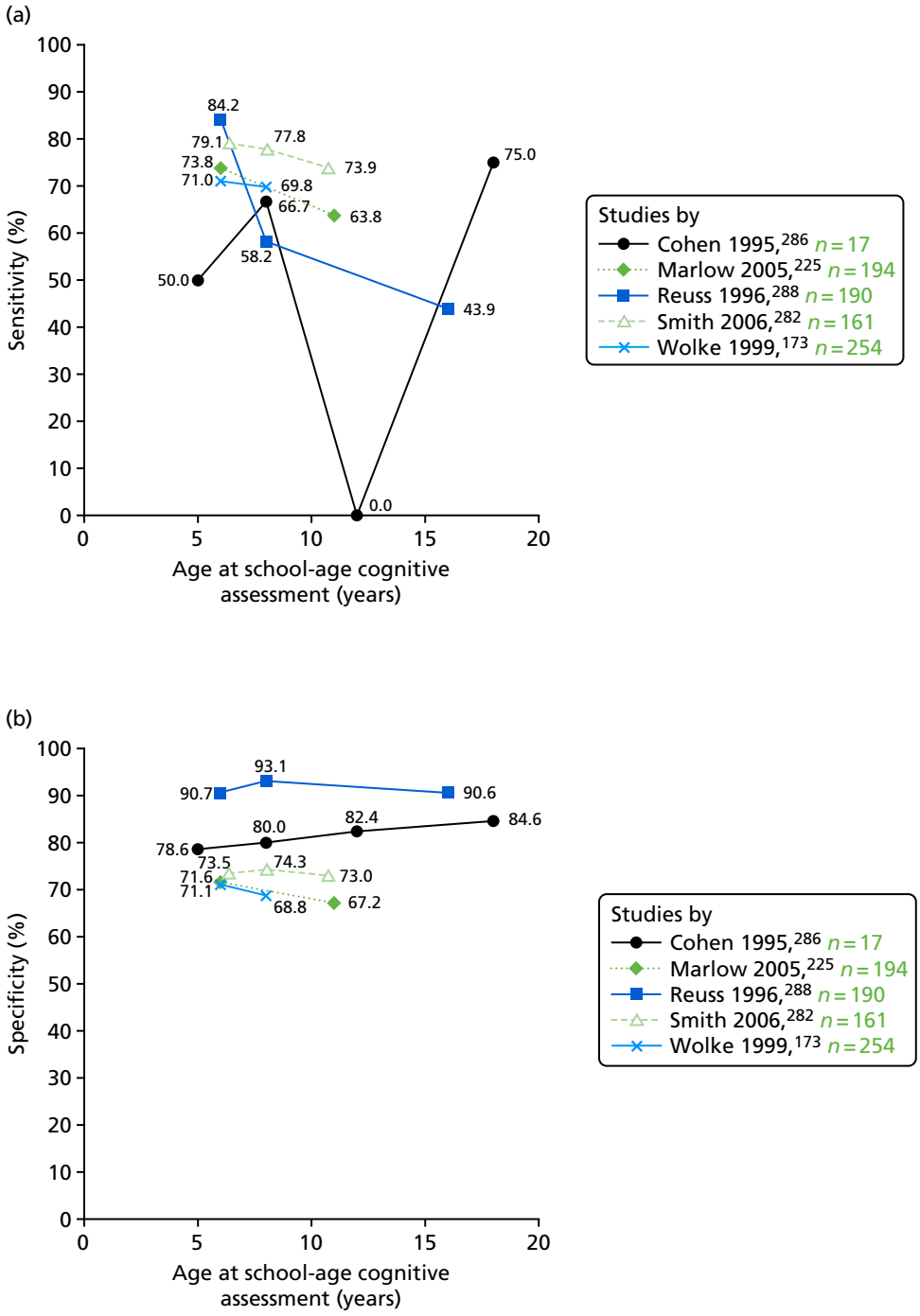


FIGURE 23 Line graphs demonstrating the change in (a) sensitivity and (b) in specificity when school-age cognitive assessments were repeated at different ages in four studies.

Funnel plot for sample size-related effects and publication bias

The funnel plot of the log-DOR against the inverse of the square root of the effect sample size is presented in *Figure 24*. Significance testing (ESS weighted regression test) confirmed that asymmetry was not present in the funnel plot ($p = 0.22$), indicating the absence of sample size-related effects in the meta-analysis.

TABLE 30 Association of study-level variables with estimated sensitivity and specificity

Study-level variable	Sensitivity		Specificity		*p*-value for joint test
	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value	
Mean gestational age (per 1-week increase)	0.84 (0.68 to 1.04)	0.11	1.29 (0.98 to 1.61)	0.04	0.11
Mean birthweight (per 100-g increase)	0.86 (0.72 to 1.03)	0.09	1.21 (1.00 to 1.48)	0.05	0.14
Mean age at early assessment (per 1-year increase)	1.51 (0.77 to 2.98)	0.22	0.79 (0.36 to 1.72)	0.54	0.35
Mean age at school-age assessment (per 1-year increase)	0.98 (0.86 to 1.11)	0.73	1.01 (0.88 to 1.17)	0.86	0.90
Mean time between assessments (per 1-year increase)	0.97 (0.86 to 1.10)	0.57	1.02 (0.89 to 1.17)	0.78	0.82
Year of birth (per 1-year increase)	0.99 (0.97 to 1.01)	0.291	0.99 (0.97 to 1.01)	0.23	0.82
Prevalence of impairment (per 1% increase)	1.02 (0.99 to 1.04)	0.16	0.97 (0.94 to 1.00)	0.01	0.02
Prevalence of severe impairment (per 1% increase)	1.05 (0.99 to 1.12)	0.12	0.93 (0.87 to 1.00)	0.02	0.03

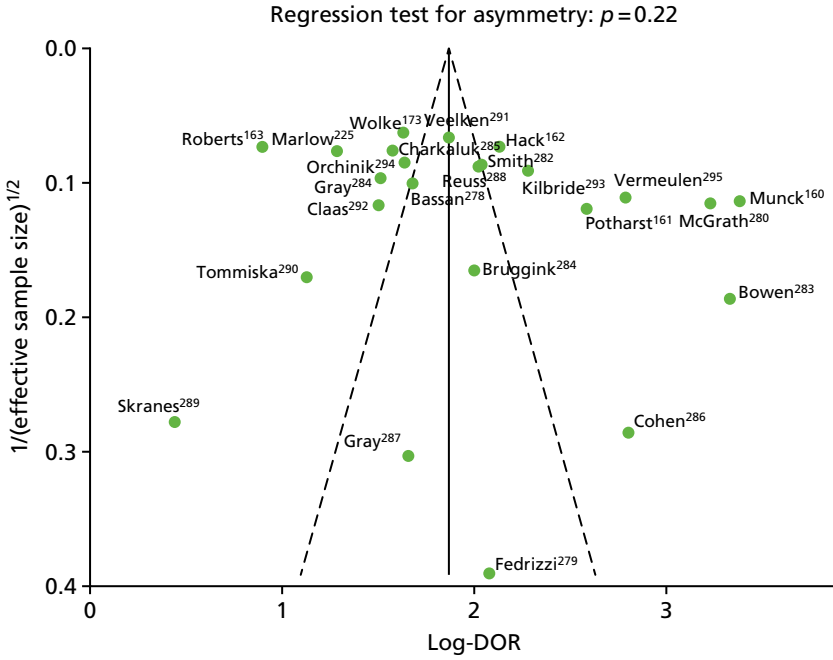


FIGURE 24 Funnel plot of the log-DOR against the inverse of the square root of the ESS, with pseudo-95% confidence limits.

Conclusions

Agreement between NHS and research-standard data

Among children who were born before 30 weeks' gestation, the agreement in classifying neurodevelopmental status at the age of 2 years between data recorded during routine NHS assessments and those obtained through a research assessment was strong in the absence of neurodevelopment impairment. However, NHS assessments lack satisfactory sensitivity for identifying children with impairment, particularly in the cognitive and language domains. Using the Bayley-III scores as the 'gold-standard' tool, approximately 30% of children with at least mild cognitive impairments and nearly 50% with at least mild language impairments were falsely

classified as having no impairment through NHS assessment. The discordance between NHS and research assessment remained, irrespective of the criteria used to categorise outcomes. This implies that the structural and content differences between the classification tools are unlikely to account for the discordances identified.

The strengths of the study include a single, carefully trained research assessor blinded to the results of the NHS assessment; the involvement of 13 hospitals serving patients from a wide range of ethnic, social, economic and cultural backgrounds; analyses that took into account the possibility of clustering by study sites and multiple births; and overestimation of abilities by the Bayley-III assessment. A limitation to the study was that the targeted sample size was not achieved. As the recruitment of participants occurred at a steady rate over the planned 2-year period, factors that may be relevant are the high population mobility in London, leading to loss to follow-up, and the large number of consultant and trainee doctors involved in outpatient follow-up clinics leading to missed opportunities to invite participation, because health professionals were unaware of the study. The sample size target was calculated with the desire to estimate the sensitivity of NHS assessment in correctly classifying children with severe impairment to a high precision, achieving a narrow 95% CI with half-widths within 10%. In addition to sample size, the precision was also dependent on the actual value of the sensitivity estimate: the lower the sensitivity, the wider the CI. As the estimated sensitivity for diagnosing severe overall impairment was low (52.0%), based on the 14.1% prevalence of severe impairment among the study participants, a sample of at least 680 children providing independent (unclustered) observations would be necessary to have achieved the intended precision. The increment in precision with increasing sample size followed the law of diminishing returns. It was therefore difficult to justify the continued provision of additional time and resources required to achieve the desired precision for the sensitivity estimate.

The study population differed from the population of very preterm children discharged from participating hospitals in that it consisted of proportionally more white children, who were less likely to have been mechanically ventilated, diagnosed with chronic lung disease (bronchopulmonary dysplasia) and/or were living in less deprived areas. It is possible, therefore, that the study population was at lower risk for neurodevelopmental impairment. The selection bias was introduced by attrition of children from routine NHS follow-up and the non-random recruitment method. In the literature, the proportions of VLBW children who were lost to follow-up or reviewed with difficulty in regional follow-up programmes were reported to be around 11–27% at 2 years[296–298] and 25% at 5 years.[299] Characteristics associated with dropping out from follow-up included non-white ethnicity, young maternal age and low socioeconomic and maternal educational status.[297–300] Ideally, a random sample of participants selected from a known sampling frame (e.g. list of all children with scheduled follow-up appointments) would provide the most representative study cohort.

The presence of selection bias may have affected the accuracy of the estimated prevalence of impairment in the population. Traditionally, sensitivity and specificity are considered to be independent of disease prevalence.[301] Consequently, the adverse effect of selection bias on the validity of this study can be regarded as minimal. However, a number of studies have shown that variation in prevalence can result in either clinical or artefactual variation in test accuracy.[302,303] As it is probably easier to diagnose impairment in severe than in mild–moderate cases, a study population with a lower spectrum of impairment might have more false-negative or false-positive results.

Inter-rater variability in outcome assignment is likely to have been one of the main reasons for the disagreement between NHS assessments and research assessments. Clinical judgement is inevitably influenced by the assessor's knowledge, experience, beliefs and preconceptions. Studies on behavioural psychology have shown that people tend to rely on judgemental heuristics (e.g. intuition), which are, by nature, unreliable, to simplify the complex task of assessing probabilities and predicting values to provide reasoning on the outcome of an event, such as the diagnosis of disability in a child.[304] The use of standardised assessment tools improves inter-rater agreement by establishing objective measures. Without using a standardised assessment, the judgement and interpretation of clinical findings may be highly variable. It is therefore unsurprising that, in a study comparing the diagnosis of cognitive impairment made

using an intelligence test with judgements by paediatricians, the agreement was only fair (κ 0.39). Even if standardised assessments were used, the agreement between the different tools in classifying impairment was uncertain. Chaudhary *et al.*[305] reported that, at 22 months, children scored 5 points higher on the BSID-II MDI than on the Griffiths Scales developmental quotient. Furthermore, the interpretation and translation of the standardised assessment scores into the NPEU/Oxford classification instruction can still be inconsistent and subject to biases and errors.

Another difference between the conduct of the NHS assessments and research assessments was the reliance of the latter on parents' reports on their child's ability, particularly for language and cognitive skills. Parents are a valuable source of information in a time-restricted appointment, especially if the child does not engage in the assessment. However, studies investigating the level of agreement of neurodevelopmental status between parent and paediatrician evaluation have reported variable results.[244–246,248]

Intrasubject (participant) variability in performance between assessments could also contribute to the discordance between NHS and research assessments. There are multiple factors, such as mood and ease of engagement of the participant, time of the day (meals/snacks or nap times) and environment, that can influence the children's performance. Preterm children have been shown to be at risk of inattention/hyperactivity[306,307] and social-emotional delays,[308] which could manifest as inability to complete a task. The testing time is also generally longer for research assessments and it was not unusual for children to become tired during testing. These issues may not have been taken into account by the assessors and, specifically, the objective scoring of a standardised assessment would not have made allowance for underperformance because of these factors.

There are several possible reasons for the higher sensitivities for diagnosing motor impairment than for cognitive or language impairments at routine NHS assessments. Important motor developmental milestones (e.g. sitting and walking) are reached at a relatively young age and parents and health professionals place great emphasis on checking that children achieve these milestones. Cerebral palsy is the most commonly quoted morbidity of preterm birth; therefore, motor assessments are regularly performed at follow-up appointments. Cognitive and language skills can be difficult to ascertain in a single setting, particularly without the use of standardised assessment tools, and can be affected by the issues of judgement and reporting bias discussed above. Furthermore, in the modified NPEU/Oxford classification, the categorising of cognitive impairment by 'number of months behind corrected age' introduced another level of variability. In addition, in the electronic '2-year outcome' form, the term 'development' was used as the heading for the cognitive domain. As a result, there was misinterpretation among the NHS assessors that the questions in that category applied to 'overall development in all domains' rather than being specific to cognitive function, potentially leading to misclassification.

The impact of inter-rater and intrasubject variability would be exacerbated by the classification of neurodevelopment skills, a continuous trait, into categories of ability or impairment. Levels of abilities or skills near the 'cut-off' point between categories are more difficult to discriminate and are at risk of being misclassified into higher or lower impairment categories.

Subgroup analyses were used to investigate whether or not the validity of the NHS assessment was affected by neonatal and sociodemographic factors, as well as factors related to the conduct of the assessments. However, given that the numbers of children with impairment ('true positives') within each subgroup were small, it was likely to that subgroup analyses were underpowered. Therefore, the possibility remains that the negative findings were a reflection of type II errors ('false negatives').

The Bayley-III was selected as the research assessment as it is the most commonly used assessment in neonatal outcome studies. However, the validity of the Bayley-III, particularly in identifying school-age outcomes, is unknown. Several studies have raised concerns that, when compared with the BSID-II, the Bayley-III was underestimating neurodevelopmental impairment.[202–205] It is, however, reassuring that validation studies of the Bayley-III showed that the scores are consistent with other revised ability tests such

as the Wechsler Preschool and Primary Scale of Intelligence – Third Edition[309] and the Preschool Language Scale – Fourth Edition.[201,310] In this study, more children were classified as being impaired using the predicted BSID-II MDI scores than using the Bayley-III scores if the same threshold were applied. Therefore, as expected, the sensitivities of the NHS assessment dropped when the predicted BSID-II MDI scores were used as the 'gold standard' instead of the Bayley-III scores.

Another issue that needs to be considered is the impact of administering the English-based Bayley-III assessment on cognitive and language scores in children whose primary language is not English. Although families who require interpretation for English were excluded, 52% of the study population was living in a bilingual environment. Studies that examined the effect of bilingualism on language acquisition have provided conflicting evidence,[311–313] but testing bias cannot be ruled out. However, for a child who is functioning in the 'severe impairment' category, communication skills are assessed by the observation of gestures and the production of consonant and/or vowel sounds, which are not language specific, and, hence, the assessment of children with severe language impairment is likely to be valid. Testing bias can also occur in NHS assessments where there is likely to be greater reliance on parental reporting.

Social communication skills of children who were born very preterm

At 24 months corrected age, children who were born before 30 weeks' gestation were rated by their parents on the Q-CHAT as having greater social communication difficulties and autistic traits than of the general population. The higher frequency in autistic traits was observed mainly in the areas of restricted, repetitive, stereotyped behaviour, communication and sensory abnormalities.

Previous studies have reported significantly higher odds of positive autism screening on the M-CHAT in children with motor, visual, hearing and cognitive impairments.[213,214] The Q-CHAT contains similar questions, leading to children with such disabilities receiving higher Q-CHAT scores. Therefore, it is likely that the distribution of Q-CHAT scores in a very preterm population would be even higher if children with cerebral palsy and severe neurosensory disabilities were included.

There was no sex difference in our study population. This may be due to insufficient statistical power, given that a sample size of 24,000 children would be required for the 0.3-point sex difference in Q-CHAT scores that we detected to be significantly different. Nevertheless, it has been suggested that the autistic phenotype seen in preterm children resembles more closely syndromic or medically explained autism, the sex ratio of which is closer to 1 : 1, than those with idiopathic autism,[213] supporting the hypothesis that autism in preterm children, rather than being a primary deficit, represents part of a 'preterm phenotype' with different aetiology.

Preterm children experience difficulties across all aspects of autistic behaviour but particularly in the categories of restricted, repetitive, stereotyped behaviour, communication and sensory abnormalities. The presence of reduced language abilities among children who were born preterm is well described.[176,177] Dysfunction in sensory modulation in preterm children, characterised by either hyposensitivity or hypersensitivity to sensory input, is a problem anecdotally recognised by parents and clinicians. It is hypothesised that exposure to the stressful environment of the neonatal intensive care unit at a critical period of brain development in the third trimester interferes with the normal maturation of the sensory system.[314,315] Sensory modulation dysfunction is thought to be negatively associated with emotional development and can affect social interactive capabilities.[316]

There is some evidence that restricted and repetitive behaviours are associated with cognitive status.[317,318] EPICure study investigators also concluded that cognitive deficits in their extremely preterm cohort accounted for the excess of repetitive and stereotyped behaviour when compared with the full-term controls.[319] Although there was no correlation between cognitive scores and subcategorical Q-CHAT scores in the restricted and repetitive behaviour domain, as the mean cognitive score of the preterm population was lower than would be expected in the general population. The potential association between cognition and restricted and repetitive behaviour could, in part, explain the higher Q-CHAT scores obtained by the participants in this category.

There were fewer differences between preterm children and the general population in response to items exploring social relatedness. Q-CHAT items exploring social relatedness may provide a higher degree of specificity for differentiating early autistic features from concurrent developmental delay in children without severe physical and neurosensory impairment compared with items in the other categories. Although parents reported a lower frequency of pretend play among the preterm children, development in joint attention (elucidated by questions on protodeclarative pointing and following a gaze) was similar in the general population. Focusing on elucidating social relatedness for autism screening in the preterm population may reduce the 'false-positive' screening rate associated with currently available screening tools.

The significant association between language ability at the age of 2 years and Q-CHAT scores was unsurprising, as four items on the Q-CHAT specifically examined language development. Furthermore, language ability was closely related to cognitive function, which, in turn, influenced performance on other Q-CHAT items. Separate cognitive and language scores were obtained from the Bayley-III assessment in this study. Language scores confounded and accounted for the association observed between cognitive scores and Q-CHAT scores.

This study also highlights the inter-relationship between ethnicity, area deprivation, language skills and Q-CHAT scores. Our findings suggest the possibility of an environmental impact of socioeconomic disadvantage on early social communication development.

The Q-CHAT, M-CHAT and the Bayley-III Social-Emotional are some of the developmental surveillance tools designed to identify toddlers at risk for developing ASD, with the aim of implementing timely intervention strategies to achieve better outcomes for these children. The M-CHAT has a sensitivity of 87% and specificity of 98% for ASD when applied in a mixed sample of children aged between 16 and 30 months.[211] The Bayley-III Social-Emotional questionnaire, using a scaled score of 6, reportedly had a sensitivity of 87.0% and specificity of 90.0% for the identification of ASD.[207] However, the predictive validity of these screening tools when applied to the preterm population has not been investigated. Furthermore, there is little understanding of the differences in properties of the available screening tools. Oosterling *et al.*[320] compared four instruments: the Early Screening of Autistic Traits Questionnaire;[321–323] the Social Communication Questionnaire;[323] the Communication and Symbolic Behaviour Scales-Developmental Profile, Infant-Toddler Checklist;[216] and key items of the Checklist for Autism in Toddlers.[324] They found that no particular tool showed superior discriminating power for distinguishing children with ASD from those without.

Meta-analysis

The specificities of early neurodevelopmental assessment in predicting later school-age cognitive outcomes were generally high, especially for severe cognitive impairment, but sensitivities were inconsistent. Early neurodevelopmental assessment has low sensitivity and high specificity for identifying school-age cognitive deficit. This means that, when a neurodevelopmental impairment was diagnosed at ages 1–3 years, the likelihood of having cognitive deficit at school age was high (low false-positive rate, or 1 – specificity). However, it would not be possible to exclude later cognitive deficit even when an early assessment demonstrated normal neurodevelopmental outcomes (high false negative; or 1 – sensitivity). The results suggest that almost half the children who were thought to have normal neurodevelopmental function at ages 1–3 years will experience cognitive difficulties at school age. Even for cases of severe cognitive deficit, the accuracy in early detection was low (meta-analytic sensitivity of 39.2%). This finding is not unexpected. Cognitive function in infancy is a poor predictor of later IQ in the general population.[325] This may reflect changes in cognitive function during childhood, unveiling of deficits in complex task performance that are non-essential in early childhood, or the increasing effect of social and environmental influences on cognition over time. Other explanations may be the impact of behaviour and attention during testing at different ages, and differences in the contents and psychometric properties of early neurodevelopmental and later cognitive assessment tools.

The internal validity of this study is influenced by the quality of the data from the included studies as well as by the methods adopted. Data quality as appraised by the QUADAS-2 tool was good, with most studies considered to be at a low risk of bias. Nevertheless, the presence of missing data from participants who were lost to follow-up over time is a common problem affecting these longitudinal studies. Incomplete outcome ascertainment can distort the result in either direction. Another source of missing data arose from the exclusion of children with severe neurosensory and motor impairment who were unable to complete the assessments. If we assume that these children had stable diagnoses of severe neurodevelopmental and cognitive deficits throughout childhood, then the impact of excluding them from the study population would be an underestimation of the sensitivity of early neurodevelopmental assessments.

An additional bias that could affect accuracy, and which was not identified through the QUADAS-2 appraisal, is the experience of the assessors. Although all included studies employed trained assessors using standard assessment tools, interobserver differences are inevitable. Neurodevelopmental and cognitive abilities exist as a continuum but, for the purpose of the study, participants were dichotomised using a 'cut-off' score into groups 'with impairment' and 'without impairment'. Interobserver variations around the 'cut-off' score would result in misclassification of outcomes. The effect of differential misclassification on the study results is difficult to predict but in general it can be expected to have a bigger impact on sensitivity, which is calculated using a small number of 'positives' in this condition of relative low prevalence, than on specificity, which is based on a large number of 'negatives'.

Participants were included in the review if they fulfilled either the gestational age or the birthweight inclusion criterion. The birthweight criterion was used in order to capture all relevant studies, as it has been common for neonatal studies to base eligibility on birthweight rather than gestational age. However, the methodological bias in using a birthweight criterion is the inclusion of more mature but growth-restricted children. Notably, in the study by Bassan *et al.*[278] all the participants were small for gestational age (birthweight < 10th percentile for gestational age). Intrauterine growth restriction is a risk factor for poor neurodevelopmental outcome.[326] The QUADAS-2 appraisal highlighted the lack of applicability of older study populations and outdated assessment tools in more than half of the included studies and, hence, raises the question on the wider generalisability of the study findings. This is, of course, a reflection of the nature of all longitudinal studies but it is a significant limitation, particularly in the context of a rapidly advancing neonatal specialty. The past couple of decades have seen an overall reduction in the proportions of survivors of very preterm birth with adverse neurodevelopmental outcomes at the age of 2 years;[231,327,328] therefore, we can expect the characteristics of the current preterm population to be different to those from past eras. Only 14 of the 24 included studies recruited participants born after 1990 and none was born in the previous 10 years (i.e. after 2004).

More importantly, the assessment tools used in the included studies, although validated and contemporary at the time of each study, have mostly been superseded by newer editions. Therefore, caution should be exercised when extrapolating results based on earlier versions of assessment tools to current practice. The timing and setting of the assessments also played a part in determining the external validity of the study findings. The early neurodevelopmental assessments were performed between 12 and 36 months and the timings matched common clinical practice. School-age assessments were mostly conducted between the ages of 5 and 8 years, when children were at the primary stages of schooling. Only three studies reported cognitive assessment during adolescence, one of which had only 20 participants. Therefore, the validity of early assessment in diagnosing cognitive deficit extending into adulthood could not be estimated from this study, although one could speculate that the sensitivity might be even poorer. As the sensitivity estimates from individual studies were based on a small number of participants with cognitive impairment, the corresponding 95% CIs were very wide. The use of a meta-analytic approach increases the sample size and improves the precision of the pooled estimate.

The review was restricted to English-language literature. There is concern that the English-language journals publish a skewed sample of studies that report positive and more noteworthy results.[329] Similarly, it is common for articles with negative or inconclusive findings to remain unpublished. The exclusion of

grey literature, including abstracts and dissertations, could have led to the omission of essential and more recent information.

Heterogeneity between studies was investigated using metaregression. This method has a few drawbacks. The statistical power to detect associations between the study estimates and the explanatory variables is related to the magnitude of the relationship between them, and is typically considered low in metaregression.[330] This was compounded by the narrow range of values available for each of the explanatory variables under evaluation. For example, the mean gestational age of the included studies ranged only between 25.9 and 33.1 weeks. Hence, a type II error could not be excluded. More importantly, metaregression is subject to ecological fallacy (or aggregation bias) (i.e. the mistaken assumption that a statistical between-study relationship based on aggregated data reflects a within-study relationship). Therefore, in order to reliably identify factors that influence the validity of early developmental assessments, it would be necessary to obtain individual patient-level data.

In conclusion, early neurodevelopmental assessment has high specificity but low sensitivity in identifying later school-age cognitive deficit.

Implications of results

Clinical relevance of results

Routine NHS assessments have low sensitivity for identifying mild to moderate neurodevelopmental impairment. This has significant clinical implications. At an individual level, children with impairment may be missed. At a population level, current documentation of 2-year outcomes during routine NHS assessments, using the standardised EPR in its present format, will underestimate the proportion of children with impairment, compared with a research-standard Bayley-III assessment. Many neonatal networks and units rely on routine follow-up for impairment rates of their graduates. The results of this study question the validity of these practices.

The findings of higher Q-CHAT scores in the preterm population suggest that suboptimal development of social communication skills exists from early childhood. The 7-point right-shift in mean Q-CHAT score of the preterm population corresponds to nearly 1 SD difference. As ASD exists on a continuum, with autism representing the extreme end of the spectrum, the results also support the likelihood that a large proportion of preterm children experience clinically significant social communication difficulties below the diagnostic threshold for ASD from a young age, when early intervention may be possible. The findings draw attention to the need for better understanding and potentially early assessment of social communication skills in the preterm population.

The results from the systematic review and meta-analysis confirm that a significant proportion of children who were born very preterm and who are assessed as having normal neurodevelopment in early childhood go on to experience cognitive difficulties later in school. The implications of this finding on current clinical practice are considerable because neurodevelopmental assessment at 2 or 3 years of age is often used as the end point for post-discharge follow-up of very preterm infants. Outcome data used in discussions with parents during the antenatal and neonatal periods are commonly based on neurodevelopmental outcomes determined in early childhood. Given these findings, it is essential to discuss potential difficulties at school that children may face, even in the absence of obvious impairment or disability at the 2-year assessment.

Reassuringly, the false-positive rate for early diagnosis of impairment was low, indicating that children with more severe impairments, who would receive greater benefit from early intervention, will be correctly identified.

Implications for health care

There are advantages in the current practice of embedding neurodevelopmental follow-up of very preterm children with neonatal services. These include the continued involvement of health professionals known to the families and local flexibility in organisation. However, this also risks regional variation in follow-up

criteria, reliability assessments and quality of data recording. Standardising the neurodevelopmental tool and ensuring that staff are trained in its use during follow-up assessment would be an obvious way of minimising some of this variability. Such a tool should have strong psychometric properties, be user friendly and, ideally, be adaptable for use in non-English-speaking patients. Since this study, the NICE guideline[331] for developmental follow-up of children and young people born preterm has been published. The guideline recommends that the PARCA-R parent-completed questionnaire is used to identify children at risk of developmental delay. Misclassification occurs during categorisation of outcomes; hence, the strategy of presenting outcome data in categories should also be further considered. Categorical outcomes are easy to interpret and to communicate, and mirror clinical practice (e.g. referral of children below a certain threshold for further assessment or intervention). However, for an individual child, the labelling of 'outcome category' is unhelpful. Besides, as shown, categories of outcomes do not remain stable over time. It is, arguably, more valuable to present the distribution of standardised scores.

Under the UK Healthy Child Programme, all children receive health visitor-led developmental screening. There has been some interest in extending the roles of health visitors to capture developmental outcome data of children who were born preterm, assessed using developmental screening tools or through questions similar to those listed on the electronic '2-year outcome forms' (NPEU/Oxford criteria).[332] The findings of this study indicate a need for caution in this approach. Even with the use of developmental screening tools, the false-negative rates (sensitivities) are unacceptably high. Other factors, such as shortage of health visitors, requirement for further training and lack of universal uptake of the screening programme, might further limit success.

Based on the findings of this study, a centralised approach to the assessment and recording of 2-year outcome data for children who were born very preterm is worthy of consideration. Typically, very preterm children are offered post-discharge appointments every 3 to 6 months; these visits might be conducted at hospitals where allied health professional support (e.g. dietetics, physiotherapy) can be sought if necessary. There may be advantages for the 2-year neurodevelopmental assessment to be organised at neonatal network level, as this could ensure that each child receives assessment by an appropriately trained team of health professionals using standardised tools, and could benefit from centralised, co-ordinated administrative support to trace and contact families.

In addition to being of high quality, the data recorded during clinical care should be complete to enable meaningful analysis. Currently, the utility of the routinely recorded electronic clinical data as a source of population-based outcome information is limited by poor data completeness. According to the National Neonatal Audit Programme report, 2-year outcome data were available from only 44% of all infants born before 30 weeks' gestation in England and Wales between July 2010 and June 2011. Strategies to reduce missing data need to be aimed at clinician engagement.

Since 2007, the American Academy of Pediatrics (AAP) has recommended ASD-specific screening at 18 months for all children to facilitate early diagnosis and to prevent delay in the initiation of early intervention.[333] The UK National Screening Committee does not currently recommend universal screening, on the basis that none of the available screening tools has sufficient reliability in identifying children at risk for ASD when applied to the general population.[334] Regardless of an ASD diagnosis, toddlers who were born preterm experience problems in current functioning that may interfere with adaptive exploration and social engagement. It is important that clinicians recognise these difficulties and the impact that they have on families. Parents require information on the social communication difficulties that preterm children experience, particularly as some of these behaviours may be amenable to specific interventions, such as speech and language, occupation and sensory integration therapies, as well as educational programmes targeted at enhancing communication, social skills instruction and reducing interfering maladaptive behaviours.[335]

Research recommendations

Improve the electronic '2-year outcome' form

In the absence of the standardised use of a single assessment tool to allow comparison of outcomes between centres, improvements to routine data recording could be sought. On the electronic '2-year outcome' form, the documentation of outcomes in the cognitive and language domains is more subjective than in the motor domains. It is possible that, by modifying the form to increase the objectivity of the items recorded, the validity of the data would be improved. For example, for the cognitive domain, it may be possible to identify a standardised set of cognitive test items, perhaps from the Bayley-III assessment or other tools that can be easily administered in a clinical setting. Language function can be ascertained by determining if a child can identify or say words from a list of commonly expressed words.

The NICE guidelines[331] for developmental follow-up of children and young people born preterm recommend the use of the validated parent questionnaire PARCA-R to identify children at risk of global developmental delay, learning disability or language problems, and for the PARCA-R scores to be documented in the NNRD. The predictive validity of the PARCA-R at the age of 2 years in identifying impairments at a later age and special educational needs to be evaluated. The guidelines also recommend using different approaches, such as e-mails or text messages, to provide enhanced developmental support. The utility of these approaches in assessment and outcome data acquisition should also be explored.

Comprehensive behavioural assessment and identification of risk factors for ASD in the preterm population

As stated in *Types of neurodevelopmental outcome measures*, preterm children are at a higher risk of a range of behavioural problems, including ADHD and internalising behaviour, that were not examined in this study. The age at emergence of these behavioural difficulties is unclear and should be examined in future studies. Future studies would also benefit from an examination of a more comprehensive set of neonatal and environmental variables in order to identify potential moderators and mediators of risk for ASD and other behavioural difficulties. It would then be possible to develop risk scores or risk prediction models that could aid in early diagnosis and initiation of interventional therapies. Future studies will also need to focus on the challenges faced in the early assessment of behavioural features of children with major functional disabilities and of children in non-English-speaking groups.

Linkage with school-age outcome data

Currently, there is no process or provision in the UK for continuing formal follow-up assessment beyond early childhood. Long-term programmes require significant manpower and financial investment and the likelihood of high attrition rates will further jeopardise success. Therefore, it is worthwhile considering other sources of school-age outcome data, for example primary care or community child health records, or educational data. UK national structures provide a unique opportunity for data linkage. Future research could investigate the utility of these data sources through linkage of neonatal data with later outcomes.

Chapter 6 Using the National Neonatal Research Database to inform economic evaluations of neonatal interventions

Abstract

Background: Computerised record linkage with EPRs is increasingly considered a means of obtaining primary or complementary resource use data for the purposes of health economic evaluation and more broadly for health technology assessments. We addressed whether or not reliable trial-based economic evaluations can be conducted utilising data from the NNRD.

Methods: The Probiotic in Preterm babies Study (PiPS) (a multicentre, double-blind, placebo-controlled, randomised trial in infants born between 23^{+0} and 30^{+6} weeks gestational age) was used as the test bed. Health-care resource utilisation data were extracted from the PiPS trial case report forms (CRFs), the NNRD and a combined data source, and were primarily valued using national tariffs for 2012–13.[336] Differences in economic outcomes were estimated (1) within trial by data source, thereby allowing us to draw comparisons between the probiotic and the placebo, and (2) by pooling data between trial arms, thereby allowing comparisons between the alternative data sources.

Results: Within-trial comparisons of resource use and costs revealed no statistically significant differences between the trial comparators for any resource input or cost category, regardless of data source. Across-trial tests of concordance in resource use and costs between comparator data sources revealed high levels of agreement for the majority of categories of resource use or cost and the total cost of neonatal care. Comparisons of cost-effectiveness outcomes between data sources revealed low probabilities of miscoverage of incremental net monetary benefit between the alternative data sources when the NNRD acted as the sole source of information.

Conclusions: This empirical investigation demonstrates proof of principle for the potential of the NNRD as a data source for neonatal trial-based economic evaluations in the UK. This has potential to reduce costs and improve the efficiency of economic evaluations. Research assessing the utility of the NNRD across a wider range of trial-based economic evaluations and alternative study designs are logical and are the important next steps.

Background

Economic evaluation involves the comparative analysis of alternative programmes or interventions in terms of their costs and consequences.[337] In order to estimate the total cost for an individual patient included in single study-based economic evaluations, such as trial-based economic evaluations, the quantity of each resource item they use is multiplied by the unit cost of that item and the product calculated. The resources used by patients, such as hospital admissions, consultations and types and quantities of drugs administered, are normally recorded for each patient over the time horizon of the study. The categories of resource use that are included in the study are determined by the perspective of the analysis. The main alternatives are to confine the perspective to the health-care system (sometimes referred to as the 'payer') or to include broader societal costs. The former perspective typically covers direct medical care, comprising the intervention being evaluated, treatment of any side effects or complications of treatment, and follow-up care. It may also include medical care not directly associated with the underlying condition, although regression modelling may be required at the analytical stage to disentangle background 'noise' that often occurs when this is

included.[338] The societal perspective also considers care provided by other sectors of the economy, costs incurred by patients, informal care provided by family and friends, and productivity losses from morbidity and premature death. Methodological guidelines for economic evaluation differ in their recommended perspective for the analysis. As a minimum, it is recommended that analysts adopt a health system perspective, which is currently considered to include the NHS and Personal Social Services in England and Wales.[339]

In single study-based economic evaluations, such as trial-based economic evaluations, many resources used can normally be recorded on study CRFs with little or no additional burden, but sometimes additional information will be required from medical records, patient questionnaires and diaries, and other sources.[340] A recent trial-based economic evaluation of neonatal extracorporeal membrane oxygenation necessitated observational research to estimate resource use associated with complications, and parent-completed questionnaires to document post-neonatal discharge hospital and community health service use.[341] Increasingly, however, computerised record linkage with data from EPRs is being considered as a means of obtaining primary or complementary resource use data for the purposes of health economic evaluation and more broadly for the purposes of health technology assessment. In principle, the successful development of systems for extracting resource utilisation data from EPRs should reduce the complexity, time and cost of conduct of trial-based economic evaluations, and offer considerable additional utility for NHS commissioning and service management.

Aims

We aimed to assess whether or not reliable health service utilisation data can be obtained from the NNRD and if these data can be used to inform future trial-based economic evaluations of neonatal interventions.

Methods

Overview

The study population for this empirical investigation comprised infant participants in the Probiotic in Preterm babies Study (PiPS). For each study infant, health-care resource utilisation was measured using three primary data sources: (1) the PiPS trial CRFs, (2) the NNRD; and (3) a data source that combined information from both PiPS trial CRFs and the NNRD.

Resource inputs captured by each data source were primarily valued using national tariffs and expressed in Great British pounds (GBP) (2012/13 prices). In our empirical investigation we sought to estimate (1) the level of agreement for hospital resource utilisation and costs between the alternative data sources and (2) the level of precision of incremental cost-effectiveness for the probiotic evaluated in PiPS using alternative data sources.

PiPS trial: design

PiPS was a multicentre, double-blind, placebo-controlled, randomised trial of probiotic administration in infants born between 23^{+0} and 30^{+6} weeks gestational age. Infants were recruited within 48 hours of birth from 24 hospitals within 60 miles of London over a 37-month period from July 2010 onwards. They were randomised to either the probiotic (given in a daily oral dose of 8.3–8.8 $\log_{10}$ colony-forming units) or the placebo (provided as an identical powder in identical sachets, until 36 weeks postmenstrual age or discharge from hospital, if sooner). There were three primary outcomes: any episode of neonatal NEC Bell stage II or III;[112] any positive blood culture of an organism not recognised as a skin commensal on a sample drawn > 72 hours after birth and < 46 weeks postmenstrual age or discharge if sooner (hereafter sepsis for brevity); and death before discharge from hospital. Secondary outcomes included a composite of the three primary outcomes. The trial was sized ($n = 1300$) to detect a 40% relative risk reduction from 15% to 9.1% for each of the primary outcomes at a two-sided significance level of 5% and with

90% power. PiPS was approved by a national research ethics committee and co-ordinated by the NPEU, University of Oxford. Further details about PiPS, sampling procedures, methodology, outcome measures and responses rates are reported in full elsewhere.[342]

PiPS trial: measurement of resource use and costs

A comprehensive profile of resource inputs was integrated at the outset into the PiPS trial CRFs. There were four main trial CRFs: (1) form 1: entry, (2) form 2: daily data collection, (3) form 3: transfer/discharge and (4) form 4: abdominal pathology. The bulk of the relevant resource inputs were captured by the second and third of these trial CRFs. The forms captured a comprehensive profile of resource use by each infant, encompassing length of stay by intensity of care, surgeries, investigations, procedures, transfers and post-mortem examinations until final hospital discharge or death (whichever was earliest). Resource inputs were primarily valued based on data collated from secondary national tariff sets[343,344] (*Table 31*). All costs were expressed in GBP and reflected values for the financial year 2012/13.

TABLE 31 Unit costs for resource use variables (£, 2012/13 prices)

Resource use variable	Unit cost	Source	Notes
Resource use variables in the PiPS data set			
Vaginal birth – cephalic	1337.31	*NHS Reference Costs 2012–13*[343]	
Vaginal birth – breech	2488.33	*NHS Reference Costs 2012–13*[343]	
Vaginal birth – other presentation	1958.75	*NHS Reference Costs 2012–13*[343]	
Caesarean section before onset of labour	2950.40	*NHS Reference Costs 2012–13*[343]	
Caesarean section after onset of labour	3690.41	*NHS Reference Costs 2012–13*[343]	
Coroner/hospital	649.66	Birthplace report[345]	Inflated to 2012/13 prices
Neonatal critical care transportation	1370.37	*NHS Reference Costs 2012–13*[343]	HRG code XA06Z
Cranial ultrasound scan	53.84	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes RA23Z and RA24Z
ROP	994.97	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA64A, PA64B and PA64C
ROP screen	134.11	*Unit Costs of Health and Social Care 2012*[344]	Assumed nurse input (20 minutes valued at £100 per hour) and consultant input (30 minutes valued at £201.55 per hour)
PDA	2422.50	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA23A and PA23B
Repair of inguinal hernia (weighted average cost)	1250.17	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA25A, PA25B, PA26A and PA26B
Insertion of ventricular reservoir	2922.72	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes AA15C, AA15D and AA15E
NEC treatment, peritoneal drainage/laparotomy no enterostomy/laparotomy with enterostomy	2458.01	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA25A and PA25B

continued

TABLE 31 Unit costs for resource use variables (£, 2012/13 prices) (*continued*)

Resource use variable	Unit cost	Source	Notes
Resource use variables in the NNRD data set			
Emergency caesarean section – not in labour	2784.30	*NHS Reference Costs 2012–13*[343]	
Emergency caesarean section – in labour	3269.02	*NHS Reference Costs 2012–13*[343]	
Elective section – not in labour	2784.30	*NHS Reference Costs 2012–13*[343]	
Elective section – in labour	3269.02	*NHS Reference Costs 2012–13*[343]	
Vaginal – forceps assisted	2248.64	*NHS Reference Costs 2012–13*[343]	
Vaginal – spontaneous	1337.31	*NHS Reference Costs 2012–13*[343]	
Vaginal – ventouse assisted	2248.64	*NHS Reference Costs 2012–13*[343]	
Coroner/hospital	649.66	Birthplace report[345]	
Neonatal critical care transportation	1370.37	*NHS Reference Costs 2012–13*[343]	
Cranial ultrasound scan	53.84	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes RA23Z and RA24Z
ROP	994.97	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA64A, PA64B and PA64C
ROP screen	134.11	*Unit Costs of Health and Social Care 2012*[344]	Assumed nurse input (20 minutes valued at £100 per hour) and consultant input (30 minutes valued at £201.55 per hour)
PDA	2422.50	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA23A and PA23B
Insertion of ventriculoperitoneal shunt	7563.54	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes AA15C, AA15D and AA15E
Inguinal herniotomy (bilateral)	1250.17	*NHS Reference Costs 2012–13*[343]	Weighted average of HRG codes PA25A, PA25B, PA26A and PA26B

The total length of stay (total inpatient hospital days) was computed as the total number of hospital days until final discharge to home or death. Information was available on time spent in the neonatal unit by level of care (normal, transitional, special, high dependency or intensive). The cost of routine neonatal care was calculated for each infant by multiplying the length of stay by intensity (intensive, high dependency, special care, transitional) by the per diem cost of the respective level of care using data from the NHS Reference Costs trusts schedule 2012/13.[343] Non-routine investigations excluded from these per diem costs were valued using a combination of primary and secondary costs. The costs of surgeries were calculated by assignment of surgical procedures to relevant Healthcare Resource Group (HRG) codes and application of unit costs from national tariffs. Transfers were recorded whenever an infant was transported between specialist hospitals for neonatal critical care, and were valued using costs from the NHS Reference Costs trusts schedule 2010/11,[336] and inflated using a health care specific pay and prices index to 2012/13 prices. Post-mortem costs were based on data from secondary sources.[345] Where costs of additional non-routine investigations excluded from per diem values for neonatal care were not available from national tariffs, clinicians were asked to identify the staff and material inputs required for these investigations. Staff time was valued using the *Unit Costs of Health and Social Care 2012*[344] tariffs.

Linkage and data extraction from the National Neonatal Research Database

In order to compare resource use, cost and cost-effectiveness estimates based on data solely extracted from the PiPS trial CRFs and data solely extracted from the NNRD, a NNRD extract was created for infants participating in the PiPS trial. The NNRD has been created through the collaborative efforts of neonatal services across the country to be a national resource. The NNRD contains a defined set of data items (the Neonatal Data Set) that have been extracted from the Badger.net neonatal EPR of all admissions to NHS neonatal units. Badger.net is managed by Clevermed Ltd, an authorised NHS hosting company. The Neonatal Data Set is an approved NHS Information Standard (SCCI1575). Contributing neonatal units are known as the UK Neonatal Collaborative.

The trial co-ordinating centre for the PiPS trial, namely the NPEU at the University of Oxford, provided the NDAU at Imperial College with the final PiPS data set. This included data on 1310 of a total of 1315 infants recruited into the PiPS trial (five infants were excluded because of withdrawals from the study). Clevermed Ltd, the NHS hosting company, which separately receives data from individual neonatal units, was able to match Badger IDs to NHS numbers for 1280 (98%) of the 1310 infants (*Figure 25*). These 1280 infants had 2360 episodes of care which were linked to episodes in the NNRD using the Badger ID and hospital name. Episodes were renumbered on both databases as necessary to provide linkage. A total of 81 episodes that were effectively the second episode of care on the NNRD were renumbered as the first episode to match the PiPS episodes, because the infants were recruited into the PiPS trial after transfer from the neonatal unit at the hospital of birth. Similarly, 103 episodes of care that existed on the PiPS database but not on the NNRD were excluded from comparison; these largely occurred in non-Badger hospitals or wards. This resulted in the exclusion of all episodes of care from 22 infants, leaving 2257 episodes of care from 1258 infants eligible for our comparative analyses (see *Figure 25*).

A comprehensive profile of resource use between randomisation into the PiPS trial until final hospital discharge or death (whichever was earliest) was compiled from the NNRD for the 1258 infants eligible for our comparative analyses. Our direct comparisons of resource use estimates between the PiPS and NNRD data sources required a further process of data manipulation to reconcile definitional and labelling differences for individual variables between the data sources, and coding differences by episodic and infant level. For example, information on surgery for PDA, medical treatment of PDA using ibuprofen or indometacin, and ROP treatment whether by laser or cryotherapy, can appear in multiple locations in the NNRD (e.g. discharge diagnoses, daily data, ad hoc forms). An overall summary of the interventions received during linkable episodes of care was generated on a 'by infant' rather than episodic level.

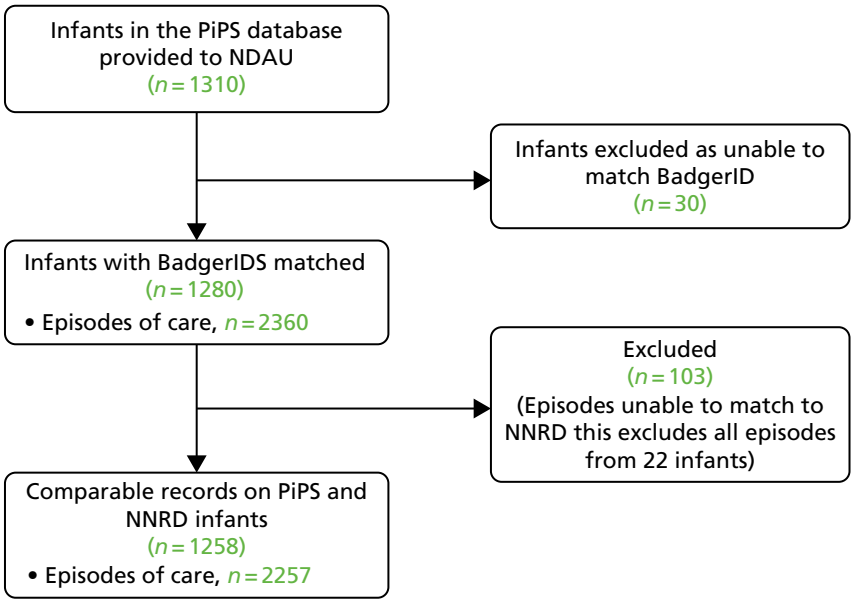


FIGURE 25 Linkage between PiPS trial data and the NNRD.

Statistical methods

A comprehensive statistical analysis plan was followed. All statistical analyses were performed using Stata or R (version 2.01).

The clinical and sociodemographic characteristics of the PiPS participants who were ($n = 1258$) and were not ($n = 52$) included in our comparative analyses of resource use, costs and cost-effectiveness were compared using the chi-squared test. Differences in resource use and costs, by category, were estimated (1) within trial by data source, thereby allowing us to draw comparisons between the probiotic and placebo arms, and (2) by pooling data between the trial arms, thereby allowing us to draw comparisons between the alternative data sources. In addition to data solely extracted from the PiPS trial CRFs and data solely extracted from the NNRD, we created a third data source for these comparative analyses. The third data source, hereafter termed the 'combined' data source for brevity, was constructed by selecting the preferred data source for each resource variable in terms of volume and granularity of information provided. It broadly followed the processes described in *Chapter 4*. The selection process for each resource variable was undertaken by the clinical investigators (KC, CB). For comparisons within trial by data source, differences in resource use and costs, by resource category, were tested using the independent-sample *t*-test for continuous variables, the chi-squared test for categorical variables and the Mann–Whitney *U*-test for medians. For comparisons between the alternative data sources, the levels of agreement in resource use and cost estimates, by category, for alternative combinations of data sources (PiPS vs. NNRD, PiPS vs. combined, NNRD vs. combined) were estimated using the Lin concordance correlation coefficient.[346] This statistic measures the agreement between two continuous variables obtained by two methods; the value of the statistic lies between 1 (perfect agreement) and –1 (perfect inverse agreement). A threshold 0.40 value for the statistic was adopted to indicate acceptable clinical or practical significance.[347] In addition, we estimated mean differences and 95% CIs to identify potential systematic biases, and the 95% limits of agreement, indicating random variation between individual measurements.[348]

We additionally performed an economic evaluation of the probiotic. For comparisons within trial by data source, the economic evaluation took the form of an incremental cost-effectiveness analysis in which we estimated the incremental costs (ΔC) and incremental effects (ΔE) attributable to the probiotic in very preterm infants, with reference to the placebo. The results were primarily expressed each in terms of an incremental cost-effectiveness ratio (ICER) ($\Delta C/\Delta E$). Estimates of incremental cost-effectiveness were made for each of the three primary clinical outcomes (any episode of NEC Bell stage II or III, any case of sepsis, death before discharge from hospital), and for the composite secondary outcome. The economic evaluation was conducted from a health system perspective.[339] The time horizon for the economic evaluation was the period between trial randomisation and final hospital discharge or death, whichever was earlier. Non-parametric bootstrapping, involving 1000 bias-corrected replications of each of the ICERs, was used to calculate uncertainty around all cost-effectiveness estimates. This was represented on four quadrant cost-effectiveness planes.[349] Decision uncertainty was addressed by estimating net benefit statistics and constructing cost-effectiveness acceptability curves (CEACs) across cost-effectiveness threshold values (λ) of between £0 and £70,000 for the health outcomes of interest. The probability that the probiotic is less costly or more effective than the placebo was based on the proportion of bootstrap replicates that had negative incremental costs or positive incremental health benefits, respectively. A series of prespecified subgroup analyses repeated all analyses by selected subgroups (sex, birthweight, gestational age, colonisation status, randomisation age) for the primary and secondary cost-effectiveness outcomes.

For comparisons of cost-effectiveness outcomes between the alternative data sources, we estimated the overall probability of miscoverage of incremental net monetary benefit based on resource use data solely extracted from the NNRD and resource use data solely extracted from the PiPS trial CRFs. In order to estimate miscoverage for incremental net monetary benefit, the bootstrap replications of each of the ICERs were rearranged on a linear scale using the formula:

$$\lambda \times \Delta E - \Delta C. \tag{7}$$

The miscoverage statistic was estimated as the percentage of bootstrap samples of incremental net monetary benefit that fell outside the CI for the reference data source.[350] For the purpose of these analyses, the combined data source acted as the referent, although we additionally assumed that the PiPS data source acted as a referent for analyses of data solely extracted from the NNRD. This was replicated for the primary and secondary outcomes of interest and for all prespecified subgroup analyses.

Finally, we performed sensitivity analyses that compared the key outputs of the economic evaluation using either resource use data and clinical outcomes extracted solely from the PiPS data set or resource use data and clinical outcomes extracted solely from the NNRD. These analyses were restricted to the two clinical outcomes used in the PiPS trial that were available in both data sources, namely (1) death before discharge from hospital and (2) sepsis. They acted as exemplars of the likely differences in economic outcomes that will be observed if we rely on the NNRD as a complete source of information (resource inputs, clinical outcomes) for an economic evaluation. The cost-effectiveness outcomes considered by these sensitivity analyses included estimates of incremental cost-effectiveness, probabilities of cost-effectiveness for the probiotic at alternative cost-effectiveness thresholds, and miscoverage of incremental net monetary benefit against the PiPS trial referent.

Results

Study population

A total of 1315 infants were recruited from 24 hospitals within 60 miles of London over 37 months, from July 2010 onwards. Data for 1310 infants were available for analysis in the PIPS trial. Of these 1310 infants, 52 infants were excluded from our empirical investigations of either because failure to match their Badger IDs to their NHS numbers or because they received part of their neonatal care in non-Badger hospitals or wards. A total of 1258 infants were therefore eligible for our comparative analyses. There were no significant differences between the baseline clinical and sociodemographic characteristics of the 1258 infants included in the analyses and the 52 infants excluded from the analyses, regardless of the use of NNRD or PiPS data (*Table 32*).

TABLE 32 Baseline clinical and sociodemographic characteristics of study participants

	Infants		
	Included		
Variable	PiPS data ($N = 1258$)	NNRD data ($N = 1258$)	Excluded[a] ($N = 52$)
Gestational age (weeks), *n* (%)			
< 28	602 (47.8)	599 (47.6)	32 (61.5)
≥ 28	656 (52.1)	658 (52.3)	20 (38.5)
Birthweight (g)			
Mean (SD)	1043 (315.9)	1042 (314.2)	993.6 (268.6)
Birthweight of ≤ 1000 g, *n* (%)	613 (48.7)	613 (48.9)	31 (59.6)
Birthweight of > 1000 g, *n* (%)	645 (51.3)	643 (51.1)	21 (40.4)
Sex, *n* (%)			
Boys	709 (56.4)	707 (56.2)	34 (67.3)
Girls	549 (43.6)	549 (43.6)	17 (33.7)
Unknown	0 (0)	1 (0.08)	0 (0)

continued

TABLE 32 Baseline clinical and sociodemographic characteristics of study participants (*continued*)

	Infants		
	Included		
Variable	PiPS data ($N = 1258$)	NNRD data ($N = 1258$)	Excluded[a] ($N = 52$)
Multiplicity, *n* (%)			
Singleton	879 (69.8)	885 (70.4)	37 (71.2)
Multiple	379 (30.2)	372 (29.6)	15 (28.8)
Apgar score at 5 minutes, *n* (%)			
0–3	38 (3.0)	38 (3.0)	2 (3.8)
4–6	177 (14.1)	169 (13.4)	5 (9.6)
7–10	1010 (80.3)	987 (78.5)	45 (86.5)
Missing	33 (2.6)	64 (5.1)	0 (0.0)
Maternal age, years			
Mean (SD)	33.8 (12.5)	33.4 (6.6)	33.7 (7.8)
Maternal ethnicity, *n* (%)			
White	707 (56.2)	684 (54.4)	29 (55.7)
Indian	58 (4.6)	28 (2.2)	3 (5.8)
Pakistani	36 (2.9)	54 (4.3)	1 (1.9)
Bangladeshi	57 (4.5)	189 (15.0)	5 (9.6)
Black African	188 (15.0)	71 (5.6)	8 (15.4)
Black Caribbean	62 (4.9)	168 (13.3)	1 (1.9)
Other	140 (11.1)	63 (5.0)	5 (9.6)
Unknown	10 (0.8)	1 (0.08)	0
Membranes ruptured > 24 hours before birth, *n* (%)			
Yes	345 (27.4)	1011 (80.4)	13 (25.0)
No	877 (69.7)	NA	37 (71.2)
Unknown	36 (2.9)	247 (19.6)	2 (3.8)
Maternal antenatal corticosteroid treatment, *n* (%)			
Any	816 (64.9)	1147 991.2)	36 (69.2)
Started < 24 hours before birth	322 (25.6)	NA	13 (25.0)
None	111 (8.8)	110 (8.7)	3 (5.8)
Unknown	9 (0.7)	1 (0.08)	0 (0.0)
Delivery by caesarean section, *n* (%)			
Yes	664 (52.9)	652 (51.8)	26 (50.0)
No	593 (47.1)	550 (43.7)	26 (50.0)
Unknown	1 (0.08)	56 (4.5)	0 (0.0)
Born in the recruiting hospital, *n* (%)			
Yes	1146 (91.1)	1168 (92.8)	46 (88.5)
No	111 (8.8)	11 (0.9)	6 (11.5)
Missing	1 (0.08)	79 (62.8)	0 (0.0)

NA, not applicable.
a Characteristics of excluded infants determined using PiPS data.

The key clinical and sociodemographic characteristics of the 1258 infants included in our empirical investigations are presented by trial arm in *Table 32*. There were no significant differences in these key characteristics between the trial arms. Furthermore, there was no evidence of clinical benefit associated with administration of the probiotic for any of the primary outcomes or the composite secondary outcome (*Table 33*).

Resource use and cost estimates: comparisons within trial by data source

Resource use measures and their values between trial randomisation and final hospital discharge (or death) are summarised by trial arm in *Tables 34* and *35* for the PiPS and NNRD data, respectively. Based on the PiPS data (see *Table 33*), the mean (SE) overall duration of hospitalisation was 75.49 (1.95) days for infants in the control arm, whereas the infants in the probiotic arm had a mean overall duration of hospitalisation of 76.60 (2.02) days. There were no statistically significant differences in the values for any resource input by trial arm. The results for the NNRD data (see *Table 35*) followed a similar pattern with no statistically significant differences in resource values by trial arm. However, the mean numbers of cranial ultrasound scans were higher in the PiPS data than in the NNRD data. *Table 36* presents the resource use measures and their values by trial arm for the combined data set.

Costs

Cost measures and their values between trial randomisation and final hospital discharge (or death) are summarised by trial arm in *Tables 37* and *38* for the PiPS and NNRD data, respectively. Based on the PiPS data (see *Table 37*), the mean (SE) total costs were estimated at £62,284 (£1876) for the control group compared with £62,799 (£1817) for the probiotic group. There were no significant differences across the cost categories by trial arm. The results for the NNRD data (see *Table 38*) followed a similar pattern with no statistically significant differences between the trial arms in cost estimates, overall and by cost category.

TABLE 33 Clinical and sociodemographic characteristics of study participants by trial arm (PiPS data)

Characteristic	Trial arm, *n* (%)		*p*-value[a]
	Placebo (*N* = 638)	*B. breve* BBG (*N* = 620)	
Male	357 (55.96)	352 (56.77)	0.770
Gestational age (weeks)			0.883
< 28	304 (47.65)	298 (48.06)	
≥ 28	334 (52.35)	322 (52.94)	
Randomisation age			0.881
< 24 hours	167 (26.18)	160 (25.81)	
≥ 24 hours	471 (73.82)	460 (74.19)	
Weight ≤ 100 g			0.990
No	327 (51.25)	318 (51.29)	
Yes	311 (48.71)	302 (48.71)	
Primary outcomes			
Death before discharge home[b]	54 (8.46)	51 (8.23)	0.879
Sepsis[c]	72 (11.29)	67 (10.81)	0.787
NEC	63 (9.87)	56 (9.03)	0.610
Secondary outcome			
Composite of primary outcomes	139 (21.79)	133 (21.45)	0.885

a Comparisons of placebo vs. *B. breve* BBG groups carried out using Student's *t*-tests for continuous variables and χ^2 test for categorical variables.
b Includes three infants who remained on paediatric wards and are analysed as survivors.
c Sepsis is defined as bloodstream infection with non-skin commensals after 72 hours postnatal age and before 46 weeks postmenstrual age.

TABLE 34 Resource use by trial arm (PiPS data)

Resource variable	Trial arm		*p*-value[a]
	Placebo (*N* = 638)	*B. breve* BBG (*N* = 620)	
Mode of delivery,* n *(%)			***0.371***
Vaginal birth – cephalic	229 (35.89)	236 (38.06)	
Vaginal birth – breech	64 (10.03)	53 (8.55)	
Vaginal birth – other presentation	6 (0.94)	5 (0.81)	
Caesarean section before onset of labour	197 (30.88)	212 (34.19)	
Caesarean section after onset of labour	141 (22.1)	114 (18.39)	
Unknown	1 (0.16)	0 (–)	
Other			
Post-mortems, *n* (%)	10 (1.57)	10 (1.61)	0.9490
Post-mortems, mean (SE)	0.02 (0)	0.02 (0.01)	0.9486
Hospital transfers, *n* (%)	310 (48.59)	312 (50.32)	0.5390
Hospital transfers, mean (SE)	0.77 (0.04)	0.80 (0.05)	0.6648
Special care, *n* (%)	580 (90.91)	565 (91.13)	0.8920
Length of special care stay (days), median (range)	30 (0–203)	30 (0–392)	0.6979
Length of special care stay (days), mean (SE)	396.83 (68.15)	416.28 (72.43)	0.7849
High-dependency care, *n* (%)	570 (89.34)	562 (90.65)	0.4410
Length of high-dependency care stay (days), median (range)	17 (0–174)	20 (0–195)	0.1202
Length of high-dependency care stay (days), mean (SE)	22.34 (0.94)	24 (0.99)	0.2257
Intensive care, *n* (%)	604 (94.67)	595 (95.97)	0.2770
Length of intensive care stay (days), median (range)	11.5 (0–339)	10 (0–378)	0.9338
Length of intensive care stay (days), mean (SD)	22.20 (1.21)	21.30 (1.07)	0.5778
Total length of stay care stay (days), median (range)	64 (2–378)	67 (2–547)	0.3804
Total length of stay (days), mean (SE)	75.49 (1.95)	76.60 (2.02)	0.6914
ROP screens, *n* (%)	583 (91.38)	571 (92.10)	0.6440
ROP screens, mean (SE)	0.91 (0.01)	0.92 (0.01)	0.6442
ROP treatment, *n* (%)	9 (1.41)	13 (2.10)	0.3530
ROP treatment, mean (SE)	0.01 (0.005)	0.02 (0.006)	0.3551
Cranial ultrasound scans, *n* (%)	637 (99.84)	619 (99.84)	0.9840
Cranial ultrasound scans, mean (SE)	1.61 (0.04)	1.63 (0.03)	0.7718
PDA surgery, *n* (%)	26 (4.08)	33 (5.32)	0.2950
PDA surgery, mean (SE)	0.04 (0.01)	0.05 (0.01)	0.5574
Hernia surgery, *n* (%)	19 (2.98)	18 (2.90)	0.9370
Hernia surgery, mean (SE)	0.03 (0.01)	0.03 (0.01)	0.9306
Reservoir surgery, *n* (%)	0	1 (0.16)	0.3100
Reservoir surgery, mean (SE)	0 (0)	0 (0)	0.3177
VP shunt surgery, *n* (%)	5 (0.78)	4 (0.65)	0.7710
VP shunt surgery, mean (SE)	0.01 (0)	0.01 (0)	0.7706
NEC surgery, *n* (%)	39 (6.11)	34 (5.48)	0.6330
NEC surgery, mean (SE)	0.06 (0.001)	0.05 (0.001)	0.6334

TABLE 34 Resource use by trial arm (PiPS data) (*continued*)

Resource variable	Trial arm		p-value[a]
	Placebo (*N* = 638)	*B. breve* BBG (*N* = 620)	
Other procedures, *n* (%)	16 (2.51)	11 (1.77)	0.3690
Other procedures, mean (SE)	0.03 (0.01)	0.02 (0.01)	0.2875

VP, ventriculoperitoneal.
a Comparisons of placebo vs. *B. breve* BBG groups carried out using Student's *t*-tests for continuous variables and χ^2 test for categorical variables. The Mann–Whitney *U*-test was used to compare medians.

TABLE 35 Resource use by trial arm (NNRD data)

Resource variable	Trial arm		p-value[a]
	Placebo (*N* = 638)	*B. breve* BBG (*N* = 620)	
Mode of delivery,* n *(%)			***0.7000***
Vaginal – spontaneous	281 (44.04)	278 (44.84)	
Vaginal – forceps assisted	11 (1.72)	12 (1.94)	
Elective section – in labour	5 (0.78)	3 (0.48)	
Elective section – not in labour	33 (5.17)	32 (5.16)	
Emergency caesarean section – in labour	139 (21.79)	112 (18.06)	
Emergency caesarean section – not in labour	158 (24.76)	170 (27.42)	
Unknown	11 (1.72)	13 (2.1)	
Other			
Post-mortems, *n* (%)	4 (0.63)	2 (0.32)	0.4330
Post-mortems, mean (SE)	0.01 (0)	0 (0)	0.4317
Hospital transfers, *n* (%)	323 (50.63)	323 (52.10)	0.6020
Hospital transfers, mean (SE)	0.89 (0.05)	0.92 (0.05)	0.6858
Normal care, *n* (%)	185 (29.0)	166 (26.77)	0.3800
Length of normal care stay (days), median (range)	0 (0 to 9)	0 (0 to 11)	0.4195
Length of normal care stay (days), mean (SE)	0.55 (0.04)	0.55 (0.05)	0.9407
Transitional care, *n* (%)	13 (2.04)	17 (2.74)	0.6701
Length of transitional care stay (days), median (range)	0 (0 to 4)	0 (0 to 3)	0.4134
Length of transitional care stay (days), mean (SE)	0.03 (0.01)	0.04 (0.01)	0.5760
Special care, *n* (%)	589 (92.32)	562 (90.65)	0.2870
Length of special care stay (days), median (range)	31 (0–105)	30.5 (0–87)	0.7810
Length of special care stay (days), mean (SE)	31 (0.7)	30.77 (0.72)	0.8192
High-dependency, *n* (%)	583 (91.38)	574 (92.58)	0.4330
Length of high-dependency care stay (days), median (range)	16 (0–256)	16 (0–175)	0.6736
Length of high-dependency care stay (days), mean (SE)	24.45 (1.15)	24.78 (1.09)	0.8322
Intensive care, *n* (%)	574 (89.97)	573 (92.42)	0.1250
Length of intensive care stay (days), median (range)	11 (0–267)	12 (0–166)	0.5336
Length of intensive care stay (days), mean (SD)	19.11 (1.04)	18.59 (0.84)	0.6980
Total length of stay care stay (days), median (range)	65 (2–337)	68 (2–277)	0.4697
Total length of stay (days), mean (SE)	75.14 (1.87)	74.73 (1.67)	0.8697

continued

TABLE 35 Resource use by trial arm (NNRD data) (*continued*)

Resource variable	Trial arm		
	Placebo ($N = 638$)	*B. breve* BBG ($N = 620$)	*p*-value[a]
ROP screens, *n* (%)	543 (85.11)	530 (85.48)	0.8510
ROP screens, mean (SE)	0.85 (0.01)	0.85 (0.01)	0.8515
ROP treatment, *n* (%)	26 (4.08)	23 (3.71)	0.7380
ROP treatment, mean (SE)	0.04 (0.001)	0.04 (0.001)	0.7377
Cranial ultrasound scans, *n* (%)	415 (65.05)	393 (63.39)	0.5390
Cranial ultrasound scans, mean (SE)	0.8 (0.03)	0.80 (0.03)	0.9876
PDA surgery, *n* (%)	22 (3.45)	29 (4.68)	0.2690
PDA surgery, mean (SE)	0.03 (0.001)	0.05 (0.001)	0.2705
Hernia surgery, *n* (%)	41 (6.43)	40 (6.45)	0.9850
Hernia surgery, mean (SE)	0.06 (0.01)	0.06 (0.01)	0.9854
Reservoir surgery, *n* (%)	0	0	–
Reservoir surgery, mean (SE)	0 (0)	0 (0)	–
VP shunt surgery, *n* (%)	5 (0.78)	7 (1.13)	0.5290
VP shunt surgery, mean (SE)	0.01 (0.003)	0.01 (0.004)	0.5301
NEC surgery, *n* (%)	34 (5.33)	26 (4.19)	0.3450
NEC surgery, mean (SE)	0.05 (0.01)	0.04 (0.01)	0.3443
Other procedures, *n* (%)	134 (21.0)	108 (17.42)	0.1070
Other procedures, mean (SE)	0.38 (0.04)	0.40 (0.04)	0.6858

VP, ventriculoperitoneal.
a Comparisons of placebo vs. *B. breve* BBG groups carried out using Student's t-tests for continuous variables and χ^2 test for categorical variables. The Mann–Whitney *U*-test was used to compare medians.

TABLE 36 Resource use by trial arm (combined data)

Resource variable	Source	Trial arm		
		Placebo ($N = 638$)	*B. breve* BBG ($N = 620$)	*p*-value[a]
Mode of delivery,* n *(%)				***0.7000***
Vaginal – spontaneous	NNRD	281 (44.04)	278 (44.84)	
Vaginal – forceps assisted	NNRD	11 (1.72)	12 (1.94)	
Elective section – in labour	NNRD	5 (0.78)	3 (0.48)	
Elective section – not in labour	NNRD	33 (5.17)	32 (5.16)	
Emergency caesarean section – in labour	NNRD	139 (21.79)	112 (18.06)	
Emergency caesarean section – not in labour	NNRD	158 (24.76)	170 (27.42)	
Unknown	NNRD	11 (1.72)	13 (2.1)	
Other				
Post-mortems, *n* (%)	PiPS	10 (1.57)	10 (1.61)	0.9490
Post-mortems, mean (SE)	PiPS	0.02 (0)	0.02 (0.01)	0.9486
Hospital transfers, *n* (%)	PiPS	310 (48.59)	312 (50.32)	0.5390
Hospital transfers, mean (SE)	PiPS	0.77 (0.04)	0.80 (0.05)	0.6648

TABLE 36 Resource use by trial arm (combined data) (*continued*)

Resource variable	Source	Trial arm: Placebo (*N* = 638)	Trial arm: *B. breve* BBG (*N* = 620)	*p*-value[a]
Normal care, *n* (%)	NNRD	185 (29.0)	166 (26.77)	0.3800
Length of normal care stay (days), median (range)	NNRD	0 (0–9)	0 (0–11)	0.4195
Length of normal care stay (days), mean (SE)	NNRD	0.55 (0.04)	0.55 (0.05)	0.9407
Transitional care, *n* (%)	NNRD	13 (2.04)	17 (2.74)	0.6701
Length of transitional care stay (days), median (range)	NNRD	0 (0–4)	0 (0–3)	0.4134
Length of transitional care stay (days), mean (SE)	NNRD	0.03 (0.01)	0.04 (0.01)	0.5760
Special care, *n* (%)	NNRD	589 (92.32)	562 (90.65)	0.2870
Length of special care stay (days), median (range)	NNRD	31 (0–105)	30.5 (0–87)	0.7810
Length of special care stay (days), mean (SE)	NNRD	31 (0.7)	30.77 (0.72)	0.8192
High-dependency care, *n* (%)	NNRD	583 (91.38)	574 (92.58)	0.4330
Length of high-dependency care stay (days), median (range)	NNRD	16 (0–256)	16 (0–175)	0.6736
Length of high-dependency care stay (days), mean (SE)	NNRD	24.45 (1.15)	24.78 (1.09)	0.8322
Intensive care, *n* (%)	NNRD	574 (89.97)	573 (92.42)	0.1250
Length of intensive care stay (days), median (range)	NNRD	11 (0–267)	12 (0–166)	0.5336
Length of intensive care stay (days), mean (SD)	NNRD	19.11 (1.04)	18.59 (0.84)	0.6980
Total length of stay care stay (days), median (range)	NNRD	65 (2–337)	68 (2–277)	0.4697
Total length of stay (days), mean (SE)	NNRD	75.14 (1.87)	74.73 (1.67)	0.8697
ROP screens, *n* (%)	PiPS	583 (91.38)	571 (92.10)	0.6440
ROP screens, mean (SE)	PiPS	0.91 (0.01)	0.92 (0.01)	0.6442
ROP treatment, *n* (%)	PiPS	9 (1.41)	13 (2.10)	0.3530
ROP treatment, mean (SE)	PiPS	0.01 (0.005)	0.02 (0.006)	0.3551
Cranial ultrasound scan, *n* (%)	PiPS	637 (99.84)	619 (99.84)	0.9840
Cranial ultrasound scans, mean (SE)	PiPS	1.61 (0.04)	1.63 (0.03)	0.7718
PDA surgery, *n* (%)	PiPS	26 (4.08)	33 (5.32)	0.2950
PDA surgery, mean (SE)	PiPS	0.04 (0.01)	0.05 (0.01)	0.5574
Hernia surgery, *n* (%)	PiPS	19 (2.98)	18 (2.90)	0.9370
Hernia surgery, mean (SE)	PiPS	0.03 (0.01)	0.03 (0.01)	0.9306
Reservoir surgery, *n* (%)	PiPS	0	1 (0.16)	0.3100
Reservoir surgery, mean (SE)	PiPS	0 (0)	0 (0)	0.3177
VP shunt surgery, *n* (%)	NNRD	5 (0.78)	7 (1.13)	0.5290
VP shunt surgery, mean (SE)	NNRD	0.01 (0.003)	0.01 (0.004)	0.5301
NEC surgery, *n* (%)	PiPS	39 (6.11)	34 (5.48)	0.6330
NEC surgery, mean (SE)	PiPS	0.06 (0.001)	0.05 (0.001)	0.6334
Other procedures, *n* (%)	NNRD	134 (21.0)	108 (17.42)	0.1070
Other procedures, mean (SE)	NNRD	0.38 (0.04)	0.40 (0.04)	0.6858

VP, ventriculoperitoneal.

a Comparisons of Placebo vs. *B. breve* BBG groups carried out using Student's *t*-tests for continuous variables and χ^2 test for categorical variables. The Mann–Whitney *U*-test was used to compare medians.

TABLE 37 Hospitalisation costs (£, 2012/13 prices) by trial arm (PiPS data)

Cost variable	Trial arm, mean (SE)		*p*-value[a]
	Placebo (*N* = 638)	*B. breve* BBG (*N* = 620)	
Delivery cost	2474.65 (37.02)	2424.95 (36.87)	0.3417
Post-mortem cost	10.18 (3.2)	10.48 (3.29)	0.9486
Hospital transfers cost	1058.92 (59.88)	1096.3 (62.06)	0.6648
Cost of special care	15,629.21 (387.46)	15,809.23 (533.66)	0.7849
Cost of high-dependency care	17,678.63 (746.1)	18,992.10 (785.94)	0.2257
Cost of intensive care	24,817.40 (1356.07)	23,810.86 (1195.95)	0.5778
ROP screen cost	122.55 (1.49)	123.51 (1.45)	0.6442
ROP treatment cost	7.82 (2.59)	11.62 (3.19)	0.3551
Ultrasound scan costs	86.84 (1.98)	87.62 (1.84)	0.7718
PDA surgery cost	98.70 (18.97)	128.91 (21.85)	0.2967
Hernia surgery cost	37.23 (8.42)	38.31 (9.12)	0.9306
Reservoir surgery cost	0 (0)	4.71 (4.71)	0.3177
VP shunt surgery cost	59.28 (26.43)	48.80 (24.34)	0.7706
NEC surgery cost	150.25 (23.33)	134.79 (22.49)	0.6334
Other procedures cost	52.13 (17.15)	48.61 (26.18)	0.9105
Total cost	62,283.80 (1875.53)	62,799.06 (1816.75)[b]	0.8436

VP, ventriculoperitoneal.
a Comparisons of placebo vs. *B. breve* BBG groups carried out using Student's *t*-test.
b Includes cost of trial intervention.

TABLE 38 Hospitalisation costs (£, 2012/13 prices) by trial arm (NNRD data)

Cost variable	Trial arm, mean (SE)		*p*-value[a]
	Placebo (*N* = 638)	*B. breve* BBG (*N* = 620)	
Delivery cost	2199.15 (34.8)	2156.65 (34.96)	0.3890
Post-mortem cost	4.07 (2.03)	2.10 (1.48)	0.4317
Hospital transfers cost	1224.31 (69.55)	1264.28 (70.1)	0.6858
Cost of normal care	261.29 (20.41)	259 (23.00)	0.9407
Cost of transitional care	10.78 (3.47)	13.56 (3.55)	0.5760
Cost of special care	15,656.13 (352.45)	15,540.38 (363.38)	0.8192
Cost of high-dependency care	19,343.06 (911.54)	19,608.54 (858.95)	0.8322
Cost of intensive care	21,362.65 (1163.14)	20,782.22 (940.36)	0.6980
ROP screen cost	114.14 (1.89)	114.64 (1.90)	0.8515
ROP treatment cost	22.58 (4.34)	20.55 (4.21)	0.7377
Ultrasound scan costs	43.12 (1.61)	43.16 (1.69)	0.9876
PDA surgery cost	83.52 (17.51)	113.29 (20.56)	0.2705
Hernia surgery cost	80.33 (12.15)	80.65 (12.34)	0.9854
Reservoir surgery cost	0 (0)	0 (0)	0.3177
VP shunt surgery cost	59.28 (26.43)	85.40 (32.12)	0.5301
NEC surgery cost	130.99 (21.88)	103.08 (19.80)	0.3443
Other procedures cost	331.41 (51.17)	344.02 (61.07)	0.8743
Total cost	60,926.82 (1805.16)	60,559.76 (1571.11)[b]	0.8781

VP, ventriculoperitoneal.
a Comparisons of Placebo vs. *B. breve* BBG groups carried out using Student's *t*-test.
b Includes cost of trial intervention.

Nevertheless, the mean total costs were lower (by > £1300 in the placebo arm and > £2000 in the probiotic arm) in the NNRD data. *Table 39* presents the cost measures and their respective values for the combined data set.

Resource use and cost estimates: comparisons across trial between data sources

Table 40 summarises the mean (SE) resource values for each resource category reported by the alternative data sources and the overall levels of agreement between combinations of data sources (Lin's coefficient), denoted ρ_c. *Table 41* summarises the mean (SE) cost values for each cost category reported by the alternative data sources and the overall levels of agreement between combinations of data sources. For these analyses, infants were pooled across trial arms. Agreement between data sources varied greatly by resource or cost category and by combination of data sources. When the PiPS and NNRD data were compared, agreement was relatively high for utilisation or cost of hospital stay by alternative levels of neonatal care, hospital transfers, ROP screens and treatment, forms of surgery [e.g. PDA, ventriculoperitoneal (VP), NEC], and total neonatal care. However, for post-mortem examinations, ultrasound scans and other procedures, the agreement levels fell below the 0.40 threshold indicating acceptable clinical or practical significance.[347] The 95% limits of agreement, exploring the amount of random variation between the PiPS and NNRD data sources, suggest that large individual differences are likely to be encountered for several categories of resource use and costs. *Tables 40* and *41* also summarise the levels of agreement between the PiPS and combined data sources and between the NNRD and combined data sources for resource use and cost values, respectively. However, in the absence of an external gold standard for resource use and cost estimates, these analyses were constrained by the inclusion of values from either PiPS or the NNRD in the combined data source.

TABLE 39 Hospitalisation costs (£, 2012/13 prices) by trial arm (combined data)

	Trial arm, mean (SE)		
Cost variable	Placebo ($N = 638$)	*B. breve* BBG ($N = 620$)	*p*-value[a]
Delivery cost	2199.15 (34.8)	2156.65 (34.96)	0.3890
Post-mortem cost	10.18 (3.2)	10.48 (3.29)	0.9486
Hospital transfers cost	1058.92 (59.88)	1096.30 (62.06)	0.6648
Cost of normal care	261.29 (20.41)	259.00 (23.00)	0.9407
Cost of transitional care	10.78 (3.47)	13.56 (3.55)	0.5760
Cost of special care	15,656.13 (352.45)	15,540.38 (363.38)	0.8192
Cost of high-dependency care	19,343.06 (911.54)	19,608.54 (858.95)	0.8322
Cost of intensive care	21,362.65 (1163.14)	20,782.22 (940.36)	0.6980
ROP screen cost	122.55 (1.49)	123.51 (1.45)	0.6442
ROP treatment cost	7.82 (2.59)	11.62 (3.19)	0.3551
Ultrasound scan costs	86.84 (1.98)	87.62 (1.84)	0.7718
PDA surgery cost	98.70 (18.97)	128.91 (21.85)	0.2967
Hernia surgery cost	37.23 (8.42)	38.31 (9.12)	0.9306
Reservoir surgery cost	0 (0)	4.71 (4.71)	0.3177
VP shunt surgery cost	59.28 (26.43)	85.40 (32.12)	0.5301
NEC surgery cost	150.25 (23.33)	134.79 (22.49)	0.6334
Other procedures cost	331.41 (51.17)	344.02 (61.07)	0.8743
Total cost	60,796.25 (1798.97)	60,454.28 (1565.73)[b]	0.8860

a Comparisons of placebo vs. *B. breve* BBG groups carried out using Student's *t*-test.
b Includes cost of trial intervention.

TABLE 40 Agreement between data sources: resource use variables

Variable	Data set, mean (standard error)			Agreement between					
				PiPS and NNRD data sets		PiPS and combined data sets		NNRD and combined data sets	
	PiPS	NNRD	Combined	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ (95% CI)	Mean difference (95% limits of agreement)
Post-mortem	0.016 (0.004)	0.005 (0.002)	0.016 (0.004)	0.148 (0.102 to 0.193)	0.011 (–0.253 to 0.275)	1	0	0.148 (0.102 to 0.193)	–0.011 (–0.275 to 0.253)
Hospital transfers	0.78 6 (0.031)	0.908 (0.036)	0.786 (0.031)	0.93 (0.923 to 0.936)	–0.122 (–0.989 to 0.746)	1	0	0.93 (0.923 to 0.936)	0.122 (–0.746 to 0.989)
Length of intensive care stay (days)	21.76 (0.81)	18.857 (0.671)	18.857 (0.671)	0.81 (0.792 to 0.828)	2.903 (–29.157 to 34.963)	0.81 (0.792 to 0.828)	2.903 (–29.157 to 34.963)	1	0
Length of high-dependency care stay (days)	23.16 (0.684)	24.61 (0.792)	24.61 (0.792)	0.842 (0.825 to 0.856)	–1.451 (–30.88 to 27.979)	0.842 (0.825 to 0.856)	–1.451 (–30.88 to 27.979)	1	0
Length of special care stay (days)	31.118 (0.65)	30.882 (0.501)	30.882 (0.501)	0.531 (0.492 to 0.568)	0.235 (–39.613 to 40.084)	0.531 (0.492 to 0.568)	0.235 (–39.613 to 40.084)	1	0
ROP screens	0.917 (0.008)	0.853 (0.01)	0.917 (0.008)	0.509 (0.469 to 0.546)	0.064 (–0.558 to 0.687)	1	0	0.509 (0.469 to 0.546)	–0.064 (–0.687 to 0.558)
ROP treatment	0.017 (0.004)	0.039 (0.005)	0.017 (0.004)	0.524 (0.487 to 0.559)	–0.021 (–0.343 to 0.3)	1	0	0.524 (0.487 to 0.559)	0.021 (–0.3 to 0.343)
Ultrasound scans	1.62 (0.025)	0.801 (0.022)	1.62 (0.025)	0.282 (0.247 to 0.316)	0.819 (–0.973 to 2.61)	1	0	0.282 (0.247 to 0.316)	–0.819 (–2.61 to 0.973)
PDA surgery	0.047 (0.006)	0.041 (0.006)	0.047 (0.006)	0.791 (0.769 to 0.811)	0.006 (–0.258 to 0.271)	1	0	0.791 (0.769 to 0.811)	–0.006 (–0.271 to 0.258)
Hernia surgery	0.03 (0.005)	0.064 (0.007)	0.03 (0.005)	0.526 (0.489 to 0.562)	–0.034 (–0.447 to 0.379)	1	0	0.526 (0.489 to 0.562)	0.034 (–0.379 to 0.447)

Variable	Data set, mean (standard error)			Agreement between					
				PiPS and NNRD data sets		PiPS and combined data sets		NNRD and combined data sets	
	PiPS	NNRD	Combined	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ (95% CI)	Mean difference (95% limits of agreement)
Reservoir surgery	0.001 (0.001)	0	0.001 (0.001)	NA	0.001 (–0.056 to 0.057)	1	0	NA (NA, NA)	–0.001 (–0.057 to 0.056)
VP shunt surgery	0.007 (0.002)	0.01 (0.003)	0.01 (0.003)	0.76 (0.736 to 0.782)	–0.002 (–0.128 to 0.124)	0.76 (0.736 to 0.782)	–0.002 (–0.128 to 0.124)	1	0
NEC surgery	0.058 (0.007)	0.048 (0.006)	0.058 (0.007)	0.722 (0.695 to 0.747)	0.01 (–0.323 to 0.343)	1	0	0.722 (0.695 to 0.747)	–0.01 (–0.343 to 0.323)
Other procedures	0.025 (0.005)	0.39 (0.029)	0.39 (0.029)	0.075 (0.059 to 0.091)	–0.365 (–2.391 to 1.661)	0.075 (0.059 to 0.091)	–0.365 (–2.391 to 1.661)	1	0

NA, not applicable.
a ρ_c = Lin's concordance correlation coefficient.
b Bland–Altman limits of agreements = mean difference ± 2 × SD of the difference.[348]

TABLE 41 Agreement between data sources: cost variables

	Data set, mean (standard error)			Agreement between					
				PiPS and NNRD data sets		PiPS and combined data sets		NNRD and combined data sets	
Variable	PiPS	NNRD	Combined	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement)
Delivery cost	2450.157 (26.125)	2178.204 (24.662)	2178.204 (24.662)	0.793 (0.773 to 0.812)	271.95 (–780.85 to 1324.755)	0.793 (0.773 to 0.812)	271.953 (–780.85 to 1324.755)	1	0
Post-mortem cost	10.329 (2.292)	3.099 (1.262)	10.329 (2.292)	0.148 (0.102 to 0.193)	7.23 (–164.055 to 178.515)	1	0	0.148 (0.102 to 0.193)	–7.23 (–178.515 to 164.055)
Hospital transfers cost	1077.343 (43.09)	1244.01 (49.357)	1077.343 (43.09)	0.93 (0.923 to 0.936)	–166.667 (–1355.215 to 1021.882)	1	0	0.93 (0.923 to 0.936)	166.667 (–1021.882 to 1355.215)
Cost of intensive care	24,321.33 (905.508)	21,076.591 (749.921)	21,076.591 (749.921)	0.81 (0.792 to 0.828)	3244.739 (–32,588.886 to 39,078.365)	0.81 (0.792 to 0.828)	3244.739 (–32,588.886 to 39,078.365)	1	0
Cost of high-dependency care	18,325.971 (541.596)	19,473.899 (626.595)	19,473.899 (626.595)	0.842 (0.825 to 0.856)	–1147.929 (–24,434.902 to 22,139.045)	0.842 (0.825 to 0.856)	–1147.929 (–24,434.902 to 22,139.045)	1	0
Cost of special care	15,717.933 (328.192)	15,599.083 (252.931)	15,599.083 (252.931)	0.531 (0.492 to 0.568)	118.85 (–20,009.027 to 20,246.728)	0.531 (0.492 to 0.568)	118.85 (–20,009.027 to 20,246.728)	1	0
ROP screen cost	123.023 (1.042)	114.388 (1.34)	123.023 (1.042)	0.509 (0.469 to 0.546)	8.635 (–74.852 to 92.122)	1	0	0.509 (0.469 to 0.546)	–8.635 (–92.122 to 74.852)
ROP treatment cost	9.688 (2.048)	21.579 (3.023)	9.688 (2.048)	0.524 (0.487 to 0.559)	–11.89 (–189.834 to 166.053)	1	0	0.524 (0.487 to 0.559)	11.89 (–166.053 to 189.834)
Ultrasound scan costs	87.223 (1.353)	43.141 (1.163)	87.223 (1.353)	0.282 (0.247 to 0.316)	44.082 (–52.378 to 140.542)	1	0	0.282 (0.247 to 0.316)	–44.082 (–140.542 to 52.378)

	Data set, mean (standard error)			Agreement between					
				PiPS and NNRD data sets		PiPS and combined data sets		NNRD and combined data sets	
Variable	PiPS	NNRD	Combined	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement[b])	ρ_c (95% CI)[a]	Mean difference (95% limits of agreement)
PDA surgery cost	113.591 (14.443)	98.189 (13.473)	113.591 (14.443)	0.791 (0.769 to 0.811)	15.402 (–624.694 to 655.498)	1	0	0.791 (0.769 to 0.811)	–15.402 (–655.498 to 624.694)
Hernia surgery cost	37.763 (6.197)	80.485 (8.654)	37.763 (6.197)	0.526 (0.489 to 0.562)	–42.721 (–558.64 to 473.197)	1	0	0.526 (0.489 to 0.562)	42.721 (–473.197 to 558.64)
Reservoir surgery cost	2.323 (2.323)	0	2.323 (2.323)	NA (NA, NA)	2.323 (–162.484 to 167.131)	1	0	NA (NA, NA)	–2.323 (–167.131 to 162.484)
VP shunt surgery cost	54.111 (17.98)	72.153 (20.737)	72.153 (20.737)	0.76 (0.736 to 0.782)	–18.041 (–971.458 to 935.375)	0.76 (0.736 to 0.782)	–18.041 (–971.458 to 935.375)	1	0
NEC surgery cost	142.634 (16.209)	117.234 (14.775)	142.634 (16.209)	0.722 (0.695 to 0.747)	25.401 (–793.334 to 844.135)	1	0	0.722 (0.695 to 0.747)	–25.401 (–844.135 to 793.334)
Other procedures cost	50.392 (15.557)	337.626 (39.726)	337.626 (39.726)	0.216 (0.182 to 0.25)	–287.234 (–2952.95 to 2378.482)	0.216 (0.182 to 0.25)	–287.234 (–2952.95 to 2378.482)	1	0
Total cost	62,537.737 (1305.81)	60,745.916 (1198.576)	60,627.711 (1194.465)	0.917 (0.908 to 0.925)	1791.822 (–34,274.766 to 37,858.409)	0.917 (0.907 to 0.925)	1910.027 (–34,174.429 to 37,994.482)	1	118.205 (–1428.536 to 1664.946)

a ρ_c = Lin's concordance correlation coefficient.
b Bland–Altman limits of agreements = mean difference ± 2 × SD of the difference.[348]

Cost-effectiveness: comparisons within trial by data source

The incremental cost-effectiveness of the probiotic is shown in *Table 42* for the 1258 infants eligible for our comparative analyses, by clinical outcome and data source. Based on data collected from the PiPS trial CRFs, the average total cost was £62,799 in the probiotic group compared with £62,284 in the placebo group, generating a mean incremental cost of £515. The incremental cost-effectiveness of the probiotic was estimated at £216,369 per death avoided, £107,613 per episode of sepsis avoided, £61,170 per episode of NEC avoided and £153,703 per composite adverse outcome avoided. The mean ICERs fell in the north-east quadrant of the cost-effectiveness plane (see *Figures 43–46*). The corresponding CEACs (figures not shown) indicate that regardless of the clinical outcome measure of interest the probability that the probiotic is cost-effective varied between 40% and 50% depending on the value of the cost-effectiveness threshold. If decision-makers are willing to pay £30,000 to avoid an adverse perinatal outcome, the probability that the probiotic is cost-effective varied between 42.6% and 47.7%.

Based on data from the NNRD, the average total cost was £60,560 in the probiotic group, compared with £60,927 in the placebo group, generating a mean incremental saving of £367. Because the probiotic was, on average, more effective than the placebo regardless of clinical outcome, the mean ICERs fell in the south-east quadrant of the cost-effectiveness plane (see *Figures 47–50*), suggesting that the probiotic dominated the placebo in health economic terms. Regardless of the clinical outcome measure of interest, the probability that the probiotic is cost-effective varied between 50% and 60% depending on the value of the cost-effectiveness threshold. If decision-makers are willing to pay £30,000 to avoid an adverse perinatal outcome, the probability that the probiotic is cost-effective varied between 56.8% and 60.5%.

Based on the combined data, the average total cost was £60,454 in the probiotic group compared with £60,796 in the placebo group, generating a mean incremental saving of £342. Because the probiotic was, on average, more effective than the placebo regardless of clinical outcome, the mean ICERs fell in the south-east quadrant of the cost-effectiveness plane (see *Figures 51–54*), suggesting that the probiotic dominated the placebo in health economic terms. Regardless of the clinical outcome measure of interest, the probability that the probiotic is cost-effective varied between 55% and 60% depending on the value of the cost-effectiveness threshold. If decision-makers are willing to pay £30,000 to avoid an adverse perinatal outcome, the probability that the probiotic is cost-effective varied between 58.5% and 60.3%.

Our estimates of within-trial incremental cost-effectiveness were replicated for each of the prespecified subgroups, namely gender, birthweight, gestational age, colonisation status and randomisation age. *Table 42* presents the cost-effectiveness outcomes by prespecified subgroup for the composite secondary outcome. *Table 63* presents the cost-effectiveness outcomes by prespecified subgroup for the death primary outcome; *Table 64* presents the cost-effectiveness outcomes by prespecified subgroup for the sepsis primary outcome; and *Table 65* presents the cost-effectiveness outcomes by prespecified subgroup for the NEC primary outcome. The probability that the probiotic is cost-effective was notably higher for girls and for infants born at ≥ 1000 g, regardless of clinical outcome measure and data source. For the death primary outcome, the probability that the probiotic is cost-effective at a £30,000 cost-effectiveness threshold varied between 78.9% and 88.0% for girls, and between 62.7% and 93.0% for infants born at ≥ 1000 g, depending on data source (see *Table 63*). For the sepsis primary outcome, the probability that the probiotic is cost-effective at a £30,000 cost-effectiveness threshold varied between 73.4% and 81.2% for girls, and between 77.3% and 96.8% for infants born at ≥ 1000 g, depending on data source (see *Table 64*). For the NEC primary outcome, the probability that the probiotic is cost-effective at a £30,000 cost-effectiveness threshold varied between 74.2% and 82.3% for girls, and between 71.5% and 94.8% for infants born at ≥ 1000 g, depending on data source (see *Table 65*). Finally, for the composite secondary outcome, the probability that the probiotic is cost-effective at a £30,000 cost-effectiveness threshold varied between 77.1% and 85.1% for girls, and between 65.9% and 92.8% for infants born at ≥ 1000 g, depending on data source (see *Table 66*).

Comparisons of cost-effectiveness outcomes between data sources

To compare the discrepancy in the cost-effectiveness results between the different data sources, agreement statistics (namely the probability of miscoverage and two-sided probability values) were

TABLE 42 Cost-effectiveness estimates for probiotic by clinical outcome (death, sepsis, NEC or composite secondary) and data source for resource use data (PiPS, NNRD, combined)

	Mean costs (95% CI)			Mean effects (95% CI)				Probability *B. breve* BBG is (%)			
	B. breve BBG (£)	Placebo (£)	Difference (£)	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Death[e]											
PiPS data	62,799.06 (59,231 to 66,367)	62,283.79 (58,601 to 65,967)	515.27 (–4611 to 5641)	0.0823 (0.0606 to 0.1039)	0.0846 (0.0629 to 0.1063)	0.0024 (–0.0282 to 0.0330)	216,369	58.8	42.0	42.8	44.2
NNRD data	60,559.76 (57,474 to 63,645)	60,926.82 (57,382 to 64,472)	–367.07 (–5072 to 4338)	0.0823 (0.0606 to 0.1039)	0.0846 (0.0629 to 0.1063)	0.0024 (–0.0282 to 0.0330)	Treatment dominates	60.8	57.8	59.4	59.9
Combined data	60,454.28 (57,379 to 63,529)	60,796.25 (57,264 to 64,329)	–341.98 (–5031 to 4347)	0.0823 (0.0606 to 0.1039)	0.0846 (0.0629 to 0.1063)	0.0024 (–0.0282 to 0.0330)	Treatment dominates	60.8	58.0	59.0	59.5
Sepsis[f]											
PiPS data	62,799.06 (59,231 to 66,367)	62,283.79 (58,601 to 65,967)	515.27 (–4611 to 5641)	0.1081 (0.0836 to 0.1326)	0.1129 (0.0882 to 0.1375)	0.0048 (–0.0299 to 0.0395)	107,613	59.6	40.6	41.8	42.6
NNRD data	60,559.76 (57,474 to 63,645)	60,926.82 (57,382 to 64,472)	–367.07 (–5072 to 4338)	0.1081 (0.0836 to 0.1326)	0.1129 (0.0882 to 0.1375)	0.0048 (–0.0299 to 0.0395)	Treatment dominates	60.9	55.0	56.7	56.8
Combined data	60,454.28 (57,379 to 63,529)	60,796.25 (57,264 to 64,329)	–341.98 (–5031 to 4347)	0.1081 (0.0836 to 0.1326)	0.1129 (0.0882 to 0.1375)	0.0048 (–0.0299 to 0.0395)	Treatment dominates	63.2	58.0	58.4	58.5
NEC											
PiPS data	62,799.06 (59,231 to 66,367)	62,283.79 (58,601 to 65,967)	515.27 (–4611 to 5641)	0.0903 (0.0677 to 0.1129)	0.0987 (0.0755 to 0.1220)	0.0084 (–0.0240 to 0.0408)	61,170	72.8	42.0	45.7	47.7
NNRD data	60,559.76 (57,474 to 63,645)	60,926.82 (57,382 to 64,472)	–367.07 (–5072 to 4338)	0.0903 (0.0677 to 0.1129)	0.0987 (0.0755 to 0.1220)	0.0084 (–0.0240 to 0.0408)	Treatment dominates	69.3	57.8	60.3	60.5
Combined data	60,454.28 (57,379 to 63,529)	60,796.25 (57,264 to 64,329)	–341.98 (–5031 to 4347)	0.0903 (0.0677 to 0.1129)	0.0987 (0.0755 to 0.1220)	0.0084 (–0.0240 to 0.0408)	Treatment dominates	69.3	58.0	59.8	60.3

continued

TABLE 42 Cost-effectiveness estimates for probiotic by clinical outcome (death, sepsis, NEC or composite secondary) and data source for resource use data (PiPS, NNRD, combined) (*continued*)

	Mean costs (95% CI)			Mean effects (95% CI)				Probability *B. breve* BBG is (%)			
	B. breve BBG (£)	Placebo (£)	Difference (£)	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Composite[g]											
PiPS data	62,799.06 (59,231 to 66,367)	62,283.79 (58,601 to 65,967)	515.27 (–4611 to 5641)	0.2145 (0.1822 to 0.2468)	0.2179 (0.1858 to 0.2499)	0.0034 (–0.0421 to 0.0489)	153,703		42	44.6	46.2
NNRD data	60,559.76 (57,474 to 63,645)	60,926.82 (57,382 to 64,472)	–367.07 (–5072 to 4338)	0.2145 (0.1822 to 0.2468)	0.2179 (0.1858 to 0.2499)	0.0034 (–0.0421 to 0.0489)	Treatment dominates		57.8	58.4	59.5
Combined data	60,454.28 (57,379 to 63,529)	60,796.25 (57,264 to 64,329)	–341.98 (–5031 to 4347)	0.2145 (0.1822 to 0.2468)	0.2179 (0.1858 to 0.2499)	0.0034 (–0.0421 to 0.0489)	Treatment dominates		57.9	58.3	59.1

a The difference in effects was inverted (i.e. negative values were given a positive sign to reflect the fact that a reduction in adverse outcomes is a positive effect).
b Based on 1000 bootstrap replicates of the data set.
c *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a £20,000 cost-effectiveness threshold.
d *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a £30,000 cost-effectiveness threshold.
e Death before discharge home – Includes three infants who remained on paediatric wards and are analysed as survivors.
f Sepsis is defined as blood stream infection with non-skin commensals after 72 hours postnatal age and before 46 weeks' postmenstrual age.
g Death, or sepsis or NEC.

estimated using a double bootstrap strategy. A detailed description of methodology has been published.[350] Briefly, an estimate of the probability of miscoverage between the incremental net benefit generated by any two data sets was obtained as follows:

1. First, for any two data sets, one data set was designated as the reference data (referent) and the other as the test data. For the analyses reported here, the PiPS data set was initially designated as the referent; however, for completeness, we also report results where the combined data set was designated as the referent.
2. Second, bootstrapping was applied to the test data to generate 500 replicates of the test data.
3. Finally, the probability of miscoverage was obtained by counting the proportion of the 500 replicates in which the 95% CIs for the incremental net benefits (test data) did not contain the referent incremental net benefit estimate.

Tables 67 and *68* summarise the agreement statistics (two-sided *p*-values and probability estimates of miscoverage) between any two data sources obtained using the above strategy. The probability of miscoverage ranged from 3.9% at a cost-effectiveness threshold of £20,000 per case of sepsis avoided to 6.4% at a cost-effectiveness threshold of £30,000 per death avoided, when the source of outcomes data were the NNRD and the combined data sets, respectively. The *p*-values ranged from 0.387 (PiPS vs. NNRD data set at a cost-effectiveness threshold of £20,000 per death avoided) to 0.571 (PiPS vs. NNRD[2] at a cost-effectiveness threshold of £30,000 per case of sepsis avoided). These *p*-values provide no evidence to suggest that the incremental net benefit estimated using one data set differs significantly from the incremental net benefit estimated from another data set. Separate analyses performed on the prespecified subgroups did not alter these findings.

Finally, sensitivity analyses that compared the key outputs of the economic evaluation using resource use and clinical outcomes data extracted solely from the NNRD are summarised in *Table 69*. These analyses were restricted to the two clinical outcomes used in the PiPS trial that were available in both data sources, namely (1) death before discharge from hospital and (2) sepsis. Notably, the mean ICER for the sepsis outcome moved from the south-east quadrant of the cost-effectiveness plane, denoting a less costly and more effective intervention when the NNRD acted as the sole source of resource use information (see *Table 41*) to the south-west quadrant of the cost-effectiveness plane, denoting a less costly and less effective intervention when the NNRD acted as the sole source of resource use and clinical outcomes.

Conclusions

This chapter outlined a study to assess whether or not reliable health service utilisation data can be obtained from the NNRD and can be used to inform future trial-based economic evaluations of neonatal interventions. The recently completed PiPS was used as the test bed for our empirical investigations. Health-care resource utilisation data were extracted from the PiPS trial CRFs, the NNRD and a combined data source, and primarily valued using national tariffs and expressed in GBP (2012/13 prices). Differences in economic outcomes were estimated (1) within trial by data source, thereby allowing us to draw comparisons between the probiotic and its comparator (placebo), and (2) by pooling data between the trial arms, thereby allowing us to draw comparisons between the alternative data sources. Within-trial comparisons of resource use and costs revealed no statistically significant differences between the trial comparators in the values for any resource input or cost category, regardless of data source. Across-trial tests of concordance in resource use and costs between comparator data sources revealed relatively high levels of agreement for the majority of categories of resource use or cost and notably for the total cost of neonatal care. Within-trial estimates of cost-effectiveness revealed relatively low probabilities of cost-effectiveness for the probiotic across a wide range of cost-effectiveness thresholds, regardless of data source. It was notable, however, that following subgroup analyses the probiotic had a high probability of cost-effectiveness for girls and for infants born at $\geq$ 1000 g, regardless of data source. Finally, comparisons of cost-effectiveness outcomes between data sources revealed low probability levels of miscoverage of incremental net monetary benefit sources when the NNRD acted as the sole source of

resource use information. However, separate sensitivity analyses revealed that probability estimates of miscoverage for incremental net monetary benefit increased for both death and sepsis outcomes when the NNRD acted as the sole source of resource use information and clinical outcomes.

A number of caveats should be borne in mind when interpreting the results of this study. First, the third of our comparator data sources, the 'combined' data source, was constructed by selecting resource components from either the PiPS trial CRFs or the NNRD on the basis of volume and granularity of information provided. Second, there are a number of features of the economic evaluation, which although not directly impinging on our comparisons across data sources, they do constrain the conclusions we can draw about the cost-effectiveness of the probiotic. For example, by adopting the recommended health system perspective (NICE, 2013[351]), our study excluded broader costs, such as those borne by family members and informal carers, which are arguably of relevance to economic evaluations of neonatal interventions. However, given the absence of evidence of significant clinical effect for the probiotic, it is unlikely that incorporation of these broader societal costs into the analysis would have had an impact on our estimates of incremental cost-effectiveness. In addition, in the absence of validated multiattribute utility measures for use in early childhood, the effectiveness of the probiotic was not measured in terms of a preference-based outcome measure, such as the quality-adjusted life-year (QALY), which may have been more useful for cost-effectiveness comparative purposes and for which accepted threshold values are available.[339] Third, our comparisons of economic outcomes generated by the NNRD and the PiPS trial CRFs were based on a clinical trial in which there was no evidence of significant clinical effect for the intervention being evaluated. Many health economists have argued that reliance on traditional rules of statistical inference surrounding a single parameter, such as clinical effectiveness, is arbitrary, and may result in inferior health-care outcomes compared with basing decisions on expected cost-effectiveness.[352] Future research should consider the application of value of information techniques using NNRD data in order to quantify any economic costs associated with incorrect policy decisions around the adoption of neonatal interventions.[353] Under these circumstances, it is entirely appropriate to estimate the joint distribution of cost and effect differences, not either of these in isolation. Nevertheless, it will be important in future research to test the validity of our results in the context of other neonatal trials with disparate clinical and economic impacts.

Implications for health care

We have demonstrated proof of principle of the potential of the NNRD as a data source for neonatal trial-based economic evaluations in the UK context. This has potential to reduce the costs and improve the efficiency of economic evaluations conducted in relation to research studies. The results of our study have important implications for health economics research in the neonatal context. We have demonstrated that the bulk of hospital resource inputs incorporated into a rigorously designed economic evaluation of a neonatal intervention in a UK context can be successfully and accurately extracted from the NNRD. These include not only resource-generating events that contribute to national reference costs for relevant Healthcare Resource Groups for neonatal care, but resource-generating events that fall outside these per diem values, for example high-cost scans, tests and blood products, surgeries, transfers and post-mortem examinations.

Research recommendations

Further research is recommended to validate our results in the context of other trials and to assess the utility of the NNRD across a wider range of economic evaluations and alternative study designs. We would have ideally wished to triangulate the resource and cost profiles generated by the PiPS trial CRFs and the NNRD data against an external gold standard, such as data extracted directly from patient notes. However, such an exercise was not within the resources of this study and it remains an area for future research.

Chapter 7 Linking the National Neonatal Research Database to other NHS data sets; feasibility and birth cohort studies

Abstract

Background: Linking routine NHS clinical and administrative data offers considerable potential for research and health service evaluations but is dependent on data completeness and quality.

Aims and methods: We examined HES and NNRD data and level of agreement. We linked data sets to create birth cohorts, ascertain mortality and identify individuals across time in relation to hospital admissions up to the age of 1 year for a key infant illness: bronchiolitis.

Results: One in ten neonates identified in HES represent admissions to neonatal specialised care as determined by a record in the NNRD. There is > 95% agreement for key items. Data quality and completeness are generally better in the NNRD than in HES. Approximately 20% of babies in HES have missing gestational age data and around 1.5% have a biologically implausible birthweight. The completeness of HES birth data varies substantially between hospitals but has improved over time. The higher mortality rates of extremely preterm babies extend throughout the first year. Most infants admitted to hospital with bronchiolitis in England are born at term and have no recognised predisposing risks.

Conclusions: Linkage between HES and the NNRD is feasible, enhances the quality and scope of birth records, and paves the way for ascertainment of lifelong health outcomes. Improved health data quality, as well as completeness, are important service goals. Reducing reliance on administrative data, promoting clinician involvement in data assurance and extracting data from electronic health records rather than recording them anew are measures that merit wider consideration.

Introduction

Potential of data set linkage

Data set linkage offers considerable potential to enlarge the scope, enhance the richness and improve the quality and completeness of information for clinical and health services research. Linked data sets provide the opportunity to address an extensive range of research questions, such as examining long-term health outcomes, identifying risk factors for re-admission, and quantifying mortality, morbidity and health-care utilisation following discharge from neonatal specialised care. Linkage also offers potential to develop cradle-to-grave data sets, a particularly useful resource for infant health research.

Information on long-term health outcomes at multiple points in time is important for evaluating interventions in pregnancy and in the neonatal period. However, although long-term assessment of outcomes is highly desirable, it is also complex and costly. Considerable demand is placed on participants when information on long-term outcomes is sought, and the costs of research-based follow-up are substantial. Most research is

funded for a finite period, typically 3–5 years. Obtaining funding for long-term follow-up studies is extremely problematic and beyond the ability of any but a few large research organisations. As a consequence, much infant research is compromised by a focus on short-term outcomes.

Birth cohorts are an important resource for epidemiological research. They have usually involved recruitment at a specific age with repeated surveys of participants conducted at intervals of years. These longitudinal studies have typically contributed to examining societal change, for example in relation to family structures, educational attainments, equity, poverty and class. Inclusion of data on potential confounders, comorbidities and clinical outcomes, in addition to core baseline characteristics (e.g. gestational age, birthweight and sex), offer the opportunity to address causal relationships through the use of statistical techniques such as instrumental variable and multivariable analyses. This is particularly valuable when investigating exposures that are not amenable to randomisation.

Population-level clinical data sets can provide information that is highly generalisable, have power to detect small effect sizes and relate directly to real-life health-care practices. Electronic health records offer the potential to reduce the cost and complexity of data acquisition. The NHS in England, with near-universal national coverage, is potentially in a unique position to assess population-based outcomes following discharge of neonates from neonatal specialised care, using information from linked data sets. This would have major utility for neonatal clinical trials, health economic evaluations, post-marketing and other surveillance, and observational birth cohort and epidemiological studies, particularly those focused on aspects of health over the life course.

Hospital Episode Statistics

The administrative database, HES, contains details of patient diagnoses and procedures on all episodes of care in English NHS trusts (acute hospital, primary care and mental health trusts), going back to 1989. HES data were conceived in 1987, following a report on the collection and use of hospital activity information by a committee chaired by Dame Edith Körner (née Lowy), daughter of a maize miller, and a refugee who fled to England following the Nazi occupation of Czechoslovakia. She became an authority on health administration and the use of computers in the health service. In 1980, she was asked to chair a national review of NHS information. After 4 years of deliberation, the Körner Committee's several recommendations were adopted, in what became the beginning of the computerisation of the NHS. For the following 20 years or so, NHS statistical information was known as 'Körner Data'.

The HES data are submitted centrally and held by the Health and Social Care Information Centre (formerly the NHS Information Centre and now known as NHS Digital). HES data were initially collated subnationally by regional health authorities. Following the abolition of these bodies in 1996 the NHS-Wide Clearing Service took over until, in 2006, responsibility for HES data storage and management was taken over by the Secondary Uses Service, which was run by the Health and Social Care Information Centre and the National Programme for IT. HES data primarily serve administrative and financial purposes. Although not designed for these purposes, HES data are also widely used for research and health service evaluations.

The HES data are divided into financial years from 1 April to 31 March in the following year and cover admissions, outpatient appointments, and accident and emergency department attendances. HES data are recorded by clerical staff through patient administration or hospital information systems. They assign ICD-10 codes to clinical diagnoses recorded in medical notes. Data are a summary of each patient episode; they are not checked and there is no review of missing data or duplicate entries. The basic unit of measurement in HES is the 'Finished Consultant Episode', defined as an episode 'where a patient completes a period of care under a consultant and is either transferred to another consultant or discharged'. HES data are stored as a collection of individual records for each period of care. A 'patient key' is derived from six HES fields (i.e. NHS number, date of birth, sex, postcode, provider code, local patient identifier). Each individual patient key is allocated to one unique pseudonymised HES identifier: the 'HESID'. As some source fields can change, the HESID can be mapped to several different patient keys. The unique HESID and the discrete patient keys

provide the means to uniquely identify a patient and track them in HES without the risk of revealing personal or sensitive information.

Hospital Episode Statistics maternity and birth data

The HES data are recorded on all births in NHS hospitals, non-NHS hospitals funded by the NHS and NHS home births in England. When a mother gives birth, her hospital admission record changes from a general inpatient admission record to a maternity record and is updated as such before it is submitted to HES. HES data contain two types of maternity record: the delivery and the birth record (both of which contain a 'baby tail' comprising an additional 19 fields). The delivery record relates to the mother and contains the same information as the general HES record, and the associated baby tail contains information about the delivery. The birth record relates to the baby and also contains general record information; the birth record contains the same information as the baby tail in the delivery record. Diagnoses and procedures recorded in the birth record refer to the baby and, conversely, diagnoses and procedures in the delivery record refer to the mother. For multiple births, separate tails for each baby appear in the delivery record, but each birth record contains only the individual baby tail.

Potential for linkage of National Neonatal Research Database with general practice records

It is estimated that > 98% of the UK population are registered with a general practitioner (GP), almost all of which use computerised record systems. During the Medicines for Neonates programme, several sources of GP records were identified and their utility for researching the health of neonates was explored. The largest and most comprehensive source of primary care data in the UK is the General Practice Research Database (GPRD),[355] which has been widely used for research, from pharmacovigilance to risk score development, and is a rich source of longitudinal patient data. Information in GP records includes demographic data, coded clinical information including diagnoses, symptoms, preventative care and prescriptions. In the UK, a standardised hierarchical classification system of Read codes is used to record medical information in patient records. Alternative sources include The Health Improvement Network from similar and overlapping practices as well as directly accessing clinical records from smaller GP networks and individual practices. From 29 March 2012, GPRD became part of the Clinical Practice Research Datalink (CPRD), funded by the Medicines and Healthcare Regulatory Agency (MHRA) and the NIHR. The CPRD contains computerised clinical records from about 5 million active patients, 12 million patients overall, from 600 primary care practices across the UK, and is a nationally representative sample of around 8% of the UK population. The CPRD aims to maximise the way anonymised clinical data from the NHS can be used for observational research, using linkage to integrate data from primary care, secondary care and disease registries, with the aim of facilitating research that is beneficial to improving public health.

Our initial intention was to seek consent from parents during the hospital admission around the time of birth. However, we realised that that this was neither practicable nor necessary given alternatives whereby anonymised linkage between hospital and GP records were becoming available. To progress this avenue, we submitted a protocol for research ethics approval, which was approved by the Independent Scientific Advisory Committee for the MHRA, which reviews all research proposals for the use of the GPRD. Access to the data was granted free of charge under the previous Medical Research Council licence scheme with the GPRD. We focused our research on an exemplar project to demonstrate proof of concept that GP records could be used to build a birth cohort of infants with bronchiolitis, a common condition for which children are admitted to hospital. A cohort was created using medical records from the GPRD database and used to examine the natural history and management of bronchiolitis in the community setting. The main findings from this analysis of primary care data, conducted independently of the Medicines for Neonates programme, were that a cohort could be created and that data were available for research. By the end of the Medicines for Neonates programme we were, however, still awaiting access to linked hospital and general practice data and so were unable to address this area. After the Medicines for Neonates programme, data from around 50% of practices became available with linked hospitals records and practice records. We therefore recommend that this is a suitable area for future research.

Aims and objectives

We aimed to develop a continuous, longitudinal birth cohort through linkage between the NNRD and HES in order to provide a resource for observational and experimental studies of early exposures and interventions on later health outcomes in specific groups of newborn infants. Our objectives were to:

1. examine NNRD and HES data completeness and quality for key variables
2. link NNRD and HES data to create a birth cohort of infants admitted to neonatal units
3. examine level of agreement between NNRD and HES data
4. conduct a proof of concept study using linked data to ascertain health outcomes.

Methods

Approvals and time line

We received HES data at intervals over the duration of the Medicines for Neonates programme. Permission to receive NHS numbers was obtained later in the programme from the Confidentiality Advisory Group of the Health Research Authority Health Research Authority. Following approval we requested permission from all NHS Trusts in England that provide neonatal specialised care to receive infant identifiers, including the NHS number, for the purpose of data linkage. This workstream was undertaken by several research personnel over the course of the Medicines for Neonates programme.

Design

We conducted two methodological studies to assess the feasibility of using administrative hospital data to build birth cohorts for child health research. We used HES data covering the financial years 2005/6–2009/10 (study 1), and HES and NNRD data for the calendar year 2010 (study 2).

We engaged with stakeholders including health professionals within NW London CLAHRC and expert patients and parents of preterm babies to get their input into grant applications and design of studies. We held educational meetings, focus groups and dissemination events to obtain feedback at all stages of the study.

Birth episodes

We identified all individual birth episodes in HES. In study 1, we used the 'admimeth' variable, which contains a code recording how the patient was admitted to hospital. We used this field to select records with an admission method coded 82 (babies born in health-care provider) or 83 (babies born outside the health-care provider, except when born at home as intended). In study 2, we also used the 'epitype' variable, which contains a code defining the reason for which the patient was admitted to hospital. We used this field to select all records with an episode type coded 3 (birth episode) or 6 (other birth event).

Duplicate records

We identified duplicate birth episodes using the HESID in study 1 and the NHS number in study 2. If episodes were identical matches for all variables, only one record was retained. When records did not contain matching information, we retained the birth episode with the most diagnostic information (number of non-empty diagnosis fields).

In study 2, we excluded babies with a missing date of birth because it was not possible to verify if the records referred to a birth episode or a subsequent hospital readmission. We also excluded records with a missing NHS number. When single NHS numbers were associated with multiple HESIDs, we excluded them and their associated fields from the analysis because it was not possible to determine which record was correct.

Data management

Babies can have more than one episode of care within their birth admission; for example, if a baby receives specialist care from a different consultant or is transferred between hospitals. These additional episodes occur within the same birth admission but, where the initial birth event would have an 'epiorder' value of 1, subsequent episodes have an 'epiorder' value of > 1. To facilitate one-to-many linkage to subsequent hospital admission records, we developed a data set consisting of one row per individual. We incorporated key information, such as diagnostic codes, into the original birth episode and dropped subsequent episodes in the birth admission. Up to nine birth tails can be recorded for each delivery, allowing information from multiple births to appear in the mother's delivery record. Identical baby tail information for each baby can be found in their mother's delivery record. Therefore, if we found that a baby's information was not recorded in the first field of a given variable, we condensed records to retain only one field for each variable. For example, if a baby was the second twin, their gestational age at birth ('gestat') may have appeared in the second field ('gestat_2') with the first field ('gestat_1') blank because in the mother's delivery record this contained the first twin's gestation. In this case, we transferred information for the gestation variable from 'gestat_2' into 'gestat_1' and then removed all additional fields (i.e. 'gestat_2' to 'gestat_9') for that variable. We identified stillbirths using the 'birth status' and 'discharge method' fields and removed these from the final cohort.

A range of exclusion criteria was developed to clean key variable fields and examine the quality of coding. The Care Quality Commission (CQC) conducted a review exploring quality indicator specifications used to assess the quality of HES maternity data from 2009 to 2010. We combined the criteria identified within the CQC review and HES inpatient cleaning rules and applied these to the HES birth fields to ensure that suspicious data and invalid records were removed. Fields validated by the CQC and related to maternity episodes were assessed using the following criteria: when values were 'not known', invalid or outside a specific range, the field was recorded as blank or '9'; dates of birth outside the birth admission were set as invalid; flags to determine finished and unfinished episodes were created.

Data completeness and quality

We examined the completeness of HES recording for baby tail fields over 5 years (2005/6–2009/10) and compared the total number of births with ONS birth registrations (study 1). We compared the proportion of missing data for each baby tail field in 2005/6 to 2009/10 using chi-squared tests. We compared the characteristics of hospitals using a cut-off point of 90% completeness of recording for key birth fields (gestational age and birthweight). To test for significant differences between hospitals with high versus low completeness of birth record fields, we used chi-squared tests to compare proportions and *t*-tests to compare mean values.

For calendar year 2010 (study 2), we analysed the completeness of recording in HES and the NNRD for the key variables (infant sex, gestational age, birthweight, multiple birth, LSOA, maternal age and ethnicity). We compared the proportion of complete data for each variable by gestational age group. We explored the distribution of birthweight by gestational age in both sources. We also examined the standardised distribution of birthweight by gestational age (birthweight *z*-score percentiles). We used the LMS growth Excel add-in program from the Medical Research Council, UK, based on the British 1990 growth reference[356] to determine gestational age and sex-specific birthweight SDSs for both HES and the NNRD. We excluded observations for infants who were above (> 99.9th centile) or below (< 0.1st centile) 4 SDSs from the population mean.

Record linkage and agreement between Hospital Episode Statistics and the National Neonatal Research Database

We re-coded data when necessary to provide a common format for linkage. Non-informative characters and punctuation were removed from the diagnosis and procedure variables. We used a deterministic approach to link the NNRD and HES records using the NHS number as a common unique identifier. We performed a one-to-one merge of records from both sources. As we expected one unique record in the NNRD to be linked to one record in HES, we considered a successful one-to-one linkage a positive match.

We created a new data set with single birth episodes and common variables from each source. Records in the NNRD that did not have a corresponding match in HES were retained separately. The data linkage rate was calculated by using the number of positive matches divided by the total number of records available for matching in the NNRD. We performed linkage in two stages. In the first stage, we utilised all records from both databases and, in the second stage, we excluded all values for birthweight by gestational age that were outside a predefined range. We compared infant sex, gestational age, birthweight, multiple birth, social deprivation, maternal age and ethnicity in linked and unlinked babies. Social deprivation was assigned using the IMD. This is based on 32,482 geographic LSOAs across England. Economic, social and housing indicators are combined to provide a score for each LSOA; a high score indicates greater deprivation. We split the birth population into IMD quintiles for comparison. We used one-way ANOVA to compare continuous variable means, and the chi-squared test to compare categorical variables. We calculated the percentage overall agreement and Cohen's kappa, a measure of agreement adjusted for the proportion of agreement that would be expected on the basis of chance. We considered kappa values above 0.80 to indicate almost perfect agreement. [260,261] Analyses were carried out using SAS.

Health outcomes (based on the exemplar condition bronchiolitis)

We created a birth cohort from HES data for all infants born in English NHS hospitals and discharged during a 12-month period (from 1 April 2007 to 31 March 2008). We included only records from live births and excluded infants born in hospitals (85/156) with poor recording (< 90% complete) of key indicators (birthweight and gestational age) to enable us to group infants into term and preterm categories. We conducted sensitivity analyses based on number of maternity beds, annual number of births, geographic location and infant death rate to compare high- and low-recording hospitals. We linked birth records to subsequent hospital admission records up to a child's first birthday, using the unique personal identifier (HESID). We identified deaths up to the age of 1 year, including out-of-hospital deaths, through linkage to ONS mortality records.

We used diagnostic information in individual birth records and any subsequent hospital admission records from the study year to group infants into categories of risk factors for severe respiratory syncytial virus (RSV) infection using ICD-10 codes, or larger subgroups using the Agency for Health Research and Quality's Clinical Classification System (CCS):[357]

- Immunodeficiency: CCS group 57 (immunity disorders – this includes ICD-10 codes D80, D81, D82, D83, D84 and D89).
- Cystic fibrosis: CCS group 56 (cystic fibrosis – this includes ICD-10 codes under E84).
- Chronic lung disease: ICD-10 codes P27 (chronic respiratory disease originating in the perinatal period) and P28 (other chronic respiratory diseases originating in the perinatal period).
- Congenital heart diseases: CCS group 213 – this includes ICD-10 codes Q20, Q21, Q22, Q23, Q24, Q25, Q26, Q27 and Q28.
- Nervous system congenital anomalies: CCS group 216 (this includes ICD-10 codes Q00 to Q07 which incorporate conditions such as spina bifida, anencephaly and other congenital malformations of the nervous system).
- Other congenital anomalies and perinatal conditions: CCS groups 224 and 217 (this includes a broad range of congenital anomalies and perinatal conditions with ICD-10 P- and Q- codes, excluding those included within other definitions listed above, such as codes for chronic lung disease).
- Down's syndrome: ICD-10 code Q90.
- Cerebral palsy: ICD-10 code G80.

If a birth record had no gestational age recorded (i.e. premature status was unknown), then the infant was assumed to be not preterm on the basis that they had similarly low intensive care unit admission rates and short length of stay at birth was with infants in the group known to be born at term.

We identified infants admitted with a primary diagnosis of acute bronchiolitis using the 'J21' ICD-10 codes (J210: acute bronchiolitis due to respiratory syncytial virus; J218: acute bronchiolitis due to other specified organisms; J219: acute bronchiolitis, unspecified). We grouped all bronchiolitis codes into a single category. We examined age at bronchiolitis admission and calculated the median length of stay for bronchiolitis admissions. We calculated the absolute risk of a bronchiolitis admission among infants with and without risk factors for severe infection. Infants without a particular risk factor condition were considered 'healthy'. We used Poisson approximation to calculate 95% CI. We calculated the relative risk (RR) of a bronchiolitis admission, with associated 95% CI for infants in each individual risk group, by comparing them with the baseline group of infants without the particular risk factor. Infants may belong to more than one of these risk groups, so we controlled for this potential confounding using Poisson regression models to calculate the adjusted RR of bronchiolitis admission for infants in each risk group. To test for significant differences between proportions we used chi-squared tests and Mann–Whitney *U*-tests to compare median values for non-normal data. Data were analysed using the SAS 9.2 software package.

We have reported our findings in line with the reporting standards for observational research (RECORD statement).[358]

Results

Completeness of Hospital Episode Statistics data (study 1)

The proportions of missing/unknown HES data by field for the period 2005/6 to 2009/10 are shown in *Table 43*. The proportion of missing or unknown data in key birth record fields decreased significantly over the 5-year period; for example, missing gestational age fell from 46.2% in 2005/6 to 18.1% in 2009/10, and birthweight from 43.9% in 2005/6 to 16.9% in 2009/10. Overall, the HES cohort captured 87% of all live births recorded by the ONS in England during the period 2005/6 to 2009/10.

We tested the effect of selecting birth records only from hospitals with high completeness of recording by creating a 2007/8 birth cohort comprising birth records only from hospitals where $\geq 90\%$ of their birth records contained complete recording of the key variables, birthweight and gestational age. The resulting cohort included 296,618 babies born at 71 hospitals across England. *Table 43* shows a comparison of characteristics of included ($n = 71$) and excluded ($n = 85$) hospitals. The mean numbers of births, maternity beds and access to neonatal intensive care (*Table 44*) were mostly similar among hospitals with high and low completeness of recording. The mean maternal age, the proportion of babies of non-white British ethnicity and the proportion of babies in the most deprived quintile were similar among the two groups of hospitals. Full details are provided in Murray *et al.*[360]

Completeness of National Neonatal Research Database and Hospital Episode Statistics data (study 2)

There were 66,403 records in the NNRD of admissions into NHS neonatal specialised care units for the period 1 January to 31 December 2010, of which 66,117 (99.6%) had complete recording of gestational age. For babies with a valid gestational age, all NNRD variables with the exception of maternal age were complete in > 90% of the records. NNRD records represented 9.7% of births identified in HES (683,556 records) for the same period. After removing duplicates and data cleaning, 651,073 babies remained. Of these, 528,671 (81.2%) had a complete recording for gestational age. For babies with a valid gestational age, HES records were complete in over 90% of cases (*Table 45*) with the exception of multiple birth and maternal age at delivery, for which completeness ranged from 84.4% to 91.7% and 77.9% to 87%, respectively, and LSOA, for which completeness was 0.01%.

TABLE 43 Completeness of recording of baby tail fields in HES birth records (2005/6–2009/10)

Baby tail fields in HES birth records (field name)	% missing or unknown					*p*-value
	05/6	06/7	07/8	08/9	09/10	
Anaesthetic given during labour or delivery (delpren)	41.9	41.8	44.8	29.6	16.5	< 0.001
Anaesthetic given post-labour or delivery (delposn)	48.1	46.0	49.6	34.8	21.6	< 0.001
Antenatal days of stay (antedur) (derived from other HES fields)	0.2	0.2	0.2	0.1	0.1	< 0.001
Baby's age in days (neodur) (derived from other HES fields)	0.1	0.0	0.0	0.0	0.0	< 0.001
Birth order (birorder)	33.9	36.9	39.4	24.9	13.7	< 0.001
Birthweight (birweit)	43.9	47.1	50.1	31.3	16.9	< 0.001
Delivery place change reason (delchang)	45.2	45.8	47.4	34.6	21.7	< 0.001
Delivery method (delmeth)	35.1	35.8	44.3	30.6	14.9	< 0.001
Delivery place (actual) (delplace)	44.0	46.8	57.0	41.3	17.9	< 0.001
Delivery place (intended) (delinten)	41.3	42.7	43.6	30.2	14.9	< 0.001
First antenatal assessment date (anasdate)	41.7	44.3	44.6	34.7	20.4	< 0.001
Gestation in weeks at first antenatal assessment (anagest)	54.5	63.9	55.2	45.6	28.3	< 0.001
Length of gestation (gestat)	46.2	54.2	48.0	34.6	18.1	< 0.001
Birth status (birstat)	43.9	47.0	48.0	32.9	16.2	< 0.001
Labour/delivery onset method (delonset)	36.2	37.7	41.1	25.5	11.5	< 0.001
Mother's age at delivery (matage)	42.4	43.3	43.0	34.5	30.5	< 0.001
Neonatal level of care (neocare)	16.1	16.0	17.1	18.4	12.4	< 0.001
Number of babies (numbaby)	31.8	33.3	36.1	23.6	11.9	< 0.001
Resuscitation method (biresus)	44.2	45.3	48.0	34.2	21.1	< 0.001
Status of person conducting delivery (delstat)	38.9	42.6	48.4	33.9	19.1	< 0.001
Total number of births	554	566	575	589	603	–
Total number of births[a]	521	749	493	684	786	–

a Total number of births after removal of duplicate episodes and data cleaning.
Reproduced with permission from Murray.[359]

TABLE 44 Comparison of maternity characteristics between hospitals with high and low completeness of birth admission records,[a] financial year 2007/8

Hospital maternity factors	Hospitals with		*p*-value
	High completeness ($n = 71$)	Low completeness ($n = 85$)	
Mean (SD) number of annual births	3957 (2011)	3465 (1997)	0.13
Mean (SD) number of maternity beds	55.1 (30.3)	55.3 (26.4)	0.96
Mean (SD) occupied maternity beds	35.4 (21.4)	35.2 (18.0)	0.95
Number (%) with neonatal intensive care	52 (73)	68 (80)	0.30
Mean (SD) number of neonatal intensive care cost	10.6 (11.7)	11.4 (10.9)	0.66
Mean maternal age (% missing data)	28.9 (18.4)	29.0 (70.1)	0.39
Proportion of births per hospital in most deprived deprivation score quintile (% missing the data)	0.472 (69.4)	0.435 (56.2)	0.64
Proportion of births of non-white British ethnicity (% missing the data)	0.527 (7.1)	0.564 (5.1)	0.64

a Low completeness of recording defined as hospitals where < 90% of birth admission records contained complete recording of birthweight and gestational age.
Reproduced with permission from Murray.[359]

TABLE 45 Complete records (%) by gestational age in HES birth records and the NNRD

	HES				NNRD			
	Gestation weeks				Gestation weeks			
Key variables	≤ 32	33–36	37–42	≥ 43	≤ 32	33–36	37–42	≥ 43
Infant sex	99.8	99.9	99.9	99.9	100.0	100.0	99.9	100.0
Birthweight	93.9	91.5	91.2	97.1	100.0	100.0	100.0	100.0
Number of babies	91.7	86.4	88.2	84.4	99.9	99.8	99.8	100.0
Maternal ethnicity	95.8	95.0	94.4	94.5	96.6	97.4	94.8	92.5
Maternal age	87.0	78.3	77.9	79.1	66.8	72.6	66.5	50.0
LSOA	0.0	0.0	0.0	0.0	96.0	95.8	93.5	90.0
Total	13,883	28,648	48,3091	3025	10,533	19,300	36,248	40

In HES, the distribution of gestational age and birthweight was discordant with a large number of unlikely outlying combinations, especially for babies born preterm (*Figure 26a*). In the NNRD there were fewer outliers (*Figure 26b*). After excluding babies with values above and below 4 SDSs, the HES distribution approximated more closely to the NNRD (*Figures 26c* and *26d*).

Record linkage, Hospital Episode Statistics and National Neonatal Research Database

We included only neonatal units from which permission had been obtained to receive identifier data at the time of study 2 ($n = 159$). When records from neonatal units that did not give permission to access NHS number were excluded from both sources, 47,345 and 650,301 babies remained eligible for record linkage in the NNRD and HES, respectively. Of 47,345 eligible NNRD records, 44,426 (93.8%) were successfully linked to HES (*Figure 27*). We combined information gained through record linkage by replacing missing values with information from either of the two data sources. After excluding babies with missing values (0.1%; step 2, see *Figure 27*) and babies with birthweight values above or below 4 SDSs (0.3%; step 3,

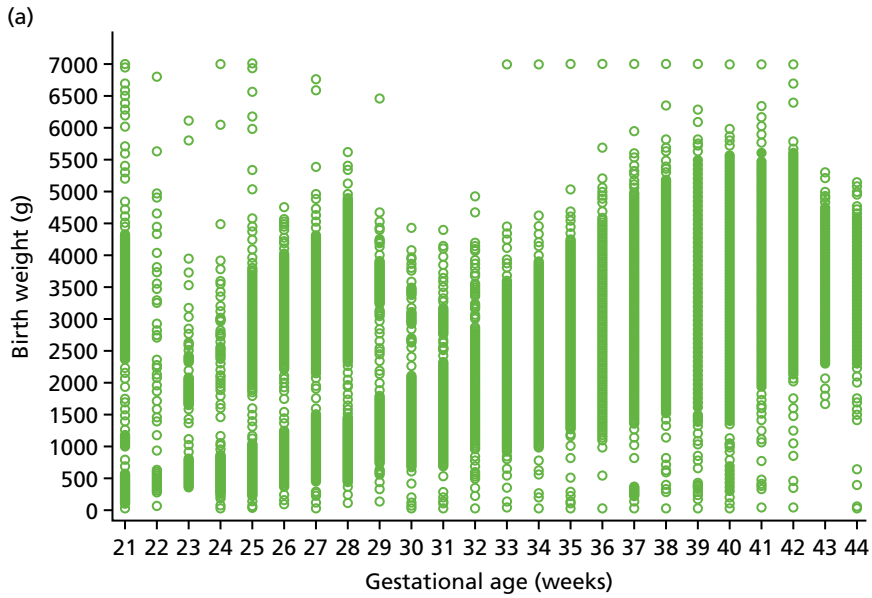


FIGURE 26 Distribution of birthweights by gestational age in HES and the NNRD. (a) All babies HES; (b) NNRD; (c) after excluding birthweights above and below 4 SDS HES; and (d) NNRD. (*continued*)

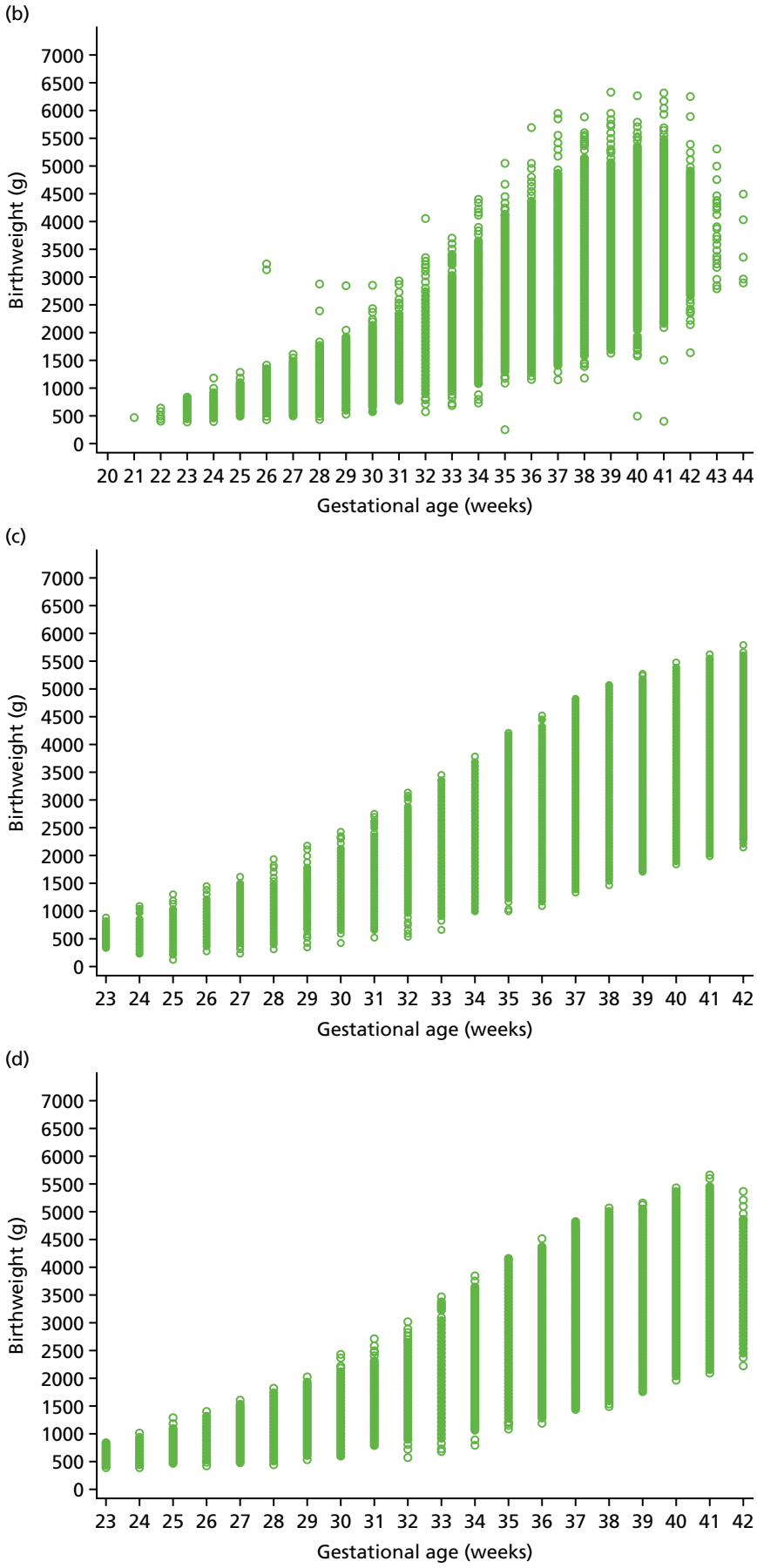


FIGURE 26 Distribution of birthweights by gestational age in HES and the NNRD. (a) All babies HES; (b) NNRD; (c) after excluding birthweights above and below 4 SDS HES; and (d) NNRD.

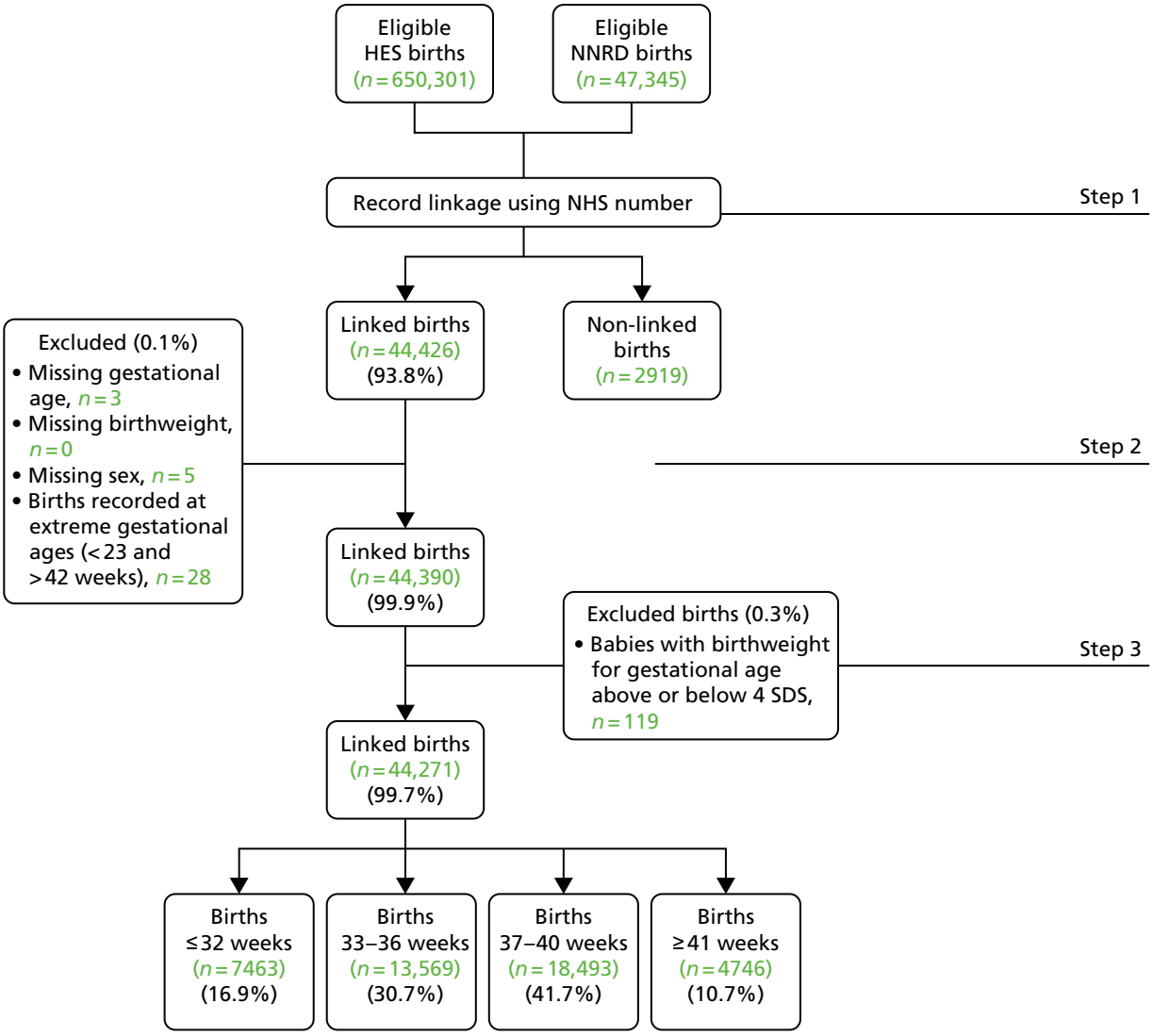


FIGURE 27 Flow diagram of record linkage.

see *Figure 27*), 44,271 records remained. Of these, 7463 (16.9%) were born before 32 weeks' gestation, 13,569 (30.7%) between 33 and 36 weeks, 18,493 (41.7%) between 37 and 40 weeks, and 4746 (10.7%) ≥ 41 weeks' gestation. In the second stage, we excluded babies with missing values and implausible birthweights before linkage (26.4% and 1.5% in HES; 0.1% and 0.2% in the NNRD), leaving 61.3% of babies in the NNRD successfully linked to HES.

Table 46 shows the comparison of characteristics of babies linked ($n = 44,426$) and unlinked ($n = 2919$). Multiplicity at birth, birthweight and gestational age differed significantly between linked and unlinked babies. The unlinked group had more multiple births (unlinked 21.3%; linked 13.9%), extremely preterm births below 32 weeks' gestational age (unlinked 24.1%; linked 16.8%), and lower birthweights (unlinked 2.527 kg; linked 2.682 kg). In a sensitivity analysis, restricted to singleton babies (unlinked 2290; linked 38,176), birthweight and gestational age remained significantly different between linked and unlinked groups.

Agreement between Hospital Episode Statistics and National Neonatal Research Database

Table 47 shows the level of agreement between HES and NNRD for the final linked birth cohort. For the key variables studied, overall agreement was > 95% (kappa coefficient 0.97 to 0.99) with the exception of gestational age (81.0%) and maternal ethnicity (86.1%) (kappa 0.71 and 0.79, respectively).

TABLE 46 Characteristics of NNRD babies linked and unlinked to HES (first linkage prior to removal of implausible values)

Fields	NNRD babies, *n* (%)		*p*-value
	Linked to HES	Not linked	
Infant sex			0.012
Male	25,182 (56.7)	1584 (54.4)	
Female	19,200 (43.3)	1330 (45.6)	
Gestational age (weeks)			< 0.0001
≤ 32	7479 (16.8)	702 (24.1)	
33–36	13,602 (30.6)	804 (27.6)	
37–42	23,303 (52.5)	1411 (48.4)	
≥ 43	20 (0.1)	0 (0)	
Birthweight, mean (SD)	2681.5 (939.9)	2527.4 (1024)	< 0.0001
Multiple birth			< 0.0001
Singleton	38,208 (80.8)	2294 (78. 7)	
Multiple births	6185 (13.9)	622 (21.3)	
Maternal age, mean (SD)	29.4 (6.3)	29.7 (6.4)	0.018
Maternal ethnicity			0.75
White British	30,108 (72.5)	1947 (72.2)	
Non-white British	11,401 (27.5)	748 (27.8)	
IMD score quintiles			0.43
1 (most deprived)	8568 (20.1)	530 (19.2)	
2	8552 (20.0)	548 (19.9)	
3	8568 (20.1)	530 (19.2)	
4	8523 (19.9)	577 (20.9)	
5 (least deprived)	8527 (20.0)	572 (20.8)	
Total	44,426	2919	

TABLE 47 Agreement between HES and the NNRD

Fields	Overall agreement (%)	Kappa coefficient	*p*-value[a]
Infant sex	99.5	0.99	< 0.0001
Gestational age	81.0	0.79	< 0.0001
Birthweight	98.1	0.98	< 0.0001
Number of babies	98.2	0.92	< 0.0001
Maternal age	99.3	0.99	< 0.0001
Maternal ethnicity	86.1	0.71	< 0.0001
LSOA	96.9	0.94	< 0.0001

a The test statistic is a *z*-score.

Admissions with bronchiolitis (study 1)

The birth cohort included 296,618 infants from 71 NHS hospitals in England; 410 infants in the cohort died during the study year; 51% (151,897/296,618) were boys, 1% (2,891) were multiple births and 7.5% (22,215) were born preterm before 37 weeks' gestation. We identified 7189 admissions to hospital over the period 1 April 2007 to 31 March 2008 with a primary diagnosis of bronchiolitis up to the age of 1 year [admissions per 1000 infants 24.2 (95% CI 23.7 to 24.8)]. Of these, 2015 (28.0%) were specifically coded as being due to RSV and the remainder were coded 'unspecific aetiology'. In total, 1529 (21.3%) infants had more than one bronchiolitis admission during their first year. The modal age for bronchiolitis admission was 1 month and the median age was 120 days (IQR 61 to 209 days). The median length of hospital stay was 1 day (IQR 0 to 3 days). The majority of infants admitted with bronchiolitis were not in recognised high-risk groups, with only 24% (1722/7189) having one or more recognised risk factors for severe infection.

Discussion

We conducted a series of analyses involving the NNRD containing point-of-care, clinician-entered health-care data, and HES maternity records containing administrative data. We found that the completeness of HES birth records varies substantially between hospitals but has improved over time. Data quality and completeness of recording were better in the NNRD than in HES for most key variables, including gestational age. We demonstrated the feasibility of record linkage between HES data and the NNRD. We also showed the feasibility of linkage of HES records across time to quantify and describe the burden of bronchiolitis, an important infectious disease of infancy. Our work provides proof of principle that routine NHS data sources may be utilised to create national longitudinal birth cohorts and that combining HES with the NNRD can substantially enhance the quality and scope of birth records. Our methods pave the way for future studies to support research, ascertainment of longer-term outcomes of babies admitted to neonatal units, and the delivery of neonatal specialised care, a high-cost, nationally commissioned clinical service.

We accept that our comparisons were limited to specific fields; the large number of missing data is a further consideration. The distribution of birthweight by gestational age revealed a large number of inconsistent values in HES compared with a more plausible distribution in the NNRD. Removal of implausible birthweights, prior linkage, resulted in a reduced NNRD to HES linkage rate comprising approximately two-thirds of the total number of babies. We performed only deterministic linkage, and probabilistic linkage is an alternative strategy. Linkage using a unique identifier, such as the NHS number, is considered highly acceptable with the greatest face validity, but combining probabilistic with deterministic linkage might have increased the linkage rate. Another limitation of our study is that we were unable to determine how accurately the HES identifier (the patient key) we used to link records across time is allocated to unique individuals.

We have shown that it is possible to identify infants up to the age of 1 year admitted to NHS hospitals and to report on the population burden of an important infant condition: bronchiolitis. Just over one-fifth of infants admitted with bronchiolitis had a further admission for the same condition during their first year. Our study has highlighted that the burden of bronchiolitis hospital admissions among infants in England predominantly affects those born at term, without any risk factors for severe infection, and the age at admission appears to be significantly lower now than previously reported. Although risk of admission is higher in known risk groups, 85% of infants admitted to hospital with bronchiolitis in England have no known predisposing risk factors. We also found that infants with Down's syndrome, cerebral palsy and cystic fibrosis appear to be at higher risk of hospital admission. Our findings in study 2, namely that HES data were more likely to be missing for preterm babies, indicate that this was a weakness of our analysis. Other important limitations are that our case definition for bronchiolitis and comorbidity was dependent on the accuracy of clinical coding and recording in diagnosis fields in HES records. We combined RSV and unspecified bronchiolitis, presenting data on all bronchiolitis admissions. Only 28% of the bronchiolitis admissions were coded as being due to RSV, and the remainder had an unspecific bronchiolitis code. We found that the median length of stay for bronchiolitis admissions was only 1 day, suggesting that improved management in the community may reduce the need for admission.

Our study extends earlier work by investigating not only the completeness but also the quality of HES data. We found that about one-fifth of babies in HES have missing gestational age data,[360–362] but a novel finding is that HES had 1.5% of recorded birthweights outside a biologically plausible range. Our analyses also showed that infants with missing HES data were more likely to have been very preterm (< 32 weeks' gestation) and have a lower birthweight. A possible explanation is that infants born very preterm are more likely to be missed as a birth in HES registrations and coded as a new admission, as they are admitted directly to a neonatal unit.

Linkage of HES to other data sources has been explored in previous studies. Dattani *et al.*[361] linked maternity HES, birth registration data and NHS 'numbers for babies' to assess the quality and completeness of these sources, using the NHS number in combination with other information as identifiers, and achieved a similar linkage rate of > 90%. Hockley *et al.*[362] applied a variety of linkage methods for data from Scotland, Wales and England and concluded that the use of a probabilistic method for data from England would not have improved the linkage rate because of the poor completeness of HES data. This supports our choice of using a deterministic method with a single identifier for linkage.

The effectiveness of routine health record linkage in adults has been demonstrated in Australia and Canada, where it has improved both data quality and utility.[363–366] Birth records have been successfully linked to hospital discharge data in Australia, with matching rates of 99%.[367] In several regions of the USA, data from birth certificates have been linked to hospital discharge records to examine maternal outcomes.[368,369] In Scandinavian countries, the assignment of unique personal identification numbers permits linkage between civil registration systems and enables the development of population-based cohorts. These have facilitated a broad array of epidemiological studies such as investigations of the impact of place of birth, familial risk factors for autism[370] and the association between prenatal exposures and ADHD in childhood.[371] In Wales, the Secure Anonymised Information Linkage databank brings together anonymised person-based electronic health and social care data. This is being combined to establish an anonymised Wales-wide Electronic Cohort for Children.[372,373] This databank has been used successfully to identify potential clinical trial participants from primary care data.[372–374] The Scottish Health Informatics Programme (SHIP) is an example of a complex database of linked EPRs, providing health and social care information from birth through to death.[375] To date, SHIP data have primarily been used to conduct pharmacovigilance and diabetes epidemiology research.[375,376] Another UK cohort is the Oxford Record Linkage Study, established in 1963 and comprising > 10 million records on around 5 million people.[377] This has been used in longitudinal research studies to identify maternal and perinatal risk factors for conditions such as inflammatory bowel disease,[378] asthma[379] and coeliac disease.[380]

Other aspects of routine NHS data are worth mentioning. Stand-alone maternity systems in around 20 hospitals are not linked to their patient administration system, from which HES data are obtained.[380–382] Some hospitals return data on birth or delivery episodes but not both, and stillbirths are neither reliably recorded in every hospital nor allocated a NHS number.[382,383] We suggest that future studies involving HES records are likely to benefit from steps to check data quality as well as completeness. Improvement in administrative data quality and completeness are important health services goals. Reducing reliance on administrative data, promoting clinician involvement in data assurance and, as is the case with the NNRD, extracting maternity data from electronic health records are measures that also merit wider consideration.

Implications for health care

We have shown that EPR data can be used to create UK birth cohorts. The sole use of the NHS administrative database, HES, to build birth cohorts will result in the inclusion of many babies with missing gestational age data and implausible birthweights. However, linking the NNRD to HES substantially enhances the quality and scope of UK birth records. Improvements in administrative and clinical data quality and completeness are important health services goals.

Research recommendations

Researchers who are planning to study the association of gestational age with specific outcomes in childhood may find it helpful to use the NNRD. If HES data are used, we recommend removing implausible birthweights for gestational age from the study cohort in order to improve the validity of the conclusions.

More exploration is needed to exploit the use of the NNRD linked to HES to understand the epidemiology and health-care resource utilisation of conditions, such as bronchiolitis in healthy infants and infants with multiple comorbidities, and to examine the long-term independent impacts of preterm births and such conditions on health in childhood.

Chapter 8 Parent involvement in the National Neonatal Research Database

Abstract

Background: Parents (or legal guardians) have a primary responsibility for contributing to the current debate on the use of clinical data in research on behalf of their infants.

Aims: We aimed to establish a PPI group of parents with experience of a baby in neonatal care. We aimed to co-design with this group a survey for parents of infants admitted to neonatal units in England in order to obtain their views.

Methods: We undertook a review of the literature on public understanding of the use of clinical data for research purposes and identified parameters of relevance to the intended survey. We established and supported a PPI group to co-design a survey. Research nurses at each of 28 participating hospitals approached potential parent participants to explain the study, provide written information (available in eight languages) and seek consent (see *Appendix 5*).

Results: The survey was completed by 1319 parents or primary carers. Overall, there was a very high level of support for the use of health data for research purposes, with parents of babies who had experienced higher intensity care more likely to say 'yes'. Over 80% and 85% of respondents respectively were very or fairly confident about data security and accuracy. We identified a high level of altruism. Nearly two-thirds agreed with 'opt out' as the default position for data-sharing.

Conclusions: There is strong parent support for sharing health data for research. The identification of effective and efficient methods to improve knowledge of potential benefits, processes and regulation are important to secure trust and confidence in the use of clinical data in research.

Background

The use of routinely recorded clinical data for research purposes is a key concern of contemporary e-health policy, research governance and public debate.[384] For infants and young children, without autonomous decision-making capacity, it is their parents (or legal guardians) whose voices represent their contribution to the debate on the values, benefits, risks and uncertainties of permitting use of their clinical data for research. In the case of the neonatal care population, at the time of the commencement of the Medicines for Neonates programme, the attitudes of parents were unknown as we had been unable to identify any prior study in this specific clinical care context.[385] Influences on the diversity of parental attitudes, the acceptability of data use and what, if anything, might be particular to the neonatal care context were unexplored.

Aims and objectives

We aimed to conduct a survey of attitudes in relation to the use of personal data in research of parents whose children were in receipt of neonatal care services across NHS sites in England. Our objectives were to:

- explore the relevance, in the newborn context, of dimensions of public attitudes to data used for research purposes that have been previously identified from the literature
- identify additional issues and concerns from the perspective of parents that are relevant to the newborn context
- identify advantages, disadvantages and preferences for communication of information and knowledge about the use of data for research purposes.

Methods

Survey instrument design, and patient (parent) and public involvement

Two strands of work informed the design of the survey instrument that was to be completed by parents (and/or primary carers) who had a baby or babies in neonatal care.

First, we undertook an initial review of literature concerning public understanding of health data use for research purposes and contemporary e-health policy and identified 10 parameters of relevance to the survey design:[385]

1. being specific about what counts as routinely collected health/patient/clinical data in the context of the investigation
2. the influence of the digital and e-format of data on patient and public understanding and attitudes
3. data use in the context of protection, promotion and prevention
4. informed consent in relation to identifiable versus de-identified data
5. personal benefit, indirect benefit and altruism
6. the framing of routine data use for research purposes within professional discourse(s)
7. informed choice may not be synonymous with informed consent
8. privacy and confidentiality are distinct but related
9. rational and emotional approaches to decision-making
10. the balance of rights and responsibilities.

Second, we established a PPI group (10 mothers and one father) with previous experience of a child in neonatal care to be supported to co-design the survey instrument. Parents were recruited following response to an advert sent to community groups, online support groups and through Bliss networks. We specifically targeted two geographical areas, in the North West and the West Midlands, where we could reasonably assume a diverse population and which had well-established Bliss support groups. Advertising and recruitment and information materials were made available in eight languages in addition to English (Urdu, Punjabi, Bangla, Mandarin, Somali, Polish, French and Spanish). The two groups met separately on the first two occasions, then jointly for the subsequent three meetings. Each meeting lasted 4 hours with suitable breaks facilitated by one of the research team.

The study was approved by the University of Manchester Research Ethics Committee (reference 10/H1013/35). The aims of the parent groups were to:

- inform the research team, based on personal experience, of the key questions, benefits and concerns associated with the routine use of babies' clinical data for research purposes and thus contribute to the content of the survey items
- provide guidance to the research team on the format of the survey including how to ask questions, in which order and why to maximise uptake, increase clarity and minimise any potential distress to parents completing the survey.

The research team's responsibility was conceived of as:

- facilitating information and experience sharing relevant to the study, in a manner that supported all who were involved
- equipping parents with additional knowledge skills, should they not already possess them, that would enable them to fully participate in the instrument design process supported by the researcher
- creating an approach to co-design that enabled confident challenge of any pre-existing assumptions, respected differences of opinion and valued equally a wide range of contributions.

The approach to parent involvement in research design was, therefore, based on a participatory model that was not merely consultation on what had been pre-designed, but rather involved an active contribution to both process and output from those involved.[385] That said, the overall aims and objectives of the research study were already set and this group of parents had not been involved at the outset. The topics covered in the meetings were (1) introductions, personal experiences and initial thoughts on data-sharing, (2) introduction to research methods, research design and the role of the research nurse, (3) asking questions and creating questions in a written format, (4) testing out the draft questionnaire and recruitment materials, and (5) evaluation of the process, outcomes and future plans.

From the parent groups' perspective, there were seven key issues that influenced the final content, format and design of the questionnaire that was used:

1. The use of a personal 'voice' throughout the questionnaire – this meant that the questionnaire was written using 'we' in the instruction sections and there was an explicit commentary throughout, which spoke directly to the person who would be filling in the questionnaire. For example: 'In this section we want to hear about your attitudes towards . . .' and 'The first three questions might seem very similar, but they are looking at slightly different circumstances so please answer all 3.'
2. Producing the questionnaire as a booklet – it was considered important that parents were not confronted with something that looked like a form because they would be so used to filling out lots of similar items in their stay in hospital. The questionnaire had to look different and less official.
3. The use of colour – colours were considered important not just so that the questionnaire looked attractive but also that the colours should be muted and gentle to create a soft impression. In addition, each of the four sections was assigned its own colour and the answers were marked against the coloured background.
4. Order of questions – the group recommended that potentially the most distressing questions should be left to the end, such as those concerning previous miscarriages or infant deaths. Furthermore, they recommended that such questions were explicitly marked 'sensitive' so that a parent would be warned in advance and could choose not to complete them if they wished.
5. Font – the font that was chosen for the questionnaire was one that was regarded as less formal looking (Comic Sans MS) to create a more welcoming feel to the questionnaire.
6. Options for completion – the parent group felt it was important to emphasise that there was not one right answer to anything and not one way to complete the questionnaire that was preferable to another. Therefore, throughout the questionnaire there were occasional reminders that there were no right or wrong answers and plenty of spaces for any additional comments. In addition, parents could choose to complete the questionnaire anonymously or leave their contact details, they could request it in a written language other than English, they could request an interpreter to complete it with them or they could state that a family member had assisted them.
7. Clarity about words and phrases – the parent group was particularly helpful in spotting jargon and suggesting simpler alternatives as well as making sure that key terms were well defined within with the questionnaire; for example:

    i. 'Data' refers to all sorts of information that is collected, from birthweight to drugs administered, to the progress your baby is making and so on. This might be recorded on a database or in paper notes. We are asking about information only, not tissue samples, etc.
    ii. 'Research' refers to the process of collecting, ordering and evaluating information in order to provide further understanding, new knowledge and/or a basis for decision-making and action or change. This might include research on the frequency of disease in babies (epidemiology), on the safety of drugs prescribed to babies (drug safety), on the impact of drugs or treatments on babies' health (clinical effectiveness), or to identify babies with certain specific diseases for inclusion in research studies.

The survey content and design was tested with the parent group at each stage of design.

Survey distribution and recruitment

With the support of the Greater Manchester, Lancashire and South Cumbria Medicines for Children Research Network, 29 NHS hospitals in England with neonatal care units were recruited as research sites (three of which were in London). One hospital dropped out of the study as it was unable to recruit a research nurse to assist with the study; therefore, data are presented from 28 hospitals.

Research ethics approval was obtained from the National Research Ethics Committee North West Cheshire (REC reference 11/NW/0765; UK Clinical Research Network Portfolio ID: 11960).

Research nurses at each site were responsible for approaching parents in person while they were still on the ward to explain the study, provide written information (available in English, Urdu, Punjabi, Bangla, Cantonese and Polish and on request any other additional language) and/or go through the information in person or through a translator (to overcome any potential literacy difficulties). They were also responsible for taking consent and distributing the survey subsequently and arranging any language support required to complete it. The principal inclusion criterion was that a parent had a child in any level of neonatal care at any of the 28 participating hospitals. The survey questionnaire could be completed by the mother and/or the father or other principal carer (i.e. someone who was not the child's other biological parent but who would play a significant role in their care and upbringing). This could be a partner of the same sex or a grandparent, for example. Participants were invited to complete the questionnaire alone, or with their partner (other carer) and/or through an interpreter if required. The research nurse was available to clarify any questions and to provide support if requested. Parents could complete the questionnaire on the ward before they left or take it home and send it in later.

The majority of the questionnaires were completed in hospital ($n = 1090$). A total of 1225 participants (92.9%) did the questionnaire on their own; 80 (6.1%) reported that they had not filled it in on their own, and there were missing data in questionnaires from 14 people (1.1%). Of those who reported that they did not fill the questionnaire in on their own, the majority ($n = 63$) reported that they had filled it in with their partner, 16 stated that a research nurse/midwife had helped them, and one person said that they had done it with an interpreter. When asked if an interpreter had assisted with the questionnaire, four people responded (three had completed it with a Polish/English interpreter and one had completed it with an Urdu/English interpreter).

Sample size

The target sample size was 1300, allowing a percentage to be estimated with a margin of error $\leq 2.5\%$ for 95% CI. Written consent to take part in the study was provided by 1722 people and 1319 completed questionnaires were received (return rate = 76.6%). The discrepancy between number of consents and number of questionnaires completed is largely explained by those parents who gave consent while in hospital but did not complete the questionnaire before discharge. A breakdown of participating sites and completed questionnaires is shown in *Table 48*.

Of the sites ($n = 27$) where neonatal unit admissions data are available for the recruitment period (November 2011 to September 2012), total admissions were 10,983 and the sample size total was 1291 (11.75% of all admissions). The sample included three special care baby units (SCBUs) (level 1), 15 local neonatal units (LNUs) (level 2) and 12 neonatal intensive care units (NICUs) (level 3).

Sample characteristics

Of the 1319 parents or carers who completed a questionnaire, 930 (70.5%) were mothers, 370 (28%) were fathers and 12 were others who identified themselves as having a primary care responsibility for the baby who was in neonatal care (including 10 who were grandparents). Data were missing in seven cases. The median age of the mothers in our sample was 30 years (range 15–52 years). This compares favourably with all mothers in the 28 sample sites ($n = 10,983$; median age 30 years; range 13–55 years) and all mothers recorded in the NNRD encompassing 167 neonatal units in England during the period 1 November 2011 to 30 September 2012 ($n = 55,731$; median age 31 years; range 12–59 years). The median age of fathers in our

TABLE 48 Participating sites

Site number	Neonatal unit designation[a]	Questionnaires received	Consent forms received	Recruitment target	Neonatal admissions 1 November 2011–30 September 2012
1	LNU	52	60	55	168
2	–	28	41	25	Data not available
3	SCBU + LNU	12	12	15	869
4	LNU	55	77	50	234
5	NICU	53	74	50	546
6	NICU	70	103	70	483
7	LNU	23	30	20	182
8	LNU	80	81	70	404
9	NICU	48	99	55	437
10	LNU	48	52	50	415
11	NICU	58	79	60	545
12	NICU	97	152	95	724
13	LNU	30	30	30	226
14	NICU	66	73	30	1069
15	LNU	20	22	30	119
16	NICU	71	72	55	669
17	NICU	58	75	55	471
18	SCBU + LNU	47	57	30	200
19	NICU	20	52	20	352
20	SCBU + LNU	28	30	20	435
21	NICU	65	108	60	425
22	NICU	14	14	25	446
23	LNU	20	26	20	223
24	LNU	35	36	30	242
25	LNU	86	87	70	297
26[b]	–	0	0	25	Site withdrew
27	LNU	82	98	80	206
28	NICU	21	28	30	303
29	LNU	32	54	30	293
Total	–	1319	1722	1255	10,983

LBU, local neonatal unit; NICU, neonatal intensive care unit; SCBU, special care baby unit.
a Neonatal unit designations in accordance with BAPM criteria.[386]
b Included here for recruitment target, but discounted in data presentation as withdrew.

sample was 32 years (range 15–52 years). Of the 1273 returns for which data are available, the baby in neonatal care was the first child for 45.6% ($n = 601$) of participants and the second child for 31.6% ($n = 417$) of participants.

The vast majority of participants described their ethnic group as white British (82.5%; 1088/1290). Over half of the sample described themselves as Christian (56%; 732/1306), with an additional one-third preferring to state that they had no religion (33%; 433/1306). Ninety-seven participants stated that they had not been born in the UK with over half ($n = 50$) arriving in this country between 2005 and 2011. A comparison of mothers' ethnicity in the sample with data from the 28 sites overall and that of the NNRD overall for the study period (1 November 2011 to 30 September 2012) reveals some differences in representativeness. The study sample has a greater proportion of mothers who are 'white British' (81.4%) than either the 28 sites overall (56.5%) or the NNRD (65.3%); mothers of Asian ethnicity (7.4%) are slightly under-represented, as are those who are black (2.3%), in comparison both with the 28 sites and with the NNRD (see *Table 70*).

Most participants described themselves as married, in a civil partnership or in a relationship (93.3%; 1209/1296). The proportion that were married or in a civil partnership (42.3%, 393/930) is consistent with the proportion recorded in NNRD for the same time period of the study (40.6%, 22,643/55,731). Two-thirds of participants were in employment either full time or part time (65.8%; 868/1316). Of the 1218 who were prepared to provide information on their educational qualifications, one-fifth (21%; 251/1218) had no qualifications beyond O level (ordinary level)/GCSE (General Certificate of Secondary Education), with five of those declaring no qualifications whatsoever. Nearly one-third of participants had a university or other higher degree (386/1218; 32%).

At the time of completing the questionnaire, the amount of time that participants' babies had been in neonatal care ranged from 1 to 217 days (mean 20.4 days; median 11 days). Parents were asked to state the highest level of care that their babies had received at any point. Just under half (609/1288; 47.5%) had experience of level 3 (intensive care), 15.5% (199) of level 2 (high-dependency care) and 30% (387/1288) of level 1 (special care). A total of 93 (7.2%) participants reported that they did not know what level of care their baby had received.

Participants displayed very high levels of satisfaction with their experiences of neonatal care; on an ordinal scale of 1 (least satisfied) to 7 (most satisfied), 84.8% ($n = 1119$) responded either 6 or 7. When asked if they thought that their level of satisfaction with care had influenced how they had responded to the survey questions about the routine use of their baby's data for research purposes, opinion was divided, with a slight majority responding 'no' (391 responded 'yes'; 511 responded 'no'; 326 responded 'possibly'; 59 responded 'don't know').

Participants were also asked about how 'included' they felt in their baby's care and scored feelings of inclusion on an ordinal scale of 1 (least included) to 7 (most included). Of 1297 available sets of data, the majority of parents (638; 49.2%) reported that they felt most included (scoring 7). The mean and median scores were 6.4 and 7.0, respectively. The question was also one that the parent group had requested to be added, because they identified this feeling of inclusion as a key marker of quality provision at the time when their baby was still on the neonatal unit.

Summary

- The sample constitutes 11.75% of admissions to the neonatal units in the study during the recruitment period (total admissions = 10,983, sample size total = 1291).
- The sample includes three SCBUs (level 1), 15 LNUs (level 2) and 12 NICUs (level 3).
- The age of mothers in the sample is consistent with mothers overall for the 28 neonatal units and for all units within the NNRD.

- One-fifth (21%; 251) of the sample had no qualifications beyond O level/GCSE and nearly one-third of participants had a university or other higher degree (386/1218; 32%).
- The study sample has a greater proportion of mothers who are 'white British' than either the 28 sites overall or those in the NNRD; mothers of Asian ethnicity are slightly under-represented, as are those who are black, in comparison both with the 28 sites and with the NNRD.
- The vast majority of participants felt positively included in their baby's care and displayed very high levels of satisfaction with their experience of neonatal care, but were divided on whether or not these experiences would influence their attitudes to routine use of their baby's data for research purposes.

Results

Willingness for data-sharing for research purposes

Parents were asked about their willingness for their baby's data to be used for research purposes in three different conditions: (1) in general, (2) if identifying information was removed, and (3) if explicit permission was asked on each occasion. Overall, there was a very high level of support for the routine use of health data for research purposes (69.4%), increasing to nearly 80% 'yes' responses if identifying information were removed. If permission was asked each time, the percentage agreement was 77%; 847 participants (68.9%) responded 'yes' to all three questions (*Table 49*).

A statistically significant association was found between participants' responses when asked about data-sharing in the different conditions offered: willingness in general and willingness if identifying information was removed; willingness in general and willingness if permission were asked; willingness for non-identifiable data to be used and willingness if permission were asked. In each case, the association was significant ($p < 0.001$), with each of the kappa values approaching 0.60.

A comparison of individual participants' responses to each question reveals that, if identifying information were removed, 128 participants changed their response from 'possibly' in the general condition to 'yes' in the non-identifiable condition, and 25 participants changed from 'no' to 'yes'. If permission were asked, 115 changed their response from 'possibly' to 'yes' in comparison with the general condition. However, of those who said 'yes' if identifying information were removed, 89 said 'possibly' instead to the condition of permission being asked each time.

The association between participants' highest level of qualification and their willingness for their babies' data to be used for research if identifying information were to be removed was found to be significant ($\chi^2_{trend} = 7.625$; $p = 0.022$). Participants who said 'yes' to using de-identified information about their baby for research were likely to be those whose highest qualification was a university or higher degree. Those with O levels or GCSEs selected 'possibly' more than those in the two other groups. Overall, 83.3% responded 'yes' to willingness for de-identified data to be shared, 12.6% responded 'possibly' and 4.2% responded 'no' (see *Table 71*).

TABLE 49 Participants' willingness for their baby's data to be used for research purposes

Participants' willingness for their baby's data to be used for research purposes	Participants' willingness for their baby's data to be used for research purposes, *n* (%)				Missing data, *n* (%)
	Yes	No	Possibly	Don't know	
In general	915 (69.4)	84 (6.4)	262 (19.9)	41 (3.1)	17 (1.3)
If identifying information was removed	1052 (79.8)	59 (4.5)	163 (12.4)	30 (2.3)	15 (1.1)
If permission was asked each time	1015 (77)	50 (3.8)	217 (12.4)	31 (2.4)	6 (0.5)

In addition, the association between participants' levels of qualification and their awareness of electronic health records prior to the study was found to be significant ($\chi^1_t = 119.26$, df = 2; $p = 0.000$). Those with O levels or GCSEs were more likely to say 'no' (66.2%) (they had not heard of electronic records prior to participation) than those in the other two groups, and those with degrees or higher degrees were more likely to say 'yes' (76.6%) (see *Table 72*).

Of those who reported that they had heard of electronic health records, the NHS was the most common source of this information ($n = 374$), followed by the media ($n = 271$).

We investigated whether or not there was an association between willingness for data to be used for research purposes (defined as 'yes', 'possibly', 'no') and whether or not parents had one child (the one in neonatal care) or more than one child (the youngest being in neonatal care) (see *Table 73*).

A chi-squared trend test showed that the association between having one child or more than one child was significant with regard to participants' willingness for their babies' data to be used for research ($\chi^1_{trend} = 9.32$, df = 1; $p = 0.002$). More people with more than one child ($n = 496$) (i.e. the child in neonatal care was not their first child) said that they would be willing for their baby's data to be used for research compared to those who had only one child ($n = 395$).

In the case of willingness for data to be used for research purposes if they were de-identified (see *Table 74*), participants who said 'yes' were more likely to be those who had more than one child ($n = 555$) (i.e. the child in neonatal care was not their only child). This association was significant ($\chi^1_{trend} = 4.14$, df = 1; $p = 0.042$).

We investigated whether or not there was an association between the highest level of neonatal care experienced and willingness for data to be used for research purposes. Level of care refers to the highest level of care a baby had experienced at any point. It does not necessarily refer to the duration of that care or to the current level of care at the point of completing the questionnaire (*Table 50*).

TABLE 50 Frequency (%) of participants' responses to questions about willingness to use baby's data for research purposes

Statement	Highest level of care experienced	Willingness, frequency (%)			Total
		Yes	Possibly	No	
Willingness for baby's data to be used for health research	1	264 (71.5)	74 (20.1)	31 (8.4)	369
	2	131 (67.2)	51 (26.2)	13 (6.7)	195
	3	443 (76.2)	111 (19.1)	27 (4.6)	581
	Total	838 (73.2)	236 (20.6)	71 (6.2)	1145
Willingness for baby's de-identified data to be used for health research	1	289 (77.1)	65 (17.3)	21 (5.6)	375
	2	155 (79.9)	28 (14.4)	11 (5.7)	194
	3	515 (87.7)	55 (9.4)	17 (2.9)	587
	Total	959 (83.0)	148 (12.8)	49 (4.2)	1156
Willingness for baby's data to be used for research if asked for permission	1	284 (75.3)	73 (19.4)	20 (5.3)	377
	2	152 (77.9)	38 (19.5)	5 (2.6)	195
	3	488 (83.0)	84 (14.3)	16 (2.7)	588
	Total	924 (79.7)	195 (16.8)	41 (3.5)	1160

According to the results of a Kruskal–Wallis test, the association between the level of a baby's care experienced and the parents' willingness for their baby's data to be used for research was found to be significant ($\chi^2_{trend} = 7.218$; $p = 0.027$). Participants whose babies had experienced level 1 care were more likely to say 'no' than those who had experienced levels 2 or 3, and those with babies who had experienced level 3 care were more likely to say 'yes' than those whose babies had not. With regard to participants' willingness for their baby's de-identified data to be used for research, the level of care was found to be significant ($\chi^2_{trend} = 19.963$; $p = 0.000$). Participants who had babies in level 3 care were more willing to say 'yes' if any identifying information were removed, whereas those who had babies in level 1 were more likely to say 'no'. The association of the level of care with participants' willingness for their baby's data to be used if asked permission was also found to be significant ($\chi^2_{trend} = 9.111$; $p = 0.011$); those with babies who had experienced level 3 care said 'yes' more than those with babies with experience of levels 1 and 2.

We also investigated willingness for babies' data to be used for research purposes in the case of participants who had lost a baby previously through termination, miscarriage or stillbirth. The question referred specifically to data that might be associated with the lost child. Of the 611 parents who responded, 315 (51.5%) said 'yes', 121 (19.8%) said 'no', 140 (22.9%) said 'possibly' and 35 (5.7%) said they did not know.

Consent

Opt-out system

Participants expressed strong support for an 'opt-out system', described as 'your baby's data would be used for research unless you actively said you didn't want this to happen'. Almost two-thirds thought that this was a 'good idea' (802/1307; 61.4%), with an additional 15.6% (203/1307) thinking that it was 'possibly' a good idea and fewer than one-fifth clearly saying 'no' (229/1307; 17.5%). Some participants expressed concern that, if an opt-out system were in place, parents might not fully understand that their data would be used unless they explicitly opted out, or that distressed parents might tick an 'opt-out' option without really understanding what it implied.

Associations between qualification-based groups and their responses regarding whether or not it would be a good idea to use an opt-out system were found to be significant ($\chi^4 = 18.768$; $p = 0.001$). Participants who had university or other higher degrees were more likely to say 'no' (25.1%) to an opt-out system, whereas participants who had O levels or GCSEs as their highest qualification were more likely to say 'possibly' (19.8%). Those with advanced levels (A levels) or vocational qualifications as their highest qualification were more likely to say 'yes' (70%) to an opt-out system (see *Table 75* and *Figure 55*).

No significant differences in response to the question concerning the opt-out system were found in relation to the number of children, the highest level of baby's care or the relationship status.

Despite strong support for an opt-out system, when asked 'Are there any occasions when it would be OK to use a baby's data for research without asking parents?', nearly two-thirds (806/1301; 62%) responded that there was no occasion on which this would be acceptable, although just over one-third responded 'yes' (211/1301; 16.2%) or 'possibly' (237/1301; 17.4%), with the remainder responding 'don't know' (57/1301; 4.4%).

If specific permission were requested

Participants were also asked about their preferences if, instead of an opt-out system, parents were to be asked specific permission (i.e. consent for the use of their baby's data) (see *Table 76*). When asked whether the way in which permission might be sought (rather than what they personally might prefer) would make a difference to whether or not they consented to the use of their baby's data for research purposes, there was no strong trend: 398 participants (30.2%) said 'yes', 486 participants (36.8%) said 'no', 356 participants (27%) said 'possibly', and the rest [79 (6%)] said that they did 'not know' or they did not answer.

If asked specifically for consent (rather than an opt-out system), half of them would prefer to be asked in writing (658/1299; 50.7%) and around one-quarter would prefer to be asked in person (349/1299; 26.9%). Some people commented that being asked in person is more personal and presents opportunities for clarification, whereas others said that being asked in writing could be useful as a record and would give them time to digest the information.

The association between the highest level of care experienced and participants' preferences (in person or in writing) if permission were requested was found to be significant ($\chi^2_{trend} = 8.84$; $p = 0.012$). Overall, more participants preferred to be asked in writing than in person, but as their experience of level of care increased, there was a significant rise in the percentage requesting to be asked in person ($\chi^2_{trend} = 8.83$, df = 1; $p = 0.003$) (see *Table 76*).

If an opt-out system were not in place, participants were asked whether or not they would be influenced if the person who was directly involved in their baby's care was the one who asked them to give permission for their baby's data to be used for research. Over half said 'yes' (668/1300; 51.4%) and a further quarter said that it would 'possibly' influence them (324/1300; 25%), whereas just under one-fifth said that it definitely would not (228/1300; 17.5%).

Influence was regarded as both positive and negative by the participants. Some participants felt that being asked by their direct carer would have made a difference because of the trust that had been built up already, but others raised concerns that it might mean that they would have less choice and they would be worried that saying 'no' would affect the care. In addition, some participants said that it would not make any difference as long as the right person asked them (e.g. the person would need to have knowledge and understanding of the research).

We investigated the response rate for each qualification-based group regarding their likeliness to agree to share their baby's data for research if the person who asked them was directly involved in their baby's care. The association here was significant ($\chi^2_{trend} = 6.060$; $p = 0.048$). Participants who said 'yes' ($n = 548$) were more likely to be those with lower levels of academic qualifications (see *Table 77*).

Data access for research purposes

Participants were told that sometimes researchers need to identify which babies would be suitable to take part in medical research and one of the ways to do this is for non-medical staff to access their baby's data. There was little objection to this occurring. Over 50% of participants (703/1306; 53.8%) reported that they would be happy for this to happen, with a further 25% saying 'possibly' (337/1306; 25.8%). One-fifth said 'no' (200/1306; 15.3%) or 'don't know' (66/1306; 5.1%).

Data security

Participants displayed very high levels of confidence about the security and accuracy of patient data held within neonatal services. Over 80% were very or fairly confident about its security (*Table 51*) and over 85% were very or fairly confident about its accuracy (*Table 52*). However, participants' qualification levels were found to be significantly associated with their responses regarding their levels of confidence about the security of the data ($\chi^2_{trend} = 27.07$, df = 2; $p = 0.000$) and the accuracy of the data ($\chi^2_{trend} = 20.95$, df = 2; $p = 0.000$). Those with degrees/higher degrees were less likely to be confident about both the security of data and the accuracy of data.

Subsequent contact as a result of research findings

Participants were asked to rate how important it would be for them to have feedback on any research for which their baby's data were used (1 being the least important and 7 being the most important). The majority of participants ($n = 629$; 47.7%) thought that it was very important to have feedback. The mean and median of the importance ratings were 5.7 and 6.0, respectively.

TABLE 51 Levels of confidence about the security of data, across qualification-based groups

Highest qualification	Level of confidence, *n* (%)					Total
	Very confident	Fairly confident	Neither confident nor unconfident	Fairly unconfident	Very unconfident	
O levels/GCSEs	82 (35.3)	112 (48.3)	35 (15.1)	3 (1.3)	0 (0)	232
A levels/vocational qualification	140 (37.5)	169 (45.3)	55 (14.7)	7 (1.9)	2 (0.5)	373
Degree/higher degree	75 (20.9)	194 (54.2)	73 (20.4)	14 (39)	2 (0.6)	358
Total	297 (30.8)	475 (49.3)	163 (16.9)	24 (2.5)	4 (0.4)	963

TABLE 52 Levels of accuracy about the security of data, across qualification-based groups

Highest qualification	Level of confidence, *n* (%)					Total
	Very confident	Fairly confident	Neither confident nor unconfident	Fairly unconfident	Very unconfident	
O levels/GCSEs	81 (35.7)	114 (50.2)	28 (12.3)	4 (1.8)	0 (0)	227
A levels/vocational qualification	125 (33.9)	196 (53.1)	40 (10.8)	8 (2.2)	0 (0)	369
Degree/higher degree	71 (19.9)	220 (61.8)	55 (15.4)	10 (2.8)	0 (0)	356
Total	277 (29.1)	530 (55.7)	123 (12.9)	22 (2.3)	0 (0)	952

When asked if they would want to be contacted if new information were discovered about their baby's condition as a result of research that used their baby's data, the majority of participants said 'yes' (1037/1299; 79.8%). A further 15.3% (199/1299) said 'possibly' and only 3.4% (44/1299) said 'no', with 1.5% (19/1299) responding that they did not know.

Research benefit for others

Participants were asked about the use of their baby's data for research that may benefit other babies in the future, but may not have any direct benefits for their own baby. On a scale of 1 (least happy) to 7 (most happy), the overwhelming majority of parents responded positively, with 65.3% (845/1294) scoring 7 and a further 10.7% (203/1294) scoring 6. The mean and median scores were 6.32 and 7, respectively. When parents' responses were investigated by the highest level of care they had experienced, the result remained consistent: there were high levels of support for research that might benefit others but not themselves.

Conclusions

This is the largest sample to date of parents with experience of babies in neonatal care who have been consulted on the issue of the use of clinical data for research purposes. Our sample constitutes 11.8% of admissions to participating neonatal units at the time of conducting the survey. The survey for parents was co-designed with parents of babies in neonatal care and represents participatory preparatory work that proved effective in gathering high numbers of participants. The commitment of the children's clinical local research network to a non-medicines study was crucial in the mobilisation of adequate numbers of research

nurses and the oversight of the process of recruitment and consent. The age of those participating is consistent with broader national data, but black mothers and those of Asian ethnicity are slightly under-represented in the sample in terms of both the neonatal units from where data were collected and the national figures.

The parents sampled displayed very high levels of satisfaction with the care they experienced. This result should be considered in the context that the majority of parents completed the questionnaire when their baby was still in hospital; it is unknown whether or not, when asked the same question at a later point and after having returned home, the result would remain the same. Although parents were explicitly reminded as part of the consent procedure that participation in the survey would not affect their baby's care, this may have been an influencing factor on high expressed levels of satisfaction. It is of note that the participants themselves were equivocal about whether or not their level of satisfaction with the care received might be a source of bias in their responses.

In broad terms, we find strong support for the sharing of routinely recorded health data for research purposes among parents with children in neonatal care, with over two-thirds of participants responding positively. The possibility of de-identified data-sharing or sharing only if explicit permission were asked raised the percentage of those saying 'yes' by around 10%, but most of those who had said 'no' in general terms remained opposed despite the introduction of these additional conditions. This conclusion is supported by results from questions that explored 'opt out' as the default position for data-sharing; nearly two-thirds agreed that this was a good idea, with less than one-fifth definitely saying 'no'.

This headline result is moderated by several factors. We also explored the bias inherent in one-third of the sample having a degree or higher degree qualification in the conclusions we draw. Key among those is highest level of educational qualifications. This was found to be statistically significant in respect of whether or not electronic records had been heard of in the first place; the group with the lowest levels of educational qualifications were the least likely to be aware of them. Educational background also had a statistically significant effect on willingness for de-identified data to be used. Those with a degree/higher degree were more likely to agree and those with the lowest qualifications were more likely than others to respond 'possibly', which indicated some element of uncertainty perhaps through limited understanding, given their limited awareness of electronic records. In addition, in the case of the acceptability of an opt-out system, despite strong support, educational background was found to have a statistically significant relationship with the responses; those with a degree/higher degree were more likely to say 'no' than those with lowest qualifications, and those with lowest qualifications were more likely to say 'possibly'.

Overall, these results point to the need to ensure that those who are less well educated are afforded every opportunity to understand the new digital NHS in order to make informed choices about the use of data and/or the implications of systems, such as opt out, to which they may be asked to subscribe. This conclusion is lent modest support by the statistically significant relationship between educational background and parents' willingness to consent to data-sharing if asked by an individual directly involved in their baby's care. Parents with the lowest level of educational qualifications were more likely than those in the other qualification groups to say yes to data-sharing if asked by someone they know who cared for their baby. Direct knowledge and understanding through an individual contact with a trusted person, rather than knowledge and understanding in the abstract, is, for this group, more influential.

Whether or not the child in neonatal care was the parents' only child also had a statistically significant effect on responses to questions about data-sharing. Parents with more than one child often said that they would be willing for their baby's data to be used for research purposes. Parents whose child in neonatal care was not their first child were more likely to say 'yes' to routine data-sharing both in general terms and if data were de-identified. These results point perhaps to the role of maternal experience in moderating attitudes. More experienced parents were less concerned about any potential difficulties caused by agreeing to the use of their baby's data and/or were more appreciative of the potential advantages of doing so.

The level of care a baby had experienced was also found to have a statistically significant effect on parental attitudes. (Level of care refers to the highest level of care experienced at any point. It does not refer to the duration of care at that level, nor the current level of care at the time of responses.) Parents whose children had experienced level 3 care (the highest level), were in general terms, more likely to say 'yes' in all three conditions associated with willingness to share data than those who had not, if data were de-identified and if permission were obtained. These three conditions were not treated in the research as alternatives, but rather attitudes to all three were sought in their own right without seeking an expressed preference between them. Overall, these results demonstrate that the greater the level of concern or the more complex the level of care experienced by parents, the stronger their willingness to permit the routine sharing of data for research purposes, regardless of how that is framed.

The data revealed a strong orientation towards the assistance of others through routine use of health data for research, even if such data-sharing would not have direct beneficial effects for the parent or their baby. When asked this question directly, over three-quarters responded in the two most positive categories of willingness for their babies' data to be used. In addition, parents who had lost a child through miscarriage, termination or stillbirth, were asked specifically about data that might be associated with that child. Over half of them expressed willingness for data to be shared for research purposes and, additionally, nearly one-quarter said 'possibly'. These results suggest that a greater emphasis could be placed on the contribution that parents' willingness to share data makes to the benefit of others. This contribution can be cast in terms of altruism, because the motivation clearly is disengaged from direct benefit to self.

There are some contradictions in the results that may be an artefact, in part, of how the questions were asked. Responses are elicited to seemingly unconnected questions in the layout of the survey, but when results are placed side by side the contradiction is revealed. However, the contradictions revealed may be real for individuals. It is perfectly possible to hold one attitude alongside another seemingly contradictory one and to be unaware of the conflict between them until prompted to consider both attitudes at the same time. For example, participants strongly supported the notion of being contacted again if new information were discovered about their baby's condition as a result of research that used their baby's data. Yet the majority of those who said 'yes' to this question were also those who said 'yes' to data-sharing if the data were de-identified. The survey asked for their attitude to de-identified data unconnected to their attitude to being contacted again as a result of discoveries linked to their willingness to share their baby's data. As the possibilities of e-health, digital records and digital data mining become even greater in the future, these will inevitably create dilemmas and contradictions in attitude and approach at an individual patient level. It is hard to think through the likely consequences of a personal response in the fast-changing world of digital health, when the possibilities of that world are unknown or ill-understood. Hence, seeming contradictions will appear in attitude and response. Moving forward, it will be important to be mindful of new dilemmas that EPRs and data-sharing for research might provoke for individuals and how to support individuals in those circumstances.

We find that there is strong support for the routine sharing of babies' data for research purposes, particularly among those whose babies have experienced more complex levels of neonatal care and are more experienced parents. The differences in degrees of willingness to share data are small, but those saying 'no' are likely to remain opposed regardless of whether data are de-identified or individual permission is sought. More experienced parents were less concerned about any potential difficulties caused by agreeing to the use of their baby's data and/or appreciated more the potential advantages of doing so. The support for routine sharing of babies' data for research purposes is overall underpinned by a strong altruistic motivation from parents to support the benefit of others regardless of benefit to self.

We acknowledge that a limitation in fully understanding the influences underlying the observed trends and associations is the lack of follow-up exploratory, qualitative inquiry with a subsample of those who participated, which might have illuminated further aspects of the conclusions drawn.

Implications for health care

The positive result in support of routine data-sharing should be tempered by a concern to ensure that those with lower educational backgrounds are afforded greater opportunities to understand the significance of digital records and data-sharing, and the possibilities such as opt out, in order to make informed decisions. Our finding that, in this group, the provision of information through individual contact with a trusted person is more influential than the provision in the abstract, indicates the important role of clinical staff in explaining the way in which clinical data may be used in research. This suggests the importance of doctors and nurses being aware of these issues, being trained in conveying information to parents, and having sufficient time for explanation.

Research recommendations

The identification of effective and efficient methods to engage the public in debate and improve their knowledge of potential benefits, processes and regulation are important to secure their trust and confidence in the use of clinical data in research.

Chapter 9 Conclusions

What we found

We established the Medicines for Neonates programme on the principle that it should be possible to employ EPR data to support, improve and advance patient care and health services. We obtained multiple approvals, including from the Caldicott Guardians of every NHS trust providing NHS neonatal services. We showed that it is possible to create a national data resource, the NNRD, from extractions from EPRs, with the support of the neonatal clinical community.

We conducted formal assessment of the utility of the NNRD in population research into neonatal NEC, mortality of very preterm babies admitted to NHS neonatal units, and in trial-based economic evaluations. We compared NNRD data against research-standard data from a NIHR multicentre clinical trial, determined the validity of 2-year neurodevelopmental status recorded in the EPR against a research assessment by a single examiner, and demonstrated that it is possible to link the NNRD to HES data to create a longitudinal patient record. We examined parent attitudes to the use of personal clinical data in research and identified both support and strong altruism.

We developed standard operating procedures to assure the quality and completeness of data held in the NNRD. These include internal consistency, logic, range and completeness checks and identification of duplicate entries. We linked multiple episodes of care across different hospitals to create a single record for each infant. We conducted a comparison of NNRD and HES data, also extracted from hospital systems. We show that for key newborn variables, rates of both missing and potentially erroneous data are substantially lower in the NNRD. Other important differences are that the NNRD contains a far wider range of data items and we have defined each data item clearly, with a comprehensive metadata set available.

We show that clinical data from EPRs held in the NNRD can be used to create UK birth cohorts. We also show that the sole use of the NHS administrative database, HES, to build birth cohorts will result in the inclusion of many babies with missing gestational age data and implausible birthweights. However, linking the NNRD to HES substantially enhances the quality and scope of data.

We performed a formal comparison of NNRD data against clinical trial data. This showed that for economic evaluations and baseline information, data in the NNRD perform as well as trial data. We identified strong parent support underpinned by altruism for the routine sharing of babies' data for research purposes.

Our overall conclusion is that we have established proof of principle that EPR data may be employed successfully to support patient care and clinical services through research and a range of health service evaluations. We show that the potential of EPR to serve as the source of data for secondary uses is substantial. In addition to the National Neonatal Audit Programme, the NNRD is now used for a growing range of outputs by NHS England, Public Health England, Department of Health and Social Care and other organisations.

Implications for health care

We have demonstrated proof of principle of the potential of the NNRD as a data source for neonatal trial-based economic evaluations in the UK context. This has potential to reduce the costs and improve the efficiency of economic evaluations conducted in relation to neonatal research studies. The creation of a national data resource from EPR data minimises the burden placed on busy clinical teams by providing a single national data source to service multiple outputs, eliminating the need for multiple individual

collections, with repetitive capture of many commonly required data items, and reduces the risk of transcription errors and other errors that arise from repeated data recording.

The reasons for the differences between NNRD and HES data merit consideration. HES data are administrative, with entry by coding clerks and no clinical oversight. Data quality checks are minimal and in effect restricted to ensuring that the format of the NHS number is valid. There is no feedback to clinical teams. In contrast, the use of EPR data de facto ensures close clinical involvement. However, we do not believe that clinician involvement in assuring data quality and completeness should be taken for granted. Considerable effort is required to secure clinician engagement. In the case of the NNRD, the pivotal factor was its use for national clinical audit that involved publication of data completeness, detailed analyses and outlier status for named neonatal units and networks. We maintain close engagement with the clinical community through regular newsletters, national meetings and a 'hot-line' for staff, for one-to-one responses to queries. Neonatal units have been important stakeholders and collaborators in outputs to date, including Department of Health and Social Care reports and peer-reviewed publications.

To assist clinical teams, we developed a web portal that enables users from individual neonatal units to view data items identified as missing or potentially erroneous so that they could make corrections in the real-time infant EPRs. We developed this feedback loop initially to assist users in ensuring only reliable complete data were used in analyses for the National Neonatal Audit Programme; however, this has been extended (e.g. for annual case-mix mortality analyses conducted by the NDAU) and the intention is to further develop this approach. The strength of this data quality feedback loop is that both the clinical records and the NNRD are updated. The traditional model involving separate databases for research or other evaluations has meant that, although the research database might be corrected, the original record used clinically remained uncorrected, with potential detriment to patient care.[387]

It is worth noting that, although improvement in completeness and accuracy is required for some key clinical outcomes, the extent of agreement we identified, even though there had been no prior notification to clinical teams that a comparison would be made, is an encouraging indication of EPR data quality. It is equally worth noting that we acted on the assumption that trial data represented the gold standard but this may not in fact be the case. A potential advantage of using the NNRD for clinical research is that the EPR system has a clear audit trail so that source data verification is assured.

The high degree of parent support that we identified for use of EPR data for research is in sharp contrast with the experience of high-profile projects, such as care.data, where public distrust has been marked. It may be of relevance that we found that more experienced parents were less concerned about any possible difficulties caused by agreeing to the use of their baby's data and evidenced greater appreciation of the potential advantages. The implication for health care of our findings of strong parent support for data-sharing, but also that provision of information through individual contact with a trusted person is more influential than the provision in the abstract, is that doctors and nurses with adequate knowledge and training in conveying such information to parents and patients from a wide range of backgrounds need sufficient time for explanation.

Research recommendations

We suggest that future research might test the roadmap that we have established to create research databases from EPR by other specialties. More work is also needed to develop and evaluate such secondary databases if they are to be reliable sources of data for research. Current regulatory processes for data linkage are challenging. We would therefore welcome initiatives to develop regulatory frameworks that are clear and straightforward to navigate. The NNRD has been developed and is currently maintained through academic endeavour; an operational challenge for health-care services is how best to develop and maintain such databases as long-term national resources.

Improvements in administrative and clinical data quality and completeness are important health services goals. To our best knowledge, there are no national processes to evaluate health services administrative and clinical data formally and systematically, or to improve quality and completeness. These are also important issues for future research.

The potential of EPRs as data sources for secondary purposes requires further research and development. At present, using a real-time system directly for health services analytics such as benchmarking, or as a source of data without further processing, would result in several difficulties. In a real-time system, data change from second to second, hence the same request, conducted again at a different point in time, is very likely to yield a different result. Data in a real-time system have not undergone any quality assurance and, in the case of the neonatal EPR, contain duplicate, erroneous and missing entries. Users are able to access only data relating to patients in their hospital, with access to data from other providers only possible with specific regulatory approval. When attempting to make comparisons across neonatal units, variation between users in the application of algorithms will lead to outputs that are not necessarily comparable (e.g. selecting on < 1500 g instead of ≤ 1500 g will produce different results). Complex algorithms are particularly problematic (e.g. ROP screening criteria that are based on birthweight, gestational age, postnatal age, postmenstrual age and age at discharge). Thus a real-time platform, although excellent in enabling the rapid sharing of data between providers and facilitating a move away from paper medical records, is not an appropriate vehicle for even simple health services evaluations and research or for providing data without further processing. Future research, for example to flag missing entries, embed prompts and alerts, and range and internal consistency checks into the EPR, might assist users in improving data accuracy and completeness.

The development of methods to improve clinician and NHS trust engagement in data quality assurance, such as incentives, and mandates might also have utility. The identification of effective and efficient methods to involve parents in helping to assure the accuracy and completeness of their babies' data and to improve their knowledge of processes, regulatory safeguards and potential benefits might assist in securing continuing trust and confidence in the use of clinical data in research.

We have shown that screening neurodevelopment assessments of very preterm children at the age of 2 years, carried out by health-care staff with a wide range of training and experience, is insufficient to identify neurocognitive impairment, hence approaches to improve NHS assessments of impairment following preterm birth require to be identified. The complete national coverage of the NNRD of all admissions to neonatal units with no gestational age, birthweight or other restrictions, offers the opportunity to acquire reliable estimates of the prevalence of conditions likely to lead to neurodisability and other impairment. Such population data would have wide utility (e.g. to examine time-trends and national variation, conduct natural history of disease research, and epidemiological surveillance of rare conditions, as we illustrated in our study of NEC).

Acknowledgements

We wish to thank Professor Michael Goldacre, Emeritus Professor of Public Health at the University of Oxford the independent chairperson, and Professor Andrew Wilkinson, Emeritus Professor of Paediatrics at the University of Oxford, the independent deputy chairperson of the programme Steering Committee. We are also grateful to Dr Matthew Hyde for assisting in the preparation of the final report and Mr Richard Colquhoun for administrative support.

Contributions of authors

Neena Modi (Professor of Neonatal Medicine and Honorary Consultant) was the lead applicant; was the lead investigator; conceived study and led grant application; was responsible for the overall programme co-ordination and delivery, report writing and organisation of programme committees; was the lead for *Chapter 1*; and was the lead supervisor of the research conducted for *Chapters 2, 3* and *5*.

Deborah Ashby (Professor of Medical Statistics and Clinical Trials) was a co-applicant and supplied statistical supervision.

Cheryl Battersby (Clinical Research Fellow) conducted the research reported in *Chapters 2* and *4*.

Peter Brocklehurst (Professor of Women's Health and Director, Birmingham Clinical Trials Unit) was a co-applicant and supplied methodological expertise.

Zoe Chivers (Head of Services, Bliss) was a co-applicant and supplied parent–public engagement expertise.

Kate Costeloe (Professor of Paediatrics) was a co-applicant and the lead supervisor for research conducted for *Chapter 4*.

Elizabeth S Draper (Professor of Perinatal and Paediatric Epidemiology) took over the co-applicant role from David Field and supplied methodological expertise.

Victoria Foster (Senior Lecturer in Social Sciences) conducted research reported in *Chapter 8*.

Jacquie Kemp (National Programme of Care Senior Manager) was a co-applicant and supplied health services expertise.

Azeem Majeed (Professor of Primary Care and Public Health) was a co-applicant and the senior supervisor of research conducted for *Chapter 7*.

Joanna Murray (PhD Student) conducted the research reported in *Chapter 7*.

Stavros Petrou (Professor of Health Economics) was a co-applicant and the lead for research conducted for *Chapter 6*.

Katherine Rogers (Research Fellow) conducted research reported in *Chapter 8*.

Shalini Santhakumaran (Statistician) supplied statistical support and conducted research reported in *Chapter 3*.

Sonia Saxena (Clinical Professor of Primary Care) was the supervisor of research conducted for *Chapter 7*.

Yevgeniy Statnikov (Data Manager) conducted data management; conducted research for *Chapter 1*; assisted with the work reported in *Chapter 4*; and assisted with work reported in *Chapter 7*.

Hilary Wong (Clinical Research Fellow) conducted the research reported in *Chapter 5*.

Alys Young (Professor of Nursing, Midwifery and Social Work) was a co-applicant and the lead supervisor for research conducted for *Chapter 8*.

Contributions of others

Felix Achana (Research Fellow) assisted with work reported in *Chapter 6*.

Richard Colquhoun (Programme Manager) supplied administrative support.

Buthaina Ibrahim (Research Assistant) contributed to work reported in *Chapter 7*.

Kamran Khan (Research Associate) assisted with work reported in *Chapter 6*.

Sam Watson (PhD student) assisted with work reported in *Chapter 6*: *Using the National Neonatal Research Database to inform economic evaluations of neonatal interventions*.

Publications

Foster V, Young A, Modi N, Brocklehurst P, Abbott J, Costeloe K, *et al.* The use of routinely collected patient data for research: a critical review. *Health* 2012;**16**:448–63.

Gale C, Santhakumaran S, Nagarajan S, Statnikov Y, Modi N, on behalf of the Neonatal Data Analysis Unit and the Medicines for Neonates Investigator Group. The impact of introducing managed clinical networks on neonatal care in England: a population-based study. *BMJ* 2012;**344**:e2105.

Blencowe H, Lee ACC, Cousens S, Bahalim A, Narwal R, Zhong N, *et al.* Beyond newborn survival: preterm birth associated impairment estimates at regional and global level for 2010. *Pediatr Res* 2013;**74**:17–23.

Foster V, Young A. Reflecting on participatory methodologies: research with parents of babies requiring neonatal care. *Int J Social Res Methodol* 2013;**18**:91–104.

Murray J, Saxena S, Modi N, Majeed A, Aylin P, Bottle A, Medicines for Neonates Investigator Group. Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts. *J Public Health* 2013;**35**:298–307.

Spencer A, Modi N. National neonatal data to support specialist care and improve infant outcomes. *Arch Dis Child Fetal Neonatal Ed* 2013;**98**:F175–80.

Battersby C, Santhakumaran S, Upton M, Radbone L, Birch J, Modi N, East of England Perinatal Networks. The impact of a regional care bundle on maternal breast milk use in preterm infants: outcomes of the East of England quality improvement programme. *Arch Dis Child Fetal Neonatal Ed* 2014;**99**:F395–401.

Cole TJ, Statnikov Y, Santhakumaran S, Pan H, Modi N, on behalf of the Neonatal Data Analysis Unit and the Preterm Growth Investigator Group. Birth weight and longitudinal growth in infants below 32 weeks' gestation: a UK population study. *Arch Dis Child Fetal Neonatal Ed* 2014;**99**:F34–4.

Murray J, Bottle A, Sharland M, Modi N, Aylin P, Majeed A, *et al.* Risk factors for hospital admission with RSV bronchiolitis in England: a population-based birth cohort study. *PLOS ONE* 2014;**9**:e89186.

Shah PK, Lee SK, Lui K, Sjörs G, Mori R, Reichman B, *et al.* The International Network for Evaluating Outcomes of very low birth weight, very preterm neonates (iNeo): a protocol for collaborative comparisons of international health services for quality improvement in neonatal care. *BMC Pediatr* 2014;**14**:110.

Watson SI, Arulampalam W, Petrou S, Marlow N, Morgan AS, Draper ES, *et al.* The effects of designation and volume of neonatal care on mortality and morbidity outcomes of very preterm infants in England: retrospective population-based cohort study. *BMJ Open* 2014;**4**:e004856.

Wong HS, Huertas-Ceballos A, Cowan FM, Modi N, on behalf of the Medicines for Neonates Investigator Group. Evaluation of early childhood social-communication difficulties in children born preterm using the Quantitative Checklist of Autism in Toddlers. *J Pediatr* 2014;**164**:26–33.

Wong HS, Santhakumaran S, Statnikov Y, Grey D, Watkinson M, Modi N, the UK Neonatal Collaborative. Retinopathy of prematurity in English neonatal units: a national population-based analysis using NHS operational data. *Arch Dis Child Fetal Neonatal Ed* 2014;**99**:F196–202.

Gale C, Modi N; WHEAT trial development group. Neonatal randomised point-of-care trials are feasible and acceptable in the UK: results from two national surveys. *Arch Dis Child Fetal Neonatal Ed* 2016;**101**:F86–7.

Gale C, Morris I, Neonatal Data Analysis Unit (NDAU) Steering Board. The UK National Neonatal Research Database: using neonatal data for research, quality improvement and more. *Arch Dis Child Educ Pract Ed* 2016;**101**:216–8.

Gemmell L, Martin L, Murphy KE, Modi N, Håkansson S, Reichman B, *et al.* Hypertensive disorders of pregnancy and outcomes of preterm infants of 24 to 28 weeks' gestation. *J Perinatol* 2016;**36**:1067–72.

Martin LJ, Reichman B, Darlow BA, Morisaki N, Modi N, Bassler D, *et al.* Country-specific vs. common birthweight-for-gestational age references to identify small for gestational age infants born at 24–28 weeks: an international study. *Paediatr Perinat Epidemiol* 2016;**30**:450–61.

Seaton S, Barker L, Draper ES, Abrams KR, Modi N, Manktelow BN, on behalf of the UK Neonatal Collaborative. Modelling neonatal care pathways for babies born preterm: an application of multistate modelling. *PLOS ONE* 2016;**1**:e0165202.

Shah PK, Lui K, Sjörs G, Mirea L, Reichman B, Modi N, *et al.* Neonatal outcomes of very low birthweight and very preterm neonates: an international comparison. *J Pediatr* 2016;**177**:144–52.

Springett A, Mann JP, Statnikov E, Modi N, Johnson N, Morris JK. Management and outcomes of neonates with Down syndrome admitted to neonatal units. *Birth Defects Res A Clin Mol Teratol* 2016;**106**:468–74.

Watson SI, Arulampalam W, Petrou S, Marlow N, Morgan AS, Draper ES, Modi N. The effects of a one-to-one nurse to patient ratio on the mortality rate in neonatal intensive care: a retrospective, longitudinal, population-based study. *Arch Dis Child Fetal Neonatal Ed* 2016;**101**:F195–200. (Ranked first of the top 10 most read papers published in ADC FNN in 2016.)

Wong HS, Santhakumaran S, Cowan FM, Modi N. Developmental assessments in preterm children: a meta-analysis. *Pediatrics* 2016;**138**:e20160251.

Battersby C, Longford N, Costeloe K, Modi N, for UK Neonatal Collaborative Necrotising Enterocolitis Study Group. Development of a gestational age-specific case-definition for neonatal Necrotising Enterocolitis. *JAMA Pediatr* 2017;**171**:256–63.

Battersby C, Longford N, Mandalia S, Costeloe K, Modi N and the UK Neonatal Collaborative Necrotising Enterocolitis (UKNC-NEC) study group. Incidence and enteral feed antecedents of severe neonatal necrotising enterocolitis in England 2012–13: a two-year, population surveillance study. *Lancet Gastroenterol Hepatol* 2017;**2**:43–51.

Darlow BA, Lui K, Kusuda S, Reichman B, Gagliardi L, Håkansson S, *et al.* International variations and trends in the treatment for retinopathy of prematurity. *Br J Opthalmol* 2017;**101**:1399–1404.

Gale C, Hyde MJ, Modi N, on behalf of the WHEAT trial development group. Research Ethics Committee decision-making in relation to an efficient neonatal trial. *Arch Dis Child Fetal Neonatal Ed* 2017;**102**:F291–8.

Helenius K, Sjörs G, Shah PS, Modi N, Reichman B, Morisaki N, *et al.* Survival in very preterm infants: an international comparison of 10 national neonatal networks pediatrics. *Pediatrics* 2017;**240**:e20172264.

Hines D, Modi N, Lee SK, Isayama T, Sjörs G, Gagliardi L, *et al.* Scoping review shows wide variation in the definitions of bronchopulmonary dysplasia in preterm infants and calls for a consensus. *Acta Paediatr* 2017;**106**:366–74.

Statnikov Y, Ibrahim B, Modi N. A systematic review of administrative and clinical databases of infants admitted to neonatal units. *Arch Dis Child* 2017;**102**:F270–6.

Achana F, Petrou S, Khan K, Gaye A, Modi N, on behalf of the Medicines for Neonates Investigators. A methodological framework for assessing agreement between cost-effectiveness outcomes estimated using alternative sources of data on treatment costs and effects for trial-based economic evaluations. *Eur J Health Econ* 2018;**19**:75–86.

Adams G, Williams C, Modi N, Xing W, Bunce C, UK Retinopathy of Prematurity Special Interest Group, Dahlmann-Noor A. Can we reduce the burden of the current UK guidelines for retinopathy of prematurity screening. *Eye* 2018;**32**:235–7.

Santhakumaran S, Statnikov Y, Gray D, Battersby C, Ashby D, Modi N, on behalf of the Medicines for Neonates Investigator Group. Survival of very preterm infants admitted to neonatal care in England 2008–2014: time trends and regional variation. *Arch Dis Child Fetal Neonatal Ed* 2018;**103**:F208–215.

Wong HS, Cowan FM, Modi N, Medicines for Neonates Investigator Group. Validity of neurodevelopmental outcomes of children born very preterm assessed during routine clinical follow-up in England. *Arch Dis Child Fetal Neonatal Ed* 2018;**103**:F479–84.

Data-sharing statement

Requests for access to data should be addressed to the corresponding author in the first instance who will convey this to the relevant lead investigator.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. Department of Health and Social Care. *Best Research for Best Health: A New National Health Research Strategy*. Department of Health and Social Care. 2006. URL: www.gov.uk/government/publications/best-research-for-best-health-a-new-national-health-research-strategy (accessed 8 November 2018).
2. Department of Health and Social Care. *The Power of Information: Putting All of Us in Control of the Health and Care Information We Need*. Department of Health and Social Care. 2012. URL: https://webarchive.nationalarchives.gov.uk/20130802094648/http://informationstrategy.dh.gov.uk/ (accessed 29 August 2018).
3. Department of Health and Social Care. *Delivering 21st Century IT Support for the NHS: National Strategic Programme*. Department of Health and Social Care. 2002. URL: https://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4008227 (accessed 29 August 2018).
4. British Association of Perinatal Medicine. *The BAPM Neonatal Dataset for the Annual Reporting of Data by Neonatal Intensive Care Units*. London: British Association of Perinatal Medicine; 1997. www.bapm.org/publications/documents/general/neonatal_dataset.pdf (accessed 8 November 2018).
5. World Health Organization. *Classifications*. Geneva: World Health Organization. URL: www.who.int/classifications/icd/icdonlineversions/en/ (accessed 8 November 2018).
6. Donovan LE, Boyle SL, McNeil DA, Pedersen SD, Dean SR, Wood S, Edwards AL. Label of gestational diabetes mellitus affects caesarean section and neonatal intensive care unit admission without conventional indications. *Can J Diabetes* 2012;**6**:58–63. https://doi.org/10.1016/j.jcjd.2012.01.003
7. Kirkby S, Genen L, Turenne W, Dysart K. Outcomes and milestone achievement differences for very low-birth-weight multiples compared with singleton infants. *Am J Perinatol* 2010;**27**:439–44. https://doi.org/10.1055/s-0030-1247597
8. Joy S, Istwan N, Rhea D, Desch C, Stanziano G. The impact of maternal obesity on the incidence of adverse pregnancy outcomes in high-risk term pregnancies. *Am J Perinatol* 2009;**26**:345–9. https://doi.org/10.1055/s-0028-1110084
9. Moore PD, Bay RC, Balcazar H, Coonrod DV, Brady J, Russ R. Use of home visit and developmental clinic services by high risk Mexican-American and white non-Hispanic infants. *Matern Child Health J* 2005;**9**:35–47. https://doi.org/10.1007/s10995-005-2449-1
10. Wariki WM, Mori R, Boo NY, Cheah IG, Fujimura M, Lee J, Wong KY. Risk factors associated with outcomes of very low birthweight infants in four Asian countries. *J Paediatr Child Health* 2013;**49**:E23–7. https://doi.org/10.1111/jpc.12054
11. Vendittelli F, Rivière O, Neveu B, Lémery D. Does induction of labor for constitutionally large-for-gestational-age fetuses identified in utero reduce maternal morbidity? *BMC Pregnancy Childb* 2014;**14**:156. https://doi.org/10.1186/1471-2393-14-156
12. Lee QY, Quek WS, Chow S, Lui K. A population study of demographic changes and outcomes of very premature multiple births infants admitted to nicu in australia and new zealand. *J Paediatr Child Health* 2013;**49**:125.
13. Fleming N, Ng N, Osborne C, Biederman S, Yasseen AS, Dy J, *et al.* Adolescent pregnancy outcomes in the province of Ontario: a cohort study. *J Obstet Gynaecol Can* 2013;**35**:234–45. https://doi.org/10.1016/S1701-2163(15)30995-6

14. Merritt TA, Goldstein M, Philips R, Peverini R, Iwakoshi J, Rodriguez A, Oshiro B. Impact of ART on pregnancies in California: an analysis of maternity outcomes and insights into the added burden of neonatal intensive care. *J Perinatol* 2014;**34**:345–50. https://doi.org/10.1038/jp.2014.17

15. Guner YS, Friedlich P, Wee CP, Dorey F, Camerini V, Upperman JS. State-based analysis of necrotizing enterocolitis outcomes. *J Surg Res* 2009;**157**:21–9. https://doi.org/10.1016/j.jss.2008.11.008

16. Kastenberg ZJ, Lee HC, Profit J, Gould JB, Sylvester KG. Effect of deregionalized care on mortality in very low-birth-weight infants with necrotizing enterocolitis. *JAMA Pediatr* 2015;**169**:26–32. https://doi.org/10.1001/jamapediatrics.2014.2085

17. Vigod SN, Kurdyak PA, Dennis CL, Gruneir A, Newman A, Seeman MV, *et al.* Maternal and newborn outcomes among women with schizophrenia: a retrospective population-based cohort study. *BJOG* 2014;**121**:566–74. https://doi.org/10.1111/1471-0528.12567

18. Ballantyne M, Sauve R, Creighton D, Saigal S, Asztalos E, Couture E, *et al.* Preterm infant journeys in a canadian regionalized health services context. *J Paediatr Child Health (Canada)* 2014;**19**:e97. https://doi.org/10.1093/pch/19.6.e35-176

19. Mirea L, Yang J, Paterson AD, Shah V, Bassil KL, Lee SK, Shah PS, Canadian Neonatal Network. Relationship of mode of conception and sex concordance with mortality/morbidity in preterm twins. *Twin Res Hum Genet* 2013;**16**:985–93. https://doi.org/10.1017/thg.2013.61

20. Baird R, Puligandla P, Skarsgard E, Laberge JM, Canadian Pediatric Surgical Network. Infectious complications in the management of gastroschisis. *Pediatr Surg Int* 2012;**28**:399–404. https://doi.org/10.1007/s00383-011-3038-6

21. Grover TR, Brozanski BS, Barry J, Zaniletti I, Asselin JM, Durand DJ, *et al.* High surgical burden for infants with severe chronic lung disease (sCLD). *J Pediatr Surg* 2014;**49**:1202–5. https://doi.org/10.1016/j.jpedsurg.2014.02.087

22. Tabano DC, Schroeder A, Sullivan K, Vaidya N. Impact of Assisted Reproductive Therapy (Art) on infant health and health care cost outcomes. *Value Health* 2014;**17**:A520. https://doi.org/10.1016/j.jval.2014.08.1621

23. de Jongh BE, Locke R, Paul DA, Hoffman M. The differential effects of maternal age, race/ethnicity and insurance on neonatal intensive care unit admission rates. *BMC Pregnancy Childb* 2012;**12**:97. https://doi.org/10.1186/1471-2393-12-97

24. Gasparović V, Gornik I, Ivanović D. Sepsis syndrome in Croatian intensive care units: piloting a national comparative clinical database. *Croat Med J* 2006;**47**:404–9.

25. Engelbrechtsen L, Nielsen EH, Perin T, Oldenburg A, Tabor A, Skibsted L, Danish Fetal Medicine Study Group. Cesarean section for the second twin: a population-based study of occurrence and outcome. *Birth* 2013;**40**:10–16. https://doi.org/10.1111/birt.12023

26. Andersson S, Petersen JP, Henriksen TB, Ebbesen F. The Danish neonatal clinical database is valuable for epidemiologic research in respiratory disease in preterm infants. *BMC Pediatr* 2014;**14**:47. https://doi.org/10.1186/1471-2431-14-47

27. Corvaglia L, Fantini MP, Aceti A, Gibertoni D, Rucci P, Baronciani D, Faldella G, 'Emilia Romagna Perinatal Network'. Predictors of full enteral feeding achievement in very low birth weight infants. *PLOS ONE* 2014;**9**:e92235. https://doi.org/10.1371/journal.pone.0092235

28. Moore T, Hennessy EM, Myles J, Johnson SJ, Draper ES, Costeloe KL, Marlow N. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. *BMJ* 2012;**345**:e7961. https://doi.org/10.1136/bmj.e7961

29. Sengupta S, Carrion V, Shelton J, Wynn RJ, Ryan RM, Singhal K, Lakshminrusimha S. Adverse neonatal outcomes associated with early-term birth. *JAMA Pediatr* 2013;**167**:1053–9. https://doi.org/10.1001/jamapediatrics.2013.2581

30. Hummler H, Lang K, Azpeitia A, Valls ISA. Short-term outcome of very low birth weight infants (VLBWI) requiring cardiopulmonary resuscitation in the delivery room. *Monatsschr Kinderheilkd* 2014;**162**:1–103.

31. Doyle TJ, Goodin K, Hamilton JJ. Maternal and neonatal outcomes among pregnant women with 2009 pandemic influenza A(H1N1) illness in Florida, 2009–2010: a population-based cohort study. *PLOS ONE* 2013;**8**:e79040. https://doi.org/10.1371/journal.pone.0079040

32. Christensen RD, Lambert DK, Henry E, Wiedmeier SE, Snow GL, Baer VL, *et al.* Is 'transfusion-associated necrotizing enterocolitis' an authentic pathogenic entity? *Transfusion* 2010;**50**:1106–12. https://doi.org/10.1111/j.1537-2995.2009.02542.x

33. Kugelman A, Reichman B, Chistyakov I, Boyko V, Levitski O, Lerner-Geva L, *et al.* Postdischarge infant mortality among very low birth weight infants: a population-based study. *Pediatrics* 2007;**120**:e788–94. https://doi.org/10.1542/peds.2006-3765

34. Imaizumi Y, Hayakawa K. Infant mortality among singletons and twins in Japan during 1999–2008 on the basis of risk factors. *Twin Res Hum Genet* 2013;**16**:639–44. https://doi.org/10.1017/thg.2012.156

35. Puopolo KM, Draper D, Wi S, Newman TB, Zupancic J, Lieberman E, *et al.* Estimating the probability of neonatal early-onset infection on the basis of maternal risk factors. *Pediatrics* 2011;**128**:e1155–63. https://doi.org/10.1542/peds.2010-3464

36. Lorch SA, Baiocchi M, Silber JH, Even-Shoshan O, Escobar GJ, Small DS. The role of outpatient facilities in explaining variations in risk-adjusted readmission rates between hospitals. *Health Serv Res* 2010;**45**:24–41. https://doi.org/10.1111/j.1475-6773.2009.01043.x

37. Morriss FH. Increased risk of death among uninsured neonates. *Health Serv Res* 2013;**48**:1232–55. https://doi.org/10.1111/1475-6773.12042

38. Pandey A, Ranjan R. Hypertensive disorders in pregnancy requiring emergency (108) transportation in the state of Gujarat (India): an epidemiological study. *J Clin Diagn Res* 2010;**4**:2017–22.

39. Wilson K, Nagy A, Green C, Boyd D, Ratnavel N, Mohinuddin S. Factors influencing early neonatal mortality in retrieved extreme preterm neonates. *Arch Dis Child* 2012;**97**:A229. https://doi.org/10.1136/archdischild-2012-302724.0799

40. Clements KM, Barfield WD, Ayadi MF, Wilber N. Preterm birth-associated cost of early intervention services: an analysis by gestational age. *Pediatrics* 2007;**119**:e866–74. https://doi.org/10.1542/peds.2006-1729

41. Merewood A, Brooks D, Bauchner H, MacAuley L, Mehta SD. Maternal birthplace and breastfeeding initiation among term and preterm infants: a statewide assessment for Massachusetts. *Pediatrics* 2006;**118**:e1048–54. https://doi.org/10.1542/peds.2005-2637

42. Boo NY, Cheah IG, Malaysian National Neonatal Registry. Risk factors associated with pneumothorax in Malaysian neonatal intensive care units. *J Paediatr Child Health* 2011;**47**:183–90. https://doi.org/10.1111/j.1440-1754.2010.01944.x

43. Lorch SA, Passarella M, Zeigler A. Challenges to measuring variation in readmission rates of neonatal intensive care patients. *Acad Pediatr* 2014;**14**(Suppl. 5):47–53. https://doi.org/10.1016/j.acap.2014.06.010

44. Galyean AM, Lagrew DC, Bush MC, Kurtzman JT. Previous cesarean section and the risk of postpartum maternal complications and adverse neonatal outcomes in future pregnancies. *J Perinatol* 2009;**29**:726–30. https://doi.org/10.1038/jp.2009.108

45. Xu X, Grigorescu V, Siefert KA, Lori JR, Ransom SB. Cost of racial disparity in preterm birth: evidence from Michigan. *J Health Care Poor Underserved* 2009;**20**:729–47. https://doi.org/10.1353/hpu.0.0180

46. Wingate MS, Alexander GR. Racial and ethnic differences in perinatal mortality: the role of fetal death. *Ann Epidemiol* 2006;**16**:485–91. https://doi.org/10.1016/j.annepidem.2005.04.001

47. Yunis KA, Khawaja M, Beydoun H, Nassif Y, Khogali M, Tamim H, National Collaborative Perinatal Neonatal Network (NCPNN). Intrauterine growth standards in a developing country: a study of singleton livebirths at 28–42 weeks' gestation. *Paediatr Perinat Epidemiol* 2007;**21**:387–96. https://doi.org/10.1111/j.1365-3016.2007.00827.x

48. Luoto R, Matomäki J, Isolauri E, Lehtonen L. Incidence of necrotizing enterocolitis in very-low-birth-weight infants related to the use of Lactobacillus GG. *Acta Paediatr* 2010;**99**:1135–8. https://doi.org/10.1111/j.1651-2227.2010.01795.x

49. Moro M, Pérez-Rodriguez J, Figueras-Aloy J, Fernández C, Doménech E, Jiménez R, *et al.* Predischarge morbidities in extremely and very low-birth-weight infants in Spanish neonatal units. *Am J Perinatol* 2009;**26**:335–43. https://doi.org/10.1055/s-0028-1110083

50. Agarwal R, Jain A, Deorari AK, Paul VK. Post-resuscitation management of asphyxiated neonates. *Indian J Pediatr* 2008;**75**:175–80. https://doi.org/10.1007/s12098-008-0026-5

51. Watson SI, Arulampalam W, Petrou S, Marlow N, Morgan AS, Draper ES, *et al.* The effects of designation and volume of neonatal care on mortality and morbidity outcomes of very preterm infants in England: retrospective population-based cohort study. *BMJ Open* 2014;**4**:e004856. https://doi.org/10.1136/bmjopen-2014-004856

52. Johnson J, Anderson B, Raker C, Wenstrom K. Elective inductions at term and adverse neonatal outcomes. *Am J Obstet Gynecol* 2009;**1**:S124–S5. https://doi.org/10.1016/j.ajog.2009.10.326

53. Tomazevic T, Ban-Frangez H, Ribic-Pucelj M, Premru-Srsen T, Verdenik I. Small uterine septum is an important risk variable for preterm birth. *Eur J Obstet Gynecol Reprod Biol* 2007;**135**:154–7. https://doi.org/10.1016/j.ejogrb.2006.12.001

54. van Heesch MM, Evers JL, Dumoulin JC, van der Hoeven MA, van Beijsterveldt CE, Bonsel GJ, *et al.* A comparison of perinatal outcomes in singletons and multiples born after in vitro fertilization or intracytoplasmic sperm injection stratified for neonatal risk criteria. *Acta Obstet Gynecol Scand* 2014;**93**:277–86. https://doi.org/10.1111/aogs.12328

55. Laws PJ, Tracy SK, Sullivan EA. Perinatal outcomes of women intending to give birth in birth centers in Australia. *Birth* 2010;**37**:28–36. https://doi.org/10.1111/j.1523-536X.2009.00375.x

56. Tunescu M, Olariu G, Man O, Olariu S, Olariu L. Chorioamnionitis and multisystem impairment to a premature with GA under 32 weeks. *J Mater-Fetal Neo M* 2014;**27**:398.

57. Doran J, McGowan JE, Alderdice F, McCall E, Craig S, Jenkins J. Regional follow up of late preterm neonatal intensive care graduates. *Nurse Res* 2012;**19**:37–43. https://doi.org/10.7748/nr2012.07.19.4.37.c9223

58. Kusuda S, Fujimura M, Uchiyama A, Totsu S, Matsunami K, Neonatal Research Network, Japan. Trends in morbidity and mortality among very-low-birth-weight infants from 2003 to 2008 in Japan. *Pediatr Res* 2012;**72**:531–8. https://doi.org/10.1038/pr.2012.114

59. De Los Santos-Garate AM, Villa-Guillen M, Villanueva-García D, Vallejos-Ruíz ML, Murguía-Peniche MT, NEOSANO's Network. Perinatal morbidity and mortality in late-term and post-term pregnancy: NEOSANO perinatal network's experience in Mexico. *J Perinatol* 2011;**31**:789–93. https://doi.org/10.1038/jp.2011.43

60. Potti S, Jain NJ, Mastrogiannis DS, Dandolu V. Obstetric outcomes in pregnant women with diabetes versus hypertensive disorders versus both. *J Matern Fetal Neonatal Med* 2012;**25**:385–8. https://doi.org/10.3109/14767058.2011.580403

61. Maheshwari R, Luig M. An audit of respiratory management and outcomes of outborn extremely preterm neonates retrieved on the first day of life. *Journal of Paediatrics and Child Health* 2012;**48**:153.

62. Lipkind HS, Duzyj C, Rosenberg TJ, Funai EF, Chavkin W, Chiasson MA. Disparities in cesarean delivery rates and associated adverse neonatal outcomes in New York City hospitals. *Obstet Gynecol* 2009;**113**:1239–47. https://doi.org/10.1097/AOG.0b013e3181a4c3e5

63. Crane JM, Keough M, Murphy P, Burrage L, Hutchens D. Effects of environmental tobacco smoke on perinatal outcomes: a retrospective cohort study. *BJOG* 2011;**118**:865–71. https://doi.org/10.1111/j.1471-0528.2011.02941.x

64. Morriss FH, Saha S, Bell EF, Colaizy TT, Stoll BJ, Hintz SR, *et al.* Surgery and neurodevelopmental outcome of very low-birth-weight infants. *JAMA Pediatr* 2014;**168**:746–54. https://doi.org/10.1001/jamapediatrics.2014.307

65. Nili F, McLeod L, O'Connell C, Sutton E, McMillan D. Outcomes of pregnancies in women with suspected antiphospholipid syndrome. *J Neonatal Perinatal Med* 2013;**6**:225–30. https://doi.org/10.3233/NPM-1370113

66. Abdel-Latif ME, Kecskés Z, Bajuk B, NSW and the ACT Neonatal Intensive Care Audit Group. Actuarial day-by-day survival rates of preterm infants admitted to neonatal intensive care in New South Wales and the Australian Capital Territory. *Arch Dis Child Fetal Neonatal Ed* 2013;**98**:F212–7. https://doi.org/10.1136/adc.2011.210856

67. Spitzer AR, Ellsbury DL, Handler D, Clark RH. The pediatrix babySteps (R) data warehouse and the pediatrix qualitySteps improvement project system-tools for 'meaningful use' in continuous quality improvement. *Clin Perinatol* 2010;**37**:49–70. https://doi.org/10.1016/j.clp.2010.01.016

68. Vogtmann C, Koch R, Gmyrek D, Kaiser A, Friedrich A. Risk-adjusted intraventricular hemorrhage rates in very premature infants: towards quality assurance between neonatal units. *Dtsch Arztebl Int* 2012;**109**:527–33. https://doi.org/10.3238/arztebl.2012.0527

69. Brown HK, Speechley KN, Macnab J, Natale R, Campbell MK. Neonatal morbidity associated with late preterm and early term birth: the roles of gestational age and biological determinants of preterm birth. *Int J Epidemiol* 2014;**43**:802–14. https://doi.org/10.1093/ije/dyt251

70. Anderberg E, Källén K, Berntorp K. The impact of gestational diabetes mellitus on pregnancy outcome comparing different cut-off criteria for abnormal glucose tolerance. *Acta Obstet Gynecol Scand* 2010;**89**:1532–7. https://doi.org/10.3109/00016349.2010.526186

71. Lisonkova S, Sheps SB, Janssen PA, Lee SK, Dahlgren L. Effect of older maternal age on birth outcomes in twin pregnancies: a population-based study. *J Perinatol* 2011;**31**:85–91. https://doi.org/10.1038/jp.2010.114

72. Heaman M, Kingston D, Brownell M, Helewa M. Predictors of prenatal and postpartum psychological distress: a population-based study in Manitoba, Canada. *Reproductive Sciences* 2014;**1**:320A.

73. Stock SJ, Ferguson E, Duffy A, Ford I, Chalmers J, Norman JE. Outcomes of elective induction of labour compared with expectant management: population based study. *BMJ* 2012;**344**:e2838. https://doi.org/10.1136/bmj.e2838

74. Jasso-Gutiérrez L, Durán-Arenas L, Flores-Huerta S, Cortés-Gallo G. Recommendations to improve healthcare of neonates with respiratory insufficiency beneficiaries of Seguro Popular. *Salud Publica Mex* 2012;**54**(Suppl. 1):57–64. https://doi.org/10.1590/S0036-36342012000700008

75. Tiblad E, Kublickas M, Ajne G, Bui, TH, Ek S, Karisson A, *et al.* Procedure-related complications and perinatal outcome after intrauterine transfusions in red cell alloimmunization in Stockholm. *Fetal Diagn Ther* 2011;**30**:266–73. https://doi.org/10.1159/000328683

76. Emilsson L, Lindahl B, Köster M, Lambe M, Ludvigsson JF. Review of 103 Swedish Healthcare Quality Registries. *J Intern Med* 2015;**277**:94–136. https://doi.org/10.1111/joim.12303

77. Rüegger C, Hegglin M, Adams M, Bucher HU, Swiss Neonatal Network. Population based trends in mortality, morbidity and treatment for very preterm- and very low birth weight infants over 12 years. *BMC Pediatr* 2012;**12**:17. https://doi.org/10.1186/1471-2431-12-17

78. Tsai WH, Hwang YS, Hung TY, Weng SF, Lin SJ, Chang WT. Association between mechanical ventilation and neurodevelopmental disorders in a nationwide cohort of extremely low birth weight infants. *Res Dev Disabil* 2014;**35**:1544–50. https://doi.org/10.1016/j.ridd.2014.03.048

79. Herrod HG, Chang CF, Steinberg SS. Variations in costs for the care of low-birth-weight infants among academic hospitals. *Clin Pediatr* 2010;**49**:443–9. https://doi.org/10.1177/0009922809341750

80. van Dommelen P, Mohangoo AD, Verkerk PH, van der Ploeg CP, van Straaten HL, Dutch NICU Neonatal Hearing Screening Working Group. Risk indicators for hearing loss in infants treated in different neonatal intensive care units. *Acta Paediatr* 2010;**99**:344–9. https://doi.org/10.1111/j.1651-2227.2009.01614.x

81. Manktelow BN, Seaton SE, Field DJ, Draper ES. Population-based estimates of in-unit survival for very preterm infants. *Pediatrics* 2013;**131**:e425–32. https://doi.org/10.1542/peds.2012-2189

82. Vogel JP, Holloway E, Cuesta C, Carroli G, Souza JP, Barrett J. Outcomes of non-vertex second twins, following vertex vaginal delivery of first twin: a secondary analysis of the WHO Global Survey on maternal and perinatal health. *BMC Pregnancy Childbirth* 2014;**14**:55. https://doi.org/10.1186/1471-2393-14-55

83. Shah A, Faundes A, Machoki M, Bataglia V, Amokrane F, Donner A, *et al.* Methodological considerations in implementing the WHO Global Survey for Monitoring Maternal and Perinatal Health. *Bull World Health Organ* 2008;**86**:126–31. https://doi.org/10.2471/BLT.06.039842

84. Soll RF, Edwards EM, Badger GJ, Kenny MJ, Morrow KA, Buzas JS, Horbar JD. Obstetric and neonatal care practices for infants 501 to 1500 g from 2000 to 2009. *Pediatrics* 2013;**132**:222–8. https://doi.org/10.1542/peds.2013-0501

85. Doyle LW, Victorian Infant Collaborative Study Group. Evaluation of neonatal intensive care for extremely low birth weight infants in Victoria over two decades: I. Effectiveness. *Pediatrics* 2004;**113**:505–9. https://doi.org/10.1542/peds.113.3.505

86. El-Sheikh A, Francis A, Gardosi J. Comparative analysis of length of stay in neonatal intensive care after 34 weeks in singleton babies with and without intrauterine growth restriction. *Arch Dis Child* 2011;**96**:Fa71. https://doi.org/10.1136/adc.2011.300161.59

87. Van Dijk JW, Anderko L, Stetzer F. The impact of Prenatal Care Coordination on birth outcomes. *J Obstet Gynecol Neonatal Nurs* 2011;**40**:98–108. https://doi.org/10.1111/j.1552-6909.2010.01206.x

88. Heller G, Günster C, Misselwitz B, Feller A, Schmidt S. Annual patient volume and survival of very low birth weight infants (VLBWs) in Germany: a nationwide analysis based on administrative data. *Z Geburtshilfe Neonatol* 2007;**211**:123–31. https://doi.org/10.1055/s-2007-960747

89. Gleissner MW, Spantzel T, Bücker-Nott HJ, Jorch G. [Risk factors of retinopathy of prematurity in infants 32 to 36 weeks gestational age.] *Z Geburtshilfe Neonatol* 2003;**207**:24–8. https://doi.org/10.1055/s-2003-37841

90. Hummler HD, Poets C, Vochem M, Hentschel R, Linderkamp O. Mortality and morbidity of very premature infants in Baden-Württemberg depending on hospital size: is the current degree of regionalization adequate? *Z Geburtshilfe Neonatol* 2006;**210**:6–11. https://doi.org/10.1055/s-2006-931508

91. Gay S, Ferdinus C, Sagot P, Gouyon JB. What are the neonatal risks in low risk pregnancies: place of paediatric organisation in birth centers development? *Arch Pediatr* 2007;**14**:1174–7. https://doi.org/10.1016/j.arcped.2007.06.028

92. Marcoux MO, Denizot S, Dassieu G, Picaud JC, Cristini C, Arnaud C, *et al.* Evidence versus experience in neonatal practice: the example of extremely premature infants. *Arch Pediatr* 2009;**16**(Suppl. 1):49–55. https://doi.org/10.1016/S0929-693X(09)75301-1

93. Schlößer RL, Frey G, Zemlin M, Misselwitz B. Mortality of very low birth weight infants during a 24 year period in Hesse a province of Germany: impact of variation in registration. *Z Geburtshilfe Neonatol* 2014;**218**:100–5. https://doi.org/10.1055/s-0034-1376992

94. Research Ethics Service. *Standard Operating Procedures for Research Ethics Committees: Version 7.4*. Leeds: NHS Digital; 2019.

95. Council for Science and Technology. *Better Use of Personal Information: Opportunities and Risks*. Council for Science and Technology. 2005. URL: http://webarchive.nationalarchives.gov.uk/20130705054945/www.bis.gov.uk/assets/cst/docs/files/cst-reports/05-2177-better-use-personal-information.pdf (accessed 19 December 2015).

96. Evans TW. Best research for best health: a new national health research strategy. *Clin Med* 2006;**6**:435–7. https://doi.org/10.7861/clinmedicine.6-5-435

97. The Academy of Medical Sciences. *Personal Data for Public Good: Using Health Information in Medical Research*. The Academy of Medical Sciences. 2013. URL: www.acmedsci.ac.uk/policy/policy-projects/personal-data (accessed 10 October 2015).

98. Department of Health and Social Care. *Toolkit for High Quality Neonatal Services*. London: Department of Health and Social Care; 2009. URL: https://webarchive.nationalarchives.gov.uk/20130123200735/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_107845 (accessed 8 November 2018).

99. Wang W, Krishnan E. Big data and clinicians: a review on the state of the science. *JMIR Med Inform* 2014;**2**:e1. https://doi.org/10.2196/medinform.2913

100. Neu J, Walker WA. Necrotizing enterocolitis. *N Engl J Med* 2011;**364**:255–64. https://doi.org/10.1056/NEJMra1005408

101. Lin PW, Stoll BJ. Necrotising enterocolitis. *Lancet* 2006;**368**:1271–83. https://doi.org/10.1016/S0140-6736(06)69525-1

102. Uauy RD, Fanaroff AA, Korones SB, Phillips EA, Phillips JB, Wright LL. Necrotizing enterocolitis in very low birth weight infants: biodemographic and clinical correlates. National Institute of Child Health and Human Development Neonatal Research Network. *J Pediatr* 1991;**119**:630–8. https://doi.org/10.1016/S0022-3476(05)82418-7

103. Guillet R, Stoll BJ, Cotten CM, Gantz M, McDonald S, Poole WK, Phelps DL, National Institute of Child Health and Human Development Neonatal Research Network. Association of H2-blocker therapy and higher incidence of necrotizing enterocolitis in very low birth weight infants. *Pediatrics* 2006;**117**:e137–42. https://doi.org/10.1542/peds.2005-1543

104. Guthrie SO, Gordon PV, Thomas V, Thorp JA, Peabody J, Clark RH. Necrotizing enterocolitis among neonates in the United States. *J Perinatol* 2003;**23**:278–85. https://doi.org/10.1038/sj.jp.7210892

105. Kamitsuka MD, Horton MK, Williams MA. The incidence of necrotizing enterocolitis after introducing standardized feeding schedules for infants between 1250 and 2500 grams and less than 35 weeks of gestation. *Pediatrics* 2000;**105**:379–84. https://doi.org/10.1542/peds.105.2.379

106. Ballard RA, Ballard PL. Antenatal hormone therapy for improving the outcome of the preterm infant. *J Perinatol* 1996;**16**:390–6.

107. Bauer CR, Morrison JC, Poole WK, Korones SB, Boehm JJ, Rigatto H, Zachman RD. A decreased incidence of necrotizing enterocolitis after prenatal glucocorticoid therapy. *Pediatrics* 1984;**73**:682–8.

108. Halac E, Halac J, Begue EF, *et al.* Prenatal and postnatal corticosteroid therapy to prevent neonatal necrotizing enterocolitis: a controlled trial. *J Pediatr* 1990;**117**(1 Pt 1):132–8. https://doi.org/10.1016/S0022-3476(05)72461-6

109. Smith LM, Qureshi N, Chao CR. Effects of single and multiple courses of antenatal glucocorticoids in preterm newborns less than 30 weeks' gestation. *J Matern Fetal Med* 2000;**9**:131–5. https://doi.org/10.1002/(SICI)1520-6661(200003/04)9:2<131::AID-MFM9>3.0.CO;2-M

110. Bajwa NM, Berner M, Worley S, Pfister RE, Swiss Neonatal Network. Population based age stratified morbidities of premature infants in Switzerland. *Swiss Med Wkly* 2011;**141**:w13212. https://doi.org/10.4414/smw.2011.13212

111. Kliegman RM, Walsh MC. Neonatal necrotizing enterocolitis: pathogenesis, classification, and spectrum of illness. *Curr Probl Pediatr* 1987;**17**:213–88. https://doi.org/10.1016/0045-9380(87)90031-4

112. Bell MJ, Ternberg JL, Feigin RD, Keating JP, Marshall R, Barton L, Brotherton T. Neonatal necrotizing enterocolitis: therapeutic decisions based upon clinical staging. *Ann Surg* 1978;**187**:1–7. https://doi.org/10.1097/00000658-197801000-00001

113. Pan H, Cole T. *LMS Growth, a Microsoft Excel add-in to access growth references based on LMS method*. Version 2.77 2014. URL: wwwhealthforallchildrencom/ (accessed 8 November 2018).

114. Breslow N, Day N. *Statistical Methods in Cancer Research: IARC Scientific Publications No. 32*. 6th edn. Lyon: International Agency for Research on Cancer (IARC); 1994.

115. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995;**57**:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

116. Battersby C, Santhakumaran S, Upton M, Radbone L, Birch J, Modi N, East of England Perinatal Networks. The impact of a regional care bundle on maternal breast milk use in preterm infants: outcomes of the East of England quality improvement programme. *Arch Dis Child Fetal Neonatal Ed* 2014;**99**:F395–401. https://doi.org/10.1136/archdischild-2013-305475

117. Palmer SR, Biffin A, Gamsu HR. Outcome of neonatal necrotising enterocolitis: results of the BAPM/CDSC surveillance study, 1981–84. *Arch Dis Child* 1989;**64**:388–94. https://doi.org/10.1136/adc.64.3.388

118. Rees CM, Eaton S, Pierro A. National prospective surveillance study of necrotizing enterocolitis in neonatal intensive care units. *J Pediatr Surg* 2010;**45**:1391–7. https://doi.org/10.1016/j.jpedsurg.2009.12.002

119. Gordon PV, Swanson JR, Attridge JT, Clark R. Emerging trends in acquired neonatal intestinal disease: is it time to abandon Bell's criteria? *J Perinatol* 2007;**27**:661–71. https://doi.org/10.1038/sj.jp.7211782

120. Llanos AR, Moss ME, Pinzòn MC, Dye T, Sinkin RA, Kendig JW. Epidemiology of neonatal necrotising enterocolitis: a population-based study. *Paediatr Perinat Epidemiol* 2002;**16**:342–9. https://doi.org/10.1046/j.1365-3016.2002.00445.x

121. Stoll BJ. Epidemiology of necrotizing enterocolitis. *Clin Perinatol* 1994;**21**:205–18. https://doi.org/10.1016/S0095-5108(18)30341-5

122. Cole CR, Hansen NI, Higgins RD, Ziegler TR, Stoll BJ, Eunice Kennedy Shriver NICHD Neonatal Research Network. Very low birth weight preterm infants with surgical short bowel syndrome: incidence, morbidity and mortality, and growth outcomes at 18 to 22 months. *Pediatrics* 2008;**122**:e573–82. https://doi.org/10.1542/peds.2007-3449

123. Holman RC, Stoll BJ, Clarke MJ, Glass RI. The epidemiology of necrotizing enterocolitis infant mortality in the United States. *Am J Public Health* 1997;**87**:2026–31. https://doi.org/10.2105/AJPH.87.12.2026

124. Holman RC, Stoll BJ, Curns AT, Yorita KL, Steiner CA, Schonberger LB. Necrotising enterocolitis hospitalisations among neonates in the United States. *Paediatr Perinat Epidemiol* 2006;**20**:498–506. https://doi.org/10.1111/j.1365-3016.2006.00756.x

125. Fellman V, Hellström-Westas L, Norman M, Westgren M, Källén K, Lagercrantz H, *et al.* One-year survival of extremely preterm infants after active perinatal care in Sweden. *JAMA* 2009;**301**:2225–33. https://doi.org/10.1001/jama.2009.771

126. Hintz SR, Kendrick DE, Stoll BJ, Vohr BR, Fanaroff AA, Donovan EF, *et al.* Neurodevelopmental and growth outcomes of extremely low birth weight infants after necrotizing enterocolitis. *Pediatrics* 2005;**115**:696–703. https://doi.org/10.1542/peds.2004-0569

127. Costeloe KL, Hennessy EM, Haider S, Stacey F, Marlow N, Draper ES. Short term outcomes after extreme preterm birth in England: comparison of two birth cohorts in 1995 and 2006 (the EPICure studies). *BMJ* 2012;**345**:e7976. https://doi.org/10.1136/bmj.e7976

128. Sankaran K, Puckett B, Lee DS, Seshia M, Boulton J, Qiu Z, Lee SK, Canadian Neonatal Network. Variations in incidence of necrotizing enterocolitis in Canadian neonatal intensive care units. *J Pediatr Gastroenterol Nutr* 2004;**39**:366–72. https://doi.org/10.1097/00005176-200410000-00012

129. Luig M, Lui K, NSW. ACT NICUS Group. Epidemiology of necrotizing enterocolitis: Part I – changing regional trends in extremely preterm infants over 14 years. *J Paediatr Child Health* 2005;**41**:169–73. https://doi.org/10.1111/j.1440-1754.2005.00582.x

130. Stoll BJ, Hansen NI, Bell EF, Shankaran S, Laptook AR, Walsh MC, *et al.* Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. *Pediatrics* 2010;**126**:443–56. https://doi.org/10.1542/peds.2009-2959

131. Battersby C, Longford N, Mandalia S, Costeloe K, Modi N, UK Neonatal Collaborative Necrotising Enterocolitis (UKNC-NEC) study group. Incidence and enteral feed antecedents of severe neonatal necrotising enterocolitis across neonatal networks in England, 2012–13: a whole-population surveillance study. *Lancet Gastroenterol Hepatol* 2017;**2**:43–51. https://doi.org/10.1016/S2468-1253(16)30117-0

132. Royal College of Paediatrics and Child Health. *UK-WHO Growth Charts: Neonatal and Infant Close Monitoring (NICM)*. URL: www.rcpch.ac.uk/resources/uk-who-growth-charts-neonatal-infant-close-monitoring-nicm (accessed 8 November 2018).

133. McLennan D, Barnes H, Noble M, Davies J, Garratt E, Dibben C. *The English Indices of Deprivation 2010*. London: Department for Communities and Local Government; 2011. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf (accessed 17 July 2019).

134. NHS Digital. *Linked HES–ONS Mortality Data*. Leeds: NHS Digital; 2018. URL: https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/linked-hes-ons-mortality-data (accessed 8 November 2018).

135. Cole TJ, Hey E, Richmond S. The PREM score: a graphical tool for predicting survival in very preterm births. *Arch Dis Child Fetal Neonatal Ed* 2010;**95**:F14–9. https://doi.org/10.1136/adc.2009.164533

136. Office for National Statistics. Newport: Office for National Statistics; 2018. URL: www.ons.gov.uk/ (accessed 8 November 2018).

137. Costeloe K, Hennessy E, Gibson AT, Marlow N, Wilkinson AR. The EPICure study: outcomes to discharge from hospital for infants born at the threshold of viability. *Pediatrics* 2000;**106**:659–71. https://doi.org/10.1542/peds.106.4.659

138. Esteve, J. *Statistical Methods in Cancer Research: Volume IV – Descriptive Epidemiology*. New York, NY: Oxford University Press; 1994.

139. Kiely JL. What is the population-based risk of preterm birth among twins and other multiples? *Clin Obstet Gynecol* 1998;**41**:3–11. https://doi.org/10.1097/00003081-199803000-00005

140. Buekens P, Wilcox A. Why do small twins have a lower mortality rate than small singletons? *Am J Obstet Gynecol* 1993;**168**:937–41. https://doi.org/10.1016/S0002-9378(12)90849-2

141. Jones HE, Ohlssen DI, Spiegelhalter DJ. Use of the false discovery rate when comparing multiple health care providers. *J Clin Epidemiol* 2008;**61**:232–40. https://doi.org/10.1016/j.jclinepi.2007.04.017

142. Bolisetty S, Legge N, Bajuk B, Lui K, New South Wales and the Australian capital territory neonatal intensive care units' data collection. Preterm infant outcomes in New South Wales and the Australian Capital Territory. *J Paediatr Child Health* 2015;**51**:713–21. https://doi.org/10.1111/jpc.12848

143. Costeloe KL, Bowler U, Brocklehurst P, Hardy P, Heal P, Juszczak E, *et al.* A randomised controlled trial of the probiotic Bifidobacterium breve BBG-001 in preterm babies to prevent sepsis, necrotising enterocolitis and death: the Probiotics in Preterm infantS (PiPS) trial. *Health Technol Assess* 2016;**20**(66). https://doi.org/10.3310/hta20660

144. Levene MI. Measurement of the growth of the lateral ventricles in preterm infants with real-time ultrasound. *Arch Dis Child* 1981;**56**:900–4. https://doi.org/10.1136/adc.56.12.900

145. Agresti A, Coull BA. Approximate is better than 'exact' for interval estimation of binomial proportions. *Am Stat* 1998;**52**:119–26. https://doi.org/10.1080/00031305.1998.10480550

146. Brown LD, Cai TT, DasGupta A. Interval estimation for a binomial proportion. *Stat Sci* 2001;**16**:101–17. https://doi.org/10.1214/ss/1009213286

147. Rogers W. Regression standard errors in clustered samples. *Stata Tech Bull* 1994;**3**.

148. Royal College of Paediatrics and Child Health. *Annual Reports 2009, 2010, 2012, 2013*. London: Royal College of Paediatrics and Child Health; 2014.

149. Köpcke F, Lubgan D, Fietkau R, Scholler A, Nau C, Stürzl M, *et al.* Evaluating predictive modeling algorithms to assess patient eligibility for clinical trials from routine data. *BMC Med Inform Decis Mak* 2013;**13**:134. https://doi.org/10.1186/1472-6947-13-134

150. De Moor G, Sundgren M, Kalra D, Schmidt A, Dugas M, Claerhout B, *et al.* Using electronic health records for clinical research: the case of the EHR4CR project. *J Biomed Inform* 2015;**53**:162–73. https://doi.org/10.1016/j.jbi.2014.10.006

151. Newsham AC, Johnston C, Hall G, Leahy MG, Smith AB, Vikram A, *et al.* Development of an advanced database for clinical trials integrated with an electronic patient record system. *Comput Biol Med* 2011;**41**:575–86. https://doi.org/10.1016/j.compbiomed.2011.04.014

152. Carlo WA, Finer NN, Walsh MC, Rich W, Gantz MG, Laptook AR, *et al.* Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med* 2010;**362**:1959–69. https://doi.org/10.1056/NEJMoa0911781

153. Horbar JD, Carpenter JH, Buzas J, Soll RF, Suresh G, Bracken MB, *et al.* Collaborative quality improvement to promote evidence based surfactant for preterm infants: a cluster randomised trial. *BMJ* 2004;**329**:1004. https://doi.org/10.1136/bmj.329.7473.1004

154. Vohra S, Reilly M, Rac VE, Bhaloo Z, Zayack D, Wimmer J, *et al.* Study protocol for multicentre randomized controlled trial of HeLP (Heat Loss Prevention) in the delivery room. *Contemp Clin Trials* 2013;**36**:54–60. https://doi.org/10.1016/j.cct.2013.06.001

155. Milligan DW. Outcomes of children born very preterm in Europe. *Arch Dis Child Fetal Neonatal Ed* 2010;**95**:F234–40. https://doi.org/10.1136/adc.2008.143685

156. Hille ET, Weisglas-Kuperus N, van Goudoever JB, Jacobusse GW, Ens-Dokkum MH, de Groot L, *et al.* Functional outcomes and participation in young adulthood for very preterm and very low birth weight infants: the Dutch project on preterm and small for gestational age infants at 19 years of age. *Pediatrics* 2007;**120**:e587–95. https://doi.org/10.1542/peds.2006-2407

157. Doyle LW, Anderson PJ. Adult outcome of extremely preterm infants. *Pediatrics* 2010;**126**:342–51. https://doi.org/10.1542/peds.2010-0710

158. Pedersen SJ, Sommerfelt K, Markestad T. Early motor development of premature infants with birthweight less than 2000 grams. *Acta Paediatr* 2000;**89**:1456–61. https://doi.org/10.1111/j.1651-2227.2000.tb02776.x

159. Paneth N, Qiu H, Rosenbaum P, Saigal S, Bishai S, Jetton J, *et al.* Reliability of classification of cerebral palsy in low-birthweight children in four countries. *Dev Med Child Neurol* 2003;**45**:628–33. https://doi.org/10.1111/j.1469-8749.2003.tb00968.x

160. Munck P, Niemi P, Lapinleimu H, Lehtonen L, Haataja L, PIPARI Study Group. Stability of cognitive outcome from 2 to 5 years of age in very low birth weight children. *Pediatrics* 2012;**129**:503–8. https://doi.org/10.1542/peds.2011-1566

161. Potharst ES, Houtzager BA, van Sonderen L, Tamminga P, Kok JH, Last BF, van Wassenaer AG. Prediction of cognitive abilities at the age of 5 years using developmental follow-up assessments at the age of 2 and 3 years in very preterm children. *Dev Med Child Neurol* 2012;**54**:240–6. https://doi.org/10.1111/j.1469-8749.2011.04181.x

162. Hack M, Taylor HG, Drotar D, Schluchter M, Cartar L, Wilson-Costello D, *et al.* Poor predictive validity of the Bayley scales of infant development for cognitive function of extremely low birth weight children at school age. *Pediatrics* 2005;**116**:333–41. https://doi.org/10.1542/peds.2005-0173

163. Roberts G, Anderson PJ, Doyle LW, Victorian Infant Collaborative Study Group. The stability of the diagnosis of developmental disability between ages 2 and 8 in a geographic cohort of very preterm children born in 1997. *Arch Dis Child* 2010;**95**:786–90. https://doi.org/10.1136/adc.2009.160283

164. Bracewell M, Marlow N. Patterns of motor disability in very preterm children. *Ment Retard Dev Disabil Res Rev* 2002;**8**:241–8. https://doi.org/10.1002/mrdd.10049

165. Pallás Alonso CR, de La Cruz Bértolo J, Medina López MC, Orbea Gallardo C, Gómez Castillo E, Simón De Las Heras R. Cerebral palsy and age of sitting and walking in children weighing less than 1500 g at birth. *An Esp Pediatr* 2000;**53**:48–52. https://doi.org/10.1016/S1695-4033(00)77413-3

166. de Kieviet JF, Piek JP, Aarnoudse-Moens CS, Oosterlaan J. Motor development in very preterm and very low-birth-weight children from birth to adolescence: a meta-analysis. *JAMA* 2009;**302**:2235–42. https://doi.org/10.1001/jama.2009.1708

167. Wood NS, Marlow N, Costeloe K, Gibson AT, Wilkinson AR. Neurologic and developmental disability after extremely preterm birth: EPICure Study Group. *N Engl J Med* 2000;**343**:378–84. https://doi.org/10.1056/NEJM200008103430601

168. Stoelhorst GM, Rijken M, Martens SE, van Zwieten PH, Feenstra J, Zwinderman AH, *et al.* Developmental outcome at 18 and 24 months of age in very preterm children: a cohort study from 1996 to 1997. *Early Hum Dev* 2003;**72**:83–95. https://doi.org/10.1016/S0378-3782(03)00011-2

169. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;**288**:728–37. https://doi.org/10.1001/jama.288.6.728

170. Hutchinson EA, De Luca CR, Doyle LW, Roberts G, Anderson PJ, Victorian Infant Collaborative Study Group. School-age outcomes of extremely preterm or extremely low birth weight children. *Pediatrics* 2013;**131**:e1053–61. https://doi.org/10.1542/peds.2012-2311

171. Jansson-Verkasalo E, Valkama M, Vainionpää L, Pääkkö E, Ilkko E, Lehtihalmes M. Language development in very low birth weight preterm children: a follow-up study. *Folia Phoniatr Logop* 2004;**56**:108–19. https://doi.org/10.1159/000076062

172. Sansavini A, Guarini A, Alessandroni R, Faldella G, Giovanelli G, Salvioli G. Are early grammatical and phonological working memory abilities affected by preterm birth? *J Commun Disord* 2007;**40**:239–56. https://doi.org/10.1016/j.jcomdis.2006.06.009

173. Wolke D, Meyer R. Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. *Dev Med Child Neurol* 1999;**41**:94–109. https://doi.org/10.1017/S0012162299000201

174. Pietz J, Peter J, Graf R, Rauterberg-Ruland I, Rupp A, Sontheimer D, Linderkamp O. Physical growth and neurodevelopmental outcome of nonhandicapped low-risk children born preterm. *Early Hum Dev* 2004;**79**:131–43. https://doi.org/10.1016/j.earlhumdev.2004.05.001

175. Vohr B. Speech and language outcomes of very preterm infants. *Semin Fetal Neonatal Med* 2014;**19**:78–83. https://doi.org/10.1016/j.siny.2013.10.007

176. Barre N, Morgan A, Doyle LW, Anderson PJ. Language abilities in children who were very preterm and/or very low birth weight: a meta-analysis. *J Pediatr* 2011;**158**:766–74.e1. https://doi.org/10.1016/j.jpeds.2010.10.032

177. van Noort-van der Spek IL, Franken MC, Weisglas-Kuperus N. Language functions in preterm-born children: a systematic review and meta-analysis. *Pediatrics* 2012;**129**:745–54. https://doi.org/10.1542/peds.2011-1728

178. Synnes AR, Anson S, Baum J, Usher L. Incidence and pattern of hearing impairment in children with ≤ 800 g birthweight in British Columbia, Canada. *Acta Paediatr* 2012;**101**:e48–54. https://doi.org/10.1111/j.1651-2227.2011.02437.x

179. D'Amore A, Broster S, Le Fort W, Curley A, East Anglian Very Low Birthweight Project. Two-year outcomes from very low birthweight infants in a geographically defined population across 10 years, 1993–2002: comparing 1993–1997 with 1998–2002. *Arch Dis Child Fetal Neonatal Ed* 2011;**96**:F178–85. https://doi.org/10.1136/adc.2009.171876

180. Ari-Even Roth D, Hildesheimer M, Maayan-Metzger A, Muchnik C, Hamburger A, Mazkeret R, Kuint J. Low prevalence of hearing impairment among very low birthweight infants as detected by universal neonatal hearing screening. *Arch Dis Child Fetal Neonatal Ed* 2006;**91**:F257–62. https://doi.org/10.1136/adc.2005.074476

181. Veen S, Sassen ML, Schreuder AM, Ens-Dokkum MH, Verloove-Vanhorick SP, Brand R, *et al.* Hearing loss in very preterm and very low birthweight infants at the age of 5 years in a nationwide cohort. *Int J Pediatr Otorhinolaryngol* 1993;**26**:11–28. https://doi.org/10.1016/0165-5876(93)90192-6

182. Royal College of Paediatrics and Child Health, Royal College of Ophthalmologists, British Association of Perinatal Medicine, Bliss. *UK Retinopathy of Prematurity Guideline May 2008*. London: Royal College of Paediatrics and Child Health; 2018. URL: www.rcophth.ac.uk/wp-content/uploads/2014/12/2008-SCI-021-Guidelines-Retinopathy-of-Prematurity.pdf (accessed 8 November 2018).

183. Dhaliwal C, Fleck B, Wright E, Graham C, McIntosh N. Incidence of retinopathy of prematurity in Lothian, Scotland, from 1990 to 2004. *Arch Dis Child Fetal Neonatal Ed* 2008;**93**:F422–6. https://doi.org/10.1136/adc.2007.134791

184. Rahi JS, Cable N, British Childhood Visual Impairment Study Group. Severe visual impairment and blindness in children in the UK. *Lancet* 2003;**362**:1359–65. https://doi.org/10.1016/S0140-6736(03)14631-4

185. Indredavik MS, Vik T, Heyerdahl S, Kulseng S, Fayers P, Brubakk AM. Psychiatric symptoms and disorders in adolescents with low birth weight. *Arch Dis Child Fetal Neonatal Ed* 2004;**89**:F445–50. https://doi.org/10.1136/adc.2003.038943

186. Farooqi A, Hägglöf B, Sedin G, Gothefors L, Serenius F. Mental health and social competencies of 10- to 12-year-old children born at 23 to 25 weeks of gestation in the 1990s: a Swedish national prospective follow-up study. *Pediatrics* 2007;**120**:118–33. https://doi.org/10.1542/peds.2006-2988

187. Elgen I, Sommerfelt K, Markestad T. Population based, controlled study of behavioural problems and psychiatric disorders in low birthweight children at 11 years of age. *Arch Dis Child Fetal Neonatal Ed* 2002;**87**:F128–32. https://doi.org/10.1136/fn.87.2.F128

188. Johnson S, Marlow N. Preterm birth and childhood psychiatric disorders. *Pediatr Res* 2011;**69**(5 Pt 2):11R–8R. https://doi.org/10.1203/PDR.0b013e318212faa0

189. Pinto-Martin JA, Levy SE, Feldman JF, Lorenz JM, Paneth N, Whitaker AH. Prevalence of autism spectrum disorder in adolescents born weighing < 2000 grams. *Pediatrics* 2011;**128**:883–91. https://doi.org/10.1542/peds.2010-2846

190. Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Autism spectrum disorders in extremely preterm children. *J Pediatr* 2010;**156**:525–31.e2. https://doi.org/10.1016/j.jpeds.2009.10.041

191. Williams JG, Higgins JP, Brayne CE. Systematic review of prevalence studies of autism spectrum disorders. *Arch Dis Child* 2006;**91**:8–15. https://doi.org/10.1136/adc.2004.062083

192. Autism and Developmental Disabilities Monitoring Network Surveillance, Centers for Disease Control and Prevention (CDC). Prevalence of autism spectrum disorders: autism and developmental disabilities monitoring network, United States, 2006. *MMWR Surveill Summ* 2009;**58**:1–20.

193. Aarnoudse-Moens CS, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics* 2009;**124**:717–28. https://doi.org/10.1542/peds.2008-2816

194. 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. *J Am Acad Child Adolesc Psychiatry* 2010;**49**:453–63.e1. https://doi.org/10.1016/j.jaac.2010.02.002

195. Johnson S, Marlow N. Developmental screen or developmental testing? *Early Hum Dev* 2006;**82**:173–83. https://doi.org/10.1016/j.earlhumdev.2006.01.008

196. Flynn JR. Searching for justice: the discovery of IQ gains over time. *Am Psychol* 1999;**54**:5–20. https://doi.org/10.1037/0003-066X.54.1.5

197. Aylward GP, Aylward BS. The changing yardstick in measurement of cognitive abilities in infancy. *J Dev Behav Pediatr* 2011;**32**:465–8. https://doi.org/10.1097/DBP.0b013e3182202eb3

198. Shribman S, Billingham K. *Healthy child programme-pregnancy and the first five years*. Department of Health and Social Care. 2009. URL: www.gov.uk/government/uploads/system/uploads/attachment_data/file/167998/Health_Child_Programme.pdf (accessed 7 November 2014).

199. The Victorian Infant Collaborative Study Group. Improved outcome into the 1990s for infants weighing 500–999 g at birth: The Victorian Infant Collaborative Study Group. *Arch Dis Child Fetal Neonatal Ed* 1997;**77**:F91–4. https://doi.org/10.1136/fn.77.2.F91

200. Vohr BR, Wright LL, Dusick AM, Perritt R, Poole WK, Tyson JE, *et al.* Center differences and outcomes of extremely low birth weight infants. *Pediatrics* 2004;**113**:781–9. https://doi.org/10.1542/peds.113.4.781

201. Bayley N. *Technical manual for the Bayley Scales of Infant and Toddler Development (Third Edition)*. 3rd edn. San Antonio, TX: Harcourt Assessment; 2006. https://doi.org/10.1037/t14978-000

202. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW, Victorian Infant Collaborative Group. Underestimation of developmental delay by the new Bayley-III Scale. *Arch Pediatr Adolesc Med* 2010;**164**:352–6. https://doi.org/10.1001/archpediatrics.2010.20

203. Lowe JR, Erickson SJ, Schrader R, Duncan AF. Comparison of the Bayley II Mental Developmental Index and the Bayley III Cognitive Scale: are we measuring the same thing? *Acta Paediatr* 2012;**101**:e55–8. https://doi.org/10.1111/j.1651-2227.2011.02517.x

204. Moore T, Johnson S, Haider S, Hennessy E, Marlow N. Relationship between test scores using the second and third editions of the Bayley Scales in extremely preterm children. *J Pediatr* 2012;**160**:553–8. https://doi.org/10.1016/j.jpeds.2011.09.047

205. Vohr BR, Stephens BE, Higgins RD, Bann CM, Hintz SR, Das A, *et al.* Are outcomes of extremely preterm infants improving? Impact of Bayley assessment on outcomes. *J Pediatr* 2012;**161**:222–8.e3. https://doi.org/10.1016/j.jpeds.2012.01.057

206. Johnson S, Moore T, Marlow N. Using the Bayley-III to assess neurodevelopmental delay: which cut-off should be used? *Pediatr Res* 2014;**75**:670–4. https://doi.org/10.1038/pr.2014.10

207. Casenhiser D, Breinbauer C, Greenspan S. *Evaluating Greenspan's social emotional growth scale/chart as a screening for autism*. Presented at the ICDL 11th Annual International Conference: Critical Factors for Optimal Outcomes for Children with Autism and Special Needs, Tyson's Corner, VA, November 2007.

208. Vohr BR, Msall ME, Wilson D, Wright LL, McDonald S, Poole WK. Spectrum of gross motor function in extremely low birth weight children with cerebral palsy at 18 months of age. *Pediatrics* 2005;**116**:123–9. https://doi.org/10.1542/peds.2004-1810

209. Haataja L, Mercuri E, Regev R, Cowan F, Rutherford M, Dubowitz V, Dubowitz L. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. *J Pediatr* 1999;**135**(2 Pt 1):153–61. https://doi.org/10.1016/S0022-3476(99)70016-8

210. Frisone MF, Mercuri E, Laroche S, Foglia C, Maalouf EF, Haataja L, *et al.* Prognostic value of the neurologic optimality score at 9 and 18 months in preterm infants born before 31 weeks' gestation. *J Pediatr* 2002;**140**:57–60. https://doi.org/10.1067/mpd.2002.119626

211. Robins DL, Fein D, Barton ML, Green JA. The modified checklist for autism in toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord* 2001;**31**:131–44. https://doi.org/10.1023/A:1010738829569

212. Limperopoulos C, Bassan H, Sullivan NR, Soul JS, Robertson RL, Moore M, *et al.* Positive screening for autism in ex-preterm infants: prevalence and risk factors. *Pediatrics* 2008;**121**:758–65. https://doi.org/10.1542/peds.2007-2158

213. Kuban KC, O'Shea TM, Allred EN, Tager-Flusberg H, Goldstein DJ, Leviton A. Positive screening on the modified checklist for autism in toddlers (M-CHAT) in extremely low gestational age newborns. *J Pediatr* 2009;**154**:535–40.e1. https://doi.org/10.1016/j.jpeds.2008.10.011

214. Moore T, Johnson S, Hennessy E, Marlow N. Screening for autism in extremely preterm infants: problems in interpretation. *Dev Med Child Neurol* 2012;**54**:514–20. https://doi.org/10.1111/j.1469-8749.2012.04265.x

215. Luyster RJ, Kuban KC, O'Shea TM, Paneth N, Allred EN, Leviton A, ELGAN Study investigators. The modified checklist for autism in toddlers in extremely low gestational age newborns: individual items associated with motor, cognitive, vision and hearing limitations. *Paediatr Perinat Epidemiol* 2011;**25**:366–76. https://doi.org/10.1111/j.1365-3016.2010.01187.x

216. Wetherby AM, Prizant BM. *Communication and Symbolic Behavior Scales Developmental Profile*: *First Normed Edition* Baltimore, MD: Paul H Brookes Publishing; 2002. https://doi.org/10.1037/t11529-000

217. Dunn W. *Infant/Toddler Sensory Profile*. San Antonio, TX: Harcourt Assessment; 2002.

218. Stephens BE, Bann CM, Watson VE, Sheinkopf SJ, Peralta-Carcelen M, Bodnar A, *et al.* Screening for autism spectrum disorders in extremely preterm infants. *J Dev Behav Pediatr* 2012;**33**:535–41. https://doi.org/10.1097/DBP.0b013e31825fd0af

219. Dudova I, Kasparova M, Markova D, Zemankova J, Beranova S, Urbanek T, Hrdlicka M. Screening for autism in preterm children with extremely low and very low birth weight. *Neuropsychiatr Dis Treat* 2014;**10**:277–82. https://doi.org/10.2147/NDT.S57057

220. Allison C, Baron-Cohen S, Wheelwright S, Charman T, Richler J, Pasco G, Brayne C. The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): a normally distributed quantitative measure of autistic traits at 18-24 months of age – preliminary report. *J Autism Dev Disord* 2008;**38**:1414–25. https://doi.org/10.1007/s10803-007-0509-7

221. National Perinatal Epidemiology Unit (NPEU). *Disability and Perinatal Care: Measurement of Health Status at Two Years: A Report of Two Working Groups Convened by the National Perinatal Epidemiology Unit and the former Oxford Regional Health Authority*. Oxford: NPEU, 1994.

222. British Association of Perinatal Medicine. *Classification of Health Status at 2 Years as a Perinatal Outcome*. British Association of Perinata Medicine. 2008. URL: www.networks.nhs.uk/nhs-networks/staffordshire-shropshire-and-black-country-newborn/documents/2_year_Outcome_BAPM_WG_report_v6_Jan08.pdf (accessed 8 November 2018).

223. Jones HP, Guildea ZE, Stewart JH, Cartlidge PH. The health status questionnaire: achieving concordance with published disability criteria. *Arch Dis Child* 2002;**86**:15–20. https://doi.org/10.1136/adc.86.1.15

224. Bohin S, Draper ES, Field DJ. Health status of a population of infants born before 26 weeks gestation derived from routine data collected between 21 and 27 months post-delivery. *Early Hum Dev* 1999;**55**:9–18. https://doi.org/10.1016/S0378-3782(99)00003-1

225. Marlow N, Wolke D, Bracewell MA, Samara M, EPICure Study Group. Neurologic and developmental disability at six years of age after extremely preterm birth. *N Engl J Med* 2005;**352**:9–19. https://doi.org/10.1056/NEJMoa041367

226. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Dev Med Child Neurol* 1997;**39**:214–23. https://doi.org/10.1111/j.1469-8749.1997.tb07414.x

227. Jahnsen R, Aamodt G, Rosenbaum P. Gross motor function classification system used in adults with cerebral palsy: agreement of self-reported versus professional rating. *Dev Med Child Neurol* 2006;**48**:734–8. https://doi.org/10.1017/S0012162206001575

228. Palisano RJ, Cameron D, Rosenbaum PL, Walter SD, Russell D. Stability of the gross motor function classification system. *Dev Med Child Neurol* 2006;**48**:424–8. https://doi.org/10.1017/S0012162206000934

229. Wood E, Rosenbaum P. The gross motor function classification system for cerebral palsy: a study of reliability and stability over time. *Dev Med Child Neurol* 2000;**42**:292–6. https://doi.org/10.1017/S0012162200000529

230. Eliasson AC, Krumlinde-Sundholm L, Rösblad B, Beckung E, Arner M, Ohrvall AM, Rosenbaum P. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Dev Med Child Neurol* 2006;**48**:549–54. https://doi.org/10.1017/S0012162206001162

231. Salt A, D'Amore A, Ahluwalia J, Seward A, Kaptoge S, Halliday S, Dorling J, East Anglian Very Low Birthweight Project Group. Outcome at 2 years for very low birthweight infants in a geographical population: risk factors, cost, and impact of congenital anomalies. *Early Hum Dev* 2006;**82**:125–33. https://doi.org/10.1016/j.earlhumdev.2005.10.016

232. Dorling JS, Field DJ. Follow up of infants following discharge from the neonatal unit: structure and process. *Early Hum Dev* 2006;**82**:151–6. https://doi.org/10.1016/j.earlhumdev.2006.01.006

233. Dawson C, Perkins M, Draper E, Johnson A, Field D. Are outcome data regarding the survivors of neonatal care available from routine sources? *Arch Dis Child Fetal Neonatal Ed* 1997;**77**:F206–10. https://doi.org/10.1136/fn.77.3.F206

234. Johnson A, King R. Can routine information systems be used to monitor serious disability? *Arch Dis Child* 1999;**80**:63–6. https://doi.org/10.1136/adc.80.1.63

235. Modi N, Carpenter T. Fetal growth and coronary heart disease. *Lancet* 1997;**349**:286–7. https://doi.org/10.1016/S0140-6736(05)64900-8

236. House of Commons Health Committee Session 1991–2. *Maternity Services Vol. 1: Report Together with Appendices and the Proceedings of the Committee*. London: HSMO; 1992.

237. Audit Commission. *Children First: A Study of Hospital Services*. (Audit Commission NHS report No. 7). London: HMSO; 1993. URL: https://webarchive.nationalarchives.gov.uk/20150423154441/http://archive.audit-commission.gov.uk/auditcommission/aboutus/publications/pages/national-reports-and-studies-archive.aspx.html (accessed 9 July 2019).

238. Clinical Standards Advisory Group. *Neonatal Intensive Care*. London: HMSO; 1993.

239. Cumberlege J. *Changing Childbirth: Part I – Report of the Expert Maternity Group. Winterton report*. London: The Stationery Office; 1993.

240. British Association of Perinatal Medicine. *Standards for Hospitals Providing Neonatal Intensive and High Dependency Care and Categories of Babies Requiring Neonatal Care*. London: British Association of Perinatal Medicine; 2001.

241. The National Audit Office. *Caring for Vulnerable Babies: The Reorganisation of Neonatal Services in England*. London: The Stationery Office; 2007.

242. National Institute for Health and Care Excellence (NICE). *Neonatal Specialist Care: Quality Standard*. London: NICE; 2010.

243. Fooks J. Four key questions that identify severe disability. *Arch Dis Child* 1999;**80**:67–8. https://doi.org/10.1136/adc.80.1.67

244. Kim MM, O'Connor KS, McLean J, Robson A, Chance G. Do parents and professionals agree on the developmental status of high-risk infants? *Pediatrics* 1996;**97**:676–81.

245. Bortolus R, Parazzini F, Trevisanuto D, Cipriani S, Ferrarese P, Zanardo V, Gruppo di Studio Metodologie nei Follow-up Pediatrici. Developmental assessment of preterm and term children at 18 months: reproducibility and validity of a postal questionnaire to parents. *Acta Paediatr* 2002;**91**:1101–7. https://doi.org/10.1111/j.1651-2227.2002.tb00106.x

246. Johnson S, Marlow N, Wolke D, Davidson L, Marston L, O'Hare A, *et al.* Validation of a parent report measure of cognitive development in very preterm infants. *Dev Med Child Neurol* 2004;**46**:389–97. https://doi.org/10.1017/S0012162204000635

247. Johnson S, Wolke D, Marlow N, Preterm Infant Parenting Study Group. Developmental assessment of preterm infants at 2 years: validity of parent reports. *Dev Med Child Neurol* 2008;**50**:58–62. https://doi.org/10.1111/j.1469-8749.2007.02010.x

248. Pritchard MA, Colditz PB, Beller EM, Queensland Optimising Preterm Infant Outcomes Group. Parents' evaluation of developmental status in children born with a birthweight of 1250 g or less. *J Paediatr Child Health* 2005;**41**:191–6. https://doi.org/10.1111/j.1440-1754.2005.00586.x

249. Da Costa D, Bann CM, Hansen NI, Shankaran S, Delaney-Black V, National Institute of Child Health and Human Development Neonatal Research Network. Validation of the Functional Status II questionnaire in the assessment of extremely-low-birthweight infants. *Dev Med Child Neurol* 2009;**51**:536–44. https://doi.org/10.1111/j.1469-8749.2009.03318.x

250. Skellern CY, Rogers Y, O'Callaghan MJ. A parent-completed developmental questionnaire: follow up of ex-premature infants. *J Paediatr Child Health* 2001;**37**:125–9. https://doi.org/10.1046/j.1440-1754.2001.00604.x

251. Fooks J, Mutch L, Yudkin P, Johnson A, Elbourne D. Comparing two methods of follow up in a multicentre randomised trial. *Arch Dis Child* 1997;**76**:369–76. https://doi.org/10.1136/adc.76.4.369

252. Martin AJ, Darlow BA, Salt A, Hague W, Sebastian L, McNeill N, Tarnow-Mordi W, *et al.* Performance of the Parent Report of Children's Abilities-Revised (PARCA-R) versus the Bayley Scales of Infant Development III. *Arch Dis Child* 2013;**98**:955–8. https://doi.org/10.1136/archdischild-2012-303288

253. Marlow N, Greenough A, Peacock JL, Marston L, Limb ES, Johnson AH, Calvert SA. Randomised trial of high frequency oscillatory ventilation or conventional ventilation in babies of gestational age 28 weeks or less: respiratory and neurological outcomes at 2 years. *Arch Dis Child Fetal Neonatal Ed* 2006;**91**:F320–6. https://doi.org/10.1136/adc.2005.079632

254. Brocklehurst P, Farrell B, King A, Juszczak E, Darlow B, Haque K, *et al.* Treatment of neonatal sepsis with intravenous immune globulin. *N Engl J Med* 2011;**365**:1201–11. https://doi.org/10.1056/NEJMoa1100441

255. Cummings SM, Savitz LA, Konrad TR. Reported response rates to mailed physician questionnaires. *Health Serv Res* 2001;**35**:1347–55.

256. Field D, Draper ES, Gompels MJ, Green C, Johnson A, Shortland D, *et al.* Measuring later health status of high risk infants: randomised comparison of two simple methods of data collection. *BMJ* 2001;**323**:1276–81. https://doi.org/10.1136/bmj.323.7324.1276

257. Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. *Gross Motor Function Classification System (GMFCS)*. URL: www.motorgrowth.canchild.ca/en/GMFCS/resources/GMFCS_English.pdf (accessed 8 July 2014).

258. Department for Communities and Local Government. *English Indices of Deprivation 2010*. 2011. URL: www.communities.gov.uk/documents/statistics/pdf/1871208.pdf (accessed 29 August 2018).

259. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. 1st edn. Washington, DC: American Educational Research Association; 1999.

260. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74. https://doi.org/10.2307/2529310

261. London Perinatal Group. *London Perinatal Networks Annual Report*. 2008. URL: www.neonatal.org.uk/documents/4391.pdf (accessed 29 August 2018).

262. Obuchowski NA, Zhou XH. Prospective studies of diagnostic test accuracy when disease prevalence is low. *Biostatistics* 2002;**3**:477–92. https://doi.org/10.1093/biostatistics/3.4.477

263. Cochran WG. Some methods for strengthening the common Chi^2 tests. *Biometrics* 1954;**10**:417–51. https://doi.org/10.2307/3001616

264. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Cochrane diagnostic test accuracy working group systematic reviews of diagnostic test accuracy group. *Ann Intern Med* 2008;**149**:889–97.

265. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535. https://doi.org/10.1136/bmj.b2535

266. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;**155**:529–36. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

267. Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol* 2005;**5**:19. https://doi.org/10.1186/1471-2288-5-19

268. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993;**12**:1293–316. https://doi.org/10.1002/sim.4780121403

269. Deeks JJ. Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;**323**:157–62. https://doi.org/10.1136/bmj.323.7305.157

270. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;**13**:313–21. https://doi.org/10.1177/0272989X9301300408

271. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;**20**:2865–84. https://doi.org/10.1002/sim.942

272. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;**58**:982–90. https://doi.org/10.1016/j.jclinepi.2005.02.022

273. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**:239–51. https://doi.org/10.1093/biostatistics/kxl004

274. Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 2002;**31**:88–95. https://doi.org/10.1093/ije/31.1.88

275. Higgins JPT, Green S. editors. *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0*. The Cochrane Collaboration; 2011. URL: https://training.cochrane.org/handbook (accessed 19 July 2019).

276. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;**58**:882–93. https://doi.org/10.1016/j.jclinepi.2005.01.016

277. Schenker N, Gentleman JF. On judging the significance of differences by examining the overlap between confidence intervals. *Am Stat* 2001;**55**:182–6. https://doi.org/10.1198/000313001317097960

278. Bassan H, Stolar O, Geva R, Eshel R, Fattal-Valevski A, Leitner Y, *et al.* Intrauterine growth-restricted neonates born at term or preterm: how different? *Pediatr Neurol* 2011;**44**:122–30. https://doi.org/10.1016/j.pediatrneurol.2010.09.012

279. Fedrizzi E, Inverno M, Botteon G, Anderloni A, Filippini G, Farinotti M. The cognitive development of children born preterm and affected by spastic diplegia. *Brain Dev* 1993;**15**:428–32. https://doi.org/10.1016/0387-7604(93)90082-J

280. McGrath MM, Sullivan MC, Lester BM, Oh W. Longitudinal neurologic follow-up in neonatal intensive care unit survivors with various neonatal morbidities. *Pediatrics* 2000;**106**:1397–405. https://doi.org/10.1542/peds.106.6.1397

281. Gray PH, Burns YR, Mohay HA, O'Callaghan MJ, Tudehope DI. Neurodevelopmental outcome of preterm infants with bronchopulmonary dysplasia. *Arch Dis Child Fetal Neonatal Ed* 1995;**73**:F128–34. https://doi.org/10.1136/fn.73.3.F128

282. Smith KE, Landry SH, Swank PR. The role of early maternal responsiveness in supporting school-aged cognitive development for children who vary in birth status. *Pediatrics* 2006;**117**:1608–17. https://doi.org/10.1542/peds.2005-1284

283. Bowen JR, Gibson FL, Leslie GI, Arnold JD, Ma PJ, Starte DR. Predictive value of the Griffiths assessment in extremely low birthweight infants. *J Paediatr Child Health* 1996;**32**:25–30. https://doi.org/10.1111/j.1440-1754.1996.tb01536.x

284. Bruggink JL, Van Braeckel KN, Bos AF. The early motor repertoire of children born preterm is associated with intelligence at school age. *Pediatrics* 2010;**125**:e1356–63. https://doi.org/10.1542/peds.2009-2117

285. Charkaluk ML, Truffert P, Marchand-Martin L, Mur S, Kaminski M, Ancel PY, Pierrat V, Epipage study group. Very preterm children free of disability or delay at age 2: predictors of schooling at age 8 – a population-based longitudinal study. *Early Hum Dev* 2011;**87**:297–302. https://doi.org/10.1016/j.earlhumdev.2011.01.033

286. Cohen SE. Biosocial factors in early infancy as predictors of competence in adolescents who were born prematurely. *J Dev Behav Pediatr* 1995;**16**:36–41. https://doi.org/10.1097/00004703-199502000-00006

287. Gray D, Woodward LJ, Spencer C, Inder TE, Austin NC. Health service utilisation of a regional cohort of very preterm infants over the first 2 years of life. *J Paediatr Child Health* 2006;**42**:377–83. https://doi.org/10.1111/j.1440-1754.2006.00876.x

288. Reuss ML, Paneth N, Pinto-Martin JA, Lorenz JM, Susser M. The relation of transient hypothyroxinemia in preterm infants to neurologic development at two years of age. *N Engl J Med* 1996;**334**:821–7. https://doi.org/10.1056/NEJM199603283341303

289. Skranes J, Vik T, Nilsen G, Smevik O, Andersson HW, Brubakk AM. Can cerebral MRI at age 1 year predict motor and intellectual outcomes in very-low-birthweight children? *Dev Med Child Neurol* 1998;**40**:256–62. https://doi.org/10.1111/j.1469-8749.1998.tb15458.x

290. Tommiska V, Heinonen K, Kero P, Pokela ML, Tammela O, Järvenpää AL, *et al.* A national two year follow up study of extremely low birthweight infants born in 1996–1997. *Arch Dis Child Fetal Neonatal Ed* 2003;**88**:F29–35. https://doi.org/10.1136/fn.88.1.F29

291. Veelken N, Stollhoff K, Claussen M. Development of very low birth weight infants: a regional study of 371 survivors. *Eur J Pediatr* 1991;**150**:815–20. https://doi.org/10.1007/BF02026720

292. Claas MJ, de Vries LS, Bruinse HW, van Haastert IC, Uniken Venema MM, Peelen LM, Koopman C. Neurodevelopmental outcome over time of preterm born children ≤ 750 g at birth. *Early Hum Dev* 2011;**87**:183–91. https://doi.org/10.1016/j.earlhumdev.2010.12.002

293. Kilbride HW, Daily DK, Claflin K, Hall RT, Maulik D, Grundy HO. Improved survival and neurodevelopmental outcome for infants less than 801 grams birthweight. *Am J Perinatol* 1990;**7**:160–5. https://doi.org/10.1055/s-2007-999471

294. Orchinik LJ, Taylor HG, Espy KA, Minich N, Klein N, Sheffield T, Hack M. Cognitive outcomes for extremely preterm/extremely low birth weight children in kindergarten. *J Int Neuropsychol Soc* 2011;**17**:1067–79. https://doi.org/10.1017/S135561771100107X

295. Vermeulen GM, Bruinse HW, de Vries LS. Perinatal risk factors for adverse neurodevelopmental outcome after spontaneous preterm birth. *Eur J Obstet Gynecol Reprod Biol* 2001;**99**:207–12. https://doi.org/10.1016/S0301-2115(01)00383-9

296. Tin W, Fritz S, Wariyar U, Hey E. Outcome of very preterm birth: children reviewed with ease at 2 years differ from those followed up with difficulty. *Arch Dis Child Fetal Neonatal Ed* 1998;**79**:F83–7. https://doi.org/10.1136/fn.79.2.F83

297. Campbell MK, Halinda E, Carlyle MJ, Fox AM, Turner LA, Chance GW. Factors predictive of follow-up clinic attendance and developmental outcome in a regional cohort of very low birth weight infants. *Am J Epidemiol* 1993;**138**:704–13. https://doi.org/10.1093/oxfordjournals.aje.a116908

298. Catlett AT, Thompson RJ, Johndrow DA, Boshkoff MR. Risk status for dropping out of developmental followup for very low birth weight infants. *Public Health Rep* 1993;**108**:589–94.

299. Callanan C, Doyle L, Rickards A, Kelly E, Ford G, Davis N. Children followed with difficulty: how do they differ? *J Paediatr Child Health* 2001;**37**:152–6. https://doi.org/10.1046/j.1440-1754.2001.00621.x

300. Wolke D, Söhne B, Ohrt B, Riegel K. Follow-up of preterm children: important to document dropouts. *Lancet* 1995;**345**:447. https://doi.org/10.1016/S0140-6736(95)90425-5

301. Gordis L. Assessing the Validity and Reliability of Diagnostic and Screening Tests. In *Epidemiology*. *5th edn.* Philadelphia, PA: Elsevier; 2014. pp. 88–115.

302. Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol* 2009;**62**:5–12. https://doi.org/10.1016/j.jclinepi.2008.04.007

303. Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. *Stat Med* 1997;**16**:981–91. https://doi.org/10.1002/(SICI)1097-0258(19970515)16:9<981::AID-SIM510>3.0.CO;2-N

304. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974;**185**:1124–31. https://doi.org/10.1126/science.185.4157.1124

305. Chaudhary T, Walch E, Herold B, Metze B, Lejeune A, Burkhardt F, Bührer C. Predictive and concurrent validity of standardized neurodevelopmental examinations by the Griffiths scales and Bayley scales of infant development II. *Klin Padiatr* 2013;**225**:8–12. https://doi.org/10.1055/s-0032-1331169

306. Brogan E, Cragg L, Gilmore C, Marlow N, Simms V, Johnson S. Inattention in very preterm children: implications for screening and detection. *Arch Dis Child* 2014;**99**:834–9. https://doi.org/10.1136/archdischild-2013-305532

307. Wilson-Ching M, Molloy CS, Anderson VA, Burnett A, Roberts G, Cheong JL, *et al.* Attention difficulties in a contemporary geographic cohort of adolescents born extremely preterm/extremely low birth weight. *J Int Neuropsychol Soc* 2013;**19**:1097–108. https://doi.org/10.1017/S1355617713001057

308. Boyd LA, Msall ME, O'Shea TM, Allred EN, Hounshell G, Leviton A. Social-emotional delays at 2 years in extremely low gestational age survivors: correlates of impaired orientation/engagement and emotional regulation. *Early Hum Dev* 2013;**89**:925–30. https://doi.org/10.1016/j.earlhumdev.2013.09.019

309. Wechsler D. *Wechlser Preschool and Primary Scale of Intelligence for Children: Third Edition*. San Antonio, TX: The Psychological Corporation; 2002. https://doi.org/10.1037/t15177-000

310. Zimmerman IL, Steiner VG, Pond RE. *Preschool Language Scale: Fourth Edition*. San Antonio, TX: The Psychological Corporation; 2002. https://doi.org/10.1037/t15140-000

311. Páez MM, Tabors PO, López LM. Dual language and literacy development of Spanish-speaking preschool children. *J Appl Dev Psychol* 2007;**28**:85–102. https://doi.org/10.1016/j.appdev.2006.12.007

312. Bialystok E, Martin MM. Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev Sci* 2004;**7**:325–39. https://doi.org/10.1111/j.1467-7687.2004.00351.x

313. Walch E, Chaudhary T, Herold B, Obladen M. Parental bilingualism is associated with slower cognitive development in very low birth weight infants. *Early Hum Dev* 2009;**85**:449–54. https://doi.org/10.1016/j.earlhumdev.2009.03.002

314. Als H. A synactive model of neonatal behavioral organization: framework for the assessment and support of the neurobehavioral development of the premature infant and his parents in the environment of the neoantal intensive care unit. *Phys Occup Ther Pediatr* 1986;**6**:3–55. https://doi.org/10.1080/J006v06n03_02

315. Bar-Shalita T, Vatine JJ, Parush S. Sensory modulation disorder: a risk factor for participation in daily life activities. *Dev Med Child Neurol* 2008;**50**:932–7. https://doi.org/10.1111/j.1469-8749.2008.03095.x

316. Bart O, Shayevits S, Gabis LV, Morag I. Prediction of participation and sensory modulation of late preterm infants at 12 months: a prospective study. *Res Dev Disabil* 2011;**32**:2732–8. https://doi.org/10.1016/j.ridd.2011.05.037

317. Bishop SL, Richler J, Lord C. Association between restricted and repetitive behaviors and nonverbal IQ in children with autism spectrum disorders. *Child Neuropsychol* 2006;**12**:247–67. https://doi.org/10.1080/09297040600630288

318. Ozonoff S, Macari S, Young GS, Goldring S, Thompson M, Rogers SJ. Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism* 2008;**12**:457–72. https://doi.org/10.1177/1362361308096402

319. Johnson S, Marlow N. Positive screening results on the modified checklist for autism in toddlers: implications for very preterm populations. *J Pediatr* 2009;**154**:478–80. https://doi.org/10.1016/j.jpeds.2008.11.028

320. Oosterling IJ, Swinkels SH, van der Gaag RJ, Visser JC, Dietz C, Buitelaar JK. Comparative analysis of three screening instruments for autism spectrum disorder in toddlers at high risk. *J Autism Dev Disord* 2009;**39**:897–909. https://doi.org/10.1007/s10803-009-0692-9

321. Dietz C, Swinkels S, van Daalen E, van Engeland H, Buitelaar JK. Screening for autistic spectrum disorder in children aged 14–15 months: II – population screening with the Early Screening of Autistic Traits Questionnaire (ESAT): Design and general findings. *J Autism Dev Disord* 2006;**36**:713–22. https://doi.org/10.1007/s10803-006-0114-1

322. Swinkels SH, Dietz C, van Daalen E, Kerkhof IH, van Engeland H, Buitelaar JK. Screening for autistic spectrum in children aged 14 to 15 months. I: the development of the Early Screening of Autistic Traits Questionnaire (ESAT). *J Autism Dev Disord* 2006;**36**:723–32. https://doi.org/10.1007/s10803-006-0115-0

323. Rutter M, Bailey A, Lord C. *Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services; 2003. URL: www.carautismroadmap.org/social-communication-questionnaire-scq/ (accessed 29 August 2018).

324. Baron-Cohen S, Wheelwright S, Cox A, Baird G, Charman T, Swettenham J, *et al.* Early identification of autism by the CHecklist for Autism in Toddlers (CHAT). *J R Soc Med* 2000;**93**:521–5. https://doi.org/10.1177/014107680009301007

325. Aylward GP. Prediction of function from infancy to early childhood: implications for pediatric psychology. *J Pediatr Psychol* 2004;**29**:555–64. https://doi.org/10.1093/jpepsy/jsh057

326. Gutbrod T, Wolke D, Soehne B, Ohrt B, Riegel K. Effects of gestation and birth weight on the growth and development of very low birthweight small for gestational age infants: a matched group comparison. *Arch Dis Child Fetal Neonatal Ed* 2000;**82**:F208–14. https://doi.org/10.1136/fn.82.3.F208

327. Wilson-Costello D, Friedman H, Minich N, Siner B, Taylor G, Schluchter M, Hack M. Improved neurodevelopmental outcomes for extremely low birth weight infants in 2000–2002. *Pediatrics* 2007;**119**:37–45. https://doi.org/10.1542/peds.2006-1416

328. Doyle LW, Roberts G, Anderson PJ, Victorian Infant Collaborative Study Group. Changing long-term outcomes for infants 500–999 g birth weight in Victoria, 1979–2005. *Arch Dis Child Fetal Neonatal Ed* 2011;**96**:F443–7. https://doi.org/10.1136/adc.2010.200576

329. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, *et al.* The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care* 2012;**28**:138–44. https://doi.org/10.1017/S0266462312000086

330. Bossuyt P, Davenport C, Deeks J, Hyde C, Leeflang M, Scholten R. Chapter 11: Interpreting Results and Drawing Conclusions. In Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 09*: London: The Cochrane Collaboration; 2013.

331. National Institute for Health and Care Excellence (NICE). *Developmental Follow-up of Children and Young People Born Preterm [NG72]*. London: NICE; 2017.

332. Amess P, Young T, Burley H, Khan Y. Developmental outcome of very preterm babies using an assessment tool deliverable by health visitors. *Eur J Paediatr Neurol* 2010;**14**:219–23. https://doi.org/10.1016/j.ejpn.2009.06.005

333. American Academy of Pediatrics. Identifying infants and young children with developmental disorders in the medical home: an algorithm for developmental surveillance and screening. *Pediatrics* 2006;**118**:405–20. https://doi.org/10.1542/peds.2006-1231

334. UK National Screening Commitee. *The UK NSC Policy on Autism Screening in Children*. 2009. URL: www.screening.nhs.uk/autism (accessed 8 November 2018).

335. Myers SM, Johnson CP, American Academy of Pediatrics Council on Children With Disabilities. Management of children with autism spectrum disorders. *Pediatrics* 2007;**120**:1162–82. https://doi.org/10.1542/peds.2007-2362

336. UK Government. *NHS Reference Costs 2010–11*. URL: https://data.gov.uk/dataset/7fa41b07-a296-47c7-b74b-40ec3a102e6a/nhs-reference-costs-2010-11 (accessed 11 July 2019).

337. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart G. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd edn. New York, NY: Oxford University Press; 2005.

338. Simon J, Gray A, Duley L, Magpie Trial Collaborative Group. Cost-effectiveness of prophylactic magnesium sulphate for 9996 women with pre-eclampsia from 33 countries: economic evaluation of the Magpie Trial. *BJOG* 2006;**113**:144–51. https://doi.org/10.1111/j.1471-0528.2005.00785.x

339. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. 2013. URL: www.nice.org.uk/process/pmg9/chapter/foreword (accessed 9 July 2019).

340. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ* 2011;**342**:d1548. https://doi.org/10.1136/bmj.d1548

341. Petrou S, Edwards L, UK Collaborative ECMO Trial. Cost effectiveness analysis of neonatal extracorporeal membrane oxygenation based on four year results from the UK Collaborative ECMO Trial. *Arch Dis Child Fetal Neonatal Ed* 2004;**89**:F263–8. https://doi.org/10.1136/adc.2002.025635

342. Costeloe K, Hardy P, Juszczak E, Wilks M, Millar MR, Probiotics in Preterm Infants Study Collaborative Group. Bifidobacterium breve BBG-001 in very preterm infants: a randomised controlled phase 3 trial. *Lancet* 2016;**387**:649–60. https://doi.org/10.1016/S0140-6736(15)01027-2

343. Department of Health and Social Care. *NHS Reference Costs 2012–13*. 2013. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/261154/nhs_reference_costs_2012-13_acc.pdf (accessed 9 July 2019).

344. Curtis L. *Unit Costs of Health and Social Care 2012*. Personal Social Services Research Unit. URL: www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2012/ (accessed 9 July 2019).

345. Schroeder E, Petrou S, Patel N, Hollowell J, Puddicombe D, Redshaw M, Brocklehurst P, Birthplace in England Collaborative Group. Cost effectiveness of alternative planned places of birth in woman at low risk of complications: evidence from the Birthplace in England national prospective cohort study. *BMJ* 2012;**344**:e2292. https://doi.org/10.1136/bmj.e2292

346. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;**45**:255–68. https://doi.org/10.2307/2532051

347. Byford S, Leese M, Knapp M, Seivewright H, Cameron S, Jones V, *et al.* Comparison of alternative methods of collection of service use data for the economic evaluation of health care interventions. *Health Econ* 2007;**16**:531–6. https://doi.org/10.1002/hec.1175

348. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10. https://doi.org/10.1016/S0140-6736(86)90837-8

349. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;**10**:212–14. https://doi.org/10.1177/0272989X9001000308

350. Achana FA, Petrou S, Khan K, Gaye A, Modi N. A methodological framework for assessing agreement between cost-effectiveness outcomes estimated using alternative sources of data on treatment costs and effects for trial-based economic evaluations. *Eur J Health Econ* 2018;**19**:75–86. https://doi.org/10.1007/s10198-017-0868-8

351. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal 2013*. London: NICE; 2013. URL: www.nice.org.uk/process/pmg9/chapter/the-reference-case (accessed 9 July 2019).

352. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999;**18**:341–64. https://doi.org/10.1016/S0167-6296(98)00039-3

353. Claxton K, Ginnelly L, Sculpher M, Philips Z, Palmer S. A pilot study on the use of decision theory and value of information analysis as part of the NHS Health Technology Assessment programme. *Health Technol Assess* 2004;**8**:(31). https://doi.org/10.3310/hta8310

354. Murray J, Bottle A, Sharland M, Modi N, Aylin P, Majeed A, *et al.* Risk factors for hospital admission with RSV bronchiolitis in England: a population-based birth cohort study. *PLOS ONE* 2014;**9**:e89186. https://doi.org/10.1371/journal.pone.0089186

355. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, Smeeth L. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;**44**:827–39.

356. Pan H, Cole T. *lmsGrowth, A Microsoft Excel Add-in to Access Growth References Based on the LMS Method*. *(Version 2.68)*. 2009. URL: www.healthforallchildren.com/ (accessed 9 July 2019).

357. (AHRQ) AfHRaQ. *Clinical Classifications Software for ICD-10 Data*. 2003. URL: www.ahrqgov/data/hcup/icd10usrgdhtm (accessed June 2012).

358. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLOS Med* 2015;**12**:e1001885. https://doi.org/10.1371/journal.pmed.1001885

359. Murray JC. *The Clinical Burden of Respiratory Syncytial Virus (RSV) Bronchiolitis Among Infants in the United Kingdom (UK)*. PhD Thesis. London: Imperial College London; 2013.

360. Murray J, Saxena S, Modi N, Majeed A, Aylin P, Bottle A, Medicines for Neonates Investigator Group. Quality of routine hospital birth records and the feasibility of their use for creating birth cohorts. *J Public Health* 2013;**35**:298–307. https://doi.org/10.1093/pubmed/fds077

361. Dattani N, Datta-Nemdharry P, Macfarlane A. Linking maternity data for England, 2005–06: methods and data quality. *Health Stat Q* 2011;**49**:53–79. https://doi.org/10.1057/hsq.2011.3

362. Hockley C, Quigley MA, Hughes G, Calderwood L, Joshi H, Davidson LL. Linking Millennium Cohort data to birth registration and hospital episode records. *Paediatr Perinat Epidemiol* 2008;**22**:99–109. https://doi.org/10.1111/j.1365-3016.2007.00902.x

363. Field K, Kosmider S, Johns J, Farrugia H, Hastie I, Croxford M, *et al.* Linking data from hospital and cancer registry databases: should this be standard practice? *Intern Med J* 2010;**40**:566–73. https://doi.org/10.1111/j.1445-5994.2009.01984.x

364. Chamberlayne R, Green B, Barer ML, Hertzman C, Lawrence WJ, Sheps SB. Creating a population-based linked health database: a new resource for health services research. *Can J Public Health* 1998;**89**:270–3.

365. Roos LL, Brownell M, Lix L, Roos NP, Walld R, MacWilliam L. From health research to social research: privacy, methods, approaches. *Soc Sci Med* 2008;**66**:117–29. https://doi.org/10.1016/j.socscimed.2007.08.017

366. Taylor LK, Travis S, Pym M, Olive E, Henderson-Smart DJ. How useful are hospital morbidity data for monitoring conditions occurring in the perinatal period? *Aust N Z J Obstet Gynaecol* 2005;**45**:36–41. https://doi.org/10.1111/j.1479-828X.2005.00339.x

367. Ford JB, Roberts CL, Taylor LK. Characteristics of unmatched maternal and baby records in linked birth records and hospital discharge data. *Paediatr Perinat Epidemiol* 2006;**20**:329–37. https://doi.org/10.1111/j.1365-3016.2006.00715.x

368. Bradford HM, Cárdenas V, Camacho-Carr K, Lydon-Rochelle MT. Accuracy of birth certificate and hospital discharge data: a certified nurse-midwife and physician comparison. *Matern Child Health J* 2007;**11**:540–8. https://doi.org/10.1007/s10995-007-0178-3

369. Lydon-Rochelle MT, Holt VL, Nelson JC, Cárdenas V, Gardella C, Easterling TR, Callaghan WM. Accuracy of reporting maternal in-hospital diagnoses and intrapartum procedures in Washington State linked birth records. *Paediatr Perinat Epidemiol* 2005;**19**:460–71. https://doi.org/10.1111/j.1365-3016.2005.00682.x

370. Lauritsen MB, Pedersen CB, Mortensen PB. Effects of familial risk factors and place of birth on the risk of autism: a nationwide register-based study. *J Child Psychol Psychiatry* 2005;**46**:963–71. https://doi.org/10.1111/j.1469-7610.2004.00391.x

371. Linnet KM, Wisborg K, Secher NJ, Thomsen PH, Obel C, Dalsgaard S, Henriksen TB. Coffee consumption during pregnancy and the risk of hyperkinetic disorder and ADHD: a prospective cohort study. *Acta Paediatr* 2009;**98**:173–9. https://doi.org/10.1111/j.1651-2227.2008.00980.x

372. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke JP, Ford DV, *et al.* The SAIL databank: linking multiple health and social care datasets. *BMC Med Inform Decis Mak* 2009;**9**:3. https://doi.org/10.1186/1472-6947-9-3

373. Welsh Electronic Cohort for Children (WECC). URL: www.swan.ac.uk/ils/research/chiral/methodologies/healthinformatics/wecc/ (accessed October 2012).

374. McGregor J, Brooks C, Chalasani P, Chukwuma J, Hutchings H, Lyons RA, Lloyd K. The Health Informatics Trial Enhancement Project (HITE): using routinely collected primary care data to identify potential participants for a depression trial. *Trials* 2010;**11**:39. https://doi.org/10.1186/1745-6215-11-39

375. Programme SHI. URL: www.scot-ship.ac.uk (accessed October 2011).

376. Govan L, Wu O, Briggs A, Colhoun HM, McKnight JA, Morris AD, *et al.* Inpatient costs for people with type 1 and type 2 diabetes in Scotland: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabetologia* 2011;**54**:2000–8. https://doi.org/10.1007/s00125-011-2176-7

377. Gill L, Goldacre M, Simmons H, Bettley G, Griffith M. Computerised linking of medical records: methodological guidelines. *J Epidemiol Community Health* 1993;**47**:316–19. https://doi.org/10.1136/jech.47.4.316

378. Roberts SE, Wotton CJ, Williams JG, Griffith M, Goldacre MJ. Perinatal and early life risk factors for inflammatory bowel disease. *World J Gastroenterol* 2011;**17**:743–9. https://doi.org/10.3748/wjg.v17.i6.743

379. Davidson R, Roberts SE, Wotton CJ, Goldacre MJ. Influence of maternal and perinatal factors on subsequent hospitalisation for asthma in children: evidence from the Oxford record linkage study. *BMC Pulm Med* 2010;**10**:14. https://doi.org/10.1186/1471-2466-10-14

380. Roberts SE, Williams JG, Meddings D, Davidson R, Goldacre MJ. Perinatal risk factors and coeliac disease in children and young adults: a record linkage study. *Aliment Pharmacol Ther* 2009;**29**:222–31. https://doi.org/10.1111/j.1365-2036.2008.03871.x

381. NHS Digital. *Hospital Episode Statistics*. Leeds: NHS Digital. URL: www.hesonline.nhs.uk (accessed 8 November 2018).

382. Abrahams C, Davy K. Linking HES maternity records with ONS birth records. *Health Stat Q* 2002;**13**:22–30.

383. Dezateux CHC, Johnson J, Joshi H, Quigley M, Rosenburg R. Millenium Cohort study: Birth registration and maternity Hospital Episode Statistics Linkage. *Centre for Longitudinal Studies Report*; 2006.

384. UK Clinical Research Collaboration (UKCRC). *UK Clinical Research Collaboration Research and Development Advisory Group to Connecting for Health: Report of Research Simulations*. 2007. URL: www.ukcrc.org/wp-content/uploads/2014/06/CfH-report-June-07-full.pdf (accessed 11 July 2019).

385. Foster V, Young A, Modi N, Brocklehurst P, Abbott J, Costeloe K, *et al.* The use of routinely collected patient data for research: a critical review. *Health* 2012;**16**:448–63. https://doi.org/10.1177/1363459311425513

386. British Association of Perinatal Medicine. *Service Standards for Hospitals Providing Neonatal Care (3rd edition)*. 2010. URL: www.bapm.org/resources/service-standards-hospitals-providing-neonatal-care-3rd-edition-2010 (accessed 9 July 2019).

387. Adams WG, Mann AM, Bauchner H. Use of an electronic medical record improves the quality of urban pediatric primary care. *Pediatrics* 2003;**111**:626–32. https://doi.org/10.1542/peds.111.3.626

388. Royal College of Obstetricians and Gynaecologists. *Antenatal Corticosteroids to Reduce Neonatal Morbidity and Mortality: Green–top Guideline No. 7*. London: Royal College of Obstetricians and Gynaecologists; 2010. URL: www.glowm.com/pdf/Antenatal%20Corticosteroids%20to%20Reduce%20Neonatal%20Morbidity.pdf (accessed 9 July 2019).

389. National Institute for Health and Care Excellence (NICE). *Antenatal Care for Uncomplicated Pregnancies Clinical Guideline [CG62]*. London: NICE; 2008. URL: www.nice.org.uk/guidance/cg62/chapter/guidance#fetal-growth-and-wellbeing (accessed 9 July 2019).

390. National Institute for Health and Care Excellence (NICE). *BNF for Children*. London: NICE; 2019. URL: https://bnfc.nice.org.uk (accessed 19 July 2019).

Appendix 1 Supplementary tables

Chapter 2

TABLE 53 Characteristics of infants born before 32 weeks' gestation with and without severe NEC

Characteristic	Number of infants born before 32 weeks' gestation, *n* (%) (*N* = 14,678)		*p*-value
	No severe NEC (*n* = 14,216)	Severe NEC (*n* = 462)	
Gestational age (weeks) (mean ± SD)	28.5 (2.27)	26.2 (2.16)	< 0.001
Birthweight (g) (mean ± SD)	1217.7 (378.9)	884.2 (310.6)	< 0.001
Birthweight SDS, *n* (%)			< 0.001
Missing	39 (0.3)	1 (0.2)	
–4 SD	6 (0.04)	0 (0)	
–3 SD	153 (1.1)	16 (3.5)	
–2 SD	931 (6.6)	46 (10.0)	
–1 SD	2998 (21.1)	108 (23.4)	
1 SD	3404 (23.9)	81 (17.5)	
2 SD	497 (3.5)	11 (2.4)	
3 SD	63 (0.4)	1 (0.2)	
4 SD	5 (0.04)	0 (0)	
Average	6120 (43.1)	198 (42.9)	
Sex, *n* (%)			0.7074
Missing	8 (0.06)	0	
Male	7807 (54.9)	261 (56.5)	
Fetus number			0.0006
Missing	1 (0.01)	1 (0.2)	
1	10,505 (73.9)	344 (74.5)	
2	3345 (23.5)	111 (24.0)	
≥ 3	365 (2.6)	6 (1.3)	
Antenatal steroids			0.8068
Missing	155 (1.1)	5 (1.1)	
Yes	12,595 (88.6)	405 (87.7)	
Maternal factors			
Chorioamnionitis	552 (3.9)	21 (4.6)	0.4632
Maternal infection[a]	725 (5.1)	21 (4.6)	0.6003
Received antibiotics in labour	3460 (23.3)	112 (24.3)	< 0.001
Pyrexia above 38 °C in labour	674 (4.7)	29 (6.3)	0.2615

continued

TABLE 53 Characteristics of infants born before 32 weeks' gestation with and without severe NEC (*continued*)

Characteristic	Number of infants born before 32 weeks' gestation, *n* (%) (*N* = 14,678)		*p*-value
	No severe NEC (*n* = 14,216)	Severe NEC (*n* = 462)	
Mode of delivery			< 0.001
Unknown	1068 (7.5)	40 (8.7)	
Emergency caesarean (not in labour)	4326 (30.4)	106 (22.9)	
Emergency caesarean (in labour)	2445 (17.2)	64 (13.9)	
Elective section (not in labour)	794 (5.6)	19 (4.1)	
Elective section (in labour)	97 (0.7)	3 (0.7)	
Vaginal	5486 (38.6)	230 (49.8)	

a UTI, other infection, group B *Streptococcus*.

TABLE 54 Parameters for the final multivariable logistic regression model showing unadjusted and adjusted odds of severe NEC for infants born before 32 weeks' gestation

Variable	Unadjusted OR for severe NEC relative to reference category	95% CI	*p*-value	Adjusted OR for severe NEC relative to reference category	95% CI	*p*-value
Gestation in weeks(+days)						
22 to 25^{+6}	13.4	10.2 to 17.9	< 0.001	13.8[a]	10.5 to 18.5	< 0.001
26^{+0} to 28^{+6}	5.8	4.4 to 7.7		5.6	4.3 to 7.5	< 0.001
29^{+0}–31^{+6}	*Reference*			*Reference*		
Birthweight SDS						
Missing	0.8	0.05 to 3.7	< 0.001	0.5[b]	0.03 to 2.7	0.6
–4 SD	0.0	–		0.0	–	–
–3 SD	3.2	1.8 to 5.4		4.1	2.3 to 7.0	< 0.001
–2 SD	1.5	1.1 to 2.1		1.9	1.3 to 2.6	< 0.001
–1 SD	1.1	0.9 to 1.4		1.1	0.9 to 1.4	0.3
1 SD	0.7	0.6 to 0.9		0.9	0.7 to 1.2	0.5
2 SD	0.7	0.6 to 1.0		1.0	0.5 to 1.8	1.0
3 SD	0.5	0.3 to 1.2		0.9	0.05 to 4.3	0.9
4 SD	0.0	–		–		–
Average	*Reference*			*Reference*		
Any steroids given						
Missing	0.9	0.7 to 1.3	0.8	1.1[c]	0.4 to 2.6	0.9
Yes	0.9	0.3 to 2.1		1.0	0.7 to 1.3	0.8
No	*Reference*			*Reference*		

a Adjusted for birthweight SDS and antenatal steroids.
b Adjusted for gestational age (completed weeks) and antenatal steroids.
c Adjusted for birthweight SDS and gestational age (completed weeks).
No significant interaction found between gestation and birthweight SDS.

TABLE 55 Multivariable logistic regression model showing unadjusted and adjusted odds of severe NEC in each booking network relative to the reference network for infants born before 32 weeks' gestation

Network of booking	Unadjusted OR of severe NEC relative to reference network	95% CI of OR	Bonferroni adjusted *p*-value	Adjusted OR of severe NEC relative to reference network[a]	95% CI of OR	Bonferroni adjusted *p*-value
Bedfordshire and Hertfordshire	0.69	0.29 to 1.4	0.9	0.71	0.30 to 1.52	0.9
Cheshire and Merseyside	1.29	0.74 to 2.22	0.9	1.54	0.87 to 2.68	0.9
Eastern	1.59	0.98 to 2.6	0.9	1.93	1.17 to 3.21	0.2
Greater Manchester	1.37	0.84 to 2.27	0.9	1.47	0.89 to 2.46	0.9
Kent	1.08	0.56 to 2.00	0.9	1.30	0.67 to 2.44	0.9
Lancashire and South Cumbria	0.91	0.43 to 1.77	0.9	1.07	0.50 to 2.12	0.9
London North Central	0.72	0.32 to 1.46	0.9	0.71	0.31 to 1.47	0.9
London North West	1.40	0.84 to 2.35	0.9	1.65	0.98 to 2.80	0.9
London South East	1.31	0.75 to 2.27	0.9	1.33	0.75 to 2.33	0.9
London South West	1.05	0.53 to 2.00	0.9	1.20	0.59 to 2.31	0.9
Midlands Central	1.63	1.00 to 2.69	0.9	1.90	1.15 to 3.16	0.3
Midlands South West	0.72	0.38 to 1.29	0.9	0.76	0.40 to 1.39	0.9
Midlands North	0.94	0.51 to 1.68	0.9	1.05	0.57 to 1.91	0.9
North Trent	1.09	0.60 to 1.92	0.9	1.41	0.77 to 2.52	0.9
Northern	0.83	0.45 to 1.47	0.9	0.85	0.46 to 1.53	0.9
Peninsula South West	0.68	0.27 to 1.48	0.9	0.80	0.32 to 1.76	0.9
South Central (North)	1.03	0.56 to 1.85	0.9	1.15	0.62 to 2.08	0.9
South Central (South)	0.74	0.38 to 1.37	0.9	0.81	0.41 to 1.53	0.9
Surrey and Sussex	1.51	0.88 to 2.58	0.9	1.63	0.93 to 2.82	0.9
Trent	0.97	0.51 to 1.80	0.9	1.09	0.56 to 2.03	0.9
Western	0.66	0.33 to 1.25	0.9	0.76	0.38 to 1.45	0.9
Yorkshire	0.79	0.45 to 1.38	0.9	0.98	0.56 to 1.73	0.9
London North East (Reference network)	Reference network	–		Reference network		–

a Adjusted for gestational age, birthweight SDS and antenatal steroids.

Chapter 3

TABLE 56 Coefficients from logistic regression model to predict death before discharge

Variable	Coefficient (SE)	*p*-value
Intercept	–4.058 (0.297)	
Male	0.315 (0.036)	< 0.001
Multiple pregnancy	–0.122 (0.062)	0.048
Antenatal steroids given	–0.726 (0.047)	< 0.001
Gestational age spline terms (/week)		
GA1	–1.193 (0.117)	< 0.001
GA2	0.95 (0.212)	< 0.001
GA3	–2.961 (1.071)	0.006
GA4	7.434 (3.329)	0.026
Birthweight (BWT) (/100 g)		
BWT	–0.09 (0.052)	0.081
BWT^2	0.045 (0.002)	< 0.001
Interactions		
GA1*BWT	–0.17 (0.026)	< 0.001
GA2*BWT	0.012 (0.039)	0.753
GA3*BWT	0.309 (0.22)	0.159
GA4*BWT	–0.657 (0.73)	0.368
$GA1*BWT^2$	–0.008 (0.001)	< 0.001
GA*multiple pregnancy	0.007 (0.028)	0.788
BWT*multiple pregnancy	–0.083 (0.022)	< 0.001

2, squared; *, multiplied by.

TABLE 57 Survival by population characteristics; percentages exclude missing values; *p*-value from chi-squared tests

	Survived to discharge, *n* (%)	Missing, *n*	*p*-value	Survived to 28 days, *n* (%)	Missing, *n*	*p*-value
Gestational age (weeks^{+days})						
22^{+0} to 23^{+6}	452 (35)	57		646 (48.4)	13	
24^{+0} to 25^{+6}	3555 (66.7)	311		4240 (76)	59	
26^{+0} to 27^{+6}	7324 (86.2)	392	$p < 0.001$	7943 (90.2)	87	$p < 0.001$
28^{+0} to 29^{+6}	12,155 (94.2)	443		12,609 (95.6)	163	
30^{+0} to 31^{+6}	19,958 (97.8)	487		20,224 (98.2)	281	
Birthweight (g)						
< 500	127 (34.8)	17		192 (50.7)	3	
500 to 999	11,748 (76.8)	772		13,256 (83.4)	167	
1000 to 1499	19,918 (95.6)	613	$p < 0.001$	20,431 (96.4)	259	$p < 0.001$
1500 to 1999	10,913 (97.9)	262		11,031 (98.1)	158	
≥ 2000	738 (94.4)	26		752 (94.9)	16	
SGA						
No	37,309 (90.4)	1406	$p < 0.001$	38,985 (92.5)	538	$p < 0.001$
Yes	6135 (85.9)	284		6677 (90.7)	65	
Sex						
Female	20,190 (90.6)	732	$p < 0.001$	21,090 (92.8)	284	$p < 0.001$
Male	23,254 (88.9)	958		24,572 (91.7)	319	
Multiplicity of pregnancy						
Singleton	31,845 (89.7)	1225	$p < 0.001$	33,506 (92.3)	417	$p < 0.001$
Twins	10,472 (89.3)	433		10,992 (91.7)	172	
Triplets or more	1127 (93.1)	32		1164 (94.8)	14	
Any antenatal steroids given						
No	4421 (82.1)	233	$p < 0.001$	4711 (85)	72	$p < 0.001$
Yes	38,327 (90.8)	1369		40,196 (93.2)	485	
Mode of delivery						
Vaginal	16,346 (85.9)	546	$p < 0.001$	17,275 (89.1)	190	$p < 0.001$
Caesarean	23,473 (93)	665		24,367 (94.9)	227	
Maternal age (years)						
< 20	3143 (88.3)	147	$p < 0.001$	3326 (90.9)	51	$p < 0.001$
20 to 24	7639 (88.5)	308		8063 (91.3)	108	
25 to 29	11,268 (90.3)	395		11,821 (92.7)	122	
30 to 34	11,890 (90)	421		12,460 (92.5)	157	
35 to 40	7171 (90.4)	246		7505 (92.8)	89	
> 40	2105 (90.7)	67		2198 (93)	23	

continued

TABLE 57 Survival by population characteristics; percentages exclude missing values; *p*-value from chi-squared tests (*continued*)

	Survived to discharge, *n* (%)	Missing, *n*	*p*-value	Survived to 28 days, *n* (%)	Missing, *n*	*p*-value
Maternal ethnicity						
British, Irish, other white	29,511 (90.2)	1004		30,810 (92.4)	364	
Mixed	670 (92.2)	21		696 (94.2)	9	
Indian, Pakistani, Bangladeshi, other Asian	5036 (89.5)	131	$p < 0.001$	5255 (92.1)	48	$p < 0.001$
Black Caribbean, Black African, other black	4205 (88.7)	106		4479 (92.8)	21	
Chinese	161 (90.4)	7		173 (95.1)	3	
Other	672 (87.7)	20		709 (90.8)	5	
Smoking						
No	28,210 (90.6)	764	$p = 0.001$	29,446 (93)	233	$p < 0.001$
Yes	7560 (89.4)	224		7912 (91.9)	71	
IMD quintile						
1 (most deprived)	12,957 (89.2)	511		13,664 (91.8)	159	
2	9638 (89.2)	342		10,146 (91.9)	111	
3	7427 (90.5)	248	$p < 0.001$	7785 (93)	82	$p < 0.001$
4	6033 (90.7)	217		6299 (92.8)	80	
5 (least deprived)	5288 (90.7)	184		5511 (92.7)	72	
Birth year						
2008	4993 (88)	426		5364 (91.4)	233	
2009	5509 (88.9)	289		5823 (91.8)	143	
2010	6384 (89.5)	255		6704 (92.1)	108	
2011	6776 (89.5)	159	$p < 0.001$	7059 (91.7)	39	$p < 0.001$
2012	6728 (89.7)	170		7032 (92.2)	36	
2013	6547 (90.7)	151		6812 (92.7)	17	
2014	6507 (91.3)	240		6868 (93.5)	27	

Chapter 5

TABLE 58 Sensitivities and specificities of the NHS data using a broader 'moderate–severe' impairment category in identifying participants with Bayley-III scores of lower than –2 SDs below the mean

Domain of development	Sensitivity (95% CI)	Specificity (95% CI)
Cognitive	71.4 (14.4 to 100.0)	91.7 (85.6 to 97.8)
Receptive communication	36.4 (0.0 to 79.8)	94.8 (93.1 to 96.5)
Expressive communication	91.7 (72.9 to 100.0)	85.2 (78.4 to 92.1)
Combined language	66.7 (39.7 to 93.6)	85.8 (78.2 to 93.4)
Fine motor	50.0 (30.2 to 100.0)	99.5 (98.4 to 100.0)
Gross motor	88.9 (63.8 to 100.0)	98.3 (95.6 to 100.0)
Combined motor	80.0 (51.3 to 100.0)	98.3 (95.6 to 100.0)
Overall	72.0 (47.9 to 96.1)	86.3 (77.8 to 94.8)

TABLE 59 Characteristics of respondents, non-respondents and non-participants born before 30 weeks' gestation in 2008–10 and discharged from the participating study sites

Characteristics	Respondents ($N = 141$)	Non-respondents ($N = 60$)	'Baseline' population ($N = 1037$)	*p*-value	
				Respondents vs. non-respondents	Respondents vs. 'baseline' population
Gestation (completed weeks), median (IQR), range	27 (26–29), 23–29	27 (26–28), 23–29	27 (26–29), 22–29	0.15	0.58
Birthweight (g), median (IQR), range	958 (810–1167), 490–1720	920 (740–1082), 560–1400	1000 (812–1200), 455–1990	0.07	0.44
Sex, *n* (%)				0.02	0.09
Female	68 (48.2)	40 (66.7)	444 (42.8)		
Male	73 (51.8)	20 (33.3)	503 (48.5)		
Missing	0 (0.0)	0 (0.0)	90 (8.7)		
Ethnicity, *n* (%)				0.12	0.03
White	66 (46.8)	21 (35.0)	364 (35.1)		
Non-white	75 (52.2)	39 (65.0)	611 (58.9)		
Missing	0 (0.0)	0 (0.0)	62 (6.0)		
Pregnancy, *n* (%)				0.67	0.25
Singleton	110 (78.0)	48 (80.0)	690 (66.5)		
Multiples	31 (22.0)	12 (20.0)	250 (24.1)		
Missing	0 (0.0)	0 (0.0)	97 (9.4)		

continued

TABLE 59 Characteristics of respondents, non-respondents and non-participants born before 30 weeks' gestation in 2008–10 and discharged from the participating study sites (*continued*)

Characteristics	Respondents (*N* = 141)	Non-respondents (*N* = 60)	'Baseline' population (*N* = 1037)	*p*-value	
				Respondents vs. non-respondents	Respondents vs. 'baseline' population
Mode of delivery, *n* (%)				0.34	0.84
Vaginal	61 (43.3)	22 (36.7)	475 (45.8)		
Caesarean	71 (50.4)	32 (53.3)	540 (52.1)		
Missing	9 (6.4)	6 (10.0)	22 (2.1)		
Maternal age (years), mean (SD)	31.5 (6.0)	31.9 (7.8)	31.0 (6.4)	0.68	0.35
IMD quintile at birth, *n* (%)				0.77	0.05
One	13 (9.2)	6 (10.0)	43 (4.2)		
Two	13 (9.2)	5 (8.3)	81 (7.8)		
Three	20 (14.2)	5 (8.3)	144 (13.9)		
Four	42 (29.8)	17 (28.3)	268 (25.8)		
Five	53 (37.6)	27 (45.0)	477 (46.0)		
Missing	0 (0.0)	0 (0.0)	24 (2.3)		
Length of mechanical ventilation (days), median (IQR), range	1 (0–3), 0–54	1 (0–7), 0–61	4 (0–18), 0–444	0.13	< 0.001
Oxygen therapy at 36 weeks' corrected age, *n* (%)				0.11	< 0.001
Yes	38 (27.0)	23 (38.3)	466 (44.9)		
No	103 (73.1)	37 (61.7)	571 (55.1)		

TABLE 60 Final multivariable model of factors associated with Q-CHAT scores

Variable	Q-CHAT score coefficient (95% CI)	*p*-value
Bayley-III language composite score (per point)	–0.23 (–0.33 to –1.39)	< 0.001
White ethnicity	–5.30 (–7.92 to –2.67)	< 0.001
IMD quintile (per quintile increase in deprivation)	0.96 (–2.00 to 0.08)	0.07

$n = 136$; $r^2 = 0.38$.

TABLE 61 Characteristics of studies included in review

Study	Characteristic of study										
	Country	Population	Years of birth	Sampling method	Sample size	Mean (SD) or median (IQR) GA (weeks)	Mean (SD) or median (IQR) BW (grams)	Early or school age	Ages at assessments (month for early, years for school-age)	Assessment tools used	Mean (SD) assessment scores[a]
Bassan 2011[278]	Israel	BW < 10th percentile for GA[b]	1992–7	SC	32	33.1 (2.2)	1182 (229)	Early	24	BSID-II	95.8 (19.1)
								School-age	6	WPPSI-R	103.4 (17.7)
Bowen 1996[283]	Australia	BW < 1000 g	1985–8	SC	45	27.6 (2.3)	864 (90)	Early	12, 36	GMDS	–
								School-age	5	S-B-IV	94.4 (11.2)
Bruggink 2010[284]	The Netherlands	'Preterm'	1992–7	SC	50	30.0 (1.9)	1184 (292)	Early	19	BSID-II	100.5 (11.2)
								School-age	8	WISC-III	92.2 (10.6)
Charkaluk 2011[285]	France	GA < 33 weeks	1997	PB	313	29.8 (2.1)	1355 (406)	Early	24	Brunet-Lezine Revised	96.7 (12.7)
								School-age	5	KABC	94.7 (18.7)
Claas 2011[292]	The Netherlands	BW ≤ 750 g and GA ≥ 24 weeks	1996–2005	SC	101	28.0 (24.8–34.4)[c]	675 (480–750)[c]	Early	24	BSID-II/GMDS	–
								School-age	5.5	WPPSI/RAKIT/ SON-R	–
Cohen 1995[286]	USA	'Preterm'[b]	1972–4	SC	20	28.1 (2.1)	1111 (187)	Early	24	BSID	103.9 (21.1)
								School-age	5, 8, 12, 18	S-B-III/WISC-R/ WAIS-R	101.8 (19.0)

continued

TABLE 61 Characteristics of studies included in review (*continued*)

Study	Characteristic of study										
	Country	Population	Years of birth	Sampling method	Sample size	Mean (SD) or median (IQR) GA (weeks)	Mean (SD) or median (IQR) BW (grams)	Early or school age	Ages at assessments (month for early, years for school-age)	Assessment tools used	Mean (SD) assessment scores[a]
Fedrizzi 1993[279]	Italy	Spastic diplegia	1984–1991	SC	11	29.6 (1.6)	1474 (321)	Early	36	GMDS	72.6 (14.5)
								School-age	6	WPPSI	76.4 (18.9)
Gray 1995[281]	Australia	GA 23–33 weeks with diagnosis of BPD	1989–1990	SC	126	28.2	1065	Early	24	GMDS	108.5
								School-age	8	WISC-III	90.5
Gray 2006[287]	New Zealand	GA < 32 weeks or BW < 1500 g	1998–2000	SC	99	27.8 (2.4)	1065 (321)	Early	24	BSID-II	86.1 (17.3)
								School-age	6	WPPSI-R	95.4 (15.2)
Hack 2005[162]	USA	BW < 1000 g	1992–5	SC	200	26.4 (2)	811 (125)	Early	20	BSID-II	75.6 (16.0)
								School-age	8	KABC	87.8 (19.0)
Kilbride 1990[293]	USA	BW < 801 g	1983–1990	MC	129	25.9 (1.6)	698 (82)	Early	12–24, 36	BSID/S-B-III	84.4 (10.0)
								School-age	5	S-B-III	85.7 (11.6)
Marlow 2005[225]	UK	GA < 26 weeks	1995	PB	212	25.0 (0.7)	748 (116)	Early	30	BSID-II	81.7 (14.5)
								School-age	6, 11	KABC	83.8 (18.0)
McGrath 2000[280]	USA	BW < 1850 g with neonatal diagnoses[b]	1985–9	SC	88	29.6 (2.2)	1200 (285)	Early	18	BSID-II	105.2 (19.0)
								School-age	8	WISC-III	96.3 (18.4)

	Characteristic of study										
Study	Country	Population	Years of birth	Sampling method	Sample size	Mean (SD) or median (IQR) GA (weeks)	Mean (SD) or median (IQR) BW (grams)	Early or school age	Ages at assessments (month for early, years for school-age)	Assessment tools used	Mean (SD) assessment scores[a]
Munck 2012[160]	Finland	BW < 1500 g	2001–4	SC	124	28.7 (2.8)	1061 (260)	Early	24	BSID-II	101.2 (16.3)
								School-age	5	WPPSI-R	99.3 (17.7)
Orchinik 2011[294]	USA	GA < 28 weeks or BW < 1000 g	2001–3	SC	139	25.9 (1.6)	818 (174)	Early	20	BSID-II	77.2 (17.3)
								School-age	6	BIA	86.3 (21.1)
Potharst 2012[161]	The Netherlands	GA < 30 weeks or BW < 1000 g	2003–4	SC	100	28.7 (1.6)	1040 (253)	Early	24, 36	BSID-II	102.0 (14.0)
								School-age	5	WPPSI-III	93.0 (17.0)
Reuss 1996[288]	USA	BW 501–2000 g[b]	1984–7	MC	231	29.2 (2.9)	1142 (223)	Early	24	BSID/S-B-III	–
								School-age	6, 9, 16	S-B-IV/WISC-III/ WASI	–
Roberts 2010[163]	Australia	GA 22–27 weeks or BW 500–999 g	1997	PB	186	26.5 (2.0)	832 (164)	Early	24	BSID-II	–
								School-age	8	WISC-R	94.4 (14.2)
Skranes 1998[289]	Norway	BW < 1500 g	1988	PB	21	29.0 (2.0)	1218 (193)	Early	12	BSID	99.0 (18.3)
								School-age	6	WPPSI	96.0 (16.4)
Smith 2006[282]	USA	BW < 1500 g from lower socioeconomic groups	1990–2	MC	161	29.7 (2.5)	1114 (267)	Early	40	S-B-IV	86.2 (10.6)
								School-age	6, 8, 10	S-B-IV	85.1 (12.4)

continued

TABLE 61 Characteristics of studies included in review (*continued*)

	Characteristic of study										
Study	Country	Population	Years of birth	Sampling method	Sample size	Mean (SD) or median (IQR) GA (weeks)	Mean (SD) or median (IQR) BW (grams)	Early or school age	Ages at assessments (month for early, years for school-age)	Assessment tools used	Mean (SD) assessment scores[a]
Tommiska 2003[290]	Finland	BW < 1000 g	1996–7	SC	72	27.1	778	Early	24	BSID-II	95.5
								School-age	5	WPPSI-R	101.0
Veelken 1991[291]	Germany	BW < 1500 g	1983–6	PB	234	29.9 (2.8)	1196 (211)	Early	18–20	GMDS	97.3 (15.9)
								School-age	9	KABC	88.3 (17.6)
Vermeulen 2001[295]	The Netherlands	GA ≤ 32 weeks or BW < 1500 g	1991–3	SC	185	29.2 (2.1)	1183 (313)	Early	18	GMDS	99.0 (13.9)
								School-age	7–10	WISC-R	100.6 (14.0)
Wolke 1999[173]	Germany	GA < 32 weeks	1985–6	PB	254	29.6 (1.5)	1298 (340)	Early	20	GMDS	90.8 (22.8)
								School-age	6, 8	KABC	88.2 (18.6)

BIA, Brief Intellectual Ability; BSID/BSID-II, Bayley Scale of Infant Development 1st or 2nd edition; BW, birthweight; GA, gestational age; KABC, Kaufman Assessment Battery for Children; MC, multicentre; PB, population based; RAKIT, Revision Amsterdam Children's Intelligence Test; S-B-III/IV, Stanford–Binet Intelligence Scale 3rd or 4th edition; SC, single-centre; SON-R, Snijders-Oomen Nonverbal Revised; WAIS-R, Wechsler Intelligence Scale for Adults-Revised; WASI, Wechsler Abbreviated Scale of Intelligence; WISC-III/R, Wechsler Intelligence Scale for Children 3rd or revised edition; WPPSI/-R, Wechsler Pre-school and Primary Scale of Intelligence 1st or revised edition.

a Where participants received multiple assessments, the mean (SD) score for the assessment performed at the oldest age was presented.

b For these studies, only participants born before 32 weeks' gestation and/or with a birthweight of < 1500 g were included in the review.

c Data presented are the medians and IQRs.

TABLE 62 Quality assessment of included studies using the QUADAS-2 appraisal tool

Study	Risk of bias				Applicability concerns			Reasons for being considered high risk for bias or applicability concerns, as judged against the standards set, with statements being numbered according to the domain it is applied to
	Patient selection [1]	Index test [2]	Reference standard [3]	Flow and timing [4]	Patient selection [5]	Index test [6]	Reference standard [7]	
Bassan 2011[278]	↑	↔	↔	↑	↑	↔	↑	[1] Inclusion criteria: birthweight below 10th percentile for gestational age [4] Final cohort represents < 30% of eligible population [5] Study population restricted to children with birthweight below 10th percentile for gestational age [7] Outdated assessment tool (WPPSI-R, 1989) used
Bowen 1996[283]	↔	↔	↑	↔	↑	↑	↑	[3] Assessors not blinded to results of developmental assessment [5] Study population was born before 1990 [6] Outdated assessment tool (GMDS, 1970) used [7] Outdated assessment tool (S-B-IV, 1986) used
Bruggink 2010[284]	?	↔	↔	↑	↔	↔	↔	[1] Recruitment/sampling method not stated [4] Final cohort represents < 30% of eligible population
Charkaluk 2011[285]	↔	↔	↔	↑	↔	↑	↑	[4] Final cohort represents < 30% of eligible population [6] Non-universal assessment tool (Brunet-Lezine Revised, a French psychometric test) used [7] Outdated assessment tool (KABC, 1983) used
Claas 2011[292]	↔	↔	↑	↔	↔	↑	↑	[3] Assessors not blinded to results of developmental assessment [6] Outdated assessment tool (GMDS, 1984) used [7] Non-universal assessment tools (RAKIT and SON-R, Dutch psychometric tests) used

continued

TABLE 62 Quality assessment of included studies using the QUADAS-2 appraisal tool (*continued*)

Study	Risk of bias				Applicability concerns			Reasons for being considered high risk for bias or applicability concerns, as judged against the standards set, with statements being numbered according to the domain it is applied to
	Patient selection [1]	Index test [2]	Reference standard [3]	Flow and timing [4]	Patient selection [5]	Index test [6]	Reference standard [7]	
Cohen 1995[286]	↔	↔	↔	↔	↑	↑	↑	[5] Study population was born before 1990 [6] Outdated assessment tool (BSID, 1969) used [7] Outdated assessment tools (S-B-III, 1973; WISC-R, 1974 and WAIS-R, 1981) used
Fedrizzi 1993[279]	↑	↔	↑	↑	↑	↑	↑	[1] Inclusion criteria: diagnosis of spastic diplegia [3] Assessors not blinded to results of developmental assessment [4] Final cohort represents < 30% of eligible population [5] Study population restricted to children with spastic diplegia [6] Outdated assessment tool (GMDS, 1970) used [7] Outdated assessment tool (WPPSI, 1967) used
Gray 1995[281]	↑	↔	↔	↔	↑	↑	↑	[1] Inclusion criteria: diagnosis of bronchopulmonary dysplasia [5] Study population restricted to children with bronchopulmonary dysplasia [6] Outdated assessment tool (GMDS, 1970) used [7] Outdated assessment tool (WPPSI-R, 1989) used
Gray 2006[287]	↔	↔	↔	↔	↔	↔	↔	
Hack 2005[162]	↔	↔	↔	↔	↔	↔	↑	[7] Outdated assessment tool (KABC, 1983) used

Study	Risk of bias				Applicability concerns			Reasons for being considered high risk for bias or applicability concerns, as judged against the standards set, with statements being numbered according to the domain it is applied to
	Patient selection [1]	Index test [2]	Reference standard [3]	Flow and timing [4]	Patient selection [5]	Index test [6]	Reference standard [7]	
Marlow 2005[225]	↔	↔	↔	↔	↔	↔	↑	[7] Outdated assessment tool (KABC, 1983) used
Kilbride 1990[293]	↔	↔	↔	↔	↑	↑	↑	[6] Outdated assessment tool (BSID, 1969) used [7] Outdated assessment (S-B-III, 1973) used
McGrath 2000[280]	↑	↔	↔	↑	↑	↔	↔	[1] Inclusion criteria: meets a priori medical criterion (not specified) [4] Final cohort represents < 30% of eligible population [5] Study population was born before 1990
Munck 2012[160]	↔	↔	↔	↔	↔	↔	↑	[7] Outdated assessment (WPPSI-R, 1989) used
Orchinik 2011[294]	↔	↔	↔	↔	↔	↔	↔	
Potharst 2012[161]	↔	↔	↔	↑	↔	↔	↔	[4] Final cohort represents < 30% of eligible population
Reuss 1996[288]	↔	↔	?	↑	↑	↑	↔	[3] Blinding of assessors not stated [4] Final cohort represents < 30% of eligible population [5] Study population was born before 1990 [6] Outdated assessment tool (BSID, 1969 and S-B-III, 1973) used
Roberts 2010[163]	↔	↔	?	↔	↔	↔	↔	[3] Blinding of assessors not stated
Skranes 1998[289]	↔	↔	↔	↑	↑	↑	↑	[4] Final cohort represents < 30% of eligible population [5] Study population was born before 1990 [6] Outdated assessment tool (BSID, 1969 and S-B-III, 1973) used [7] Outdated assessment tool (WPPSI, 1967) used

continued

TABLE 62 Quality assessment of included studies using the QUADAS-2 appraisal tool (*continued*)

Study	Risk of bias				Applicability concerns			Reasons for being considered high risk for bias or applicability concerns, as judged against the standards set, with statements being numbered according to the domain it is applied to
	Patient selection [1]	Index test [2]	Reference standard [3]	Flow and timing [4]	Patient selection [5]	Index test [6]	Reference standard [7]	
Smith 2006[282]	↑	↔	?	↑	↑	↑	↑	[1] Inclusion criteria: from middle to lower socioeconomic groups [3] Blinding of assessors not stated [4] Final cohort represents < 30% of eligible population [5] Study population restricted to children from middle to lower socioeconomic groups [6] Outdated assessment tool (S-B-IV, 1986) used [7] Outdated assessment tool (S-B-IV, 1986) used
Tommiska 2003[290]	↔	↔	?	↔	↔	↔	↑	[3] Blinding of assessors not stated [7] Outdated assessment tool (WPPSI-R, 1989) used
Veelken 1991[291]	↔	↔	↔	↑	↑	↑	↑	[4] Final cohort represents < 30% of eligible population [5] Study population was born before 1990 [6] Outdated assessment tool (GMDS, 1970) used [7] Outdated assessment tool (KABC, 1983) used
Vermeulen 2001[295]	↔	↔	↔	↔	↔	↑	↑	[6] Assessment tool developed before 1990 (GMDS, 1970) used [7] Assessment tool developed before 1990 (WISC-R, 1974) used

Study	Risk of bias				Applicability concerns			Reasons for being considered high risk for bias or applicability concerns, as judged against the standards set, with statements being numbered according to the domain it is applied to
	Patient selection [1]	Index test [2]	Reference standard [3]	Flow and timing [4]	Patient selection [5]	Index test [6]	Reference standard [7]	
Wolke 1999[173]	↔	↔	↔	↔	↑	↑	↑	[5] Study population was born before 1990 [6] Outdated assessment tool (GMDS, 1970) used [7] Outdated assessment tool (KABC, 1983) used

↔ = low risk; ↑ = high risk; ? = unclear risk.
Abbreviation for assessment tools are followed by year of publication:
BSID = Bayley Scale of Infant Development, 1969
GMDS = Griffiths Mental Development Scales,1970 and 1984
KABC = Kaufman Assessment Battery for Children, 1983
RAKIT = Revision Amsterdam Children's Intelligence Test, 1987
S-B-III/IV = Stanford–Binet Intelligence Scale 3rd edition, 1973 and 4th edition, 1986
SON-R = Snijders-Oomen Nonverbal Revised, 1998 and 2003
WAIS-R = Wechsler Intelligence Scale for Adults-Revised, 1981
WISC-R = Wechsler Intelligence Scale for Children-Revised, 1974
WPPSI = Wechsler Pre-school and Primary Scale of Intelligence-Revised, 1989.

Chapter 6

TABLE 63 Cost-effectiveness estimates based on death primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
PiPS data set											
Colonisation (yes)	61,354.61 (57,739.13 to 64,970.09)	55,071.11 (49,211.64 to 60,930.58)	6283.5 (–601.64 to 13,168.63)	0.029 (0.014 to 0.0439)	0.0327 (0.0089 to 0.0565)	0.0037 (–0.0244 to 0.0319)	1,686,947.4	60.6	4.9	5.5	5.7
Colonisation (no)	86,147.5 (72,490.12 to 99,804.87)	69,272.01 (64,240.99 to 74,303.03)	16,875.49 (2320.93 to 31,430.04)	0.1463 (0.0698 to 0.2228)	0.0523 (0.0294 to 0.0753)	–0.094 (–0.1739 to –0.0141)	–179,526.72	0.3	0.1	0.1	0.0
Gestational age < 28 years	85,643.61 (79,893.44 to 91,393.79)	85,485.04 (79,268.07 to 91,702)	158.58 (–8309.9 to 8627.06)	0.1309 (0.0926 to 0.1692)	0.1711 (0.1287 to 0.2134)	0.0402 (–0.0169 to 0.0973)	3946.6786	90.6	50.0	56.4	61.3
Gestational age ≥ 28 years	42,427.22 (39,409.85 to 45,444.6)	41,882.82 (39,068 to 44,697.65)	544.4 (–3582.07 to 4670.87)	0.0373 (0.0166 to 0.058)	0.006 (–0.0023 to 0.0143)	–0.0313 (–0.0536 to –0.009)	–17,404.544	0.0	39.0	28.8	23.8
Randomisation age ≤ 24 hours	60,552.92 (54,496.09 to 66,609.75)	59,907.42 (53,228.36 to 66,586.49)	645.5 (–8370.89 to 9661.88)	0.0625 (0.025 to 0.1)	0.0838 (0.0418 to 0.1259)	0.0213 (–0.035 to 0.0777)	30,259.012	77.6	41.9	44.7	45.7
Randomisation age > 24 hours	64,119.34 (59,745.71 to 68,492.97)	63,634.34 (59,196.12 to 68,072.56)	485 (–5746.09 to 6716.08)	0.0891 (0.0631 to 0.1152)	0.0849 (0.0597 to 0.1101)	–0.0042 (–0.0404 to 0.032)	–115,345.2	39.0	44.7	43.7	43.3
Sex (male)	64,813.38 (60,104.75 to 69,522.02)	62,142.4 (57,119.77 to 67,165.02)	2670.99 (–4213.63 to 9555.61)	0.0938 (0.0633 to 0.1242)	0.0784 (0.0505 to 0.1063)	–0.0153 (–0.0566 to 0.026)	–174,361.95	22.9	21.1	18.0	17.1
Sex (female)	61,078.55 (55,495.49 to 66,661.62)	63,314.87 (57,795.57 to 68,834.17)	–2236.32 (–10,087.01 to 5614.37)	0.0672 (0.0372 to 0.0971)	0.0925 (0.0586 to 0.1264)	0.0254 (–0.0199 to 0.0706)	–88,174.182	86.6	71.8	76.6	78.9

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Weight < 1000 g (yes)	87,432.57 (81,654.37 to 93,210.78)	85,122.31 (78,960.42 to 91,284.21)	2310.26 (–6137.03 to 10,757.54)	0.1424 (0.103 to 0.1818)	0.164 (0.1228 to 0.2051)	0.0216 (–0.0354 to 0.0786)	106,941.31	76.8	28.6	31.5	32.9
Weight < 1000 g (no)	40,184.67 (37,715.64 to 42,653.71)	41,294.42 (38,612.23 to 43,976.61)	–1109.74 (–4755.32 to 2535.84)	0.0252 (0.0079 to 0.0424)	0.0092 (–0.0012 to 0.0195)	–0.016 (–0.0361 to 0.0041)	69,433.126	5.5	71.4	65.6	62.7
NNRD data set											
Colonisation (yes)	59,017.58 (55,765.98 to 62,269.17)	53,138.61 (47,903.19 to 58,374.03)	5878.97 (–284.03 to 12,041.96)	0.029 (0.014 to 0.0439)	0.0327 (0.0089 to 0.0565)	0.0037 (–0.0244 to 0.0319)	1,578,342.3	58.4	2.9	3.2	3.3
Colonisation (no)	81,452.58 (72,626.84 to 90,278.33)	67,767.97 (62,941.77 to 72,594.18)	13,684.61 (3625.48 to 23,743.74)	0.1463 (0.0698 to 0.2228)	0.0523 (0.0294 to 0.0753)	–0.094 (–0.1739 to –0.0141)	–145,581.17	0.5	0.7	0.3	0.1
Gestational age < 28 years	81,834.95 (77,038.62 to 86,631.28)	81,626.86 (75,675.75 to 87,577.97)	208.08 (–7435.24 to 7851.41)	0.1309 (0.0926 to 0.1692)	0.1711 (0.1287 to 0.2134)	0.0402 (–0.0169 to 0.0973)	5178.8003	93	48.4	55.8	59.7
Gestational age ≥ 28 years	41,194.96 (38,659.84 to 43,730.09)	42,334.79 (39,428.07 to 45,241.5)	–1139.82 (–4996.75 to 2717.1)	0.0373 (0.0166 to 0.058)	0.006 (–0.0023 to 0.0143)	–0.0313 (–0.0536 to –0.009)	36,440.48	0.2	70.1	59.1	53.4
Randomisation age ≤ 24 hours	59,180.96 (53,370.47 to 64,991.46)	58,124.43 (51,774.27 to 64,474.58)	1056.54 (–7550.8 to 9663.88)	0.0625 (0.025 to 0.1)	0.0838 (0.0418 to 0.1259)	0.0213 (–0.035 to 0.0777)	49,527.488	78.6	38.7	42.1	43.6
Randomisation age > 24 hours	61,266.6 (57,611.07 to 64,922.13)	62,096.83 (57,835.16 to 66,358.49)	–830.22 (–6444.9 to 4784.46)	0.0891 (0.0631 to 0.1152)	0.0849 (0.0597 to 0.1101)	–0.0042 (–0.0404 to 0.032)	197,448.99	42.4	60.3	59.6	59.5
Sex (male)	63,399.49 (58,920.37 to 67,878.61)	61,730.88 (56,698.18 to 66,763.58)	1668.6 (–5068.65 to 8405.86)	0.0938 (0.0633 to 0.1242)	0.0784 (0.0505 to 0.1063)	–0.0153 (–0.0566 to 0.026)	–108,926.51	23.8	33.0	30.1	28.6

continued

TABLE 63 Cost-effectiveness estimates based on death primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (female)	57,220.04 (53,157.04 to 61,283.04)	60,200.92 (55,254.06 to 65,147.77)	–2980.88 (–9382.39 to 3420.63)	0.0672 (0.0372 to 0.0971)	0.0925 (0.0586 to 0.1264)	0.0254 (–0.0199 to 0.0706)	–117,530.82	86.2	81.4	86.2	88.0
Weight < 1000 g (yes)	83,083.85 (78,193.05 to 87,974.64)	80,400.47 (74,646.03 to 86,154.91)	2683.38 (–4868.67 to 10,235.43)	0.1424 (0.103 to 0.1818)	0.164 (0.1228 to 0.2051)	0.0216 (–0.0354 to 0.0786)	124,213.13	75.7	25.9	29.6	31.3
Weight < 1000 g (no)	39,497.7 (37,562.52 to 41,432.88)	42,660.06 (39,475.43 to 45,844.69)	–3162.35 (–6888.86 to 564.15)	0.0252 (0.0079 to 0.0424)	0.0092 (–0.0012 to 0.0195)	–0.016 (–0.0361 to 0.0041)	197,858.35	5.9	95.5	93.8	92.3
Combined data set											
Colonisation (yes)	59,082.53 (55,813.78 to 62,351.27)	53,064.15 (47,856.46 to 58,271.83)	6018.38 (–130.17 to 12,166.93)	0.029 (0.014 to 0.0439)	0.0327 (0.0089 to 0.0565)	0.0037 (–0.0244 to 0.0319)	1,615,770.8	58.4	2.5	2.7	2.8
Colonisation (no)	82,009.26 (73,172.95 to 90,845.58)	67,938.27 (63,094.89 to 72,781.64)	14,071 (3994.36 to 24,147.64)	0.1463 (0.0698 to 0.2228)	0.0523 (0.0294 to 0.0753)	–0.094 (–0.1739 to –0.0141)	–149,691.66	0.5	0.6	0.3	0.1
Gestational age < 28 years	82,110.69 (77,277.46 to 86,943.93)	81,796.8 (75,847.03 to 87,746.58)	313.89 (–7351.61 to 7979.4)	0.1309 (0.0926 to 0.1692)	0.1711 (0.1287 to 0.2134)	0.0402 (–0.0169 to 0.0973)	7812.1565	93.0	46.9	54.4	59.2
Gestational age ≥ 28 years	41,205.5 (38,669.57 to 43,741.43)	42,375.67 (39,459.53 to 45,291.81)	–1170.17 (–5034.72 to 2694.39)	0.0373 (0.0166 to 0.058)	0.006 (–0.0023 to 0.0143)	–0.0313 (–0.0536 to –0.009)	37,410.579	0.2	71.0	59.7	53.6
Randomisation age ≤ 24 hours	59,476.77 (53,591.8 to 65,361.74)	58,194.19 (51,854.17 to 64,534.22)	1282.58 (–7367.79 to 9932.94)	0.0625 (0.025 to 0.1)	0.0838 (0.0418 to 0.1259)	0.0213 (–0.035 to 0.0777)	60,123.538	78.6	37	40.3	41.9
Randomisation age > 24 hours	61,349.73 (57,679.36 to 65,020.1)	62,210.76 (57,941.56 to 66,479.97)	–861.04 (–6491.11 to 4769.04)	0.0891 (0.0631 to 0.1152)	0.0849 (0.0597 to 0.1101)	–0.0042 (–0.0404 to 0.032)	204,777.36	42.4	60.6	59.7	59.7

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (male)	63,431.42 (58,926.17 to 67,936.66)	61,714.56 (56,686.93 to 66,742.19)	1716.86 (–5034.01 to 8467.73)	0.0938 (0.0633 to 0.1242)	0.0784 (0.0505 to 0.1063)	–0.0153 (–0.0566 to 0.026)	–112,076.61	23.8	32.4	29.5	27.8
Sex (female)	57,497.39 (53,404.44 to 61,590.33)	60,454.11 (55,484.46 to 65,423.76)	–2956.72 (–9394.86 to 3481.42)	0.0672 (0.0372 to 0.0971)	0.0925 (0.0586 to 0.1264)	0.0254 (–0.0199 to 0.0706)	–116,578.36	86.2	80.6	85.7	87.8
Weight < 1000 g (yes)	83,364.15 (78,439.38 to 88,288.91)	80,535.46 (74,783.56 to 86,287.36)	2828.68 (–4743.48 to 10,400.85)	0.1424 (0.103 to 0.1818)	0.164 (0.1228 to 0.2051)	0.0216 (–0.0354 to 0.0786)	130,939.25	75.7	24.7	27.8	29.3
Weight < 1000 g (no)	39,500.58 (37,565.42 to 41,435.75)	42,731.41 (39,531.03 to 45,931.79)	–3230.83 (–6970.78 to 509.13)	0.0252 (0.0079 to 0.0424)	0.0092 (–0.0012 to 0.0195)	–0.016 (–0.0361 to 0.0041)	202,142.56	5.9	95.8	94.3	93.0

a The difference in effects was inverted, i.e. negative values were given a positive sign, to reflect the fact that a reduction in adverse outcomes is a positive effect. *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a:
b Based on 1000 bootstrap replicates of the data set.
c GBP £20,000 cost-effectiveness threshold.
d GBP £30,000 cost-effectiveness threshold.

TABLE 64 Cost-effectiveness estimates based on sepsis primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
PiPS data set											
Colonisation (yes)	61,354.61 (57,739.13 to 64,970.09)	55,071.11 (49,211.64 to 60,930.58)	6283.5 (–601.64 to 13,168.63)	0.087 (0.0618 to 0.1121)	0.0981 (0.0583 to 0.138)	0.0112 (–0.0359 to 0.0583)	562,315.79	69.9	4.9	6.2	7.8
Colonisation (no)	86,147.5 (72,490.12 to 99,804.87)	69,272.01 (64,240.99 to 74,303.03)	16,875.49 (2320.93 to 31,430.04)	0.2073 (0.1196 to 0.2951)	0.1212 (0.0876 to 0.1548)	–0.0861 (–0.1801 to 0.0078)	–195,987.42	3.2	0.1	0.0	0.0
Gestational age < 28 years	85,643.61 (79,893.44 to 91,393.79)	85,485.04 (79,268.07 to 91,702)	158.58 (–8309.9 to 8627.06)	0.1913 (0.1466 to 0.2359)	0.1743 (0.1317 to 0.217)	–0.0169 (–0.0787 to 0.0448)	–9365.0001	29.7	50.0	46.5	45.4
Gestational age ≥ 28 years	42,427.22 (39,409.85 to 45,444.6)	41,882.82 (39,068 to 44,697.65)	544.4 (–3582.07 to 4670.87)	0.0311 (0.0121 to 0.05)	0.0569 (0.032 to 0.0817)	0.0258 (–0.0054 to 0.0571)	21,075.912	94.9	39.0	47.5	52.2
Randomisation age ≤ 24 hours	60,552.92 (54,496.09 to 66,609.75)	59,907.42 (53,228.36 to 66,586.49)	645.5 –8370.89 to 9661.88)	0.075 (0.0342 to 0.1158)	0.1377 (0.0855 to 0.19)	0.0627 (–0.0036 to 0.129)	10,290.953	97.3	41.9	51.4	57.9
Randomisation age > 24 hours	64,119.34 (59,745.71 to 68,492.97)	63,634.34 (59,196.12 to 68,072.56)	485 (–5746.09 to 6716.08)	0.1196 (0.0899 to 0.1492)	0.104 (0.0765 to 0.1316)	–0.0155 (–0.056 to 0.025)	–31,227.186	23.1	44.7	41.6	40.3
Sex (male)	64,813.38 (60,104.75 to 69,522.02)	62,142.4 (57,119.77 to 67,165.02)	2670.99 (–4213.63 to 9555.61)	0.1165 (0.083 to 0.15)	0.1176 (0.0842 to 0.1511)	0.0012 (–0.0462 to 0.0485)	2,283,311.2	51.6	21.1	24.4	25.2

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (female)	61,078.55 (55,495.49 to 66,661.62)	63,314.87 (57,795.57 to 68,834.17)	–2236.32 (–10,087.01 to 5614.37)	0.097 (0.0616 to 0.1325)	0.1068 (0.0707 to 0.1429)	0.0097 (–0.0408 to 0.0603)	–229,445.08	63.8	71.8	73.5	73.4
Weight < 1000 g (yes)	87,432.57 (81,654.37 to 93,210.78)	85,122.31 (78,960.42 to 91,284.21)	2310.26 (–6137.03 to 10,757.54)	0.1887 (0.1446 to 0.2329)	0.1833 (0.1403 to 0.2263)	–0.0055 (–0.0671 to 0.0562)	–422,970.59	42.1	28.6	27.5	27.0
Weight < 1000 g (no)	40,184.67 (37,715.64 to 42,653.71)	41,294.42 (38,612.23 to 43,976.61)	–1109.74 (–4755.32 to 2535.84)	0.0314 (0.0123 to 0.0506)	0.0459 (0.0232 to 0.0685)	0.0144 (–0.0153 to 0.0441)	–76,931.903	82.0	71.4	75.3	77.3
NNRD data set											
Colonisation (yes)	59,017.58 (55,765.98 to 62,269.17)	53,138.61 (47,903.19 to 58,374.03)	5878.97 (–284.03 to 12,041.96)	0.087 (0.0618 to 0.1121)	0.0981 (0.0583 to 0.138)	0.0112 (–0.0359 to 0.0583)	526,114.09	65.6	2.9	4.9	5.4
Colonisation (no)	81,452.58 (72,626.84 to 90,278.33)	67,767.97 (62,941.77 to 72,594.18)	13,684.61 (3625.48 to 23,743.74)	0.2073 (0.1196 to 0.2951)	0.1212 (0.0876 to 0.1548)	–0.0861 (–0.1801 to 0.0078)	–158,929.43	2.5	0.7	0.3	0.3
Gestational age < 28 years	81,834.95 (77,038.62 to 86,631.28)	81,626.86 (75,675.75 to 87,577.97)	208.08 (–7435.24 to 7851.41)	0.1913 (0.1466 to 0.2359)	0.1743 (0.1317 to 0.217)	–0.0169 (–0.0787 to 0.0448)	–12,288.679	31.5	48.4	45.3	44.4

continued

TABLE 64 Cost-effectiveness estimates based on sepsis primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Gestational age ≥ 28 years	41,194.96 (38,659.84 to 43,730.09)	42,334.79 (39,428.07 to 45,241.5)	–1139.82 (–4996.75 to 2717.1)	0.0311 (0.0121 to 0.05)	0.0569 (0.032 to 0.0817)	0.0258 (–0.0054 to 0.0571)	–44,127.348	94.9	70.1	76.7	79.3
Randomisation age ≤ 24 hours	59,180.96 (53,370.47 to 64,991.46)	58,124.43 (51,774.27 to 64,474.58)	1056.54 (–7550.8 to 9663.88)	0.075 (0.0342 to 0.1158)	0.1377 (0.0855 to 0.19)	0.0627 (–0.0036 to 0.129)	16,844.074	97.0	38.7	48.6	53.4
Randomisation age > 24 hours	61,266.6 (57,611.07 to 64,922.13)	62,096.83 (57,835.16 to 66,358.49)	–830.22 (–6444.9 to 4784.46)	0.1196 (0.0899 to 0.1492)	0.104 (0.0765 to 0.1316)	–0.0155 (–0.056 to 0.025)	53,454.985	23.9	60.3	56.4	54.9
Sex (male)	63,399.49 (58,920.37 to 67,878.61)	61,730.88 (56,698.18 to 66,763.58)	1668.6 (–5068.65 to 8405.86)	0.1165 (0.083 to 0.15)	0.1176 (0.0842 to 0.1511)	0.0012 (–0.0462 to 0.0485)	1,426,418.6	50.5	33.0	33.8	33.9
Sex (female)	57,220.04 (53,157.04 to 61,283.04)	60,200.92 (55,254.06 to 65,147.77)	–2980.88 (–9382.39 to 3420.63)	0.097 (0.0616 to 0.1325)	0.1068 (0.0707 to 0.1429)	0.0097 (–0.0408 to 0.0603)	–305,836.34	65.9	81.4	81.8	81.2
Weight < 1000 g (yes)	83,083.85 (78,193.05 to 87,974.64)	80,400.47 (74,646.03 to 86,154.91)	2683.38 (–4868.67 to 10,235.43)	0.1887 (0.1446 to 0.2329)	0.1833 (0.1403 to 0.2263)	–0.0055 (–0.0671 to 0.0562)	–491,283.52	44.6	25.9	25.6	25.5
Weight < 1000 g (no)	39,497.7 (37,562.52 to 41,432.88)	42,660.06 (39,475.43 to 45,844.69)	–3162.35 (–6888.86 to 564.15)	0.0314 (0.0123 to 0.0506)	0.0459 (0.0232 to 0.0685)	0.0144 (–0.0153 to 0.0441)	–219,227.05	83.8	95.5	96.2	96.6

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Combined data set											
Colonisation (yes)	59,082.53 (55,813.78 to 62,351.27)	53,064.15 (47,856.46 to 58,271.83)	6018.38 (–130.17 to 12,166.93)	0.087 (0.0618 to 0.1121)	0.0981 (0.0583 to 0.138)	0.0112 (–0.0359 to 0.0583)	538,590.27	65.6	2.5	4.6	5.0
Colonisation (no)	82,009.26 (73,172.95 to 90,845.58)	67,938.27 (63,094.89 to 72,781.64)	14,071 (3994.36 to 24,147.64)	0.2073 (0.1196 to 0.2951)	0.1212 (0.0876 to 0.1548)	–0.0861 (–0.1801 to 0.0078)	–163,416.8	2.5	0.6	0.3	0.3
Gestational age < 28 years	82,110.69 (77,277.46 to 86,943.93)	81,796.8 (75,847.03 to 87,746.58)	313.89 (–7351.61 to 7979.4)	0.1913 (0.1466 to 0.2359)	0.1743 (0.1317 to 0.217)	–0.0169 (–0.0787 to 0.0448)	–18,537.321	31.5	46.9	44.3	43.3
Gestational age ≥ 28 years	41,205.5 (38,669.57 to 43,741.43)	42,375.67 (39,459.53 to 45,291.81)	–1170.17 (–5034.72 to 2694.39)	0.0311 (0.0121 to 0.05)	0.0569 (0.032 to 0.0817)	0.0258 (–0.0054 to 0.0571)	–45,302.084	94.9	71.0	77.5	80.1
Randomisation age ≤ 24 hours	59,476.77 (53,591.8 to 65,361.74)	58,194.19 (51,854.17 to 64,534.22)	1282.58 (–7367.79 to 9932.94)	0.075 (0.0342 to 0.1158)	0.1377 (0.0855 to 0.19)	0.0627 (–0.0036 to 0.129)	20,447.743	97.0	37.0	46.5	51.7
Randomisation age > 24 hours	61,349.73 (57,679.36 to 65,020.1)	62,210.76 (57,941.56 to 66,479.97)	–861.04 (–6491.11 to 4769.04)	0.1196 (0.0899 to 0.1492)	0.104 (0.0765 to 0.1316)	–0.0155 (–0.056 to 0.025)	55,438.983	23.9	60.6	57.7	55.4

continued

TABLE 64 Cost-effectiveness estimates based on sepsis primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (male)	63,431.42 (58,926.17 to 67,936.66)	61,714.56 (56,686.93 to 66,742.19)	1716.86 (–5034.01 to 8467.73)	0.1165 (0.083 to 0.15)	0.1176 (0.0842 to 0.1511)	0.0012 (–0.0462 to 0.0485)	1,467,669.9	50.5	32.4	33.7	33.9
Sex (female)	57,497.39 (53,404.44 to 61,590.33)	60,454.11 (55,484.46 to 65,423.76)	–2956.72 (–9394.86 to 3481.42)	0.097 (0.0616 to 0.1325)	0.1068 (0.0707 to 0.1429)	0.0097 (–0.0408 to 0.0603)	–303,357.87	65.9	80.6	81.5	81.2
Weight < 1000 g (yes)	83,364.15 (78,439.38 to 88,288.91)	80,535.46 (74,783.56 to 86,287.36)	2828.68 (–4743.48 to 10,400.85)	0.1887 (0.1446 to 0.2329)	0.1833 (0.1403 to 0.2263)	–0.0055 (–0.0671 to 0.0562)	–517,886.44	44.6	24.7	25.0	24.8
Weight < 1000 g (no)	39,500.58 (37,565.42 to 41,435.75)	42,731.41 (39,531.03 to 45,931.79)	–3230.83 (–6970.78 to 509.13)	0.0314 (0.0123 to 0.0506)	0.0459 (0.0232 to 0.0685)	0.0144 (–0.0153 to 0.0441)	–223,973.96	83.8	95.8	96.6	96.8

a The difference in effects was inverted, i.e. negative values were given a positive sign, to reflect the fact that a reduction in adverse outcomes is a positive effect. *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a:
b Based on 1000 bootstrap replicates of the data set.
c GBP £20,000 cost-effectiveness threshold
d GBP £30,000 cost-effectiveness threshold.

TABLE 65 Cost-effectiveness estimates based on NEC primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
PiPS data set											
Colonisation (yes)	61,354.61 (57,739.13 to 64,970.09)	55,071.11 (49,211.64 to 60,930.58)	6283.5 (–601.64 to 13,168.63)	0.0663 (0.0441 to 0.0884)	0.0514 (0.0218 to 0.081)	–0.0149 (–0.0518 to 0.0221)	–423,110.58	22.7	4.9	4.7	4.7
Colonisation (no)	86,147.5 (72,490.12 to 99,804.87)	69,272.01 (64,240.99 to 74,303.03)	16,875.49 (2320.93 to 31,430.04)	0.2073 (0.1196 to 0.2951)	0.1102 (0.078 to 0.1424)	–0.0971 (–0.1906 to –0.0037)	–173,751.56	1.7	0.1	0.0	0.0
Gestational age < 28 years	85,643.61 (79,893.44 to 91,393.79)	85,485.04 (79,268.07 to 91,702)	158.58 (–8309.9 to 8627.06)	0.1443 (0.1044 to 0.1842)	0.1645 (0.1228 to 0.2061)	0.0202 (–0.0375 to 0.0779)	7858.813	76.2	50.0	51.7	53.3
Gestational age ≥ 28 years	42,427.22 (39,409.85 to 45,444.6)	41,882.82 (39,068 to 44,697.65)	544.4 (–3582.07 to 4670.87)	0.0404 (0.0189 to 0.0619)	0.0389 (0.0182 to 0.0597)	–0.0015 (–0.0313 to 0.0284)	–375,313.36	44.4	39.0	38.5	38.4
Randomisation age ≤ 24 hours	60,552.92 (54,496.09 to 66,609.75)	59,907.42 (53,228.36 to 66,586.49)	645.5 (–8370.89 to 9661.88)	0.0812 (0.0389 to 0.1236)	0.0599 (0.0239 to 0.0959)	–0.0214 (–0.0769 to 0.0342)	–30,206.019	23.4	41.9	39.2	37.8
Randomisation age > 24 hours	64,119.34 (59,745.71 to 68,492.97)	63,634.34 (59,196.12 to 68,072.56)	485 (–5746.09 to 6716.08)	0.0935 (0.0669 to 0.1201)	0.1125 (0.084 to 0.1411)	0.019 (–0.02 to 0.0581)	25,461.469	81.5	44.7	49.7	51.6
Sex (male)	64,813.38 (60,104.75 to 69,522.02)	62,142.4 (57,119.77 to 67,165.02)	2670.99 (–4213.63 to 9555.61)	0.0938 (0.0633 to 0.1242)	0.098 (0.0672 to 0.1289)	0.0043 (–0.0391 to 0.0476)	622,721.24	56.3	21.1	24.2	24.6
Sex (female)	61,078.55 (55,495.49 to 66,661.62)	63,314.87 (57,795.57 to 68,834.17)	–2236.32 (–10,087.01 to 5614.37)	0.0858 (0.0523 to 0.1194)	0.0996 (0.0646 to 0.1347)	0.0138 (–0.0347 to 0.0623)	–161,779.72	72.6	71.8	74.0	74.2
Weight < 1000 g (yes)	87,432.57 (81,654.37 to 93,210.78)	85,122.31 (78,960.42 to 91,284.21)	2310.26 (–6137.03 to 10,757.54)	0.149 (0.1088 to 0.1892)	0.1608 (0.1199 to 0.2016)	0.0118 (–0.0455 to 0.069)	196,365.53	66.1	28.6	31.3	32.2

continued

TABLE 65 Cost-effectiveness estimates based on NEC primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Weight < 1000 g (no)	40,184.67 (37,715.64 to 42,653.71)	41,294.42 (38,612.23 to 43,976.61)	–1109.74 (–4755.32 to 2535.84)	0.0346 (0.0145 to 0.0547)	0.0398 (0.0186 to 0.0609)	0.0052 (–0.024 to 0.0344)	–214,893.58	63.7	71.4	71.8	71.5
NNRD data set											
Colonisation (yes)	59,017.58 (55,765.98 to 62,269.17)	53,138.61 (47,903.19 to 58,374.03)	5878.97 (–284.03 to 12,041.96)	0.0663 (0.0441 to 0.0884)	0.0514 (0.0218 to 0.081)	–0.0149 (–0.0518 to 0.0221)	–395,870.86	21.8	2.9	2.4	2.6
Colonisation (no)	81,452.58 (72,626.84 to 90,278.33)	67,767.97 (62,941.77 to 72,594.18)	13,684.61 (3625.48 to 23,743.74)	0.2073 (0.1196 to 0.2951)	0.1102 (0.078 to 0.1424)	–0.0971 (–0.1906 to –0.0037)	–140,898	2.4	0.7	0.4	0.4
Gestational age < 28 years	81,834.95 (77,038.62 to 86,631.28)	81,626.86 (75,675.75 to 87,577.97)	208.08 (–7435.24 to 7851.41)	0.1443 (0.1044 to 0.1842)	0.1645 (0.1228 to 0.2061)	0.0202 (–0.0375 to 0.0779)	10,312.272	75.3	48.4	51.5	53.1
Gestational age ≥ 28 years	41,194.96 (38,659.84 to 43,730.09)	42,334.79 (39,428.07 to 45,241.5)	–1139.82 (–4996.75 to 2717.1)	0.0404 (0.0189 to 0.0619)	0.0389 (0.0182 to 0.0597)	–0.0015 (–0.0313 to 0.0284)	785,806.24	46.7	70.1	68.7	67.8
Randomisation age ≤ 24 hours	59,180.96 (53,370.47 to 64,991.46)	58,124.43 (51,774.27 to 64,474.58)	1056.54 (–7550.8 to 9663.88)	0.0812 (0.0389 to 0.1236)	0.0599 (0.0239 to 0.0959)	–0.0214 (–0.0769 to 0.0342)	–49,440.75	22.1	38.7	36.1	34.9
Randomisation age > 24 hours	61,266.6 (57,611.07 to 64,922.13)	62,096.83 (57,835.16 to 66,358.49)	–830.22 (–6444.9 to 4784.46)	0.0935 (0.0669 to 0.1201)	0.1125 (0.084 to 0.1411)	0.019 (–0.02 to 0.0581)	–43,585.177	84.7	60.3	64.8	66.3
Sex (male)	63,399.49 (58,920.37 to 67,878.61)	61,730.88 (56,698.18 to 66,763.58)	1668.6 (–5068.65 to 8405.86)	0.0938 (0.0633 to 0.1242)	0.098 (0.0672 to 0.1289)	0.0043 (–0.0391 to 0.0476)	389,023.25	55.5	33.0	34.8	36.2
Sex (female)	57,220.04 (53,157.04 to 61,283.04)	60,200.92 (55,254.06 to 65,147.77)	–2980.88 (–9382.39, 3420.63)	0.0858 (0.0523 to 0.1194)	0.0996 (0.0646 to 0.1347)	0.0138 (–0.0347 to 0.0623)	–215,642.53	73.2	81.4	81.8	82.3

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Weight < 1000 g (yes)	83,083.85 (78,193.05 to 87,974.64)	80,400.47 (74,646.03 to 86,154.91)	2683.38 (–4868.67 to 10,235.43)	0.149 (0.1088 to 0.1892)	0.1608 (0.1199 to 0.2016)	0.0118 (–0.0455 to 0.069)	228,080.04	64.9	25.9	28.6	30.3
Weight < 1000 g (no)	39,497.7 (37,562.52 to 41,432.88)	42,660.06 (39,475.43 to 45,844.69)	–3162.35 (–6888.86 to 564.15)	0.0346 (0.0145 to 0.0547)	0.0398 (0.0186 to 0.0609)	0.0052 (–0.024 to 0.0344)	–612,366.06	66.1	95.5	95.3	94.6
Combined data set											
Colonisation (yes)	59,082.53 (55,813.78 to 62,351.27)	53,064.15 (47,856.46 to 58,271.83)	6018.38 (–130.17 to 12,166.93)	0.0663 (0.0441 to 0.0884)	0.0514 (0.0218 to 0.081)	–0.0149 (–0.0518 to 0.0221)	–405,258.48	21.8	2.5	2.1	2.3
Colonisation (no)	82,009.26 (73,172.95 to 90,845.58)	67,938.27 (63,094.89 to 72,781.64)	14,071 (3994.36 to 24,147.64)	0.2073 (0.1196 to 0.2951)	0.1102 (0.078 to 0.1424)	–0.0971 (–0.1906 to –0.0037)	–144,876.26	2.4	0.6	0.4	0.4
Gestational age < 28 years	82,110.69 (77,277.46 to 86,943.93)	81,796.8 (75,847.03 to 87,746.58)	313.89 (–7351.61 to 7979.4)	0.1443 (0.1044 to 0.1842)	0.1645 (0.1228 to 0.2061)	0.0202 (–0.0375 to 0.0779)	15,555.935	75.3	46.9	50.8	52.6
Gestational age ≥ 28 years	41,205.5 (38,669.57 to 43,741.43)	42,375.67 (39,459.53 to 45,291.81)	–1170.17 (–5034.72 to 2694.39)	0.0404 (0.0189 to 0.0619)	0.0389 (0.0182 to 0.0597)	–0.0015 (–0.0313 to 0.0284)	806,725.57	46.7	71	69.1	67.8
Randomisation age ≤ 24 hours	59,476.77 (53,591.8 to 65,361.74)	58,194.19 (51,854.17 to 64,534.22)	1282.58 (–7367.79 to 9932.94)	0.0812 (0.0389 to 0.1236)	0.0599 (0.0239 to 0.0959)	–0.0214 (–0.0769 to 0.0342)	–60,018.243	22.1	37	34.7	33.4
Randomisation age > 24 hours	61,349.73 (57,679.36 to 65,020.1)	62,210.76 (57,941.56 to 66,479.97)	–861.04 (–6491.11 to 4769.04)	0.0935 (0.0669 to 0.1201)	0.1125 (0.084 to 0.1411)	0.019 (–0.02 to 0.0581)	–45,202.854	84.7	60.6	65.3	66.8
Sex (male)	63,431.42 (58,926.17 to 67,936.66)	61,714.56 (56,686.93 to 66,742.19)	1716.86 (–5034.01 to 8467.73)	0.0938 (0.0633 to 0.1242)	0.098 (0.0672 to 0.1289)	0.0043 (–0.0391 to 0.0476)	400,273.6	55.5	32.4	34.3	35.3

continued

TABLE 65 Cost-effectiveness estimates based on NEC primary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (female)	57,497.39 (53,404.44 to 61,590.33)	60,454.11 (55,484.46 to 65,423.76)	–2956.72 (–9394.86 to 3481.42)	0.0858 (0.0523 to 0.1194)	0.0996 (0.0646 to 0.1347)	0.0138 (–0.0347 to 0.0623)	–213,894.98	73.2	80.6	81.6	82.1
Weight < 1000 g (yes)	83,364.15 (78,439.38 to 88,288.91)	80,535.46 (74,783.56 to 86,287.36)	2828.68 (–4743.48 to 10,400.85)	0.149 (0.1088 to 0.1892)	0.1608 (0.1199 to 0.2016)	0.0118 (–0.0455 to 0.069)	240,430.54	64.9	24.7	27.3	28.9
Weight < 1000 g (no)	39,500.58 (37,565.42 to 41,435.75)	42,731.41 (39,531.03 to 45,931.79)	–3230.83 (–6970.78 to 509.13)	0.0346 (0.0145 to 0.0547)	0.0398 (0.0186 to 0.0609)	0.0052 (–0.024 to 0.0344)	–625,625.58	66.1	95.8	95.4	94.8

a The difference in effects was inverted, i.e. negative values were given a positive sign, to reflect the fact that a reduction in adverse outcomes is a positive effect. *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a:
b Based on 1000 bootstrap replicates of the data set.
c GBP £20,000 cost-effectiveness threshold.
d GBP £30,000 cost-effectiveness threshold.

TABLE 66 Cost-effectiveness estimates based on composite secondary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
PiPS data set											
Colonisation (yes)	61,354.61 (57,739.13 to 64,970.09)	55,071.11 (49,211.64 to 60,930.58)	6283.5 (–601.64 to 13,168.63)	0.1408 (0.1098 to 0.1718)	0.1449 (0.0977 to 0.192)	0.0041 (–0.0524 to 0.0605)	1,542,695.3	60.1	4.9	6.5	8.2
Colonisation (no)	86,147.5 (72,490.12 to 99,804.87)	69,272.01 (64,240.99 to 74,303.03)	16,875.49 (2320.93 to 31,430.04)	0.4024 (0.2963 to 0.5086)	0.2011 (0.1599 to 0.2423)	–0.2013 (–0.3152 to –0.0875)	–83,817.078	0.0	0.1	0.0	0.0
Gestational age < 28 years	85,643.61 (79,893.44 to 91,393.79)	85,485.04 (79,268.07 to 91,702)	158.58 (–8309.9 to 8627.06)	0.3423 (0.2884 to 0.3962)	0.3651 (0.311 to 0.4193)	0.0228 (–0.0535 to 0.0992)	6940.0532	70.3	50	52.7	54.4
Gestational age ≥ 28 years	42,427.22 (39,409.85 to 45,444.6)	41,882.82 (39,068 to 44,697.65)	544.4 (–3582.07 to 4670.87)	0.0963 (0.0641 to 0.1285)	0.0838 (0.0541 to 0.1136)	–0.0124 (–0.0563 to 0.0314)	–43,758.509	28.5	39	35.7	34.2
Randomisation age ≤ 24 hours	60,552.92 (54,496.09 to 66,609.75)	59,907.42 (53,228.36 to 66,586.49)	645.5 (–8370.89 to 9661.88)	0.175 (0.1161 to 0.2339)	0.2275 (0.164 to 0.2911)	0.0525 (–0.0341 to 0.1392)	12,284.642	89.2	41.9	49.7	54.7
Randomisation age > 24 hours	64,119.34 (59,745.71 to 68,492.97)	63,634.34 (59,196.12 to 68,072.56)	485 (–5746.09 to 6716.08)	0.2283 (0.1899 to 0.2666)	0.2144 (0.1774 to 0.2515)	–0.0138 (–0.0672 to 0.0395)	–35,084.968	29	44.7	42.2	41.3

continued

TABLE 66 Cost-effectiveness estimates based on composite secondary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (male)	64,813.38 (60,104.75 to 69,522.02)	62,142.4 (57,119.77 to 67,165.02)	2670.99 (–4213.63 to 9555.61)	0.2301 (0.1861 to 0.2741)	0.2157 (0.173 to 0.2584)	–0.0144 (–0.0757 to 0.0468)	–185,133.34	30.6	21.1	21.0	20.3
Sex (female)	61,078.55 (55,495.49 to 66,661.62)	63,314.87 (57,795.57 to 68,834.17)	–2236.32 (–10,087.01 to 5614.37)	0.194 (0.1467 to 0.2414)	0.2206 (0.1722 to 0.2691)	0.0266 (–0.0412 to 0.0944)	–84,038.267	76.7	71.8	75.7	77.1
Weight < 1000 g (yes)	87,432.57 (81,654.37 to 93,210.78)	85,122.31 (78,960.42 to 91,284.21)	2310.26 (–6137.03 to 10,757.54)	0.3543 (0.3004 to 0.4083)	0.3666 (0.313 to 0.4201)	0.0123 (–0.0638 to 0.0883)	188,517.73	62	28.6	30.3	32.1
Weight < 1000 g (no)	40,184.67 (37,715.64 to 42,653.71)	41,294.42 (38,612.23 to 43,976.61)	–1109.74 (–4755.32 to 2535.84)	0.0818 (0.0516 to 0.1119)	0.0765 (0.0477 to 0.1053)	–0.0053 (–0.047 to 0.0364)	209,054.08	36.3	71.4	68.4	65.9

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
NNRD data set											
Colonisation (yes)	59,017.58 (55,765.98 to 62,269.17)	53,138.61 (47,903.19 to 58,374.03)	5878.97 (–284.03 to 12,041.96)	0.1408 (0.1098 to 0.1718)	0.1449 (0.0977 to 0.192)	0.0041 (–0.0524 to 0.0605)	1,443,377.1	52.2	2.9	4.7	5.7
Colonisation (no)	81,452.58 (72,626.84 to 90,278.33)	67,767.97 (62,941.77 to 72,594.18)	13,684.61 (3625.48 to 23,743.74)	0.4024 (0.2963 to 0.5086)	0.2011 (0.1599 to 0.2423)	–0.2013 (–0.3152 to –0.0875)	–67,968.65	0.0	0.7	0.2	0.2
Gestational age < 28 years	81,834.95 (77,038.62 to 86,631.28)	81,626.86 (75,675.75 to 87,577.97)	208.08 (–7435.24 to 7851.41)	0.3423 (0.2884 to 0.3962)	0.3651 (0.311 to 0.4193)	0.0228 (–0.0535 to 0.0992)	9106.6827	73.9	48.4	51.5	55.0
Gestational age ≥ 28 years	41,194.96 (38,659.84 to 43,730.09)	42,334.79 (39,428.07 to 45,241.5)	–1139.82 (–4996.75 to 2717.1)	0.0963 (0.0641 to 0.1285)	0.0838 (0.0541 to 0.1136)	–0.0124 (–0.0563 to 0.0314)	91,618.665	26.7	70.1	65.0	61.8
Randomisation age ≤ 24 hours	59,180.96 (53,370.47 to 64,991.46)	58,124.43 (51,774.27 to 64,474.58)	1056.54 (–7550.8 to 9663.88)	0.175 (0.1161 to 0.2339)	0.2275 (0.164 to 0.2911)	0.0525 (–0.0341 to 0.1392)	20,107.313	90.8	38.7	46.6	51.7

continued

TABLE 66 Cost-effectiveness estimates based on composite secondary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Randomisation age > 24 hours	61,266.6 (57,611.07 to 64,922.13)	62,096.83 (57,835.16 to 66,358.49)	–830.22 (–6444.9 to 4784.46)	0.2283 (0.1899 to 0.2666)	0.2144 (0.1774 to 0.2515)	–0.0138 (–0.0672 to 0.0395)	60,058.773	31.3	60.3	57.3	55.5
Sex (male)	63,399.49 (58,920.37 to 67,878.61)	61,730.88 (56,698.18 to 66,763.58)	1668.6 (–5068.65 to 8405.86)	0.2301 (0.1861 to 0.2741)	0.2157 (0.173 to 0.2584)	–0.0144 (–0.0757 to 0.0468)	–115,655.56	31.5	33.0	31.9	30.8
Sex (female)	57,220.04 (53,157.04 to 61,283.04)	60,200.92 (55,254.06 to 65,147.77)	–2980.88 (–9382.39 to 3420.63)	0.194 (0.1467 to 0.2414)	0.2206 (0.1722 to 0.2691)	0.0266 (–0.0412 to 0.0944)	–112,017.9	79.9	81.4	84.4	85.1
Weight < 1000 g (yes)	83,083.85 (78,193.05 to 87,974.64)	80,400.47 (74,646.03 to 86,154.91)	2683.38 (–4868.67 to 10,235.43)	0.3543 (0.3004 to 0.4083)	0.3666 (0.313 to 0.4201)	0.0123 (–0.0638 to 0.0883)	218,964.77	63.6	25.9	28.5	30.7
Weight < 1000 g (no)	39,497.7 (37,562.52 to 41,432.88)	42,660.06 (39,475.43 to 45,844.69)	–3162.35 (–6888.86 to 564.15)	0.0818 (0.0516 to 0.1119)	0.0765 (0.0477 to 0.1053)	–0.0053 (–0.047 to 0.0364)	595,725.68	42.7	95.5	93.9	92.2

Subgroup variable	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Combined data set											
Colonisation (yes)	59,082.53 (55,813.78 to 62,351.27)	53,064.15 (47,856.46 to 58,271.83)	6018.38 (–130.17 to 12,166.93)	0.1408 (0.1098 to 0.1718)	0.1449 (0.0977 to 0.192)	0.0041 (–0.0524 to 0.0605)	1,477,605.1	52.2	2.5	4.1	5.3
Colonisation (no)	82,009.26 (73,172.95 to 90,845.58)	67,938.27 (63,094.89 to 72,781.64)	14,071 (3994.36 to 24,147.64)	0.4024 (0.2963 to 0.5086)	0.2011 (0.1599 to 0.2423)	–0.2013 (–0.3152 to –0.0875)	–69,887.747	0.0	0.6	0.2	0.1
Gestational age < 28 years	82,110.69 (77,277.46 to 86,943.93)	81,796.8 (75,847.03 to 87,746.58)	313.89 (–7351.61 to 7979.4)	0.3423 (0.2884 to 0.3962)	0.3651 (0.311 to 0.4193)	0.0228 (–0.0535 to 0.0992)	13,737.319	73.9	46.9	51.1	53.8
Gestational age ≥ 28 years	41,205.5 (38,669.57 to 43,741.43)	42,375.67 (39,459.53 to 45,291.81)	–1170.17 (–5034.72 to 2694.39)	0.0963 (0.0641 to 0.1285)	0.0838 (0.0541 to 0.1136)	–0.0124 (–0.0563 to 0.0314)	94,057.689	26.7	71.0	65.2	61.6
Randomisation age ≤ 24 hours	59,476.77 (53,591.8 to 65,361.74)	58,194.19 (51,854.17 to 64,534.22)	1282.58 (–7367.79 to 9932.94)	0.175 (0.1161 to 0.2339)	0.2275 (0.164 to 0.2911)	0.0525 (–0.0341 to 0.1392)	24,409.129	90.8	37.0	45.4	49.9
Randomisation age > 24 hours	61,349.73 (57,679.36 to 65,020.1)	62,210.76 (57,941.56 to 66,479.97)	–861.04 (–6491.11 to 4769.04)	0.2283 (0.1899 to 0.2666)	0.2144 (0.1774 to 0.2515)	–0.0138 (–0.0672 to 0.0395)	62,287.872	31.3	60.6	58.0	56.1

continued

TABLE 66 Cost-effectiveness estimates based on composite secondary outcome; analyses by data source (PiPS, NNRD, combined) and prespecified trial population subgroup (*continued*)

	Mean costs (£) (95% CI)			Mean effects (95% CI)				Probability (%) *B. breve* BBG is			
Subgroup variable	*B. breve* BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]	ICER (£)	More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Sex (male)	63,431.42 (58,926.17 to 67,936.66)	61,714.56 (56,686.93 to 66,742.19)	1716.86 (–5034.01 to 8467.73)	0.2301 (0.1861 to 0.2741)	0.2157 (0.173 to 0.2584)	–0.0144 (–0.0757 to 0.0468)	–119,000.26	31.5	32.4	31.0	30.3
Sex (female)	57,497.39 (53,404.44 to 61,590.33)	60,454.11 (55,484.46 to 65,423.76)	–2956.72 (–9394.86 to 3481.42)	0.194 (0.1467 to 0.2414)	0.2206 (0.1722 to 0.2691)	0.0266 (–0.0412 to 0.0944)	–111,110.12	79.9	80.6	83.5	84.7
Weight < 1000 g (yes)	83,364.15 (78,439.38 to 88,288.91)	80,535.46 (74,783.56 to 86,287.36)	2828.68 (–4743.48 to 10,400.85)	0.3543 (0.3004 to 0.4083)	0.3666 (0.313 to 0.4201)	0.0123 (–0.0638 to 0.0883)	230,821.67	63.6	24.7	27.9	29.3
Weight < 1000 g (no)	39,500.58 (37,565.42 to 41,435.75)	42,731.41 (39,531.03 to 45,931.79)	–3230.83 (–6970.78 to 509.13)	0.0818 (0.0516 to 0.1119)	0.0765 (0.0477 to 0.1053)	–0.0053 (–0.047 to 0.0364)	608,624.88	42.7	95.8	94.4	92.8

a The difference in effects was inverted, i.e. negative values were given a positive sign, to reflect the fact that a reduction in adverse outcomes is a positive effect. *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a:
b Based on 1000 bootstrap replicates of the data set.
c GBP £20,000 cost-effectiveness threshold.
d GBP £30,000 cost-effectiveness threshold.

TABLE 67 Assessment of the agreement between the cost-effectiveness estimates from the different data sources for resource use or resource use and clinical outcomes based on incremental net benefit at a cost-effectiveness threshold of £20,000 per case avoided

Comparator data sets		Data source, mean net benefit (95% CI)		Agreement statistics		
PiPS vs. NNRD[a]	Outcome	PiPS	NNRD[a]	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	–468 (–5574 to 4638)	415 (–4259 to 5089)	882 (–1118 to 2883)	0.387	0.061
	Sepsis	–420 (–5776 to 4937)	463 (–4472 to 5397)	882 (–1118 to 2883)	0.387	0.064
	NEC	–347 (–5673 to 4979)	536 (–4383 to 5454)	882 (–1118 to 2883)	0.387	0.06
	Composite	–448 (–5860 to 4964)	434 (–4561 to 5429)	882 (–1118 to 2883)	0.387	0.058
				Agreement statistics		
Combined vs. PiPS	Outcome	Combined	PiPS	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	390 (–4269 to 5048)	–468 (–5574 to 4638)	–857 (–2858 to 1144)	0.401	0.047
	Sepsis	438 (–4481 to 5357)	–420 (–5776 to 4937)	–857 (–2858 to 1144)	0.401	0.044
	NEC	510 (–4393 to 5414)	–347 (–5673 to 4979)	–857 (–2858 to 1144)	0.401	0.049
	Composite	409 (–4571 to 5389)	–448 (–5860 to 4964)	–857 (–2858 to 1144)	0.401	0.049
				Agreement statistics		
Combined vs. NNRD[a]	Outcome	Combined	NNRD[a]	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	390 (–4269 to 5048)	415 (–4259 to 5089)	25 (–1976 to 1144)	0.401	0.039
	Sepsis	438 (–4481 to 5357)	463 (–4472 to 5397)	25 (–1976 to 1144)	0.401	0.048
	NEC	510 (–4393 to 5414)	536 (–4383 to 5454)	25 (–1976 to 1144)	0.401	0.045
	Composite	409 (–4571 to 5389)	434 (–4561 to 5429)	25 (–1976 to 1144)	0.401	0.049
				Agreement statistics		
PiPS vs. NNRD[e]	Outcome	PiPS	NNRD[e]	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	–468 (–5574 to 4638)	415 (–4259 to 5089)	882 (–1118 to 2883)	0.387	0.061
	Sepsis	–420 (–5776 to 4937)	301 (–4542 to 5145)	721 (–1382 to 2824)	0.502	0.056

a NNRD data set acted as source of resource use information.
b Mean difference (95% CIs) = difference between the mean bootstrap net benefit at a cost-effectiveness threshold of £20,000 per case avoided estimated from the two data sets.
c *p*-value is a two-sided probability that the difference between the mean incremental net benefits at a cost-effectiveness threshold of £20,000 per case avoided is greater or less than zero.
d Miscoverage probabilities = the proportion of bootstrap incremental net monetary benefit estimates for the comparator data source that fell outside the respective CI for the referent data source.
e NNRD acted as source of both resource use and clinical outcome information (note that there was for no information for NEC clinical outcome and hence the composite secondary outcome in the NNRD[e] data set).

TABLE 68 Assessment of the agreement between the cost-effectiveness estimates from the different data sources for resource use or resource use and clinical outcomes based on incremental net benefit at a cost-effectiveness threshold of £30,000 per case avoided

Comparator data sets		Data source, mean net benefit (95% CI)		Agreement statistics		
PiPS vs. NNRD[a]	Outcome	PiPS	NNRD[a]	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	–444 (–5571 to 4684)	439 (–4257 to 5134)	882 (–1118 to 2883)	0.387	0.057
	Sepsis	–372 (–5876 to 5133)	511 (–4577 to 5599)	882 (–1118 to 2883)	0.387	0.06
	NEC	–263 (–5719 to 5194)	620 (–4440 to 5680)	882 (–1118 to 2883)	0.387	0.059
	Composite	–415 (–6024 to 5194)	468 (–4733 to 5668)	882 (–1118 to 2883)	0.387	0.056
				Agreement statistics		
Combined vs. PiPS	Outcome	Combined	PiPS	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	413 (–4267 to 5094)	–444 (–5571 to 4684)	–857 (–2858 to 1144)	0.401	0.049
	Sepsis	486 (–4587 to 5558)	–372 (–5876 to 5133)	–857 (–2858 to 1144)	0.401	0.049
	NEC	595 (–4451 to 5640)	–263 (–5719 to 5194)	–857 (–2858 to 1144)	0.401	0.048
	Composite	443 (–4744 to 5629)	–415 (–6024 to 5194)	–857 (–2858 to 1144)	0.401	0.049
				Agreement statistics		
Combined vs. NNRD[a]	Outcome	Combined	NNRD[a]	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	413 (–4267 to 5094)	439 (–4257 to 5134)	25 (–1976 to 1144)	0.401	0.043
	Sepsis	486 (–4587 to 5558)	511 (–4577 to 5599)	25 (–1976 to 1144)	0.401	0.046
	NEC	595 (–4451 to 5640)	620 (–4440 to 5680)	25 (–1976 to 1144)	0.401	0.042
	Composite	443 (–4744 to 5629)	468 (–4733 to 5668)	25 (–1976 to 1144)	0.401	0.052
				Agreement statistics		
PiPS vs. NNRD[e]	Outcome	Combined	NNRD[a]	Mean difference (95% CI)[b]	*p*-value[c]	% miscoverage[d]
	Death	–444 (–5571 to 4684)	439 (–4257 to 5134)	882 (–1118 to 2883)	0.387	0.057
	Sepsis	–372 (–5876 to 5133)	268 (–4671 to 5207)	640 (–1573 to 2853)	0.571	0.051

a NNRD data set acted as source of resource use information.
b Mean difference (95% CIs) refers to the difference between the mean bootstrap net benefit at a cost-effectiveness threshold of £30,000 per case avoided estimated from the two data sets.
c *p*-value is a two-sided probability that the difference between the mean incremental net benefits at a cost-effectiveness threshold of £30,000 per case avoided is greater or less than zero.
d Miscoverage probabilities = the proportion of bootstrap incremental net monetary benefit estimates for the comparator data source that fell outside the respective CI for the referent data source.
e NNRD acted as source of both resource use and clinical outcome information (note that there was no information for NEC clinical outcome and hence the composite secondary outcome in the NNRD[e] data set).

TABLE 69 Sensitivity analysis for cost-effectiveness estimates for probiotic by clinical outcome (death, sepsis); NNRD data source for resource use

NNRD clinical outcome	Mean costs (£) (95% CI)			Mean effects (95% CI)			ICER (£)	Probability (%) *B. breve* BBG is			
	B. breve BBG	Placebo	Difference	*B. breve* BBG	Placebo	Difference[a]		More effective[b]	Less costly[b]	Cost-effective[b,c]	Cost-effective[b,d]
Death[e]	60,559.76 (57,480.38 to 63,639.13)	60,926.82 (57,388.7 to 64,464.94)	–367.07 (–5057.57 to 4323.44)	0.0823 (0.0606 to 0.1039)	0.0846 (0.063 to 0.1062)	0.0024 (–0.0282 to 0.0329)	–154,137.2 (SE)	60.8	57.8	59.4	59.9
Sepsis[f]	60,559.76 (57,480.38 to 63,639.13)	60,926.82 (57,388.7 to 64,464.94)	–367.07 (–5057.57 to 4323.44)	0.0613 (0.0424 to 0.0802)	0.058 (0.0399 to 0.0761)	–0.0033 (–0.0295 to 0.0229)	111,347.5 (SW)	39.3	57.8	56.9	56.1

a The difference in effects was inverted, i.e. negative values were given a positive sign, to reflect the fact that a reduction in adverse outcomes is a positive effect. *B. breve* BBG was considered to be 'cost-effective' if it had positive net benefit at a:
b Based on 1000 bootstrap replicates of the data set.
c GBP £20,000 cost-effectiveness threshold.
d GBP £30,000 cost-effectiveness threshold.
e Death before discharge home – includes three infants who remain on paediatric wards and are analysed as survivors.
f Sepsis is defined as blood stream infection with non-skin commensals after 72 hours postnatal age and before 46 weeks' postmenstrual age.

Chapter 8

TABLE 70 Participating mothers' ethnicity in comparison with the NNRD

Mothers' ethnicity	Participating mothers, *n* (%) (*n* = 930)	NNRD, *n* (%)	
		28 sites (*n* = 10,983)	All neonatal units (*n* = 55,731)
Do not wish to answer	19 (2%)	–	1 (0.0%)
White: British	757 (81.4%)	6202 (56.5%)	36,409 (65.3%)
White: Irish	6 (0.6%)	38 (0.3%)	253 (0.5%)
White: Gypsy or Irish traveller	3 (0.3%)	0 (0%)	0 (0%)
White: any other white background	29 (3.1%)	532 (4.8%)	4164 (7.5%)
White and black Caribbean	15 (1.6%)	43 (0.4%)	246 (0.4%)
White and black African	2 (0.2%)	30 (0.3%)	101 (0.2%)
White and Asian	6 (0.6%)	34 (0.3%)	129 (0.2%)
Any other mixed background	3 (0.3%)	33 (0.3%)	209 (0.4%)
Asian/Asian British: Indian	18 (1.9%)	247 (2.2%)	1978 (3.5%)
Asian/Asian British: Pakistani	30 (3.2%)	857 (7.8%)	2337 (4.2%)
Asian/Asian British: Bangladeshi	7 (0.8%)	293 (2.7%)	821 (1.5%)
Asian/Asian British: Chinese	1 (0.1%)	46 (0.4%)	249 (0.4%)
Any other Asian background	5 (0.5%)	184 (1.7%)	1317 (2.4%)
Black: African	7 (0.8%)	308 (2.8%)	2157 (3.9%)
Black: Caribbean	13 (1.4%)	71 (0.6%)	749 (1.3%)
Any other black background	1 (0.1%)	53 (0.5%)	340 (0.6%)
Arab	2 (0.2%)	–	–
Any other	2 (0.2%)	158 (1.4%)	816 (1.5%)
Not stated	–	1479 (13.5%)	2120 (3.8%)
Missing data	4 (0.4%)	375 (3.4%)	1335 (2.4%)
Total	930	10,983	55,731

TABLE 71 Highest level of qualification by willingness for de-identified data to be shared

Highest level of qualification	Willingness, *n* (%)			
	Yes	Possibly	No	Total, *n*
O levels/GCSEs	191 (80.6%)	34 (14.3%)	12 (5.1%)	237
A levels/vocational qualification	319 (81%)	54 (13.7%)	21 (5.3%)	394
Degree/higher degree	331 (87.3%)	39 (10.3%)	9 (2.4%)	379
Total	841 (83.3%)	127 (12.6%)	42 (4.2%)	1010

TABLE 72 Awareness of electronic health records prior to the study by qualification groups

	Awareness, *n* (%)		
Highest level of qualification	Yes	No	Total, *n*
O levels/GCSEs	77 (33.8%)	151 (66.2%)	228
A levels/vocational qualification	185 (47.6%)	204 (52.4%)	389
Degree/higher degree	281 (76.6%)	86 (23.4%)	367
Total	543 (55.2%)	441 (44.8%)	984

TABLE 73 Comparison between the willingness of those with only one child and those with more than one child for their baby's data to be used for research purposes

	Willingness, *n* (%)			
Group	Yes	Possibly	No	Total, *n*
More than one child	496 (76.7%)	118 (18.2%)	33 (5.1%)	647
Only one child	395 (68.8%)	135 (23.7%)	43 (7.5%)	570
Total	888 (73%)	253 (20.8%)	76 (6.2%)	1217

TABLE 74 Comparison between the willingness of those with one child and those with more than one child for de-identified data about their baby to be used for research

	Willingness, *n* (%)			
Group	Yes	Possibly	No	Total, *n*
More than one child	555 (85.1%)	71 (10.9%)	26 (4%)	652
Only one child	463 (80.2%)	85 (14.7%)	29 (5%)	577
Total	1018 (82.8%)	156 (12.7%)	55 (4.5%)	1229

TABLE 75 Acceptability of an opt-out system by highest level of qualification

	Acceptability, *n* (%)			
Highest level of qualification	Yes	Possibly	No	Total, *n*
O levels/GCSEs	154 (66.4%)	46 (19.8%)	32 (13.8%)	232
A levels/vocational qualification	268 (70%)	51 (13.3%)	64 (16.7%)	383
Degree/higher degree	217 (59.3%)	57 (15.6%)	92 (25.1%)	366
Total	639 (65.1%)	154 (15.7%)	188 (19.2%)	981

TABLE 76 Association between level of care experienced and expressed preference for how to be asked if specific permission for data-sharing is requested

	Expressed preference, *n* (%)		
Level of care[a]	In person	In writing	Total, *n*
1	88 (28.9%)	216 (71.1%)	304
2	54 (34.6%)	102 (65.4%)	156
3	175 (39.5%)	268 (60.5%)	443
Total	317 (35.1%)	586 (64.9%)	903

a 'Level of care' refers to highest level of care experienced at any point in their baby's care.

TABLE 77 Relationship between highest level of qualification and willingness to share data is influenced by being asked by someone directly involved in the baby's care

	Willingness, *n* (%)			
Highest level of qualification	Yes	Possibly	No	Total, *n*
O levels/GCSEs	143 (62.2%)	58 (25.2%)	29 (12.6%)	230
A levels/vocational qualification	215 (56%)	102 (26.6%)	67 (17.4%)	384
Degree/higher degree	190 (53.1%)	97 (27.1%)	71 (19.8%)	358
Total	548 (56.4%)	257 (26.4%)	167 (17.2%)	972

Appendix 2 Supplementary figures

Chapter 3

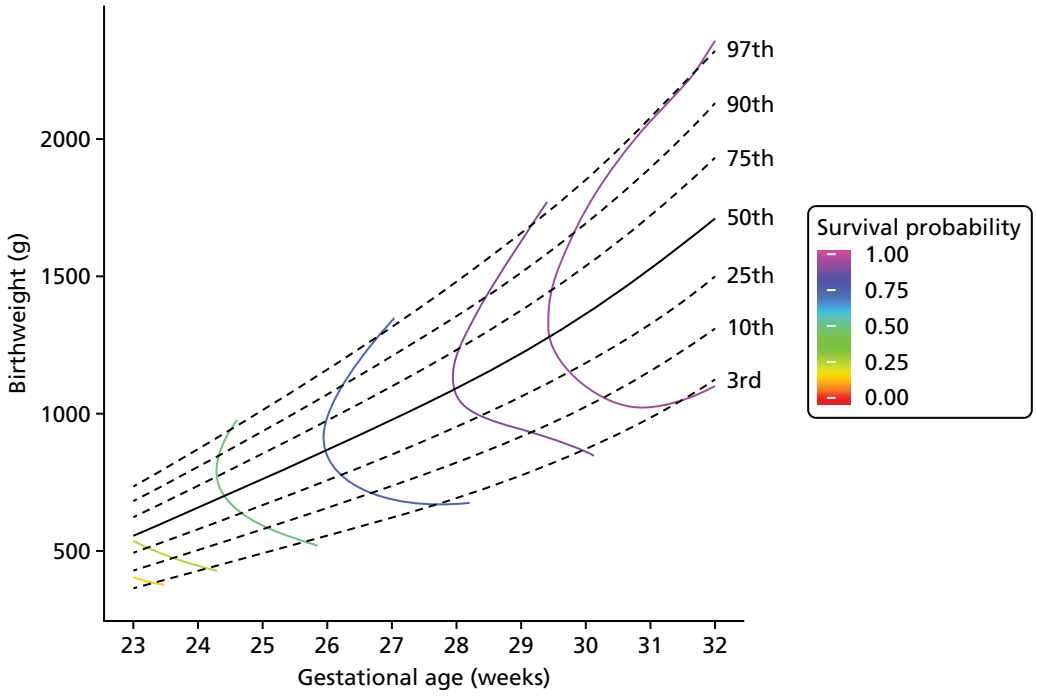


FIGURE 28 Isosurv plots for survival prediction: survival probability and birthweight centiles (singleton birth girls, antenatal steroids not received).

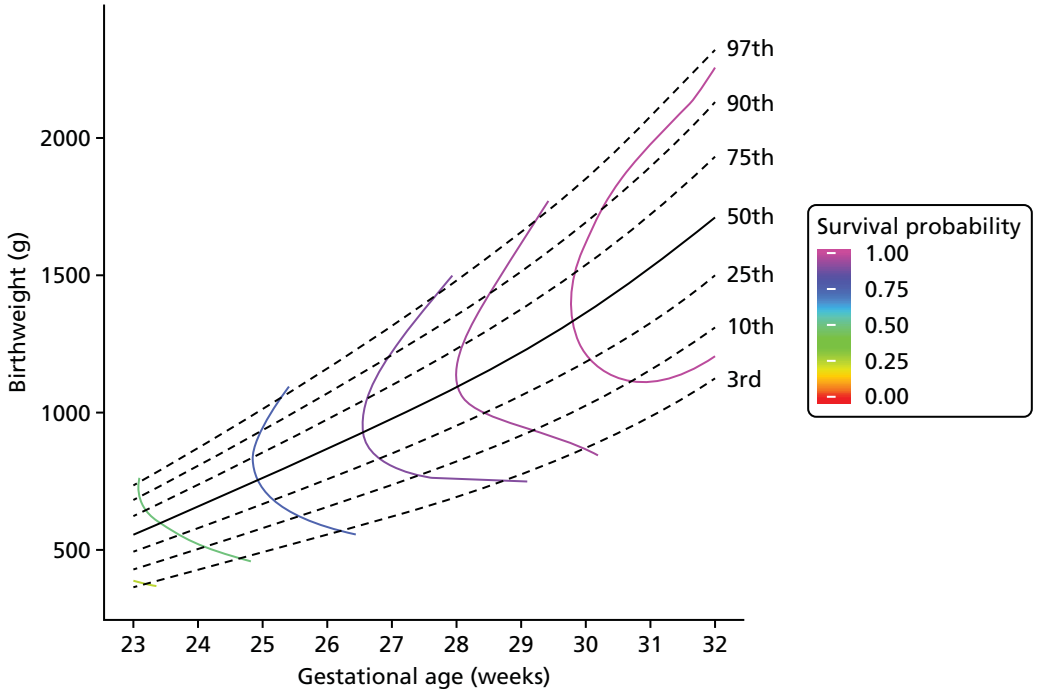


FIGURE 29 Isosurv plots for survival prediction: survival probability and birthweight centiles (singleton birth girls, antenatal steroids received).

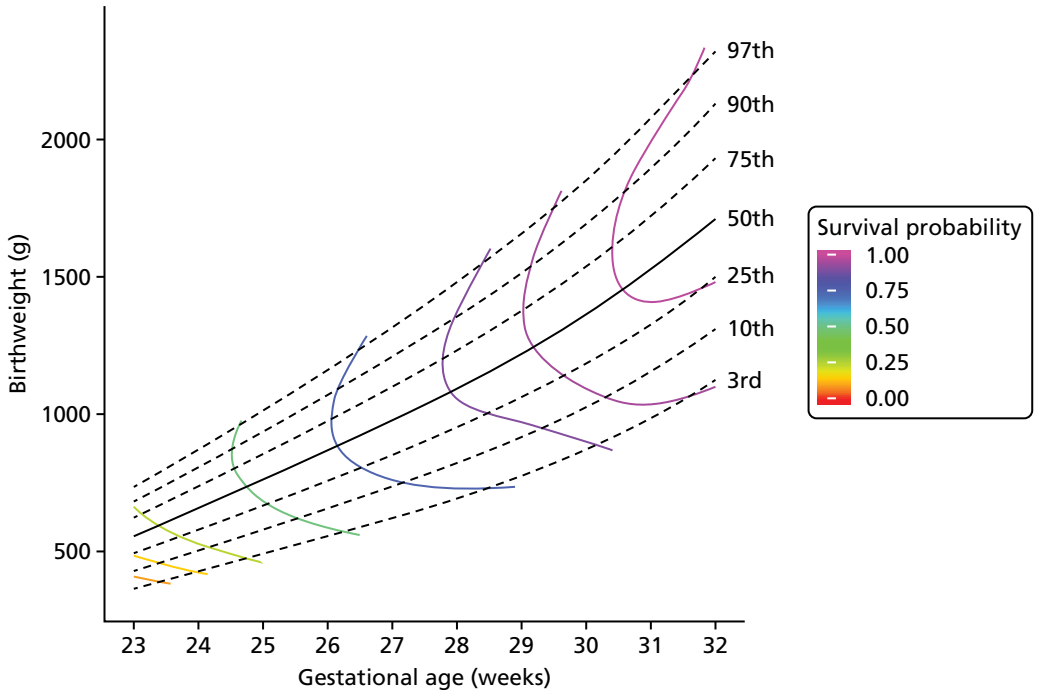


FIGURE 30 Isosurv plots for survival prediction: survival probability and birthweight centiles (multiple birth girls, antenatal steroids not received).

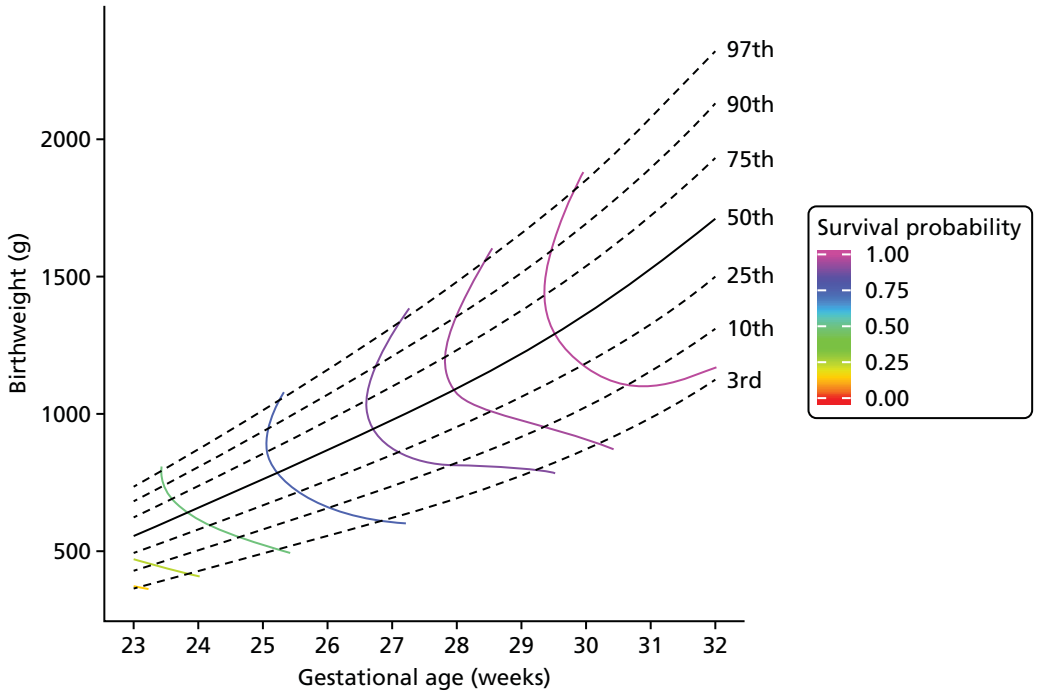


FIGURE 31 Isosurv plots for survival prediction: survival probability and birthweight centiles (multiple birth girls, antenatal steroids received).

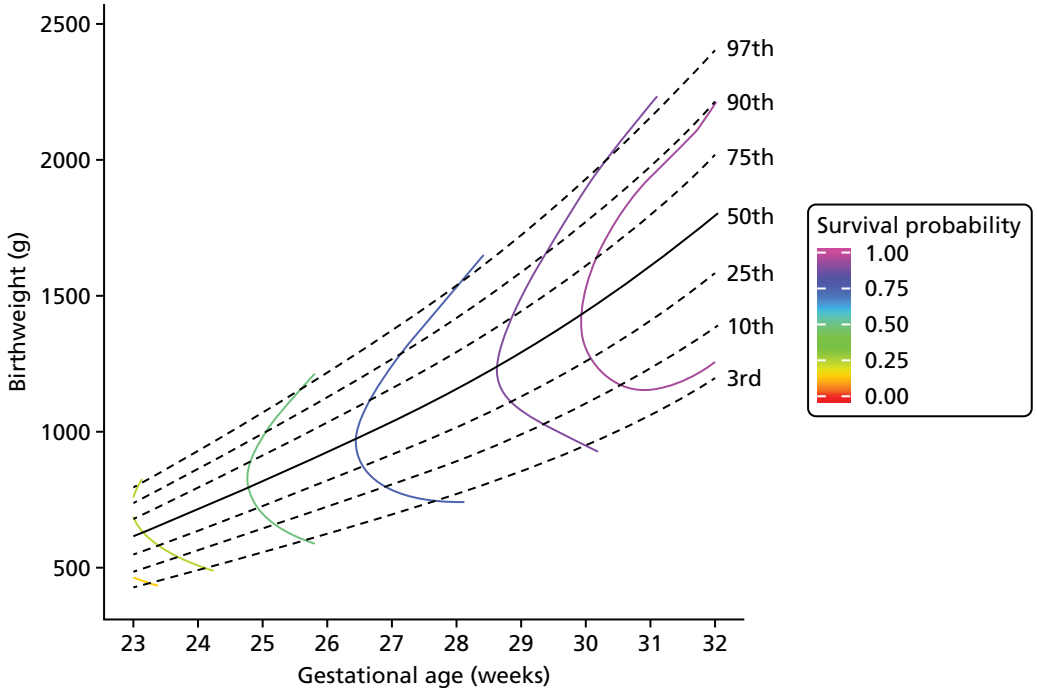


FIGURE 32 Isosurv plots for survival prediction: survival probability and birthweight centiles (singleton birth boys, antenatal steroids not received).

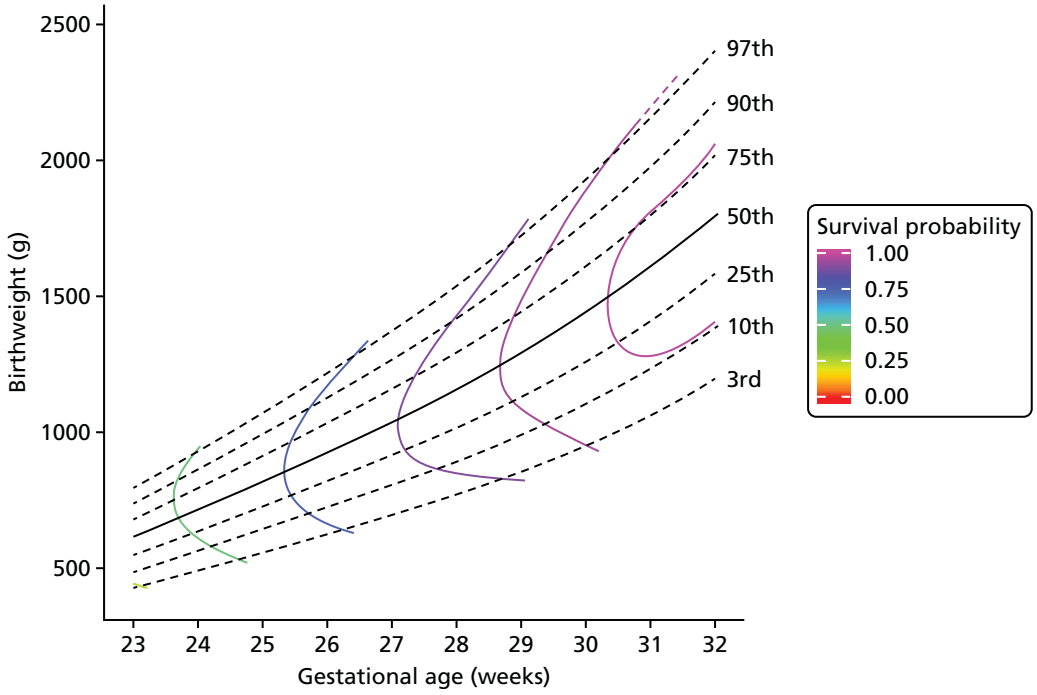


FIGURE 33 Isosurv plots for survival prediction: survival probability and birthweight centiles (singleton birth boys, antenatal steroids received).

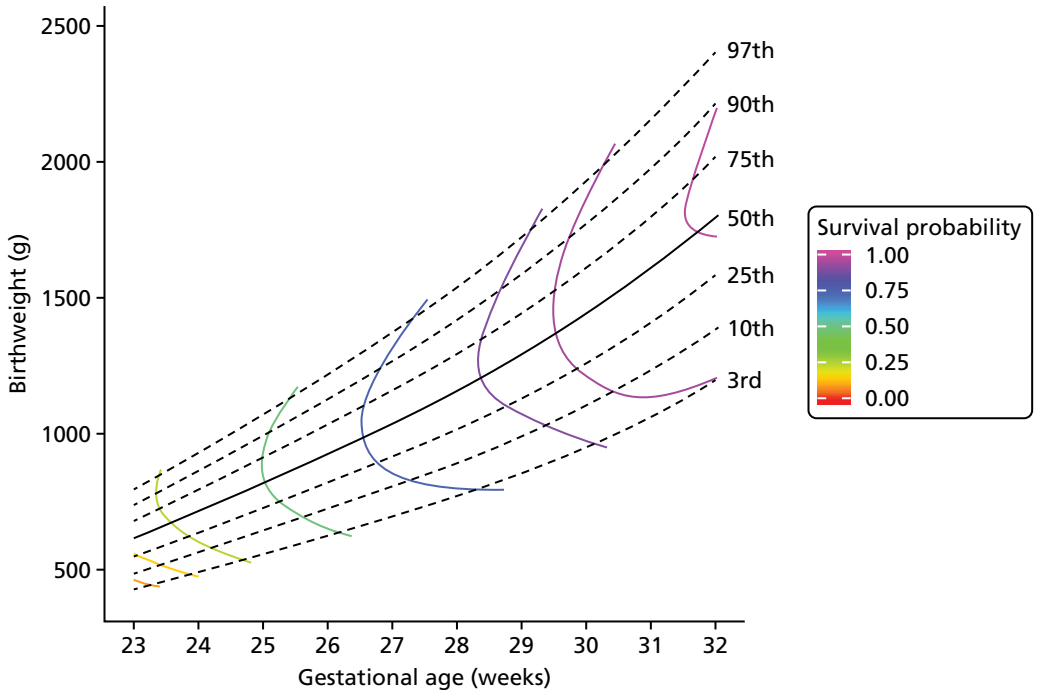


FIGURE 34 Isosurv plots for survival prediction: survival probability and birthweight centiles (multiple birth boys, antenatal steroids not received).

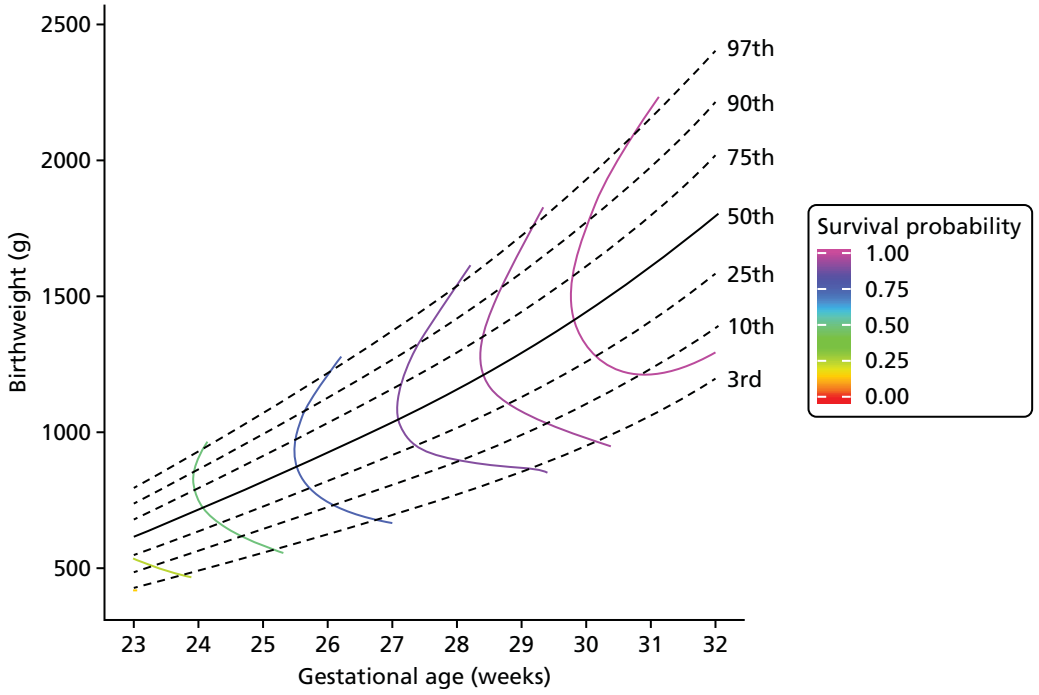


FIGURE 35 Isosurv plots for survival prediction: survival probability and birthweight centiles (multiple birth boys, antenatal steroids received).

Chapter 5

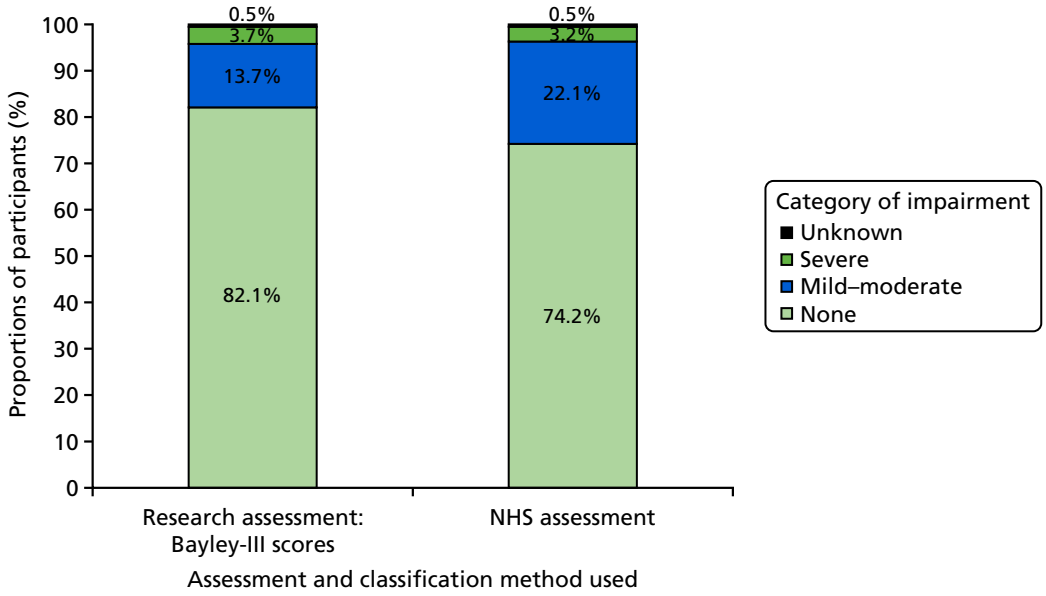


FIGURE 36 Classification of the severity of cognitive impairment of the children by research and NHS assessments.

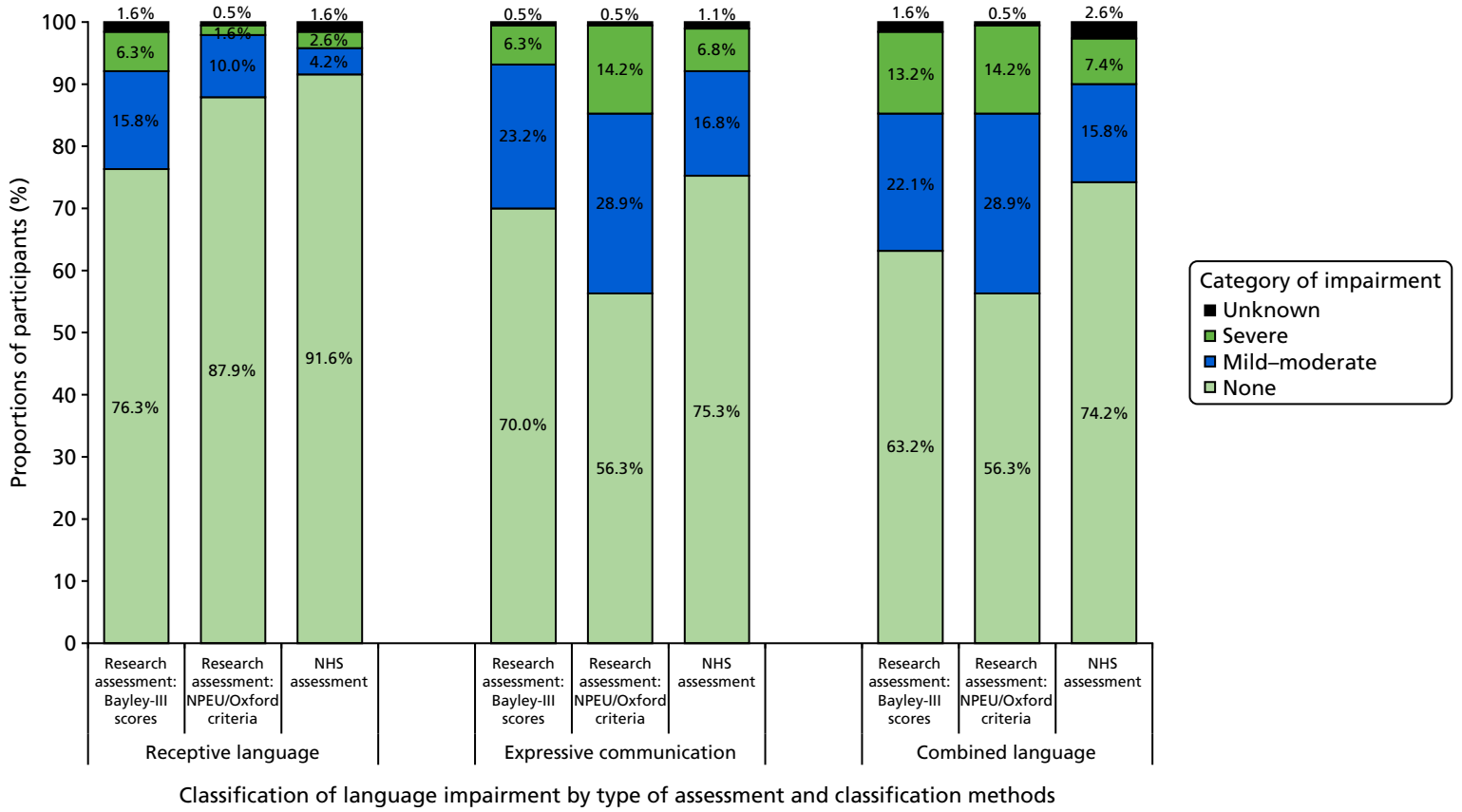


FIGURE 37 Classification of the severity of receptive communication, expressive communication and overall language impairment of the participants based on the Bayley-III and modified NPEU/Oxford classification by research assessment, and by NHS assessments.

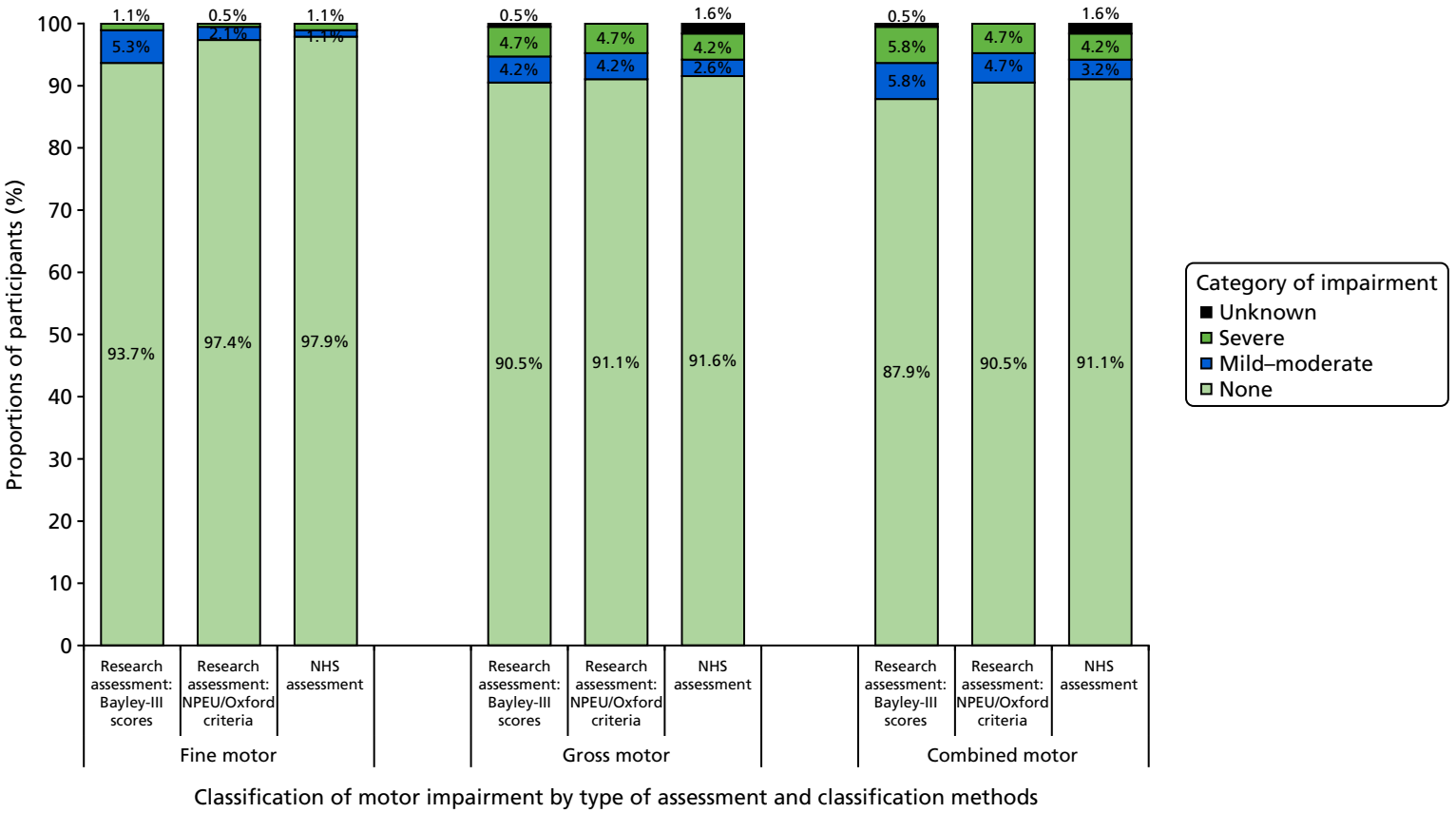


FIGURE 38 Classification of the severity of fine motor, gross motor and overall motor impairment of the participants based on the Bayley-III and modified NPEU/Oxford classification by research assessment, and by NHS assessments.

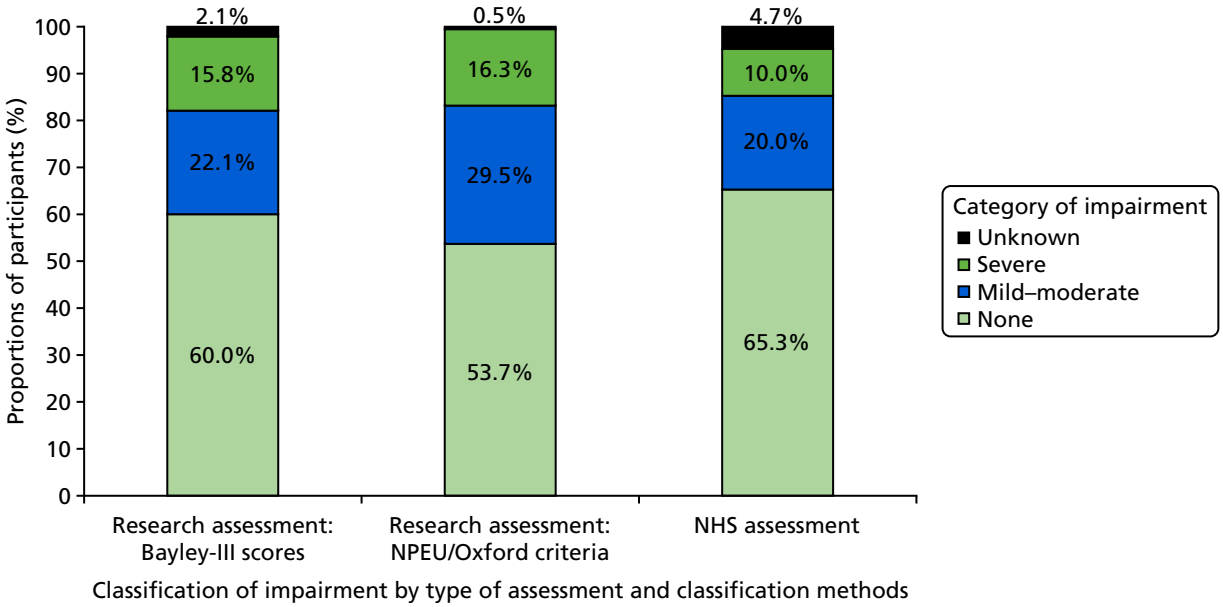


FIGURE 39 Classification of the neurodevelopmental outcome of participants by the severity of the worst impairment in the cognitive, language and motor domains through research and NHS assessments.

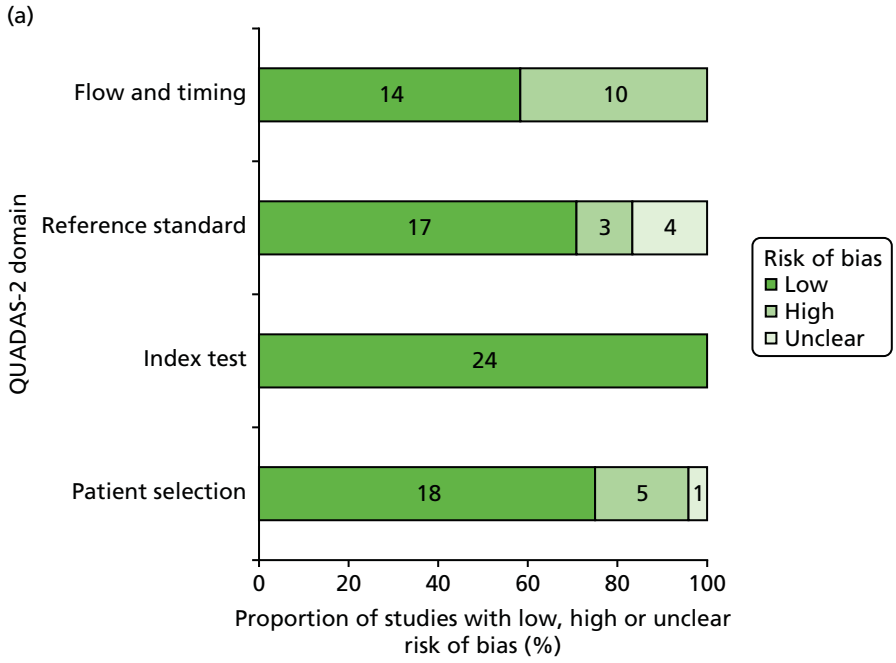


FIGURE 40 Proportions of studies with low, high or unclear risk of bias and concerns regarding applicability. Bar charts are annotated with the number of studies in each category. (*continued*)

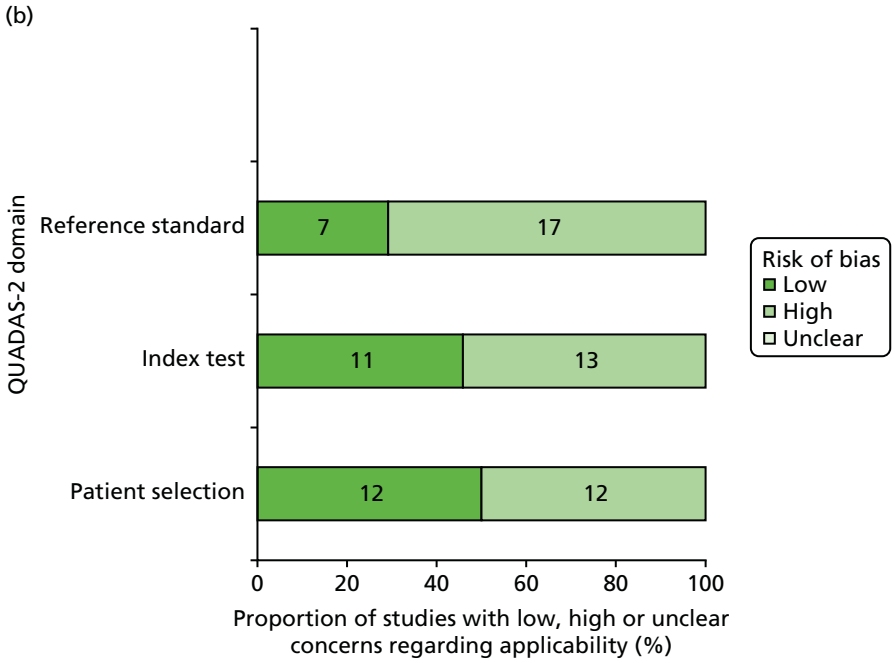


FIGURE 40 Proportions of studies with low, high or unclear risk of bias and concerns regarding applicability. Bar charts are annotated with the number of studies in each category.

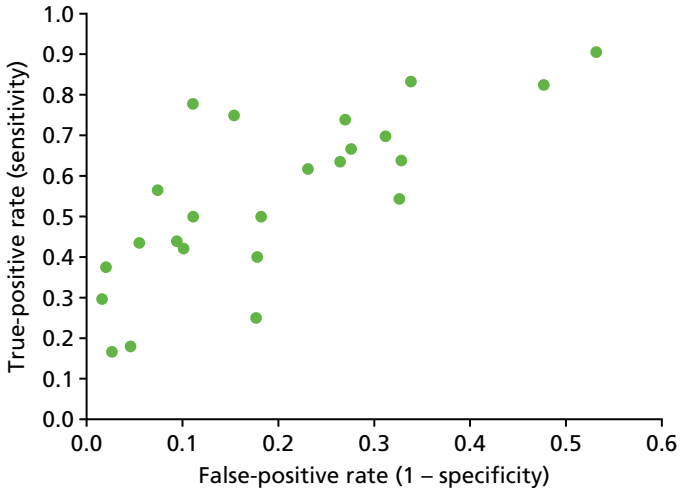


FIGURE 41 Scatterplot of the true-positive rate (sensitivity) against the false-positive rate (1 – specificity). The true-positive rate and the false-positive rate are expressed as proportions; each marker represents the result from one study. Spearman's rho –0.76; $p < 0.001$.

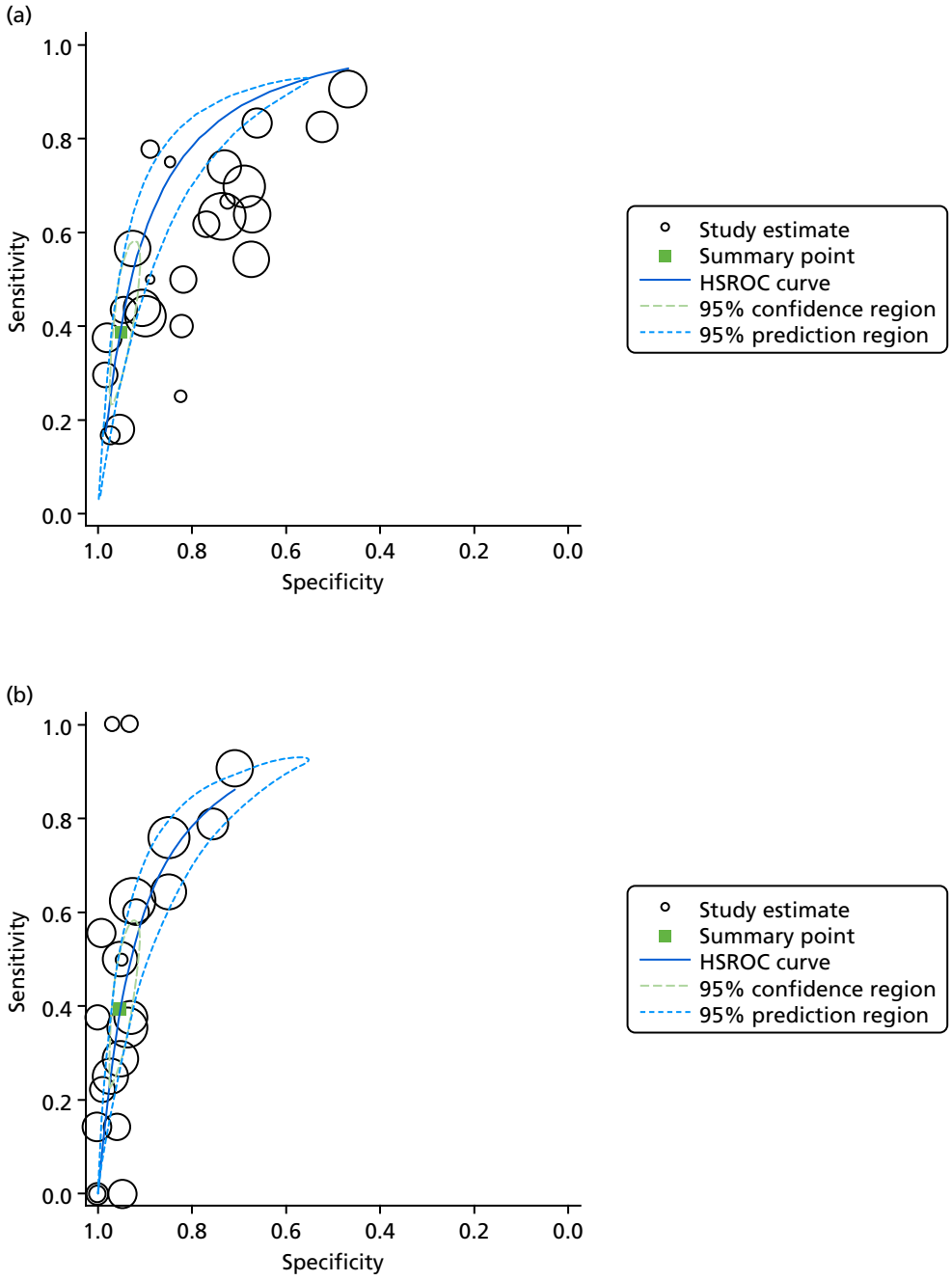


FIGURE 42 Hierarchical summary receiver operator characteristic (HSROC) curves for the pooled sensitivity and specificity of early developmental assessment in identifying (a) any impairment and (b) severe impairment. Each marker display the study estimates from one included study and is scaled according to the sample size of the study.

Chapter 6

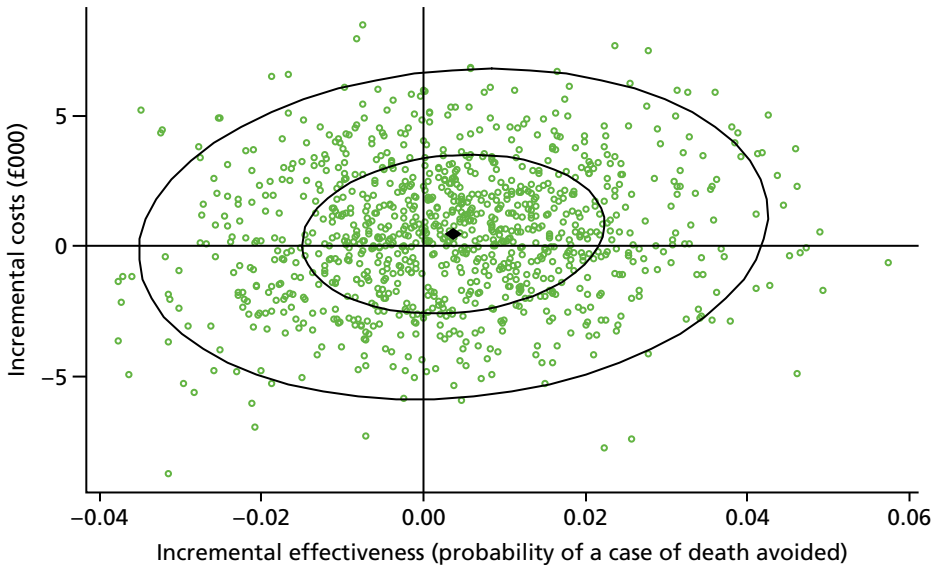


FIGURE 43 Cost-effectiveness plane: death as primary outcome – PiPS data.

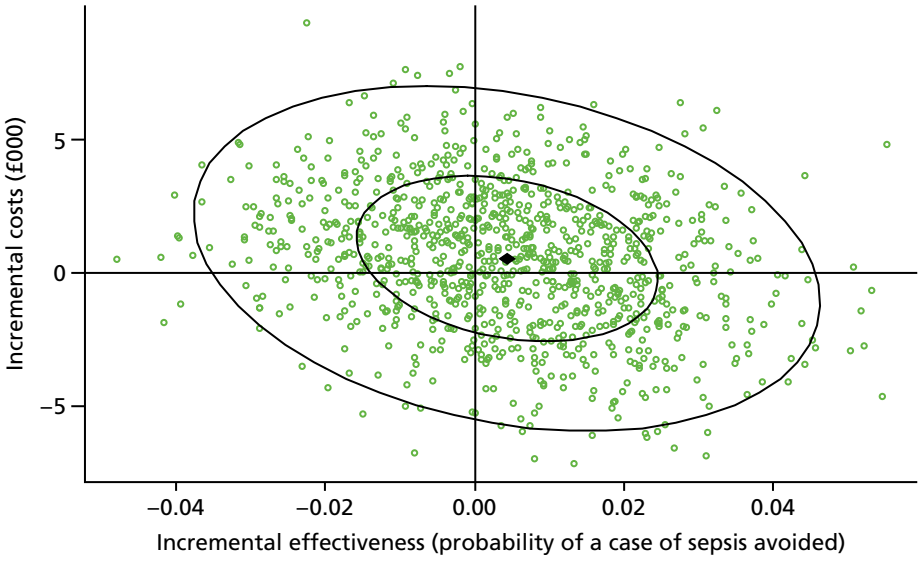


FIGURE 44 Cost-effectiveness plane: sepsis as primary outcome – PiPS data.

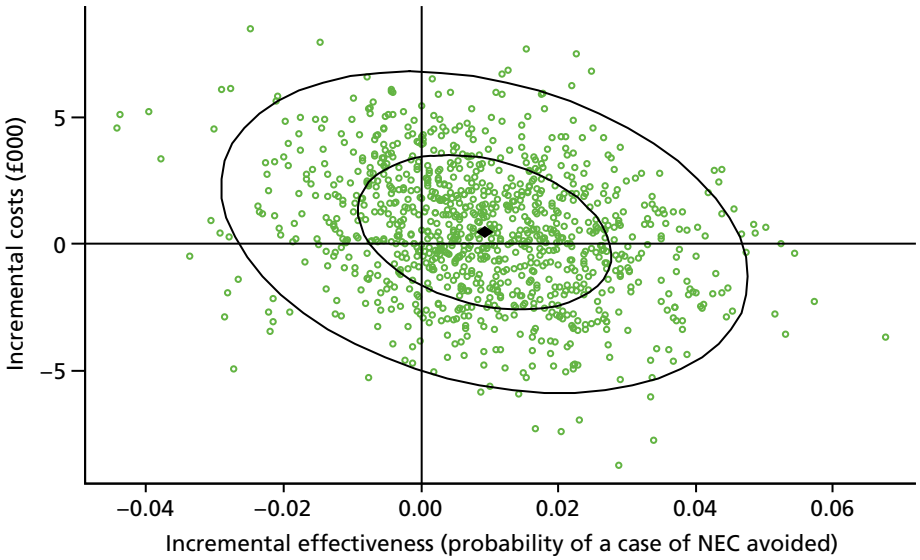


FIGURE 45 Cost-effectiveness plane: NEC as primary outcome – PiPS data.

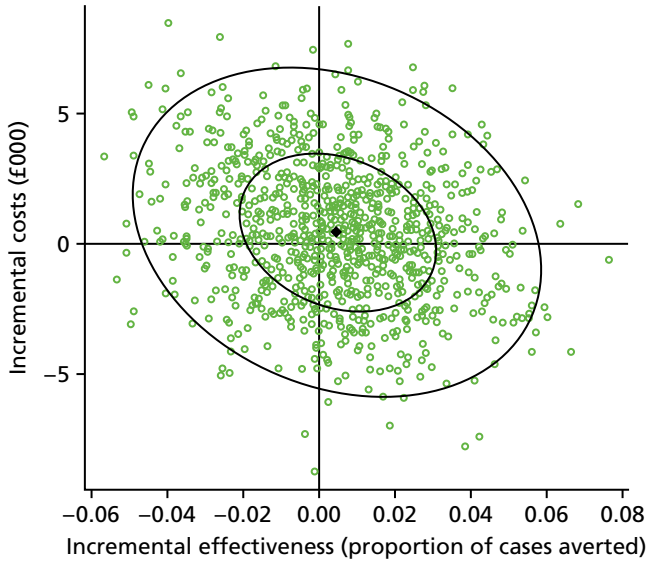


FIGURE 46 Cost-effectiveness plane: composite secondary outcome – PiPS data.

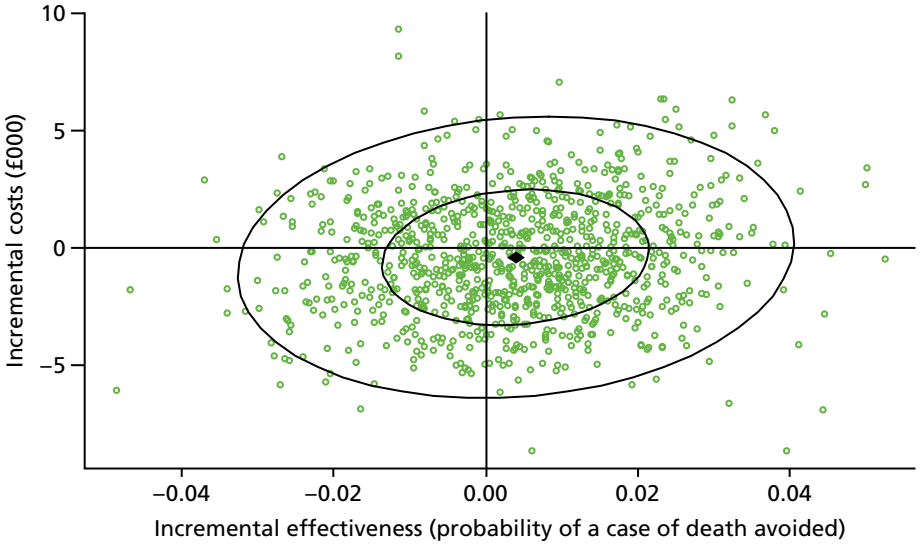


FIGURE 47 Cost-effectiveness plane: death as primary outcome – NNRD data.

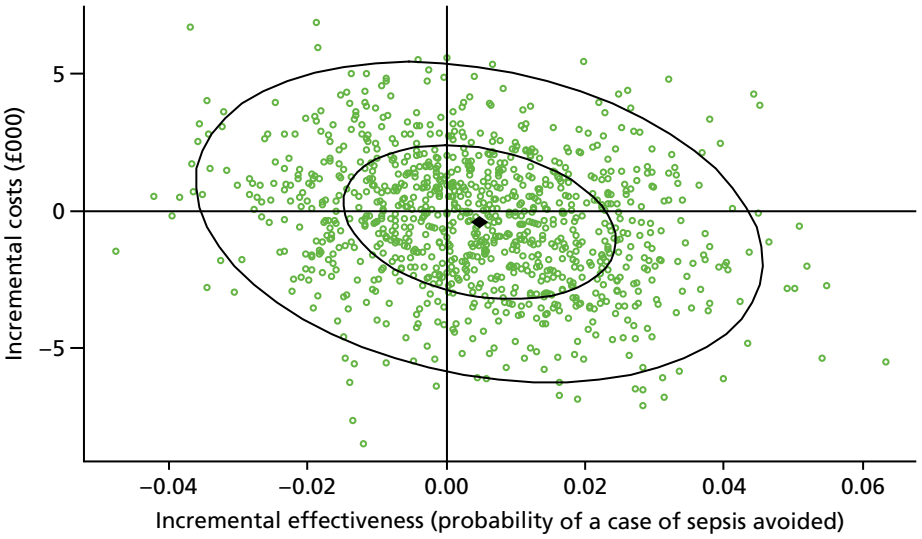


FIGURE 48 Cost-effectiveness plane: sepsis as primary outcome – NNRD data.

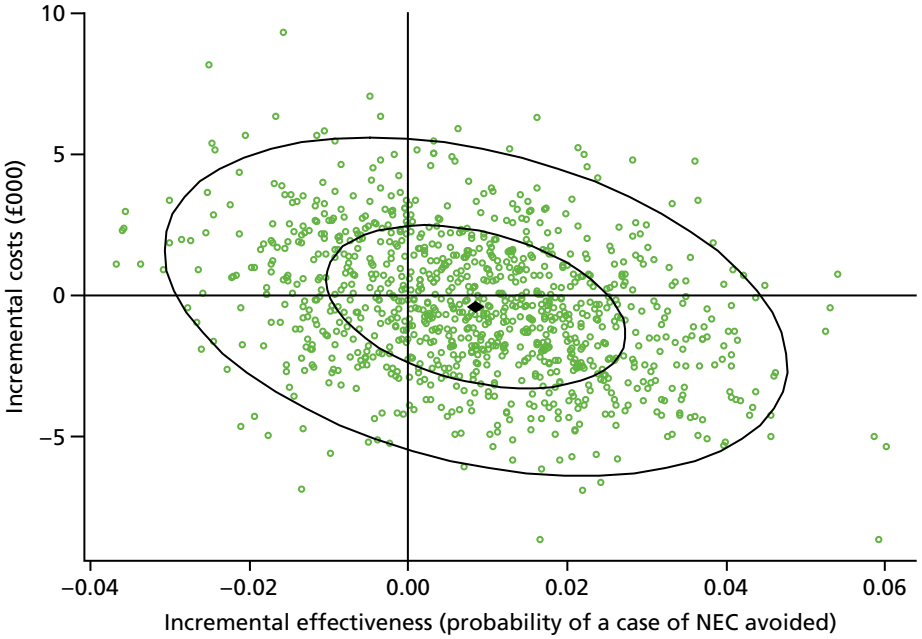


FIGURE 49 Cost-effectiveness plane: NEC as primary outcome – NNRD data.

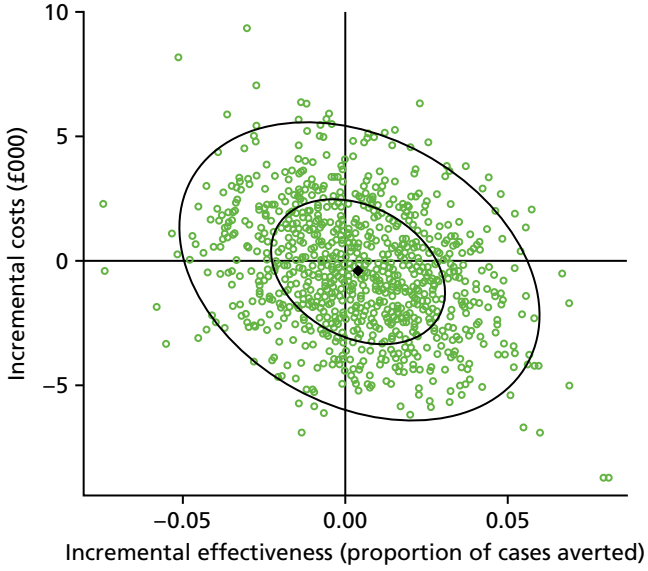


FIGURE 50 Cost-effectiveness plane: composite secondary outcome – NNRD data.

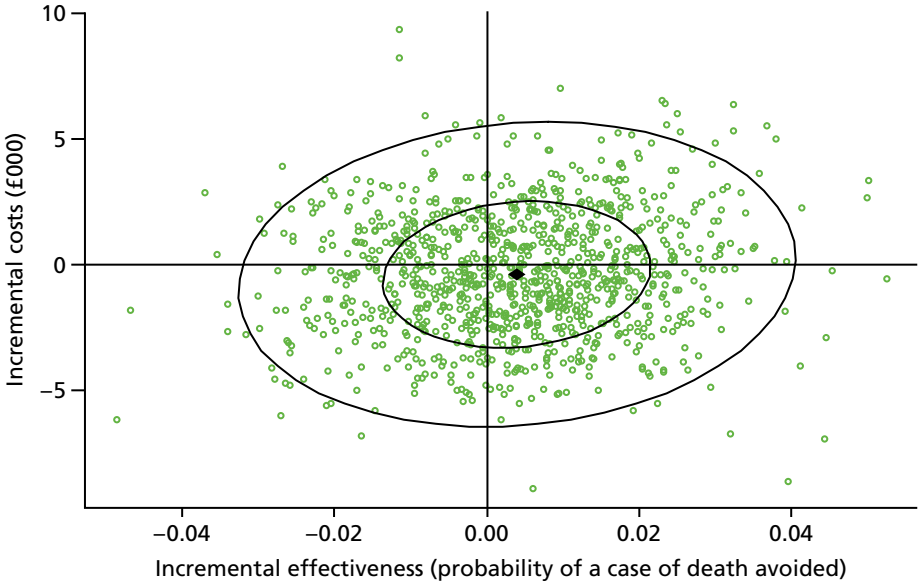


FIGURE 51 Cost-effectiveness plane: death as primary outcome – combined data.

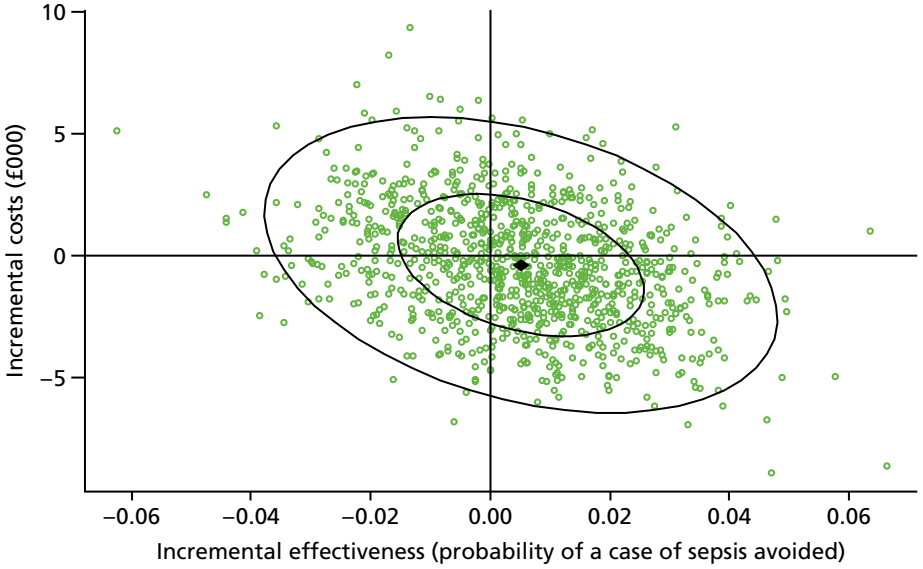


FIGURE 52 Cost-effectiveness plane: sepsis as primary outcome – combined data.

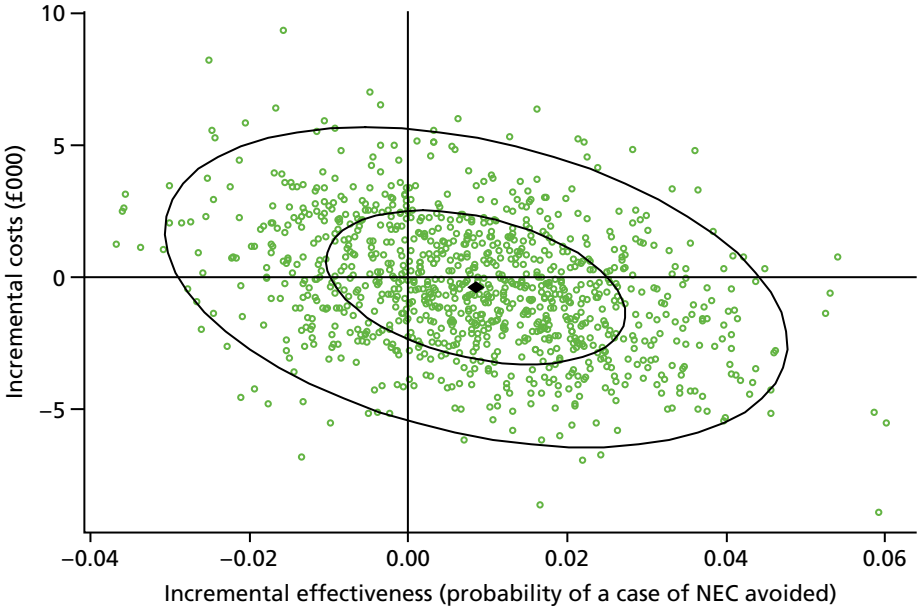


FIGURE 53 Cost-effectiveness plane: NEC as primary outcome – combined data.

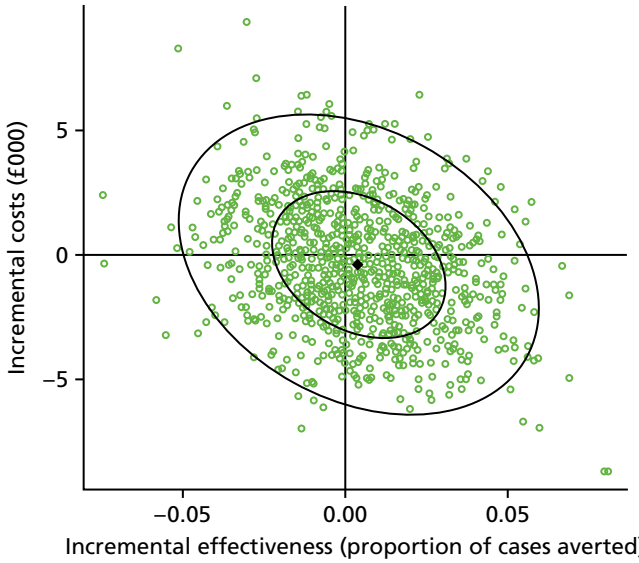


FIGURE 54 Cost-effectiveness plane: composite secondary outcome – combined data.

Chapter 8

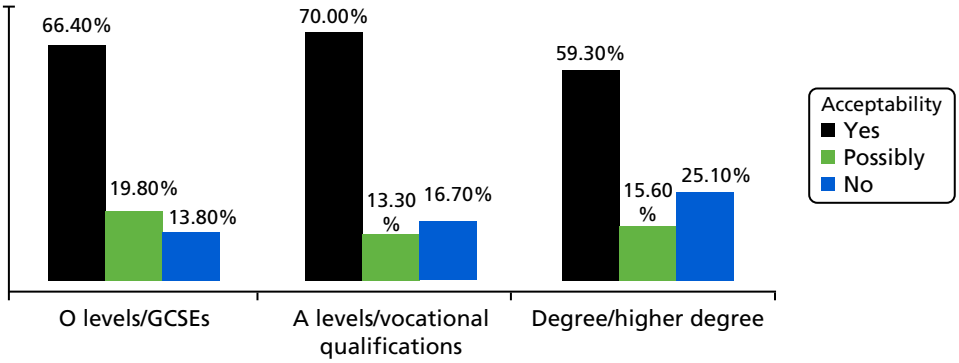


FIGURE 55 Acceptability of opt-out system by highest educational qualification. Acceptability of an opt-out system to use routinely collected health data for research purposes by highest educational qualification.

Appendix 3 Neonatal Data Set ISB1595 release 1 version 22

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Demographics and birth information (baby)	National identification baby	NationalIDBaby	R	NHS NUMBER (BABY)	n10	England and Wales – NHS number format	Baby's unique national identifier in a neonatal episode	Used to identify infant
NNUEpisodes	Demographics and birth information (baby)	National identification baby	NationalIDBaby (ENCRYPTED)	R	n/a	n10	England and Wales – NHS number format ENCRYPTED	Baby's unique national identifier in a neonatal episode encrypted using MD5	Used to identify infant
NNUEpisodes	Demographics and birth information (baby)	National identification baby status	NationalIDBabyStatus	M	NHS NUMBER STATUS INDICATOR CODE (BABY)	n2	01 number present and verified 02 number present but not traced 03 trace required 04 trace attempted – no match or multiple match found 05 trace needs to be resolved – (NHS number or PATIENT detail conflict) 06 trace in progress 07 number not present and trace not required 08 trace postponed (baby < 6 weeks old)	Whether or not the NHS number of the baby has been verified. Can be derived from patient demographic system by system provider	Used to identify whether the baby's NHS number has been verified by the NHS care records service
NNUEpisodes	Demographics and birth information (baby)	Scotland CHI number	BabyCHINumber	R	COMMUNITY HEALTH INDEX NUMBER (BABY)	n10	Scotland – CHI number format	Baby's unique national identifier in a neonatal episode	Used to identify infant
NNUEpisodes	Demographics and birth information (baby)	Scotland CHI number	BabyCHINumber (ENCRYPTED)	R	n/a	n10	Scotland – CHI Number format ENCRYPTED	Baby's unique national identifier in a neonatal episode encrypted using MD5	Used to identify infant
NNUEpisodes	Demographics and birth information (baby)	Northern Ireland H&C number	BabyHCNumber	R	HEALTH AND CARE NUMBER (BABY)	n10	Northern Ireland – H&C number format	Baby's unique national identifier in a neonatal episode	Used to identify infant
NNUEpisodes	Demographics and birth information (baby)	Northern Ireland H&C number	BabyHCNumber (ENCRYPTED))	R	n/a	n10	Northern Ireland – H&C number format ENCRYPTED	Baby's unique national identifier in a neonatal episode encrypted using MD5	Used to identify infant

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Demographics and birth information (baby)	Unique system identification	AnonPatientID	M	BABY LOCAL PATIENT IDENTIFIER (NATIONAL NEONATAL DATA SET)	an50	ID utilises capital letters of the English alphabet only	A unique ID that will only identify the baby if used by a user with permission to see the record of that baby	Used to identify infant in locations of care for the purpose of communication where patient identifiers are not included
NNUEpisodes	Demographics and birth information (baby)	Date and time of birth	DateTimeofBirth	M	DATE TIME OF BIRTH (BABY)	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The calendar date and time of birth of the baby in co-ordinated universal time	Used for calculating anonymised times in minutes, identify verification of baby, and secondary data linkages
NNUEpisodes	Demographics and birth information (baby)	Place of birth NHS code (location of baby's birth)	PlaceofBirthNHSCode	M	SITE CODE (OF ACTUAL PLACE OF DELIVERY) or ORGANISATION CODE (OF ACTUAL PLACE OF DELIVERY)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home) ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)	Place at which the birth took place as recorded	Used to conduct data analysis on the organisation
NNUEpisodes	Demographics and birth information (baby)	Birthweight (g)	Birthweight	R	BIRTHWEIGHT	n4	Accepted range is between 001–9998 9999 unknown	Birthweight at the time of delivery in grams	Used to identify risk factor on admission to neonatal care
NNUEpisodes	Demographics and birth information (baby)	Birth length (cm)	Birthlength	O	BIRTH LENGTH	nn.n	99.9 unknown	Length measured just after birth	Used to assess growth development
NNUEpisodes	Demographics and birth information (baby)	Birth head circumference	BirthHead Circumference	O	BIRTH HEAD CIRCUMFERENCE	nn.n	99.9 unknown	Occipitofrontal circumference measured at birth, in centimetres to one decimal place	Used to monitor outcomes according to head circumference
NNUEpisodes	Demographics and birth information (baby)	Gestation age in weeks	GestationWeeks	R	GESTATION LENGTH (AT DELIVERY)	n2	10–49	The best obstetric estimate at the time of delivery in weeks. This will normally be based on the postmenstrual age but, if appropriate, may be modified on the basis of antenatal ultrasound. Where gestation at delivery is	Used to identify risk factor on admission to neonatal care

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
								not known, this is based on the postnatal estimate of maturity	
NNUEpisodes	Demographics and birth information (baby)	Gestation age days	GestationDays	R	GESTATION LENGTH (REMAINING DAYS AT DELIVERY)	*n*	0–6 9 unknown	Specify, if known, the number of days between whole weeks in the gestation period	Used to identify risk factor on admission to neonatal care
NNUEpisodes	Demographics and birth information (baby)	Sex of the baby (phenotypic)	SexPhenotype	R	PERSON PHENTOTYPIC SEX	n1	1 male 2 female 9 indeterminate/intersex	The sex of the baby. 'Not known' is an option if information is missing or not recorded. 'Not specified' is an option for instances where the sex cannot be determined at birth. This option can be changed later if the chromosomal sex of the baby has been determined as follows: male (XY) or female (XX) or remain 'not specified' if the genotypic sex is not defined as XX/XY or is still not known	Used to aggregate by sex
NNUEpisodes	Demographics and birth information (baby)	Sex of the baby (genotypic)	SexGenotype	P	PERSON GENOTYPIC SEX (NATIONAL NEONATAL DATA SET)	n1	1 male 2 female 9 indeterminate/intersex X genotypic sex Unknown	Specify the genotypic sex of the infant when the phenotypic sex of the baby is indeterminate and requires genetic testing. Male (XY) or female (XX) or remain 'not specified' if the genotypic sex is not defined as XX/XY or is still not known	Used to aggregate by sex
NNUEpisodes	Demographics and birth information (baby)	Baby's blood group	BabyBloodGroup	O	BLOOD GROUP (BABY)	an2	National codes: A blood group A B blood group B AB blood group AB O blood group O	The blood group of a baby established as a result of a clinical investigation using the ABO classification system for human blood	Used to monitor implementation of anti-D prophylaxis guidance and outcomes for mothers who are rhesus negative and their babies

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							77 baby not tested 99 not known		
NNUEpisodes	Demographics and birth information (baby)	Baby's rhesus factor	BabyRhesusFactor	O	RHESUS GROUP (BABY) – note default codes are 777 and 999 not 077 and 099	an3	POS RhD-positive NEG RhD-negative 777 Baby not tested 999 not known	An indication of whether or not a baby has the rhesus factor (or RhD antigen) on the surface of their red blood cells, using the Rh system. This is indicated in association with the baby's blood group, established as a result of a clinical investigation, by RhD-positive (does have the RhD antigen) or RhD-negative (does not have the antigen)	Used to monitor implementation of anti-D prophylaxis guidance and outcomes for mothers who are rhesus negative and their babies
NNUEpisodes	Demographics and birth information (baby)	Worst base deficit within 12 hours after birth (mmol/l)	WorstBaseWithin12	R	BASE DEFICIT CONCENTRATION (WORST WITHIN 12 HOURS AFTER BIRTH)	nn.n	99.9 unknown	The worst base deficit concentration, an amount added to 1 l of the baby's blood at 40 mmHg pCO_2 to return the pH to normal, recorded within 12 hours of birth	Used to monitor outcomes according to base excess
NNUEpisodes	Parents	National identification mother	NHSNumberMother	R	NHS NUMBER (MOTHER)	n10	England and Wales – NHS number format	Mother's unique national identifier in England and Wales	Used to link children with pregnancy and mother
NNUEpisodes	Parents	National identification mother	NHSNumberMother (ENCRYPTED)	R	n/a	n10	England and Wales – NHS number format ENCRYPTED	Mother's unique national identifier in England and Wales encrypted using MD5	Used to link children with pregnancy and mother
NNUEpisodes	Parents	National identification mother status	NHSNumberMotheStatus	M	NHS NUMBER STATUS INDICATOR CODE (MOTHER)	n2	01 number present and verified 02 number present but not traced 03 trace required 04 trace attempted – no match or multiple match found	Whether or not the NHS number of the mother has been verified. Can be derived from patient demographic system by system provider	Used to identify whether or not the baby's NHS number has been verified by the NHS care records service

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							05 trace needs to be resolved – NHS number or patient detail conflict 06 trace in progress 07 number not present and trace not required 08 trace postponed (baby < 6 weeks old)		
NNUEpisodes	Parents	Scotland CHI number for mother	MotherCHINumber	R	COMMUNITY HEALTH INDEX NUMBER (MOTHER)	n10	Scotland – CHI number format	Mother's unique national identifier in Scotland	Used to link children with pregnancy and mother
NNUEpisodes	Parents	Scotland CHI number for mother	MotherCHINumber (ENCRYPTED)	R	COMMUNITY HEALTH INDEX NUMBER (MOTHER)	n10	Scotland – CHI number format ENCRYPTED	Mother's unique national identifier in Scotland encrypted using MD5	Used to link children with pregnancy and mother
NNUEpisodes	Parents	Northern Ireland H&C number for mother	MotherHCNumber	R	HEALTH AND CARE NUMBER (MOTHER)	n10	Northern Ireland – H&C number format	Mother's unique national identifier in Northern Ireland	Used to link children with pregnancy and mother
NNUEpisodes	Parents	Northern Ireland H&C number for mother	MotherHCNumber (ENCRYPTED)	R	HEALTH AND CARE NUMBER (MOTHER)	n10	Northern Ireland – H&C number format ENCRYPTED	Mother's unique national identifier in Northern Ireland. Encrypted using MD5	Used to link children with pregnancy and mother
NNUEpisodes	Parents	Birth year mother	BirthYearMother	R	YEAR OF BIRTH (MOTHER)	n4	Year of date	The calendar year of mother's birth from the mother's date of birth	Used to derive ages for comparison
NNUEpisodes	Parents	Postcode mother	PostCodeMother	R	POSTCODE OF USUAL ADDRESS (MOTHER)	an8	No-use NHS defaults: ZZ99 3VZ No fixed abode ZZ99 3WZ At sea ZZ99 3WZ In the air ZZ99 3WZ Inadequately described/specified ZZ99 3WZ Information refused ZZ99 3WZ Not collected	A UK postcode of the mother's residence at time of delivery	Used to derive PCT and other geographical areas, including Sure Start areas, for aggregation to compare outcomes and plan services

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							ZZ99 3WZ Not known ZZ99 3WZ Not stated/ specified		
NNUEpisodes	Parents	Postcode mother (LSOA)	PostCodeMotherLSOA	R	n/a (derived on receipt)	an8	Derived on mother's postcode. LSOA-equivalent of mother's postcode	An LSOA-equivalent of UK postcode of the mother's residence at time of delivery	Used to derive PCT and other geographical areas, including Sure Start areas, for aggregation to compare outcomes and plan services
NNUEpisodes	Parents	Mother's education	MumEducation	P	QUALIFICATION ATTAINMENT LEVEL MOTHER (NATIONAL NEONATAL DATA SET)	n2	00 No qualifications 01 1–4 O levels/CSEs/ GCSEs (any grades), Entry Level, Foundation Diploma 02 NVQ Level 1, Foundation GNVQ, Basic Skills 03 5 + O levels(passes)/ CSE (grade 1)/GCSEs (grades A*–C), School Certificate, 1A level/2–3 AS levels/VCEs, Higher Diplomas 04 NVQ Level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/General Diploma, RSA Diploma 05 2+ A levels/VCEs, 4+ AS levels, Higher School Certificate, Progression/Advanced Diploma 06 NVQ Level 3, Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma	Specify the current educational attainment of the mother	Used as a factor in socioeconomic analysis

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							07 Degree (for example BA, BSc), higher degree (for example MA, PhD, PGCE) 08 NVQ Level 4–5, HNC, HND, RSA Higher Diploma, BTEC Higher Level 09 Professional qualifications (for example teaching, nursing, accountancy) 10 Other vocational/ work-related qualifications 11 Foreign qualifications 99 unknown		
NNUEpisodes	Parents	Mother's occupation	MumOccupation	O	OCCUPATION MOTHER (SNOMED CT)	n18	Snomed CT for Concept ID 14679004	Mother's description of her occupation	Used to derive mother's occupational category
NNUEpisodes	Parents	Mother's ethnicity	MumEthnicity	R	ETHNIC CATEGORY (MOTHER)	An2	White A, British B, Irish C, any other white background mixed D, white and black Caribbean E, white and black African F, white and Asian G, any other mixed background Asian or Asian British H, Indian J, Pakistani K, Bangladeshi L, any other Asian background black or black British M, Caribbean N, African P, any other black background, Other ethnic groups R, Chinese S, any other ethnic group Z, not stated, 99 not known	Mother's declared ethnicity based on the NHS (England) standard codes for ethnic group	Used to compare outcomes according to ethnicity

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Parents	GPPractisecode	GPPractiseCode	R	GENERAL MEDICAL PRACTICE CODE (PATIENT REGISTRATION (MOTHER))	an8	NHS organisation code V81997 – no registered GP practice V81998 – GP practice code not applicable V81999 – GP practice code not known	Please specify mother's GP at time of delivery	Required for aggregation by GP/area
NNUEpisodes	Parents	Birth year father	BirthYearDad	R	YEAR OF BIRTH (FATHER)	n4	Year of date	The calendar year of father's birth from the father's date of birth	Used to derive ages for comparison
NNUEpisodes	Parents	Father's ethnicity	DadEthnicity	R	ETHNIC CATEGORY (FATHER)	An2	White A, British B, Irish C, any other white background mixed D, white and black Caribbean E, white and black African F, white and Asian G, any other mixed background Asian or Asian British H, Indian J, Pakistani K, Bangladeshi L, any other Asian background black or black British M, Caribbean N, African P, any other black background other ethnic groups R, Chinese S, any other ethnic group Z, not stated 99, not known	Biological father's declared ethnicity based on the NHS (England) standard codes for ethnic group of father	Used to compare outcomes according to ethnicity
NNUEpisodes	Parents	Parents consanguineous	ParentsConsaguinous	R	PARENTS CONSANGUINOUS INDICATOR	an1	N no Y yes 9 unknown	Records if parents are consanguineous – first cousins	Used to determine association with congenital anomaly
NNUEpisodes	Antenatal (pregnancy details)	Mother antenatally booked indicator	BookingIndicator	P	MOTHER ANTENATALLY BOOKED INDICATOR	an1	N no Y yes	Specify if the mother was booked for delivery of the infant	Used to assess the use of default codes in site code
NNUEpisodes	Antenatal (pregnancy details)	Intended place of delivery NHS code	BookingNHSCode	M	SITE CODE (OF INTENDED PLACE OF DELIVERY) or ORGANISATION CODE (OF INTENDED PLACE OF DELIVERY)	an9	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home)	Place at which mother was first booked for her confinement. The first intended place of delivery by the health-care professional in consultation with the woman	Used to aggregate by geographical area

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)		
NNUEpisodes	Pregnancy, labour and delivery	Mother's number of previous pregnancies	NumberOfPrevious Pregnancies	R	PREGNANCY TOTAL PREVIOUS PREGNANCIES	n2	0–29 99 unknown	Number of known pregnancies for the mother previous to the current pregnancy	Used to monitor outcome
NNUEpisodes	Labour and delivery	Birth order	BirthOrder	R	BIRTH ORDER (MATERNITY SERVICES SECONDARY USES)	n2	NN UU unknown	The numbered order in which babies are delivered in a multiple pregnancy independent of 'numbering' before delivery	Used to monitor outcomes comparing singleton and multiple pregnancies
NNUEpisodes	Labour and delivery	Fetus total	FetusTotal	R	NUMBER OF FETUSES (NOTED DURING PREGNANCY EPISODE)	n2	99 unknown	Total number of fetuses noted at any time in the pregnancy which resulted in delivery of a live or stillborn baby. This excludes fetus papyraceous and fetuses reabsorbed in utero and not delivered	Used to monitor outcomes comparing singleton and multiple pregnancies
NNUEpisodes	Antenatal (pregnancy details)	Medical problems prior to pregnancy of mother	ProblemsMedicalMother	R	MATERNITY COMPLICATING MEDICAL DIAGNOSIS TYPE (NATIONAL NEONATAL DATA SET)	an60	01 hypertension 02 cardiac disease 03 renal disease 04 mental health disease 06 haematological disease 07 central nervous system disease 08 diabetes 09 autoimmune disease 10 cancer	List of maternal problems that were present prior to this pregnancy	Used to monitor different targets for complicated/ uncomplicated pregnancies

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							12 infection disease: hepatitis A 13 infection disease: hepatitis B 14 infection disease: hepatitis C 16 endocrine disease 17 respiratory disease 18 gastrointestinal disease 19 musculoskeletal disease 0 gynaecological problems		
NNUEpisodes	Antenatal (pregnancy details)	Problems (obstetric) during pregnancy with mother	ProblemsObstPregnancy Mother	P	MATERNITY OBSTETRIC DIAGNOSIS TYPE (CURRENT PREGNANCY)	an60	01 pre-eclampsia 02 haemolytic anaemia 03 eclampsia 05 liver cholestasis of pregnancy 06 gestational diabetes mellitus 07 gestational hypertension 08 gestational proteinuria 09 antepartum haemorrhage 11 feto-maternal haemorrhage 19 placenta praevia 20 severe pre-eclampsia	List of maternal obstetric problems encountered relating to this pregnancy	To monitor outcomes for mothers and babies where complicating or risk factors are present

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Antenatal (pregnancy details)	Problems (infectious or medical condition) during pregnancy with mother	ProblemsInfct PregnancyMother	R	MATERNITY MEDICAL DIAGNOSIS TYPE (CURRENT PREGNANCY)	n2	01 rubella 02 varicella 03 Group B *Streptococcus* 04 asymptomatic bacteriuria 05 toxoplasmosis 08 ruberculosis 09 cytomegalovirus 10 parvovirus 11 malaria 13 cardiac disease 14 renal disease 15 mental health disorder 16 thromboembolic disorder 17 haematological disorder 18 central nervous system (CNS) disorder 19 diabetes 20 autoimmune disease 21 cancer 22 infectious hepatitis A 23 hepatitis B 24 hepatitis C 25 endocrine disorder 26 respiratory disease	The infections disease or medical condition diagnosed within this pregnancy	To monitor outcomes for mothers and babies where complicating or risk factors are present

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							27 gastrointestinal disorder 28 musculoskeletal disorder 29 gynaecological problems		
NNUEpisodes	Antenatal (pregnancy details)	Mother's blood group	MumBloodGroup	R	BLOOD GROUP (MOTHER)	an2	A blood group A B blood group B AB blood Group AB O blood Group O 77 mother not tested 99 not known	The blood group of a mother established as a result of a clinical investigation using the ABO classification system for human blood	Used to monitor implementation of anti-D prophylaxis guidance and outcomes for mothers who are rhesus negative and their babies
NNUEpisodes	Antenatal (pregnancy details)	Mother's rhesus factor	MumBloodRhesus	R	RHESUS GROUP (MOTHER)	an3	POS RhD-positive NEG RhD-negative 777 mother not tested 999 not known	An indication of whether or not a mother has the rhesus factor (or RhD antigen) on the surface of their red blood cells, using the Rh system. This is indicated in association with the mother's blood group, established as a result of a clinical investigation, by: RhD-positive (does have the RhD antigen) or RhD-negative (does not have the antigen)	Same as for blood group
NNUEpisodes	Antenatal (pregnancy details)	Mother's haemoglobino pathology status	MumHaemoglobinopathy	O	HAEMOGLOBINOPATHY INVESTIGATION RESULT CODE FOR NATIONAL NEONATAL DATA SET (MOTHER)	an250	1 sickle cell disease 2 sickle cell trait 3 sickle cell C disease 4 thalassemia major 5 thalassemia minor 9 unknown	Presence of known problem with a haemoglobinopathy in mother	Used to determine screening protocols for sickle cell disease and Thalassemia

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Antenatal (pregnancy details)	Smoking in pregnancy	SmokingInPregnancy	R	MOTHER CURRENT SMOKER AT BOOKING INDICATOR	an1	N no Y yes 9 unknown	Mother's smoking at the time of booking in this pregnancy	Used to compare outcomes for babies of mothers who smoke
NNUEpisodes	Antenatal (pregnancy details)	Number of cigarettes mother smoked during pregnancy	NoCigarettes	O	GIGARETTES PER DAY (MOTHER AT BOOKING)	n3	0–999	The number of cigarettes that the mother smoked on average, per day, at the time of booking for this pregnancy	Used to compare outcomes for babies of mothers who smoke
NNUEpisodes	Antenatal (pregnancy details)	Were steroids given during pregnancy?	SteroidsAntenatalGiven	R	STEROIDS GIVEN DURING PREGNANCY TO MATURE FETAL LUNGS INDICATOR	an1	Derived N no Y yes 9 unknown	Administration of any dose of steroid to mother (dexamethasone or betamethasone), at any time during pregnancy, with the intention of maturing fetal lungs	Used to compare outcomes for babies
NNUEpisodes	Antenatal (pregnancy details)	Number of antenatal steroid courses given	SteroidsAntenatalCourses	R	ANTENATAL STEROID COURSE COMPLETION STATUS CODE	n1	1 – complete: a full course of steroids at any time during pregnancy with the intention of maturing the fetal lungs 2 – incomplete: at least one injection of steroids given at any time during pregnancy with the intention of maturing the fetal lungs 3 – not given 9 – unknown	A complete course of steroids is defined by the RCOG guideline[388] as two 12 mg doses of betamethasone, given intramuscularly, 24 hours apart. Some units may use another regimen, including a different steroid. A complete course is one which complies with the local protocol. The time between the course of steroids and delivery of the baby does not matter. Enter the course as complete if the mother received the requisite course of steroids at any time, or in any unit, during the pregnancy. An incomplete course is where mother has received at least one injection of steroids	Used to compare outcomes for babies

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
								but has not gone on to complete the course as defined by the local protocol	
NNUEpisodes	Antenatal (pregnancy details)	If any, which steroids were given?	SteroidsAntenatalDrug	O	STEROID TYPE GIVEN TO MOTHER (SNOMED CT DM+D)	n18	1 betamethasone 2 dexamethasone 3 other dm+d selection	The name of the steroid given to mother presenting in preterm labour	Used to compare outcomes for babies
NNUEpisodes	Antenatal (pregnancy details)	Mother's rubella antibody status	MumRubellaStatus	O	INVESTIGATION RESULT CODE (MOTHER RUBELLA SCREENING)	n1	01 Rubella antibodies detected (> 10 IU/ml) 02 Rubella susceptible (< 10 ul/mol) 77 Not tested 99 Unknown	Result of test on mother for rubella antibody	Used to monitor implementation of screening guidelines and take up of services, and outcomes for babies
NNUEpisodes	Antenatal (pregnancy details)	Mother's date of last menstrual period	MumLMP	M	LAST MENSTRUAL PERIOD DATE	an10	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	Date of the first day of the mother's last menstrual period in this pregnancy in co-ordinated universal time	Used as a guide for calculation of timing of tests and other interventions, and gestational age at birth for those requiring critical care
NNUEpisodes	Antenatal (pregnancy details)	Mother's estimated date of delivery	MumEDD	M	ESTIMATED DATE OF DELIVERY (AGREED)	an10	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	Mother's agreed estimated date of delivery for this pregnancy is the last menstrual date if the dating ultrasound scan (performed in accordance with NICE guidelines[389]) agrees within 7 days, or if the difference is > 7 days, the date calculated from the dating ultrasound scan will be taken. If the dating ultrasound scan is unavailable then the last menstrual period date is used and if that is unavailable then a clinical assessment date is used. In co-ordinated universal time	Used as a guide for calculation of timing of tests and other interventions, and gestational age at birth for those requiring critical care

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Labour and delivery	Meconium stained liquor	MeconiumStainedLiquor	O	MECONIUM PRESENT IN LIQUOR INDICATOR	an1	N no Y yes 9 unknown	Confirm if there was presence of meconium in the liquor following rupture of the membranes or at delivery	Used to monitor the incidence of complications of delivery
NNUEpisodes	Labour and delivery	Medications administered during labour	DrugsInLabour	O	MEDICATION GIVEN DURING LABOUR (SNOMED CT DM+D)	n18	dm+d code for any drug	List of drugs given to mother during labour	Used to compare outcomes according to methods employed to induce labour
NNUEpisodes	Labour and delivery	Mother's onset of labour	Onsetoflabour	R	LABOUR OR DELIVERY ONSET METHOD CODE (NATIONAL NEONATAL DATA SET)	an2	01 spontaneous (where the labour or delivery onset method is 'Spontaneous') 02 induced (where the labour or delivery onset method is 'Surgical induction', 'Medical induction', or 'Combination of surgical induction and medical induction') 03 none (where the labour or delivery onset method is 'Caesarean section carried out before the onset of labour or a planned elective caesarean section carried out immediately following onset of labour') 09 not known	Specify the status of mother's labour	Used to monitor delays in delivery and outcomes for mothers and babies
NNUEpisodes	Labour and delivery	Date and time of rupture of membranes	DateROM	R	RUPTURE OF MEMBRANES DATE TIME	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The date and time when membranes were ruptured in this pregnancy in co-ordinated universal time	Used to monitor implementation of guidance
NNUEpisodes	Labour and delivery	Maternal pyrexia in labour > 38C	MaternalPyrexiaIn Labour38c	O	SIGNIFICANT MATERNAL PYREXIA IN LABOUR INDICATOR	an1	N no Y yes 9 unknown	Details the development of significant pyrexia by the mother during labour	Used to monitor the incidence of complications of delivery

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Labour and delivery	Intrapartum antibiotics given	IntrapartumAntibiotics Given	O	INTRAPARTUM ANTIBIOTICS GIVEN INDICATOR	an1	N no Y yes 9 unknown	Details if mother was given antibiotics during labour	Used to monitor the incidence of complications of delivery
NNUEpisodes	Labour and delivery	Presentation of fetus at delivery	PresentationOfFetusAt Delivery	R	PRESENTATION AT DELIVERY	an2	01 cephalic 02 breech 03 transverse/oblique 04 not known XX other	Presentation of the fetus at delivery	Used to monitor changes in intended plan of care
NNUEpisodes	Labour and delivery	Mode of delivery	ModeOfDelivery	R	MODE OF DELIVERY	an1	1 emergency caesarean section 2 elective caesarean section 3 vaginal – instrument assisted 4 vaginal – spontaneous 9 not known	Specify the mode of delivery	Used to compare outcomes and variance in practice
NNUEpisodes	Labour and delivery	Mother's labour status at time of caesarean	ModeofDelivery Caesarean	P	IN LABOUR BEFORE CAESAREAN SECTION INDICATOR	n1	Y mother in labour before caesarean delivery N mother not in labour before caesarean delivery	An indication of whether or not the mother had established labour onset before delivery of the baby by caesarean section. If the mode of delivery is caesarean then this item is required	Used to compare outcomes and variance in practice
NNUEpisodes	Labour and delivery	Mother's mode of delivery instrument	ModeofDeliveryInstrument	P	DELIVERY INSTRUMENT TYPE	an1	1 forceps 2 ventouse 3 other	Specify the instrument used during delivery of the infant	Used to monitor delays in delivery and outcomes for mothers and babies
NNUEpisodes	Labour and delivery	Spontaneous respiration time of onset	SpontaneousRespiration Time	O	TIME BETWEEN DELIVERY AND SPONTANEOUS RESPIRATION CODE	n1	1 < 1 minutes 2 1–1.5 minutes 3 1.6–2 minutes 4 2.1–3 minutes	Recorded time at which the infant's first gasp was observed following birth in the delivery suite	Used to monitor neonatal outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							5 3.1–4 minutes 6 4.1–5 minutes 7 > 5 minutes		
NNUEpisodes	Labour and delivery	APGAR at 1 minute	apgar_1min	R	APGAR SCORE (1 MINUTE)	n2	0–10 Apgar score 99 unknown	Apgar score at 1 minute of age as determined by immediate examination of the baby following delivery for appearance, pulse, grimace, activity, and respiration. Each of these five criteria is scored between a 0 and 2, with each score contributing to a cumulative total from 0 to 10	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	APGAR at 5 minutes	apgar_5min	R	APGAR SCORE (5 MINUTES)	n2	0–10 Apgar score 99 unknown	Apgar score at 5 minute of age as determined by immediate examination of the baby following delivery for appearance, pulse, grimace, activity, and respiration. Each of these five criteria is scored between a 0 and 2, with each score contributing to a cumulative total from 0 to 10	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	APGAR at 10 minutes	apgar_10min	R	APGAR SCORE (10 MINUTES)	n2	0–10 Apgar score 99 unknown	Apgar score at 10 minute of age as determined by immediate examination of the baby following delivery for appearance, pulse, grimace, activity, and respiration. Each of these five criteria is scored between a	Used to monitor neonatal outcomes including cooling

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
								0 and 2, with each score contributing to a cumulative total from 0 to 10	
NNUEpisodes	Labour and delivery	Methods of resuscitation	MethodsOfResuscitation	R	NEONATAL RESUSCITATION METHOD (NATIONAL NEONATAL DATA SET)	an3	00 none 10 stimulation 11 positioning managing airways 12 oxygen 13 suction 14 bag and face mask IPPV 15 intubation 16 cardiac massage	Interventions used during resuscitation or stabilisation immediately after delivery of the baby	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	Drugs used during resuscitation	DrugsforResuscitation	O	NEONATAL RESUSCITATION DRUG (SNOMED CT DM+D)	n18	dm+d code for any drug	If medication was administered at resuscitation please select relevant medications	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	Admission: time of cord clamping	CordClamp	O	UMBILICAL CORD CLAMPED IMMEDIATELY AFTER BIRTH INDICATOR	an1	N no Y yes 9 unknown	Indicate if the cord was clamped immediately after birth	
NNUEpisodes	Labour and delivery	Admission: time of cord clamping	TimeofCordClamp	O	TIME BETWEEN DELIVERY AND UMBILICAL CORD CLAMPING	n4	0–3600 seconds 9999 unknown	Please indicate the time of blood cord clamping in seconds from birth	
NNUEpisodes	Labour and delivery	Admission: stripping of blood from cord	StripBloodCord	O	UMBILICAL CORD MILKING PERFORMED INDICATOR	an1	N no Y yes U unknown	Indicate if the umbilical cord was stripped or milked of blood to enhance placental-infant transfusion at birth	
NNUEpisodes	Labour and delivery	Admission: cord artery pH	CordPhArterial	O	UMBILICAL CORD BLOOD PH LEVEL (ARTERIAL)	n.nn	6.00–8.00 9.99 unknown	The pH of cord arterial blood taken after delivery	Used to monitor neonatal outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Labour and delivery	Admission: cord venous pH	CordVenousPH	O	UMBILICAL CORD BLOOD PH LEVEL (VENOUS)	n.nn	6.00–8.00 9.99 unknown	The pH of cord venous blood taken after delivery	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	Admission: cord artery pCO_2	CordArterialPCO_2	O	UMBILICAL CORD BLOOD PARTIAL PRESSURE CARBON DIOXIDE (ARTERIAL)	n.nn	5.0–8.50 KPa 9.99 unknown	The partial pressure of CO_2 value of the blood taken from the umbilical cord artery at birth	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	Admission: cord lactate	CordlLactate	O	UMBILICAL CORD BLOOD LACTATE LEVEL	nn.nn	mmol/l 99 99 unknown	The lactate results of the umbilical lactate value of the blood taken from the umbilical cord at birth	Used for cooling intervention
NNUEpisodes	Labour and delivery	Admission: cord artery base excess	CordArterialBaseExcess	O	UMBILICAL CORD BLOOD BASE EXCESS CONCENTRATION (ARTERIAL)	an3	–30 – 30 mmol 99 unknown	Base excess concentration of arterial cord blood taken after the delivery	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	Admission: cord venous base excess	CordVenousBaseExcess	O	UMBILICAL CORD BLOOD BASE EXCESS CONCENTRATION (VENOUS)	an3	–30 – 30 mmol 99 unknown	Base excess concentration of venous cord blood taken after the delivery	Used to monitor neonatal outcomes
NNUEpisodes	Labour and delivery	Was surfactant given during resuscitation?	SurfactantGiven Resuscitation	R	SURFACTANT GIVEN INDICATOR (DURING RESUSCITATION)	an1	N no Y yes 9 unknown	Surfactant given during resuscitation	Used to monitor neonatal outcomes
NNUEpisodes	Admission details	Admission: date and time	AdmitTime	M	CRITICAL CARE START DATE AND TIME	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The calendar date and time, co-ordinated universal time, on which an inpatient stay commences an episode of care in a neonatal unit	Used for the NCCMDS, calculate the anonymised daily dates
NNUEpisodes	Admission details	Episode of care	CriticalCareIdentifier	R	EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL)	n2	The number of this episode of care for this baby	The EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) is used to sequentially identify each CRITICAL CARE PERIOD within a Neonatal Critical Care Spell. The first CRITICAL CARE PERIOD identifier commences at 1; subsequent CRITICAL	Used to ascribe outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
								CARE PERIODS during the same period of care (within the same or different Health-Care Providers) are then incremented by 1. For example, a Neonate is admitted to the Neonatal Intensive Care Unit at Trust A, starting a CRITICAL CARE PERIOD and generating EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) 1. The Neonate is then transferred to a different Health-Care Provider, Trust B (ending the CRITICAL CARE PERIOD at Trust A), which generates EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) 2. The Neonate may then return to Trust A (ending the CRITICAL CARE PERIOD at Trust B), generating EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) 3	
NNUEpisodes	Admission details	Hospital baby admitted to	ProviderNHSCode	M	SITE CODE (OF ADMITTING NEONATAL UNIT) or ORGANISATION CODE (OF ADMITTING NEONATAL UNIT)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home) ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)	This is the code for the hospital recording information on this patient. It is a code that identifies an organisation uniquely. For NHS organisations it is a code that is managed by the Corporate Data Administration section of the Department of Health and Social Care to identify most organisations that exchange information within the NHS or	Used to ascribe outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
								return information to the Centre. Examples of organisations that can be identified this way are NHS Trusts and Health Authorities	
NNUEpisodes	Admission details	Location baby admitted from	AdmitFromNHSCode	R	SITE CODE (ADMITTED FROM TO NEONATAL UNIT) or ORGANISATION CODE (ADMITTED FROM TO NEONATAL UNIT)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home) ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)	The place from which a baby was admitted into this episode of care. If the baby is admitted to the neonatal unit from its own local labour ward or theatres, then the value entered is the NHS code of this hospital	Used to analyse transfer of patients
NNUEpisodes	Admission details	Hospital baby admitted from location detail	AdmissionSource	O	LOCATION IN HOSPITAL TYPE (BABY ADMITTED FROM)	an2	1 labour and delivery ward 2 operating theatre 3 children's ward 4 postnatal ward 5 neonatal intensive care unit/special care baby unit 6 other	The exact location at the hospital from which a baby was admitted into this episode of care. Specialist care baby Unit. Neonatal unit	Used to analyse transfer of patients
NNUEpisodes	Admission details	Admission: reason for admit	ReasonForAdmit	R	PRIMARY CATEGORY OF CARE REQUIRED ON ADMISSION TO NEONATAL CRITICAL CARE	n2	10 medical intensive care 11 medical high-dependency care 12 medical special care 13 surgical care 14 cardiac care 15 tertiary specialist investigation	Specify the type of clinical service the infant is being admitted to receive, including if the service is part of back transfer (returning to hospital from which an infant was transferred for care to location of the previous episode of care). Type of care is identified using BAMP	Used to analyse transfer of patients

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							16 back transfer for continuing medical intensive care 17 back transfer for continuing medical high-dependency care 18 back transfer for continuing medical special care 19 social care 20 transitional care 99 unknown	(www.BAPM.org) classification	
NNUEpisodes	Admission details	Admission temperature status	AdmitTempStatus	M	TEMPERATURE RECORDED AFTER ADMISSION TO NEONATAL CRITICAL CARE INDICATOR	an1	N no Y yes 9 unknown	Specify the temperature measured after admission. A prompt to verify that an admission temperature was recorded after admission	Used in the National Neonatal Audit Programme
NNUEpisodes	Admission details	Temperature at admission	AdmitTemperature	M	TEMPERATURE (ON ADMISSION TO NEONATAL CRITICAL CARE)	nn.*n*	77.7 not recordable	Baby's axillary/skin temperature in degrees Celsius measured within 60 minutes of admission to this episode of care	Used in the National Neonatal Audit Programme
NNUEpisodes	Admission details	Admission temperature date and time	AdmissionTempDateTime	M	OBSERVATION DATE AND TIME (TEMPERATURE)	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The date and time, co-ordinated universal time, at which the admission temperature was measured	Used in the National Neonatal Audit Programme
NNUEpisodes	Admission details	Admission blood pressure	AdmitBP	R	MEAN ARTERIAL BLOOD PRESSURE (ON ADMISSION TO NEONATAL CRITICAL CARE)	n3	10–150 999 unknown	Specify the mean blood pressure of the baby on admission to this episode of care (in mmHg)	Used to assess clinical condition of infant on admission to care as well as transfers
NNUEpisodes	Admission details	Admission heart rate	AdmitHR	R	HEART RATE (ON ADMISSION TO NEONATAL CRITICAL CARE)	n3	50–350 999 unknown	Specify the heart rate of the baby on admission to this episode of care (per minute)	Used to assess clinical condition of infant on admission to care as well as transfers

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Admission details	Respiratory rate at admission	AdmissionRR	O	RESPIRATORY RATE (ON ADMISSION TO NEONATAL CRITICAL CARE)	n3	10–200 999 unknown	Specify the respiratory rate of the baby on admission to this episode of care (per minute)	Used to assess clinical condition of infant on admission to care as well as transfers
NNUEpisodes	Admission details	Oxygen saturation at admission	AdmissionOxygen Saturation	O	OXYGEN SATURATION (ON ADMISSION TO NEONATAL CRITICAL CARE)	n3	10–100 999 unknown	Specify the oxygen saturation of the baby on admission to this episode of care (in %)	Used to assess clinical condition of infant on admission to care as well as transfers
NNUEpisodes	Admission details	Blood glucose concentration (mmol/l) at admission	BloodGlucose	O	BLOOD GLUCOSE CONCENTRATION (ON ADMISSION TO NEONATAL CRITICAL CARE)	nn.*n*	0.0–50.0 99.9 unknown	Specify the blood glucose concentration of the baby on admission to this episode of care (in mmol/l)	Used to assess clinical condition of infant on admission to care as well as transfers
NNUEpisodes	Admission details	Clinical diagnosis at admission	DiagnosisAtAdmission	R	DIAGNOSIS (ICD ON ADMISSION TO NEONATAL CRITICAL CARE) and/or DIAGNOSIS (SNOMED CT ON ADMISSION TO NEONATAL CRITICAL CARE)	an6/n18	ICD-10 and/or Snomed CT	Specify the clinical reasons for admission of the baby from the list of ICD	Used to compare outcomes for babies
NNUEpisodes	Admission details	Was Vitamin K permission given	VitaminKPermission	O	PARENTAL CONSENT TO ADMINISTER VITAMIN K INDICATOR	an1	Y parental consent to administer vitamin K given N parental consent to administer vitamin K not given 9 not known if parental consent given	An indication of whether parental consent was given to administer vitamin K to the baby	Used to monitor the uptake of vitamin K prophylaxis
NNUEpisodes	Admission details	Was vitamin K indicator	VitaminKIndicator	O	VITAMIN K ADMINISTERED INDICATOR	an1	Y vitamin K given N vitamin K not given	Specify if the baby had received any dose of vitamin K following delivery	Used to monitor the uptake of vitamin K prophylaxis
NNUEpisodes	Admission details	Route of administration of Vitamin K	VitaminKRoute	O	VITAMIN K ROUTE OF ADMINISTRATION	*n*	1 intramuscular injection 2 intravenous injection 3 oral administration 9 Route of administration unknown	Specify the route of administration of the first dose of vitamin K given following delivery	Used to monitor the uptake of vitamin K prophylaxis

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Admission details	Admission: designation of member of staff completing admission form	AdmissionStaffDes	O	CARE PRFESSIONAL JOB ROLE CODE (COMPLETING NEONATAL INTENSIVE CARE UNIT ADMISSION FORM)	an5	NHS Data Dictionary Codes from JOB ROLE Codes	The professional designation of the person completing the admission form	Used to monitor data quality
NNUEpisodes	Admission details	Consultation with parents	ConsultationWithParents	M	PARENTS SEEN BY SENIOR STAFF MEMBER WITHIN 24 HOURS INDICATOR	an1	N no Y yes 9 unknown	Parents seen by senior member of staff within 24 hours of admission of the baby to the neonatal unit in this episode of care	Used in the National Neonatal Audit Programme
NNUEpisodes	Discharge details	Discharge date and time	DischDateTime	M	CRITICAL CARE DISCHARGE DATE AND TIME	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The date and time, co-ordinated universal time, on which an inpatient completes this episode of care, either because of death, transfer to another ward or hospital, or because of discharge home	Used to measure length of stay and to calculate the anonymised version of this field
NNUEpisodes	Discharge details	Discharge destination	DischargeDestination	M	DESTINATION ON DISCHARGE FROM NEONATAL CRITICAL CARE	n2	1 home with parent(s) 2 ward in same hospital 3 died 4 social/foster care 5 transferred to another hospital 6 hospice	The destination of the baby at discharge from this episode of care	Used to compare outcomes for babies
NNUEpisodes	Discharge details	Discharge reason	DischargeReason	R	TRANSFERRED FOR FURTHER CARE TYPE (NATIONAL NEONATAL DATA SET)	an2	10 transferred to another hospital for continuing care 11 transferred to another hospital for specialist care 12 transferred to another hospital for surgical care 13 transferred to another hospital for cardiac care 99 unknown	The destination of the baby at discharge from this episode of care. If the discharge destination is a transfer to another location then this item is required	Used to compare outcomes for babies

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Discharge details	Discharge destination ward	DischargeWard	O	WARD TYPE DISCHARGED TO (NATIONAL NEONATAL DATA SET)	an1	1 postnatal ward 2 transitional care 3 other neonatal unit 4 paediatric Intensive care unit	Specify the type or ward the baby will be discharged to	Used to compare outcomes for babies, a National Neonatal Audit Programme filter
NNUEpisodes	Discharge details	Discharge hospital NHS code	DischargeHospitalNHS Code	M	SITE CODE (RECEIVING) or ORGANISATION CODE (RECEIVING)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home) ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)	The hospital to which a baby is being transferred from this episode of care. Record if discharge reason is specified	Used to compare outcomes for babies
NNUEpisodes	Discharge details	Date of death and time	DateofDeath	R	PERSON DEATH DATE AND TIME (DURING NEONATAL CRITICAL CARE PERIOD	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The date and time, co-ordinated universal time, on which an inpatient had died in this episode of care as stated on the death certificate. If the discharge destination indicates the infant that died this item is required	Used in survival analyses and to calculate anonymised dates
NNUEpisodes	Discharge details	Cause of death	Causeofdeath	R	DEATH CAUSE ICD CODE (DURING NEONATAL CRITICAL CARE PERIOD)	an6	ICD-10	Specify the major reasons for death of the baby from the list of ICD as corresponding with death certificate	Used to compare outcomes for babies
NNUEpisodes	Discharge details	Discharge: If baby died was a post-mortem done?	IfPostMortemDone	O	POST-MORTEM CARRIED OUT INDICATOR	an1	N no Y yes 9 unknown	Specify if a post-mortem was done	Used to monitor neonatal death

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Discharge details	Discharge: post-mortem, was consent sought?	PostMortemConsent	O	PARENTAL CONSENT TO POST-MORTEM INDICATOR	an1	N no Y yes 9 unknown	Confirm if consent was obtained from the parents for the post-mortem	Used to monitor neonatal death
NNUEpisodes	Discharge details	Discharge: if NEC diagnosed, did post-mortem confirm this?	PostmortemConfirmation	O	POST-MORTEM CONFIRMED-NECROTISING ENTEROCOLITIS INDICATOR	an1	N no Y yes 9 unknown	If a NEC diagnosis was made at any point at admission, specify if the post-mortem confirmed it	Used to monitor neonatal death
NNUEpisodes	Discharge details	Discharge: Oxygen	DischargeOxygen	O	RECEIVING OXYGEN THERAPY ON DISCHARGE INDICATOR	*n*	Derived from daily data item N no Y yes	Item specifies if the baby is receiving and is dependent on oxygen therapy on discharge home	Used to compare outcomes for babies
NNUEpisodes	Discharge details	Unit responsible for 2 year follow-up	Locationforfollowup	O	SITE CODE (TWO YEAR NEONATAL OUTCOMES ASSESSMENT RESPONSIBILITY) or ORGANISATION CODE (TWO YEAR NEONATAL OUTCOMES ASSESSMENT RESPONSIBILITY)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home) ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)	Specify the name of the hospital that is responsible for undertaking the 2-year follow-up of this baby after discharge from neonatal care	Used to plan 2-year follow-up
NNUEpisodes	Discharge details	Clinical diagnoses at discharge	DiagnosesAtDischarge	R	DIAGNOSIS (ICD RECORDED ON DISCHARGE FROM NEONATAL CRITICAL CARE) and/or DIAGNOSIS (SNOMED CT ON DISCHARGE FROM NEONATAL CRITICAL CARE)	an16/n18	ICD-10 and/or Snomed CT can be derived from daily records	Specify all the applicable diagnoses for this baby if not already recorded in the daily diagnoses	Used to monitor infant care, evaluate health status and outcomes
NNUEpisodes	Discharge details	Procedures performed	ProcedureAtDischarge	R	PROCEDURE (OPCS RECORDED ON DISCHARGE FROM NEONATAL CRITICAL CARE) and/or PROCEDURE (SNOMED CT RECORDED ON DISCHARGE FROM NEONATAL CRITICAL CARE)	an4/n18	OPCS coded and/or Snomed CT can be derived from daily records	Specify the procedures performed in this episode of neonatal care if not already recorded in the daily procedures	Used to monitor infant care, evaluate health status and outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUEpisodes	Discharge details	Date of procedures performed	ProcedureAtDischargeDate	R	PROCEDURE DATE AND TIME (DURING NEONATAL CRITICAL CARE PERIOD)	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS) Can be derived from daily records	The date and time, co-ordinated universal time, on which an inpatient had a surgical procedure in this episode of care	Used in survival analyses and calculate anonymised dates
NNUEpisodes	Clinical trials	This episode of care: research study enrolment	ResearchStudyEnrol	O	CLINICAL TRIAL NAME	an250	Free–text entry	If this baby was enrolled in one or more research studies, in this episode of care, specify the study. Do not record any patient identifiable details such as contact details, address, and patient names	Used to monitor participation in research studies where applicable
NNUEpisodes	Clinical trials	This episode of care: research study medication received	ResearchDrugs	O	CLINICAL TRIAL MEDICATION ADMINISTERED NAME	an250	Free–text entry	Specify the medications the baby has received as part of the research study. Do not record any patient identifiable details such as contact details, address, and patient names	Used to monitor participation in research studies where applicable
Daily	General information	General information: date of day of care	ActiveDate	M	NEONATAL CRITICAL CARE DAILY CARE DATE	an19	DateTime coding (e.g. 1997–07–16T19:20:30) Date (an10 CCYY-MM-DD) Time (an8 HH:MM:SS)	The clinical data for a 24-hour period in a neonatal unit; a new day of care begins at midnight following the previous day of neonatal care	Used in commissioning and to determine level of care
Daily	General information	General information: day provider NHS code	DayProviderNHSCode	M	SITE CODE (OF ADMITTING NEONATAL UNIT) or ORGANISATION CODE (OF ADMITTING NEONATAL UNIT)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home)	The code for the hospital recording information on this day and episode for the patient. Can be derived by system provider	Used in commissioning to ascribe daily activity

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							ZZ888 – not applicable (intended to deliver at non-NHS organisation) ZZ203 – not known (intended place of delivery not known)		
Daily	General information	General information: weight today	DayWorkingWeight	R	PERSON WEIGHT IN GRAMS	n4	001–9998 g 9999 unknown	The weight, in grams, of the baby on this day of care	Used to monitor development of infant
Daily	General information	General information: head circumference today	DayHeadCirc	O	PERSON HEAD CIRCUMFERENCE IN CENTIMETRES	nn.n	99.9 unknown	Head circumference is in centimetres to one decimal	Used to monitor development of infant
Daily	General information	General information: length today	DayLength	O	PERSON LENGTH IN CENTIMETRES	nn.n	99.9 unknown	Length measured, in centimetrers, on this day of care	Used to monitor development of infant
Daily	General information	General information: location of care	LocationofCare	R	LOCATION OF HIGHEST LEVEL OF CARE	n2	01 neonatal unit 02 transitional care 11 postnatal ward 12 other area	This is the 'highest' location of care on this day	Used in the National Neonatal Audit Programme
Daily	General information	General information: receiving 1 : 1 nursing	OneToOneNursing	R	PATIENT RECEIVING ONE TO ONE NURSING CARE INDICATOR	an1	N no Y yes	Specify if the baby received 1 : 1 nursing on this day of care	Used to determine level of care
Daily	General information	General information: carer resident	CarerResident	R	CARER RESIDENT INDICATION CODE	*n*	1 carer not resident 2 carer resident – not caring for baby 3 carer resident – caring for baby	Specify the detail of the carer resident	Used to determine level of care
Daily	General information	Clinical diagnoses on day of care	DiagnosesDaily	R	DIAGNOSIS (ICD ON NEONATAL CRITICAL CARE DAILY CARE DATE) and/or DIAGNOSIS (SNOMED CT ON NEONATAL CRITICAL CARE DAILY DATE)	an16/n18	ICD-10 and/or Snomed CT	Specify all the applicable diagnoses for this baby	Used to monitor infant care and evaluate health status and outcomes
Daily	General information	General information: any surgical procedures	SurgicalProcedure	R	PROCEDURE (OPCS DURING NEONATAL CRITICAL CARE PERIOD) or PROCEDURE (SNOMED CT DURING NEONATAL CRITICAL CARE PERIOD)	an4/n18	OPCS coded and/or Snomed CT	Surgical procedure on the date and time specified	Used to monitor health status and outcomes, and calculate anonymised versions

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
Daily	General information	General information: baby transported	TransportedDay	R	PERSON ACCOMPANYING TRANSPORTED PATIENT	n1	1 with nurse (non-ANNP) 2 with nurse (ANNP) 3 with doctor 4 with paramedics 5 with parent	If baby has been transported today, specify with whom the transport took place. ANNP is an Advanced Neonatal Nurse Practitioner	Used in neonatal transport data set
Daily	Respiratory	Respiration: respiratory support device	RespiratorySupport	R	RESPIRATORY SUPPORT DEVICE TYPE (NATIONAL NEONATAL DATA SET)	n1	1 endotracheal tube 2 tracheostomy 3 nasal cannula 4 nasopharyngeal cannula 5 face mask	Type of respiratory support device any time during the 24-hour period (00.00–23.59)	Used to determine level of care
Daily	Respiratory	Respiration: respiratory support mode	ModeofRespiratorySupport	R	RESPIRATORY SUPPORT MODE (NATIONAL NEONATAL DATA SET)	n1	1 positive pressure ventilation 2 high frequency oscillatory ventilation 3 high frequency jet ventilation 4 CPAP 5 BiPAP/SiPAP 6 high flow	Mode of respiratory support via endotracheal tube. Conventional ventilation includes intermittent mandatory ventilation, synchronised intermittent mandatory ventilation, assist/control ventilation, pressure support ventilation, pressure targeted, volume targeted, hybrid	Used to determine level of care
Daily	Respiratory	Respiration: nitric oxide given	PulmonaryVasodilator	R	NITRIC OXIDE GIVEN INDICATOR	an1	N no Y yes	Specify if the baby is receiving nitric oxide on this day of care	Used to determine level of care
Daily	Respiratory	Respiration: chest drain present	ChestDrain	R	CHEST DRAIN IN SITU INDICATOR	an1	N no Y yes	Specify if the baby has a chest drain present on this day of care	Used to determine level of care
Daily	Respiratory	Respiration: tracheostomy tube present	DayTracheostomyTube	R	TRACHEOSTOMY TUBE IN SITU INDICATOR	an1	N no Y yes	Specify if the baby has a tracheostomy tube insert present on this day of care	Used to determine level of care

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
Daily	Respiratory	Respiration: repogle tube in situ	ReplogleTube	R	REPLOGLE TUBE IN SITU INDICATOR	an1	N no Y yes	Specify if the baby has a repogle tube insert present on this day of care	Used to determine level of care
Daily	Respiratory	Respiration: surfactant given today	DaySurfactantGiven	R	SURFACTANT GIVEN INDICATOR (ON NEONATAL CRITICAL CARE DAILY CARE DATE)	an1	N no Y yes	Records if baby received any dose of surfactant in this day while in the neonatal unit. Surfactant given at delivery/resuscitation is a data item collected separately	Used to determine health status of infant
Daily	Respiratory	Respiration: maximum oxygen supplementation today	OygenPerc	P	FRACTION OF INSPIRED OXYGEN PERCENTAGE	N3	Percentage fraction of inspired oxygen	Specify the maximum fraction of inspired oxygen percentage that the baby received on this day of care	Used to monitor oxygen dependency
Daily	Cardiovascular	Cardiovascular: i.v. infusion pulmonary vasodilator	PulmonaryVasodilator	R	CONTINUOUS INFUSION OF PULMONARY VASODILATOR RECEIVED INDICATOR	an1	N no Y yes	Specify if the baby is receiving a continuous infusion of a pulmonary vasodilator on this day of care	Used to determine level of care
Daily	Cardiovascular	Cardiovascular: Inotropes given	Inotropesgiven	R	INOTROPE INFUSION RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby is receiving an inotrope infusion on this day of care. For list of inotropic drugs refer to the *British National Formulary*[390]	Used to determine level of care
Daily	Cardiovascular	Cardiovascular: prostaglandin infusion	Prostaglandin	R	PROSTAGLANDIN INFUSION RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby is receiving an inotrope infusion on this day of care. For list of inotropic drugs refer to the *British National Formulary*[390]	Used to determine level of care
Daily	Cardiovascular	Cardiovascular: treatment for PDA	SurgeryforPDA	R	TREATMENT TYPE FOR PATENT DUCTUS ARTERIOSUS	*n*	1 indometacin/ indomethacin 2 ibuprofen 3 surgery 9 not applicable	If baby has PDA, specify treatment a baby has had for PDA on this day of care	Used to determine level of care

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
Daily	Gastrointestinal	Gastrointestinal: peritoneal dialysis	PeritonealDialysis	R	PERITONEAL DIALYSIS OR HAEMOFILTRATION RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby is receiving peritoneal dialysis or haemofiltration on this day of care	Used to determine level of care
Daily	Gastrointestinal	Gastrointestinal: haemofiltration	Haemofiltration	R	PERITONEAL DIALYSIS OR HAEMOFILTRATION RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby is receiving peritoneal dialysis or haemofiltration on this day of care	Used to determine level of care
Daily	Gastrointestinal	Gastrointestinal: treatment for NEC	DayNEC	R	TREATMENT TYPE FOR NECROTISING ENTEROCOLITIS	n	Derived 1 medical 2 surgical 9 not applicable – no treatment given	Specify if a baby received NEC treatment on this day of care. This item can be derived from the surgical procedures item	Used to determine level of care
Daily	Gastrointestinal	Gastrointestinal: rectal washouts (> 3/day)	Rectalwashout	R	MORE THAN THREE RECTAL WASHOUTS RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby has had more than three rectal washout procedures on this day of care	Used to determine level of care
Daily	Gastrointestinal	Gastrointestinal: stoma in situ	StomaInSitu	R	STOMA PRESENT INDICATOR	an1	N no Y yes	Specify if a baby has any stoma in place on this day of care	Used to determine level of care
Daily	Blood transfusions	Transfusions: type of transfusion today	DayTransfusion	R	BLOOD TRANSFUSION TYPE	n1	1 partial (dilutional) exchange transfusion 2 full exchange transfusion today	If a baby has received a transfusion on this day of care, specify type of transfusion	Used to determine level of care
Daily	Blood transfusions	Transfusions: blood products	BloodProductsTrans	R	BLOOD TRANSFUSION PRODUCT TYPE	n1	1 packed red cells or whole blood transfusion 2 fresh-frozen plasma 3 cryoprecipitate 4 platelets 5 albumin	Specify the blood products used in the transfusion procedure on this day of care	Used to determine level of care
Daily	Neurology	Neurology: central tone	Centraltone	R	CENTRAL TONE STATUS	n1	1 normal	Specify if any changes were observed in central tone by	Used in the National Neonatal Audit Programme

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							2 increased 3 decreased	monitoring the baby's neuromotor functions during this 24-hour period (00.00–23.59), for example Floppy limbs would mean decreased central tone	
Daily	Neurology	Neurology: consciousness status	Consciousness	R	NEONATAL CONSCIOUSNESS STATUS	n1	0 normal 1 hyper aler 2 lethargic 3 comatose	The consciousness status of the baby during this 24-hour period (00.00–23.59)	Used in the National Neonatal Audit Programme
Daily	Neurology	Neurology: convulsion today	Convulsion	R	SIEIZURE OCCURRED INDICATOR	an1	N no Y yes	This records if the baby had any seizures (clinical or noted on erg/cam monitoring) on this day	Used in the National Neonatal Audit Programme
Daily	Neurology	Neurology: neonatal abstinence syndrome	NeonatalAbstinence Syndrome	R	NEONATAL ABSTINENCE SYNDROME OBSERVED INDICATOR	an1	N no Y yes	Specify is neonatal abstinence syndrome is observed on this day of care	Used to determine level of care
Daily	Neurology	Neurology: surgery for ventricular-peritoneal shunt	SurgeryVPShunt	R	n/a	n1	Derived 0 no 1 yes	Derived item from surgical procedures today; 'yes' is if a VP shunt procedure is found in the field and 'no' if it has not	Used to determine level of care
Daily	Neurology	Neurology: EEG/ CFM	EEGCFM	R	BRAIN ACTIVITY SCAN PERFORMED INDICATOR	an1	N no Y yes	Specify if this baby has had a brain activity scan on this day of care using either eithelectro-encephalogram or bedside with a cerebral function monitor	Used to determine level of care
Daily	Neurology	Neurology: therapeutic hypothermia	Therapeutichypothermia	R	THERAPEUTIC HYPOTHERMIC INDUCED INDICATOR	an1	N no Y yes	This records if the baby was being cooled today as part of management of suspected hypoxic ischaemic damage to the brain. Includes both passive and active cooling	Used in the National Neonatal Audit Programme

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
Daily	Neurology	Neurology: hypoxic ischaemic encephalopathy diagnosis	HIEDiagnosis	R	HYPOXIC ISCHAEMIC ENCEPHALOPATHY GRADE (HIGHEST ON NEONATAL CRITICAL CARE DAILY CARE DATE)	n	0 none 1 1-mild 2 2-moderate 3 3-severe	The highest Hypoxic Ischaemia Encephalopathy diagnosis on this day of care	Used to determine level of care
Daily	Ophthalmology	ROP screen	ROPScreen	R	RETINOPATHY OF PREMATURITY SCREENING PERFORMED INDICATOR	an1	N no Y yes	Specify if the baby has had a screen for ROP on this day of care	Used in the National Neonatal Audit Programme
Daily	Ophthalmology	ROP surgery	ROPSurgery	R	Derived from procedures	n2	Derived 0 no 1 yes	Derived item from any major surgery today; 'yes' if a ROP surgical procedure is recorded and 'no' if not	Used in the National Neonatal Audit Programme
Daily	Fluids and feeding	Fluids and feeding: vascular lines in situ	LinesIn	R	VASCULAR LINE TYPE IN SITU	n2	1 peripheral arterial line 2 umbilical arterial line 3 umbilical venous line 4 percutaneous central venous line (long line) 5 surgically inserted central venous line 6 peripheral venous line 9 not applicable/no lines in situ	Specify any line that is in situ for any time during this day	Used to determine level of care
Daily	Fluids and feeding	Fluids and feeding: parenteral nutrition today (partial or total)	ParenteralNutrition	R	PARENTERAL NUTRITION RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby has received any parenteral nutrition on this day of care	Used to determine level of care
Daily	Fluids and feeding	Fluids and feeding: intravenous glucose and electrolyte solutions	GlucoseElectrolytes	R	INTRAVENOUS INFUSION OF GLUCOSE AND ELECTROLYTE SOLUTION RECEIVED INDICATOR	an1	N no Y yes	Specify if a baby has received an intravenous infusion of a glucose and electrolyte solution on this day of care	Used to determine level of care

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
Daily	Fluids and feeding	Fluids and feeding: enteral feeding given	DayEnteralFeed	R	ENTERAL FEED TYPE GIVEN	an250	1 suckling at the breast 2 mother's fresh expressed breast milk 3 mother's frozen expressed breast milk 4 donor expressed breast milk 5 breast milk fortifier 6 formula 9 not applicable (nil by mouth)	Types of milk given during this 24-hour period (00.00–23.59). Record all the types of milk given during the day. Nil by mouth is single choice and only selected if no enteral feed for the whole of the 24-hour period	Used in the National Neonatal Audit Programme
Daily	Fluids and feeding	Fluids and feeding: formula type	DayFormulaTypedmd	R	FORMULA MILK TYPE (SNOMED CT DM+D)	max n18	dm+d virtual therapeutic moiety listing	If enteral feeding is formula, specify the name of the formula. List will be maintained by d-dm+d	Used to monitor development of infant
Daily	Fluids and feeding	Fluids and feeding: Formula type	DayFormulaType	R	FORMULA MILK TYPE (bespoke)	an2	10 Nutriprem 1 11 Nutriprem 2 12 Neocate 13 Nutramigen 14 Pepti Junior 15 Infatrini 16 SMA High energy 17 Aptamil First milk 18 Cow & Gate First infant milk 19 SMA First infant milk 20 Enfamil A.R. 21 Aptamil Preterm 22 SMA Gold Prem 1	If enteral feeding is formula, specify the name of the formula. This is a bespoke list used instead of dm+d listing if it is unavailable. Multiple selections allowed	Used to monitor development of infant

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							23 SMA Gold Prem 2 24 Aptamil Pepti 1 25 Aptamil Pepti 2 26 Pregestimil 27 Caprilon 28 Wysoy 29 Infasoy 88 Other		
Daily	Fluids and feeding	Fluids and feeding: measured volume milk	Totalvolume	R	TOTAL VOLUME OF MILK RECEIVED	an4	000.0–999.9	The volume of milk a baby has been fed on this day of care in ml	Used to monitor development of infant
Daily	Fluids and feeding	Fluids and feeding: method of feeding	Methodfeeding	O	ENTERAL FEEDING METHOD	n2	1 breast 2 bottle 3 cup 4 nasogastric tube 5 orogastric tube 6 gastrostomy 7 nasojejunal tube 8 other	The method of feeding at this 24-hour period (00.00–23.59). Record all methods of feeding given during the day. No enteral feeds is single choice and only selected if no milk given for the whole of the 24-hour period	Used in the National Neonatal Audit Programme
Daily	Infections	Infection: suspected sepsis	DaySuspectedSepsis	R	SEPSIS SUSPECTED INDICATOR	an1	N no Y yes	Specify if you believe the baby may have a blood stream infection on this day of care	Used in the National Neonatal Audit Programme
Daily	Jaundice	Jaundice: phototherapy	Phototherapy	R	PHOTOTHERAPY RECEIVED INDICATOR	an1	N no Y yes	Specify if this baby has received a phototherapy treatment for jaundice on this day of care	Used to determine level of care
Daily	General information	General information: level of care (2001 definition)	LevelOfCare2001(derived)	O	Derived based on the BAPM categories of care 2001	n1	Derived 1 intensive care	Applying the BAPM 2001 categories of care definition	Used in commissioning

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							2 high-dependency care 3 special care		
Daily	General information	General Information: level of care (2011 definition)	LevelOfCare2011(derived)	O	Derived based on the BAPM categories of care 2011	n1	Derived 1 intensive care 2 high-dependency care 3 special care 4 normal care	Applying the BAPM 2011 categories of care definition	Used in commissioning
Daily	Medication	Drugs given: medications given on this day	DailyDrugs	R	MEDICATION GIVEN DURING NEONATAL CRITICAL CARE DAILY CARE DATE (SNOMED CT DM+D)	n18	dm+d code for any drug	Specify the exact medications the baby has received on this day of care	Used for multiple purposes including National Neonatal Audit Programme, monitor outcomes, and derive items for level of care where necessary
NNUAdhoc	Infection cultures	Culture indicator – blood	CultureIndicatorBLD	P	INFECTION CULTURE TEST INDICATOR (BLOOD)	an1	N no Y yes	Specify if at least one blood culture was taken in this episode of care	Used in the National Neonatal Audit Programme
NNUAdhoc	Infection cultures	Culture indicator – CSF	CultureIndicatorCSF	P	INFECTION CULTURE TEST INDICATOR (CEREBROSPINAL FLUID)	an1	N no Y yes	Specify if at least one CSF culture was taken in this episode of care	Used in the National Neonatal Audit Programme
NNUAdhoc	Infection cultures	Culture indicator – urine	CultureIndicatorURN	P	INFECTION CULTURE TEST INDICATOR (URINE)	an1	N no Y yes	Specify if at least one urine culture was taken in this episode of care	Used in the National Neonatal Audit Programme
NNUAdhoc	Infection cultures	Date and time future sample taken – year	CultureDateTimeYear	R	SAMPLE COLLECTION YEAR AND MONTH	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID	Used in the National Neonatal Audit Programme
NNUAdhoc	Infection cultures	Type of culture sample taken	Sampletype	R	SAMPLE TYPE (NATIONAL NEONATAL DATA SET)	n	1 blood 2 urine (suprapubic) 3 urine (catheterisation) 4 urine (clean catch) 5 cerebrospinal fluid	Specify the type of culture samples taken at the relevant date and time	Used in the National Neonatal Audit Programme

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUAdhoc	Infection cultures	Clinical signs at time of culture	Clinicalsigns	R	CLINICAL SIGN OBSERVED AT SAMPLE COLLECTION	an250	01 increased oxygen requirement or ventilator support 02 lethargy/irritability/ poor handling 03 temperature instability 04 ileus/onset of poor feed tolerance 05 fall in urine output 06 impaired peripheral perfusion (capillary refill time > 3 seconds/pallor/ mottling/core-peripheral temperature gap > 2 °C) 07 increase in apnoea/ bradycardia 08 hypotension 09 glucose intolerance 10 metabolic acidosis/ base deficit < 10 mmol/l	Specify the clinical signs observed when the culture sample was taken at the relevant date and time	Used in the National Neonatal Audit Programme
NNUAdhoc	Infection cultures	Result of culture	Pathogen	R	SAMPLE TEST RESULT ORGANISM TYPE (SNOMED CT)	n18	Organisms list in Snomed CT (Concept ID 410607006)	Specify the culture results if there was growth or no growth by selecting the relevant options in the list of codes related to this item	Used in the National Neonatal Audit Programme
NNUAdhoc	Infection cultures	Sensitivity of culture	Sensitivity	O	SAMPLE ANTIBIOTIC SENSITIVITY RESULT (SNOMED CT DM+D)	n18	dm+d virtual therapeutic moiety listing	The sensitivity results obtained from the microbiology report of the culture sample taken at the relevant date and time	Used in the National Neonatal Audit Programme
NNUAdhoc	Abdominal X-rays	Abdominal X-ray indicator	ADXIndicator	P	ABDOMINAL X-RAY PERFORMED INDICATOR	an1	N no Y yes	Specify if at least one abdominal X-ray was performed in this episode of care	Used to monitor health status and outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUAdhoc	Abdominal X-rays	Date and time of abdominal X-ray – year	AbdominalXrayDateTime Year	R	PROCEDURE YEAR AND MONTH (ABDOMINAL X-RAY)	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	Was X-ray performed to investigate abdominal signs?	InvestigateAbdsigns	R	ABDOMINAL X-RAY PERFORMED TO INVESTIGATE ABDOMINAL SIGNS INDICATOR	an1	N no Y yes	Specify if the X-ray was performed to investigate abdominal signs	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	X-ray appearance?	XrayAppearance	R	CONDITION SEEN IN ABDOMEN DURING X-RAY	an50	1 pneumatosis 2 air in the liver 3 pneumoperitoneum 4 fixed loop 5 gasless 9 none of the above	Specify if any of these appear in this abdomen X-ray	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	Clinical findings abdominal X-ray	abdominalxrayfindings	R	ABDOMINAL X-RAY PERFORMED REASON	an50	01 abdominal distension 02 abdominal tenderness 03 increased/bilious aspirates 04 abdominal discolouration 05 abdominal mass 06 bloody stools 07 mucusy stools 09 none of the above	Specify the clinical reasons that resulted in this abdominal X-ray taking place	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	This episode of care: Baby transferred out for management of NEC	TransferredOut ManagementNEC	R	TRANSFERRED FROM NEONATAL INTENSIVE CARE UNIT FOR NECROTISING ENTEROCOLITIS MANAGEMENT INDICATOR	an1	N no Y yes	Specify if the baby was transferred out of the neonatal unit for further management of NEC following this abdominal X-ray	Used to monitor health status and outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUAdhoc	Abdominal X-rays	Was laparotomy for NEC required?	necLaparotomy	R	LAPAROTOMY FOR NECROTISING ENTERCOLITIS INDICATION CODE	*n*	0 laparotomy not required 1 laparotomy required but patient too ill to carry it out 2 laparotomy required and carried out	Specify if a laparotomy for NEC was required based on this abdominal X-ray	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	If laparotomy done, did visual inspection confirm NEC	laparotomyConfirm	R	VISUAL INSPECTION CONFIRMED NECROTISING ENTEROCOLITIS DURING LAPAROTOMY INDICATOR	an1	N no Y yes	Specify if a visual inspection confirmed NEC from a laparotomy carried out following this abdominal X-ray	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	Was NEC histology confirmed?	necHistologyConfirmed	R	HISTOLOGY-CONFIRMED NECROTISING ENTEROCOLITIS FOLLOWING LAPAROTOMY INDICATOR	*n*	0 not confirmed 1 yes confirmed 9 no histological inspection/not applicable	Specify if histology confirmed NEC based on the laparotomy carried out following this abdominal X-ray	Used to monitor health status and outcomes
NNUAdhoc	Abdominal X-rays	Was peritoneal drain inserted after abdominal X-ray?	WasPeritonealdrain Inserted	R	PERITONEAL DRAIN INSERTED FOLLOWING ABDOMINAL X-RAY INDICATOR	an1	N no Y yes	Specify if a peritoneal drain was inserted after this abdominal X-ray	Used to monitor health status and outcomes
NNUAdhoc	ROP screening	ROP indicator	ROPindicator	M	RETINOPATHY OF PREMATURITY SCREENING PERFORMED INDICATOR	an1	N no Y yes	Specify if a ROP screening was performed in this episode of care	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	Date and time of ROP test – year	ROPtestDateTimeYear	R	PROCEDURE YEAR AND MONTH (RETINOPATHY OF PREMATURITY SCREENING)	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	Hospital performing test	HospitalPerforminTest	R	SITE CODE (OF RETINOPATHY OF PREMATURITY SCREENING) or ORGANISATION CODE (OF RETINOPATHY OF PREMATURITY SCREENING)	an20	Use organisation code and site code ZZ201 – not applicable (intended to deliver at home) ZZ888 – not applicable (intended to deliver at non-NHS organisation)	The code for the hospital that is recording the ROP screening at this event in this episode of care	Used in the National Neonatal Audit Programme

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
							ZZ203 – not known (intended place of delivery not known)		
NNUAdhoc	ROP screening	ROP stage – left eye	ROPStageLeft	R	RETINOPATHY OF PREMATURITY STAGE (LEFT EYE)	an1	0 no ROP 1 stage 1 ROP 2 stage 2 ROP 3 stage 3 ROP 4 stage 4 ROP 5 stage 5 ROP A aggressive posterior ROP	The ROP stage for the left eye at the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP stage – right eye	ROPStageRight	R	RETINOPATHY OF PREMATURITY STAGE (RIGHT EYE)	an1	0 no ROP 1 stage 1 ROP 2 stage 2 ROP 3 stage 3 ROP 4 stage 4 ROP 5 stage 5 ROP A aggressive posterior ROP	The ROP stage for the right eye at the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP clock hours max stage – left eye	ROPClockHourLeft	R	RETINOPATHY OF PREMATURITY CLOCK HOURS MAXIMUM STAGE (LEFT EYE)	n2	0–12	Number of clock hours affected by maximum stage of ROP in left eye – shown as number from 0 to 12	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP clock hours max stage – right eye	ROPClockHourRight	R	RETINOPATHY OF PREMATURITY CLOCK HOURS MAXIMUM STAGE (RIGHT EYE)	n2	0–12	Number of clock hours affected by maximum stage of ROP in right eye – shown as number from 0 to 12	Used in the National Neonatal Audit Programme

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUAdhoc	ROP screening	ROP maximum zone – left eye	ROPMaxZoneLeft	R	RETINOPATHY OF PREMATURITY CLOCK HOURS MAXIMUM ZONE (LEFT EYE)	n	0 no ROP 1 zone 1 2 zone 2 3 zone 3	The ROP maximum zone for the left eye at the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP maximum zone – right eye	ROPMaxZoneRight	R	RETINOPATHY OF PREMATURITY CLOCK HOURS MAXIMUM ZONE (RIGHT EYE)	n	0 no ROP 1 zone 1 2 zone 2 3 zone 3	The ROP maximum zone for the right eye at the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP plus disease – left eye	ROPPlusDiseaseLeft	R	RETINOPATHY OF PREMATURITY PLUS DISEASE STATUS (LEFT EYE)	n	0 no plus disease 1 pre-plus disease 2 plus disease	The plus disease status of the left eye of the baby at the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP plus disease – right eye	ROPPlusDiseaseRight	R	RETINOPATHY OF PREMATURITY PLUS DISEASE STATUS (RIGHT EYE)	n	0 no plus disease 1 pre-plus disease 2 plus disease	The plus disease status of the right eye of the baby at the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	ROP screening	ROP outcome	ROPOutcome	R	RETINOPATHY OF PREMATURITY SCREENING OUTCOME STATUS CODE	n	0 No ROP 1 ROP Follow-up screening required 2 ROP diagnosed, treatment needed 3 ROP diagnosed, transferred out of neonatal unit for treatment	Specify the outcome of the relevant ROP screening	Used in the National Neonatal Audit Programme
NNUAdhoc	Cranial ultrasound scan	Cranial ultrasound scan indicator	CranialUSSIndicator	P	CRANIAL ULTRASOUND SCAN INDICATOR	an1	N no Y yes	Specify if a cranial ultrasound scan test was performed in this episode of care	Used in the National Neonatal Audit Programme
NNUAdhoc	Cranial ultrasound scan	Date and time of cranial ultrasound scan test – year	CarnialUSSDateTimeYear	R	PROCEDURE YEAR AND MONTH (CRANIAL ULTRASOUND SCAN)	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID	Used to monitor health status and outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings (left): IVH	LeftIVH	O	INTRAVENTRICULAR HAEMORRHAGE GRADE (LEFT SIDE)	n2	0 no IVH seen 1 grade 1 (germinal matrix) IVH 2 grade 2 IVH 3 grade 3 IVH 4 grade 4 IVH	Most severe grade of intraventricular haemorrhage seen on left side	Used to monitor health status and outcomes
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings (left): porencephalic cyst(s)	LeftPorencephalic	O	PORENCEPHALIC CYST VISIBLE DURING CRANIAL ULTRASOUND SCAN INDICATOR (LEFT SIDE)	an1	Y porencephalic cyst visible on left side N no porencephalic cyst on left side	Records if there was a porencephalic cyst visible on left side	Used to monitor health status and outcomes
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings (left): ventricular dilatation	LeftDilation	O	VENTRICULAR DILATATION DIAGNOSED DURING CRANIAL ULTRASOUND SCAN INDICATOR (LEFT SIDE)	an1	Y yes ventricular dilatation on right side N no, ventricle on right side not dilated	Records if clinical diagnosis made of ventricular dilatation on left side	Used to monitor health status and outcomes
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings (right): IVH	RightIVH	O	INTRAVENTRICULAR HAEMORRHAGE GRADE (RIGHT SIDE)	n1	0 no IVH seen 1 grade 1 (germinal matrix) IVH 2 grade 2 IVH 3 grade 3 IVH 4 grade 4 IVH	Most severe grade of intraventricular haemorrhage seen on right side	Used to monitor health status and outcomes
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings (right): Porencephalic cyst(s)	RightPorencephalic	O	PORENCEPHALIC CYST VISIBLE DURING CRANIAL ULTRASOUND SCAN INDICATOR (RIGHT SIDE)	an2	Y porencephalic cyst visible on right side N no porencephalic cyst on right side	Records if there was a porencephalic cyst visible on right side	Used to monitor health status and outcomes
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings (right): ventricular dilatation	RightDilation	O	VENTRICULAR DILATATION DIAGNOSED DURING CRANIAL ULTRASOUND SCAN INDICATOR (RIGHT SIDE)	an2	Y yes ventricular dilatation on right side N No, ventricle on right side not dilated	Records if clinical diagnosis made of ventricular dilatation on right side	Used to monitor health status and outcomes
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings: cystic PVL	PVL	O	CYSTIC PERIVENTRICULAR LEUKOMALACIA OBSERVED DURING CRANIAL ULTRASOUND SCAN INDICATOR	an2	Y yes, cystic PVL on scan N no cystic PVL seen on scan	Records if there is any evidence of cystic PVL on the scan	Used to monitor health status and outcomes

Category	Category detail	Item name	Field ID	Data dictionary item status	NHS data dictionary item name	Format/length	Coding/allowed values	Description	Purpose
NNUAdhoc	Cranial ultrasound scan	Cranial scan findings: post haemorrhagic hydrocephalus	hydorcephalus	O	POST HAEMORRHAGIC HYDROCEPHALUS OBSERVED DURING CRANIAL ULTRASOUND SCAN INDICATOR	an2	Y yes, post haemorrhagic hydrocephalus on scan N no post haemorrhagic hydrocephalus seen on scan	Records if there is post haemorrhagic hydrocephalus evident on scan	Used to monitor health status and outcomes
NNUAdhoc	Biochemical screening	Blood spot test indicator	BloodSpotTestIndicator	M	NEWBORN BLOOD SPOT TEST PERFORMED INDICATOR	an1	N no Y yes	Specify if a blood spot test was performed	
NNUAdhoc	Biochemical screening	Date of blood spot test – year	BloodSpotTestYear	R	BLOOD SPOT CARD COMPLETION YEAR AND MONTH	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID	Used to monitor health status and outcomes
NNUAdhoc	Newborn hearing screening	Hearing test indicator	HearingTestIndicator	P	NEWBORN HEARING SCREENING PERFORMED INDICATOR	an1	N no Y yes	Specify if a hearing test was performed in this episode of care	Used to monitor health status and outcomes
NNUAdhoc	Newborn hearing screening	Date and time of hearing test – year	HearingtestDateTimeYear	R	PROCEDURE YEAR AND MONTH (NEWBORN HEARING SCREENING)	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID	Used to monitor health status and outcomes
NNUAdhoc	Newborn hearing screening	This episode of care: hearing screen result (left ear)	HearingTestLeft	O	NEWBORN HEARING SCREENING OUTCOME LEFT EAR (NATIONAL NEONATAL DATA SET)	n	1 passed 2 fail 9 not done (default)	Specify the result of the hearing screening for the left ear	Used to monitor health status and outcomes
NNUAdhoc	Newborn hearing screening	This episode of care: hearing screen result (right ear)	HearingTestRight	O	NEWBORN HEARING SCREENING OUTCOME RIGHT EAR (NATIONAL NEONATAL DATA SET)	n	1 passed 2 fail 9 not done (default)	Specify the result of the hearing screening for the right ear	Used to monitor health status and outcomes
NNUAdhoc	Two Year Follow-up	Special questions – why was child difficult to test?	SpecialQuestions DifficultToTestReason	R	CHILD DIFFICULT TO TEST REASON CODE	an1	A child was tired B poor attention C difficult to engage D other reason	Specify the reason the child was difficult to test	Use in assessing two year health outcomes following discharge from neonatal care

BA, Bachelor of the Arts; BETC, Business and Technology Education Council; BSc, Bachelor of Science; dm + d, Dictionary of Medicines and Devices; GNVQ, General National Vocational Qualification; H&C, health and care; HNC, Higher National Certificate; HND, Higher National Diploma; MD5, message-digest algorithm 5; NCCMDS, National Critical Care Minimum Data Set; NVQ, National Vocational Qualification; ONC, Ordinary National Certificate; OND, Ordinary National Diploma; PCT, primary care trust; PGCE, Postgraduate Certificate in Education; PVL, periventricular leucomalacia; RSA, Royal Society of Arts; VCE, Victorian Certificate of Education.

Appendix 4 National Information Governance Board Confidentiality Advisory Group Approval

NIGB

Ethics and Confidentiality Committee

[redacted]

5th Floor, Skipton House
80 London Road
London
SE1 6LH
Tel: [redacted]
Email: [redacted]

27 January 2012

Dear Professor [redacted]

ECC 8-05(f)/2010 - A National Neonatal Research Database

Thank you for your application for approval under the Health Service (Control of Patient Information) Regulations 2002 to process patient identifiable information outside the common law duty of confidentiality. This application was considered by the Department of Health on behalf of the Secretary of State.

The role of the ECC is to review applications submitted under these Regulations and to provide advice to the Secretary of State for Health (SofS). Following consideration of the advice, reproduced below, the Department of Health on behalf of the Secretary of State has determined that the application should be approved.

Context

This application proposed to set up a national neonatal research database to be used as a research resource. Support under the Health Service (Control of Patient Information) Regulations 2002 was sought to enable routinely collected patient identifiable data to be populated on this research database. In addition to routinely collected clinical data, the application also included requested use of NHS Number, hospital ID, date of birth, date of death, postcode, gender, ethnicity, maternal NHS number and ethnicity, and postcode of infant at two years old.

ECC Advice

This application was considered by the Ethics and Confidentiality Committee at its meeting on 01 December 2010. Members agreed that this application set out a worthy purpose, would provide a number of benefits and the additional clarity provided following queries was welcomed. In particular, the Committee noted that user involvement had been very good and the continued engagement was to be commended.

Members agreed that as a whole, consent would not be practicable due to the large numbers involved (60,000). However, the Committee was mindful that decisions taken under section 251 could not be inconsistent with the Data Protection Act (DPA) 1998. In particular, Members focused upon compliance with the fair processing aspect of the first principle of this Act. While section 251 permits access to patient identifiable information without consent, Members are clear that the

NIGB

Ethics and Confidentiality Committee

parents of the cohort should be informed where reasonably possible, thus fulfilling compliance with the fairness aspect of the first principle of the DPA. In addition, the sixth principle permits a number of rights, one of which is the right for data subjects to request that the processing of personal data should cease if demonstrated to cause damage and distress. As such, all applicants should have a mechanism in place to manage this aspect, and Members requested details on how these aspects would be implemented within the application. It was suggested that those involved in patient engagement could be consulted as this would aid in developing an appropriate mechanism.

Members highlighted that a principle of the Committee is that there should not be long-term retention of identifiable data without consent. Members could not identify a clear justification to support this long-term retention, and requested that this issue be investigated by the applicant in order to reduce the identifiability. In particular, Members queried why the HESID could not be used as the key identifier once linkage with HES data had been carried out. It was recommended that the applicant discuss this feasibility with the HES team at the NHS Information Centre.

Members also noted the onward disclosure aspect to the application and queried what controls were in place to ensure that recipients would not seek to render the information more identifiable. Members expected that a data access policy would be in place to handle this situation and requested sight of this, and/or any other controls in place.

Based upon the considerations above, the Committee agreed to recommend provisional support under the Regulations to this study. This was subject to the following clarifications, specific and standard conditions of support. The NIGB Office has now received your response to the request for clarification and to the specific conditions of approval as stated in bold below.

Request for clarification

1. Please investigate options to reduce the identifiability of information in line with the Committee's views on long-term retention of identifiable data without consent. You might find it helpful to contact the NHS Information Centre to explore this option further. Please also consider reducing identifiability of the data items. **We agree with the suggestion of the ECC that the infant HES ID is used as the key identifier once linkage with HES data is carried out. We propose to remove and do not intend to retain either infant or mother's NHS number after HES linkage has been carried out. We will retain the unique anonymous infant ID issued by Clevermed, the NHS hosting company for the electronic records.**

2. Please provide the relevant documentation to support any onward disclosure, such as a data access policy that is in place. **The NDAU data access policy (Neonatal Data Analysis Unit Data Management Version 1/07 October 2010) is provided. (Received by NIGB Office on 01 September 2011)**

Specific conditions of support

1. Confirmation of satisfactory security arrangements. **(Confirmation received from the Department of Health Security Review Team on 08 September 2011)**

2. Provision of a favourable opinion from a research ethics committee. **(Received on 01 September 2011)**

NIGB

Ethics and Confidentiality Committee

3. Provision of fair processing information to inform parents about this activity and to facilitate a mechanism to permit dissent to be registered and managed. **(Received on 25 November 2011)**

4. This approval covers only data generated within England and Wales; Scotland and Northern Ireland fall outwith the remit of section 251. **(Accepted by applicant on 01 September 2011)**

After reviewing your responses above, the Committee agreed that the minimum criteria under the Regulations appeared to have been met, and therefore advised recommending support to the Secretary of State for Health.

Conclusion

As the conditions of support have now been accepted and met, this letter provides final confirmation that the Secretary of State for Health had approved your application to process patient identifiable information outside the common law duty of confidentiality.

Annual Review

Please note that your approval is subject to submission of an annual review report to show how you have met the above conditions or report plans, and action towards meeting them. It is also your responsibility to submit this report on the anniversary of your final approval (the date of this letter) and to report any changes such as to the purpose or design of the proposed activity, or to security and confidentiality arrangements.

The Register of approved applications will be updated on our website to include this application, http://www.nigb.nhs.uk/s251/registerapp

Please do not hesitate to contact me if you have any queries following this letter, I would be grateful if you could quote the above reference number in all future correspondence.

Yours sincerely

[redacted]

NIGB Deputy Operations Manager

NIGB

Ethics and Confidentiality Committee

Standard conditions

The approval provided by the Secretary of State for Health is subject to the following standard conditions.

The applicant will ensure that:

1. The specified patient identifiable information is only used for the purpose(s) set out in the application.
2. Confidentiality is preserved and that there is no disclosure of information in aggregate or patient level form that may inferentially identify a person, nor will any attempt be made to identify individuals, households or organisations in the data.
3. Requirements of the Statistics and Registration Services Act 2007 are adhered to regarding publication when relevant.
4. All staff with access to patient identifiable information have contractual obligations of confidentiality, enforceable through disciplinary procedures.
5. All staff with access to patient identifiable information have received appropriate ongoing training to ensure they are aware of their responsibilities.
6. Activities are consistent with the Data Protection Act 1998.
7. Audit of data processing by a designated agent of the Secretary of State is facilitated and supported.
8. The wishes of patients who have withheld or withdrawn their consent are respected.
9. The NIGB Office is notified of any significant changes (purpose, data flows, security arrangements) to the application.
10. An annual report is provided no later than 12 months from the date of your final confirmation letter. Details are available on the NIGB website.

NIGB

Ethics and Confidentiality Committee
On behalf of the Secretary of State of Health

[redacted]

5th Floor
Skipton House
80 London Road
London
SE1 6LH.
Tel: [redacted]

5 March 2013

[redacted]

ECC 8-05(f)/2010 – National Neonatal Research Database

Thank you for the provision of your annual review report under the Health Service (Control of Patient Information) Regulations 2002.

The purpose of the annual review is to provide an update against the conditions of approval where applicable, confirm progress of the study, review the need to process confidential patient information, and ensure the minimum amount of identifiable information is being used.

The NIGB Ethics and Confidentiality Committee provides advice to the Secretary of State for Health on whether an application under these Regulations should or should not be approved, along with any applicable conditions, and the Secretary of State for Health (SofS) considers this advice and takes the final decision.

Secretary of State decision

Having considered the annual review, and following advice from the NIGB Ethics and Confidentiality Committee, the Secretary of State for Health has approved the continued processing of this application for the specified purposes for a further 12 months.

Context

This research application from Chelsea & Westminster NHS Foundation Trust proposed to set up a national neonatal research database to be used as a research resource. Section 251 support was sought to enable routinely collected patient identifiable data to be populated on this research database. In addition to routinely collected clinical data, the application also included requested use of NHS Number, hospital ID, date of birth, date of death, postcode, gender, ethnicity, maternal NHS number and ethnicity, and postcode of infant at two years old.

ECC Advice

Following assessment of the annual review documentation, it was advised that there was a continued justification for access to confidential patient information as stated in the original application. The high public interest in the activity continuing was also noted.

In particular, the following was noted:

1. There had been no changes to the existing security policy.

NIGB

Ethics and Confidentiality Committee
On behalf of the Secretary of State of Health

2. There is still a continued need to access confidential patient information as specified within the original application.

3. The continued patient and service user engagement was noted.

4. As a whole, it was recommended that the approval in place for the purposes set out in the application should continue for a further 12 months. It was noted that the IG toolkit score for Chelsea and Westminster was shown to be 94%, however this is shown to be an unsatisfactory score. Based on the security documentation supplied at the time of original application consideration, no further action is required, but in line with our published procedures all those submitting amendments/annual reviews/applications after 01 April 2013 will be required to provide evidence of a satisfactory IG Toolkit score.

Annual Review

Please note that your approval is subject to submission of an annual review report to show how you have met the conditions or report plans, and action towards meeting them. It is also your responsibility to submit this report on the anniversary of your final approval and to report any changes such as to the purpose or design of the proposed activity, or to security and confidentiality arrangements. An annual review should be provided no later than 12 months from the date of this letter.

IMPORTANT CHANGES

Administration of applications

Please note that the current administration of applications made under these Regulations by the NIGB Ethics and Confidentiality Committee is due to transfer to the Health Research Authority by 01 April 2013, therefore please be advised that arrangements might have changed by the time the next annual review is due. Such arrangements will be communicated once confirmed.

Security review

Please note that due to a change in Department of Health policy, all bodies processing NHS data will be expected to provide up to date assurance of their security arrangements via the Information Governance Toolkit instead of system level security policy submission. Details on this change are available here http://www.nigb.nhs.uk/s251/security%20review. Please note that prior to your next annual review you will need to have provided a relevant IG Toolkit submission to the IG Toolkit Team. Any queries on this aspect should be directed to [redacted] so as to ensure there are no delays to any future continuing approval.

Yours sincerely,

[redacted]

National Information Governance Board for Health and Social Care

NIGB

Ethics and Confidentiality Committee
On behalf of the Secretary of State of Health

Standard conditions

The approval provided by the Secretary of State for Health is subject to the following standard conditions.

The applicant will ensure that:

1. The specified patient identifiable information is only used for the purpose(s) set out in the application.
2. Confidentiality is preserved and that there is no disclosure of information in aggregate or patient level form that may inferentially identify a person, nor will any attempt be made to identify individuals, households or organisations in the data.
3. Requirements of the Statistics and Registration Services Act 2007 are adhered to regarding publication when relevant.
4. All staff with access to patient identifiable information have contractual obligations of confidentiality, enforceable through disciplinary procedures.
5. All staff with access to patient identifiable information have received appropriate ongoing training to ensure they are aware of their responsibilities.
6. Activities are consistent with the Data Protection Act 1998.
7. Audit of data processing by a designated agent of the Secretary of State is facilitated and supported.
8. The wishes of patients who have withheld or withdrawn their consent are respected.
9. The NIGB Office is notified of any significant changes (purpose, data flows, security arrangements) to the application.
10. An annual report is provided no later than 12 months from the date of your final confirmation letter. Details are available on the NIGB website.
11. Any breaches of security around this particular flow of data should be reported to the NIGB within 10 working days, along with remedial actions taken.

National Information Governance Board for Health and Social Care

NHS
Health Research Authority
Confidentiality Advisory Group

Skipton House
80 London Road
London
SE1 6LH

Tel: [redacted]
Email [redacted]

[redacted]

[redacted]

[redacted]

13 February 2014

Study title:	**National Neonatal Research Database**
CAG reference:	**ECC 8-05(f)/2010**
Protocol number:	**1**
IRAS Project ID:	**22304/158892/11/253**
REC number:	**10/H0803/151**

Thank you for the provision of an annual review report for your research application, submitted for approval under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002 to process patient identifiable information without consent. Approved applications enable the data controller to provide specified information to the applicant for the purposes of the relevant activity, without being in breach of the common law duty of confidentiality, although other relevant legislative provisions will still be applicable.

The role of the Confidentiality Advisory Group (CAG) is to review applications submitted under these Regulations and to provide advice to the Health Research Authority on whether an application should be approved, and if so, any relevant conditions. The purpose of the annual review is to provide an update against the conditions of approval where applicable, confirm progress of the study, review the need to process confidential patient information, and ensure the minimum amount of identifiable information is being used.

Health Research Authority approval decision

The Health Research Authority, having considered the advice from the Confidentiality Advisory Group as set out below, has approved the continued processing of this application for the specified purposes for a further 12 months from the anniversary of your original final approval outcome letter, therefore until 27 January 2015

Context

Purpose of application

This research application from Chelsea & Westminster NHS Foundation Trust proposed to set up a national neonatal research database to be used as a research resource. Section 251 support was sought to enable routinely collected patient identifiable data to be populated on this research database.

Confidential patient information requested

In addition to routinely collected clinical data, the application also included requested use of NHS Number, hospital ID, date of birth, date of death, postcode, gender, ethnicity, maternal, NHS number and ethnicity, and postcode of infant at two years old.

Confidentiality Advisory Group advice

1. A satisfactory Information Governance Toolkit was confirmed with IG Toolkit Helpdesk, 02/02/2014.

2. There is still a continued need to access confidential patient information as specified within the original application.

3. As a whole, it was recommended that the approval in place for the purposes set out in the application should continue for a further 12 months from the anniversary of the original final approval outcome letter, to the date specified below.

Annual Review

Please note that your approval is subject to submission of an annual review report to show how you have met the conditions or report plans, and action towards meeting them. It is also your responsibility to submit this report on the anniversary of your final approval and to report any changes such as to the purpose or design of the proposed activity, or to security and confidentiality arrangements. We are also streamlining the process to facilitate the service we provide to applicants. This means that annual reviews will be batched and reviewed on the last day of the preceding month before the date of approval. An annual review should therefore be provided no later than 31 December 2014 and preferably 4 weeks before this date.

Please do not hesitate to contact me if you have any queries following this letter. I would be grateful if you could quote the above reference number in all future correspondence.

Yours sincerely

[redacted]

Email: [redacted]

Enclosures: Standard conditions of approval

Copy to: NRESCommittee.London-WestLondon@nhs.net

NHS
Health Research Authority
Confidentiality Advisory Group

Standard conditions of approval

The approval provided by the Health Research Authority is subject to the following standard conditions.

The applicant will ensure that:

1. The specified patient identifiable information is only used for the purpose(s) set out in the application.
2. Confidentiality is preserved and there are no disclosures of information in aggregate or patient level form that may inferentially identify a person, nor will any attempt be made to identify individuals, households or organisations in the data.
3. Requirements of the Statistics and Registration Services Act 2007 are adhered to regarding publication when relevant.
4. All staff with access to patient identifiable information have contractual obligations of confidentiality, enforceable through disciplinary procedures.
5. All staff with access to patient identifiable information have received appropriate ongoing training to ensure they are aware of their responsibilities.
6. Activities are consistent with the Data Protection Act 1998.
7. Audit of data processing by a designated agent is facilitated and supported.
8. The wishes of patients who have withheld or withdrawn their consent are respected.
9. The Confidentiality Advice Team is notified of any significant changes (purpose, data flows, data items, security arrangements) prior to the change occurring.
10. An annual report is provided no later than 12 months from the date of your final confirmation letter.
11. Any breaches of confidentiality / security around this particular flow of data should be reported to CAG within 10 working days, along with remedial actions taken / to be taken.

NHS
Health Research Authority

[redacted]

Skipton House
80 London Road
London
SE1 6LH

Tel [redacted]
Email [redacted]

3rd March 2015

[redacted]

Study title: National Neonatal Research Database
CAG reference: ECC 8-05(f)/2010
Protocol number: 1
IRAS Project ID: 22304/158892/11/253
REC number: 10/H0803/151

Thank you for the provision of an annual review report for your research application, submitted for approval under Regulation 5 of the Health Service (Control of Patient Information) Regulations 2002 to process patient identifiable information without consent. Approved applications enable the data controller to provide specified information to the applicant for the purposes of the relevant activity, without being in breach of the common law duty of confidentiality, although other relevant legislative provisions will still be applicable.

The role of the Confidentiality Advisory Group (CAG) is to review applications submitted under these Regulations and to provide advice to the Health Research Authority on whether an application should be approved, and if so, any relevant conditions. The purpose of the annual review is to provide an update against the conditions of approval where applicable, confirm progress of the study, review the need to process confidential patient information, and ensure the minimum amount of identifiable information is being used.

Health Research Authority approval decision

The Health Research Authority, having considered the advice from the Confidentiality Advisory Group as set out below, has approved the continued processing of this application for the specified purposes for a further 12 months from the anniversary of your original final approval outcome letter, therefore until 27th January 2016.

Context

Purpose of application

This research application from Chelsea & Westminster NHS Foundation Trust proposed to set up a national neonatal research database to be used as a research resource. Section 251 support was sought to enable routinely collected patient identifiable data to be populated on this research database.

Confidential patient information requested

In addition to routinely collected clinical data, the application also included requested use of NHS Number, hospital ID, date of birth, date of death, postcode, gender, ethnicity,

maternal, NHS number and ethnicity, and postcode of infant at two years old. Following initial linkage to HES data at the Health and Social Care Information Centre (HSCIC) HESID only would be used for continued linkage and NHS number would not be retained following initial linkage.

Confidentiality Advice Team advice

Security arrangements

A satisfactory Information Governance Toolkit score of 87% was noted.

Study progress

It was noted that the specified conditions of approval continued to be met. In line with the original approval there had been a number of publications published and the database was being used for a growing number of service evaluations. Anonymised information was provided for analysis purposes.

Steps taken to anonymise the information or obtain consent from individuals

It was noted that the carers of infants on neonatal units were provided with a parent information leaflet explaining that their baby's electronic record would be used to populate the NNRD. If the parents wished to opt-out they needed to notify a member of the neonatal staff who then informed the supplier(s) of the NNRD to prevent the records from flowing to the NNRD.

Project changes

It was confirmed that there had not been any changes to the data controller, purpose, scope, data flows, data sources or identifiable data items requested.

Practicable alternatives/exit strategy

The applicant advised that as linkage of NNRD/HES data was ongoing, receipt of the limited number of patient identifiers was still needed. The final product from the NNRD/HES data linkage would be a pseudonymised database containing lined NNRD and HES data.

Patient/Service user involvement

The current involvement of users in a current project funded by the Academy of Medical Sciences was noted.

Justification for ongoing support

It was noted that in line with the original outcome, NHS number would not be retained beyond the establishment of linkage with HES data. It was confirmed that there was still a continued need to access confidential patient information as specified within the original application.

Confidentiality Advice Team advice conclusion

As a whole, it was recommended that the approval in place for the purposes set out in the application should continue for a further 12 months from the anniversary of the original final approval outcome letter, to the date specified below.

Annual Review

Please note that your approval is subject to submission of an annual review report to show how you have met the conditions or report plans, and action towards meeting them. It is also your responsibility to submit this report on the anniversary of your final approval and to report any changes such as to the purpose or design of the proposed activity, or to security and confidentiality arrangements. An annual review should be provided 4 weeks before the date indicated above.

Please do not hesitate to contact me if you have any queries following this letter. I would be grateful if you could quote the above reference number in all future correspondence.

Yours sincerely

[redacted]

[redacted]

On behalf of the Health Research Authority

Email: [redacted]

Enclosures: Standard conditions of approval

NHS
Health Research Authority

Standard conditions of approval

The approval provided by the Health Research Authority is subject to the following standard conditions.

The applicant will ensure that:

1. The specified patient identifiable information is only used for the purpose(s) set out in the application.
2. Confidentiality is preserved and there are no disclosures of information in aggregate or patient level form that may inferentially identify a person, nor will any attempt be made to identify individuals, households or organisations in the data.
3. Requirements of the Statistics and Registration Services Act 2007 are adhered to regarding publication when relevant.
4. All staff with access to patient identifiable information have contractual obligations of confidentiality, enforceable through disciplinary procedures.
5. All staff with access to patient identifiable information have received appropriate ongoing training to ensure they are aware of their responsibilities.
6. Activities are consistent with the Data Protection Act 1998.
7. Audit of data processing by a designated agent is facilitated and supported.
8. The wishes of patients who have withheld or withdrawn their consent are respected.
9. The Confidentiality Advice Team is notified of any significant changes (purpose, data flows, data items, security arrangements) prior to the change occurring.
10. An annual report is provided no later than 12 months from the date of your final confirmation letter.
11. Any breaches of confidentiality / security around this particular flow of data should be reported to CAG within 10 working days, along with remedial actions taken / to be taken.

Appendix 5 Patient information leaflets

PiPS
Probiotic in Preterm babies Study

Parent Information Leaflet Form

Introduction

We understand and appreciate that this is a very difficult time for you and that it may not seem a good moment to be talking about research. However we think it is important that you know about a study that this hospital is taking part in for babies born prematurely.

Short Title: PiPS: Probiotic in Preterm babies Study

Formal Title: The probiotic Bifidobacterium breve strain BBG−001 administered early to preterm infants to prevent infection, necrotising enterocolitis and death

Summary

This is a brief description of a research study designed to test whether giving probiotics to premature babies helps to protect them against serious illnesses. You have been given this leaflet because your baby has been, or may be, born 10 or more weeks early and we want to give you the opportunity to think about whether you would like your baby to take part in this study.

Children and adults have many bacteria in their gut that do not cause disease and are important for health. It is possible that if we give similar bacteria (probiotics) by mouth to premature babies soon after birth that other bacteria that can cause illness may be prevented from becoming established in the gut.

The number of babies that have been given probiotics is small; early results of studies are encouraging and probiotics appear to be safe, however we cannot be confident about this until they have been more widely used.

This hospital is one of about 20 helping in a study funded by the NHS through the Health Technology Assessment programme which will involve 1300 very premature babies and give a clear answer about whether probiotics are helpful or not.

The rest of this leaflet explains the study in more detail and describes what being in the study would mean for you and your baby. If, after reading this leaflet and discussing the study with the doctors and nurses in the neonatal unit, you decide to take part we will ask you to sign a consent form. Your baby will then enter the study and will receive either probiotic or an inactive substance that looks the same; this inactive substance is called a placebo. Probiotic or placebo will be continued once a day until shortly before you go home. Neither you nor the staff on the unit will know whether your baby is actually receiving probiotic.

The study does not involve any additional blood tests and should not cause your baby any pain or discomfort. We will collect information about your baby's progress in hospital but we do not currently have any plan to see you again after you have gone home.

Whether or not you decide to let your baby take part is entirely up to you. If you decide not to take part this will not affect the high standard of care your baby receives.

PiPS_ISRCTN No: 05511098_REC Ref: 09/H0604/30_Patient Information Leaflet_Version 3.1 dated 20Jan2010

What problem are we trying to help?

Babies born as early as yours are at increased risk of infections and of other complications of prematurity; one of the most important of these is the illness necrotising enterocolitis that affects the gut. Usually these infections can be successfully treated with antibiotics but such illnesses often prolong a baby's stay in hospital and may increase the likelihood of life-long health problems. Occasionally they are so serious that the baby may need surgery or may even die.

There are a number of different ways in which the body protects us against infection; one of the most important is through the millions of bacteria that live in our gut. These are not germs that cause disease but are helpful bacteria that are essential for good health. Babies who are born at the expected time rapidly gain these bacteria from their mother and other members of the family. Babies born early have to be separated from their family and have few of these helpful bacteria; instead they are likely to have many other bacteria in their gut. Usually these other bacteria do not cause problems but they may cause infections that can be difficult to treat and be involved in the development of serious complications such as necrotising enterocolitis. They may also make it more difficult to digest milk which is very important for your baby's long term health.

It is possible that giving probiotics soon after birth will make the bacteria living in the gut of premature babies more like those of full-term babies and decrease the risk of them getting serious infections and necrotising enterocolitis.

What are probiotics and how much do we know about their use in newborn premature babies?

Probiotics are live micro-organisms, usually bacteria, that are taken by mouth and then multiply and live in the gut. There are lots of different probiotic bacteria, many of them have names beginning with Lactobacillus and Bifidobacterium and are contained in live yoghurt and a range of freely available health products.

There have been a number of studies giving probiotics to premature babies, the results suggest that giving probiotic might help babies to digest milk and to grow better, to have fewer episodes of serious infection and less necrotising enterocolitis. There have been no reports of complications. However the studies have all been small and none so far has been in the UK. In order to be clear whether or not probiotics are helpful and to be confident of their safety they need to be studied more widely.

Some studies have used just one type of probiotic and others have used mixtures. This study will use a single probiotic called Bifidobacterium breve strain BBG-001 (BBG). For the rest of this leaflet when we talk about probiotic this is the one we mean. The same probiotic has been given routinely to many thousands of newborn babies in Japan with no reported complications.

All of the earlier studies have mixed the probiotic in the baby's milk feeds; this has meant that babies who the doctors decide should not be fed have not been included in the studies. We think it is probably important, if probiotics are to be helpful, that they are given early, before other bacteria that may cause disease become established in the gut. In this study we plan to start probiotics early whether or not milk feeds have been started.

What is the purpose of the study?

The purpose of this study is to find out whether giving BBG to babies born 10 or more weeks early, reduces episodes of blood stream infection and necrotising enterocolitis. We will also study whether there is increased survival and whether babies are likely to leave hospital sooner if they receive probiotic.

Why has my baby been chosen?

All newborn babies are at some risk of infection and of necrotising enterocolitis but this risk is much greater in very premature babies, we are therefore inviting parents of babies born 10 or more weeks early to take part in this study.

This hospital is one of about 20 in England involved with this study. We are aiming to include 1300 babies; we need this number to be confident of finding out whether probiotics are helpful or not.

Does my baby have to take part?

You do not have to agree to your baby taking part in this study. If you decide not to take part it will not affect in any way the quality of care you and your baby receive. Similarly if you decide that you would like your baby to take part and then change your mind your baby can be taken out of the study at any time without you having to give a reason.

What will happen to my baby if I agree to take part?

Because we are studying the effect of giving probiotic early we are asking you to make your initial decision about whether your baby should take part within 48 hours of birth. We realise that this may put you under increased stress and apologise for this; we would not do this if we didn't believe it was important. We will discuss the study with you again during the course of your baby's stay in hospital to make sure that you understand what is happening and that you continue to agree to your baby taking part. There will always be someone available with whom you can discuss the study, sometimes this will be by phone.

If you agree that you would like your baby to take part in this study, your baby will be put into one of two groups; one group will receive probiotic and the other will receive a dummy product that looks the same, this dummy product is called a placebo. Your baby will have a 50/50 chance of being put into either of these groups. The allocation of your baby to a group will rely on chance (rather like tossing a coin). Neither you nor the staff caring for your baby will know which group your baby is in. This is the only way we can be sure that we test probiotic fairly. The first dose of probiotic or placebo will be given as soon as is practicable for the ward staff after you have signed the consent form; this may not be until the following morning.

The probiotic and the placebo are supplied to us as granules which we mix with fluid, we then put a few drops down the baby's feeding tube. We will do this once each day until your baby reaches 36 weeks of gestation (36 + 0 days). Because it is important that nobody knows which product your baby is receiving we mix both probiotic and placebo with a very dilute preparation of a special infant formula called Neocate so that they still look the same. Neocate is an infant formula that is very easy to digest and is made especially for babies with gut problems; it is not made from cow's milk. For this study Neocate is being used at 1/8 of full strength. This does not provide any significant nutrition for your baby and is so dilute that it cannot pose any risk to the gut even in those babies that the doctors decide should not be fed. This will in no way reduce your chances later of successfully breast feeding your baby.

If your baby is unwell the doctor in charge locally will decide whether or not doses are missed out.

If your baby is discharged home earlier than 36 weeks the probiotic or placebo will be stopped a few days before. If your baby is transferred to a different hospital before 36 weeks we will aim to continue to give the product; if the new hospital is not already involved in the study we will provide training to the staff to enable this to happen.

PiPS_ISRCTN No: 05511098_REC Ref: 09/H0604/30_Patient Information Leaflet_Version 3.1 dated 20Jan2010

If your baby sucks well and is able to have the feeding tube removed earlier than 36 weeks the probiotic or placebo will be given directly into the mouth with a syringe once a day before a feed.

Two weeks after birth and again at 36 weeks (if your baby is still in hospital) we will collect a sample of your baby's stool. These samples will be sent to the microbiology laboratory at Barts and the London Hospital, London E1 where they will be tested to check whether or not your baby has been successfully colonised with probiotic and what other bacteria have colonised the gut.

If you agree the remaining stool sample will then be deep frozen and stored for later testing in a related study for which we have not yet secured funding. The additional tests are designed to help us understand the effects of probiotics.

No extra blood tests or injections are necessary and all other aspects of your baby's care will be entirely at the discretion of the local doctors and nurses.

Unless there is a specific medical reason why not, it is hoped that mothers of babies in the study will provide breast milk for their babies since human breast milk promotes the multiplication of BBG.

What are the possible side effects of the treatment?

There are no reported side effects associated with the probiotic being used for this study. However all babies in the study will be monitored very closely throughout the study by the staff on the Neonatal Intensive Care Unit.

What information will be collected about me and my baby?

We will need to collect standard clinical information about your pregnancy, the condition of your baby at birth and progress throughout the hospital stay. This information will be collected from the baby's written and electronic case record. The study will not involve you in any interviews or questionnaires. In order to get accurate results from all samples taken by the medical staff to check for infection, we will contact the hospital microbiology laboratory directly since the detail needed for the study is not always available in the case notes.

After your baby has completed the study, records maintained by the NHS Information Centre and NHS Central Register maybe used to help us contact you in future and to provide information about your baby's health status.

What are the possible disadvantages of taking part?

We believe that this intervention is safe and that there are no disadvantages for you in taking part in this study whichever group your baby is in.

We will need to collect information about you and your baby.

What if new information becomes available?

There is currently a lot of interest in the use of probiotics for premature babies and other studies in other countries using slightly different probiotics are underway. We will be monitoring any results emerging from these studies closely and will inform you if any important new information becomes available during the course of the study that might make you change your mind about your baby's involvement.

What if something goes wrong?

The chance of anything going wrong as a result of taking part in this study is very small. However we are required to tell you the following:

If your baby is harmed and this is due to someone's negligence, then you may have grounds for legal action for compensation against Queen Mary, University of London in respect of any harm arising out of the participation in the Clinical Trial or the NHS in respect of any harm which has resulted from the clinical procedure being undertaken.

Will my taking part in this study be kept confidential?

Your GP will be told that you took part in the study.

Your details and the information collected for the study will be kept securely and will only be seen by the study organisers and people from the regulatory authorities whose job is to ensure that studies such as this are carried out safely. They may also need to look at your baby's notes to check that the information collected for the study is correct. Information about you or your baby will not be used for any purpose other than to answer these research questions.

Although we currently have no plans to collect any further information about your baby after discharge from hospital we will retain your contact details in case anything emerges from this or any other study of probiotics that makes it important that we contact you again. The NHS has a central register (based at the General Register Office) that would be able to tell us if you have left the NHS and through which we would be able to locate you.

What will happen to the results of the research?

At the end of the study the results will be analysed and published in an international journal. We will send you a copy of the final results of the study. A copy of the full journal article can be requested from the National Perinatal Epidemiology Unit. You and your baby will not be identified in any report or publication arising from the study.

Who is organising and funding the research?

The study is being run jointly by Barts and the London School of Medicine at Queen Mary, University of London and by the National Perinatal Epidemiology Unit, University of Oxford.

The study is funded by the NHS through the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme.

Who has reviewed the study?

All research that involves NHS patients or staff, information from NHS records or uses NHS premises or facilities must be approved by a NHS Research Ethics Committee before it goes ahead. Approval cannot guarantee that you will not come to any harm if you take part. However approval does mean that the Committee is satisfied that your rights will be respected, that the risks have been reduced to a minimum and that balanced against the possible benefits it is reasonable for babies born as early as yours to take part. The committee has also checked that we are giving you sufficient information to make an informed decision about taking part.

Thank you for reading this leaflet. The doctor or nurse who gave it to you will be pleased to discuss it with you and to provide further information if that would be helpful. Alternatively the contact details of the study's Principal Investigator in your NHS hospital and the Study Co-ordinator are provided below.

What if I have any concerns?

If at any stage you have any concern or query about this study or the way it has been carried out, you should contact the Principal Investigator (the name and contact details are below), or you may contact the hospital complaints department.

Information is also available on the study website at: www.npeu.ox.ac.uk/pips

If you would like to contact an independent organisation to discuss the inclusion of babies in research studies without reference to this particular study we suggest that you contact the premature baby charity Bliss. Their address is:

Bliss, 9 Holyrood Street, London SE1 2EL
Freephone Family Support Helpline: [redacted]
Website: www.bliss.org.uk

Name and contact details of local contact:

Name and contact details of Study Co-ordinator:

[redacted] (PiPS Trial Co-ordinator)

National Perinatal Epidemiology Unit, Clinical Trials Unit
University of Oxford
Old Road Campus
Headington
Oxford
OX3 7LF

[redacted]

NPEU Clinical Trials Unit, National Perinatal Epidemiology Unit, University of Oxford, Old Road Campus, Headington, Oxford, OX3 7LF
Tel: +44 (0)1865 617924 Fax: +44 (0)1865 289740 Email: pips@npeu.ox.ac.uk www.npeu.ox.ac.uk/pips

PiPS_ISRCTN No: 05511098_REC Ref: 09/H0604/30_Patient Information Leaflet_Version 3.1 dated 20Jan2010

(To be printed on headed paper)

INFORMATION SHEET FOR PARENTS AND GUARDIANS

Title: Evaluating the reliability of standardised two-year neurodevelopmental data collected during NHS follow-up in children born preterm (REC reference: 10/H07020/35)

Short title: Reliability of two-year neurodevelopmental assessment in preterm infants

We would like to invite your child to take part in a research study.
This information sheet explains why this study is being carried out and what it involves. Please read it carefully and discuss it with others if you wish. Ask us if there is anything that is not clear. Take time to decide whether or not you would like your child to take part.

1. What is the purpose of the study
The long-term health outcome of children born preterm is important. All preterm infants are seen after their discharge from hospital and have their development checked. The doctors carrying out this assessment use different methods that may not be comparable. If children participate in a research study, they may have a repeat assessment.

The purpose of this study is to establish how reliable the routine check is in comparison to a detailed assessment carried out by specially trained examiner. Our secondary goal is to be able to reduce the requirement for a second assessment if children participate in research studies in the future.

2. Why has my child been chosen
We are approaching the families of children born in London at less than 30 weeks gestation (more than 10 weeks before their due date), who are being seen in a hospital follow-up clinic.

3. Does my child have to take part?
This is entirely up to you. Your decision will not affect the standard of care your child receives.

4. What is involved in joining the study?
We will invite you to bring your child for a special appointment at your local hospital, at a time convenient to you. During this session, we will examine movement, speech and problem-solving skills through play and also examine muscle tone. The assessment will be done by a specially trained doctor and does not involve any invasive, unpleasant or painful testing. Most children find the assessment enjoyable and fun. We will also ask you to fill in two questionnaires about your child's social skills and behaviour. The session will last approximately 90 minutes and your child will be with you the whole time. The study will not require you to do anything else and your child's involvement will stop after this appointment. We will reimburse the cost of travel for the hospital visit.

We will compare the results of this assessment with the check carried out in your child's normal hospital follow-up clinic. To do this, we will require your permission to look at your child's medical records.

In the future, we hope to assess the reliability of the assessment at 2 years in predicting future development. We would therefore also like to ask for your permission to contact you in the future.

5. Are there any benefits from taking part?
Your child will receive a detailed assessment of his/her skills and development. We will send a copy of the results to your child's hospital paediatrician and your GP.

There is a possibility that your child may be found to have difficulties in a particular area. If so, your paediatrician will explain what this means, and arrange for further assessments and help to improve your child's abilities should this be necessary.

Some difficulties only become apparent as a child grows older. Therefore your child should continue to receive assessments as recommended by your paediatrician. If you are concerned about your child's development at any time it is important you discuss this with your paediatrician, GP or health visitor.

6. Are there any disadvantages of taking part

The tests used in this study are used around the world to assess children. They involve play and puzzles in a structured way. Most children find them fun but if your child becomes bored or tired we can have a break or stop if you wish.

7. What if there is a problem?

We are required to inform you that there are no special compensation arrangements in the unlikely event of something going wrong. If you are harmed due to someone's negligence, you may have grounds for a legal action. If you have concerns about the way you have been treated during this study, you should inform the Chief Investigator, Professor Neena Modi (contact information below). The normal NHS complaint mechanisms are also available to you and you may contact the Imperial Academic Health Sciences Centre Joint Research Office.

8. What will happen if I do not want my child to continue with the study?

If you do not wish your child to carry on with the study, you can withdraw him/her at any time, without giving reason. This will not affect the care your child receives.

9. Will details about my child be kept confidential?

All information collected about your child during the course of the study will be kept confidential.

10. What will happen to the results of the research study?

The results will be presented at medical conferences and published in scientific journals. No participant will be identified in any presentation or publication.

11. Who is funding and organising this study?

The study is funded by the National Institute for Health Research as a component of the Medicines for Neonates Programme. It is led by the Section of Neonatal Medicine at Imperial College London.

12. Who has reviewed this study?

This study and information sheet has been reviewed and approved by *(Research Ethics Committee)*.

Thank you for reading this. Please do not hesitate to contact us if you have any questions.

[redacted]
Email: [redacted]

Dr [redacted]
Tel: [redacted]
E-mail: [redacted]

Dr [redacted]
E-mail: [redacted]

Professor [redacted]
[redacted]

Reliability of two-year assessment
Study Information Sheet/Version 1/180310

UK Neonatal Collaborative Necrotising Enterocolitis Study

(REC reference: 11/LO/1430)

UK Neonatal Collaborative
NECROTISING ENTEROCOLITIS STUDY

This study is a component of the NIHR funded "Medicines for Neonates" Programme*, which aims to use operational electronic data to support neonatal services and research. The "**UK Neonatal Collaborative NEC (UKNC-NEC) Study group**" are neonatal teams committed to entering high quality electronic data.

STUDY BACKGROUND

Preventive strategies for Necrotising Enterocolitis (NEC) remain elusive; feeding practices are widely believed to influence susceptibility but have not been tested adequately in randomised controlled trials. Most baseline incidence data come from small studies using varying case-definitions. Surveillance of NEC is also important for assessing temporal trends and geographical variation.

STUDY AIMS

1) To establish an objective case-definition for NEC that is applicable nationally and internationally
2) To determine the incidence of NEC over a large geographically defined population
3) To identify enteral-feed related antecedents that precede the onset of NEC.

METHODS

This study will utilise anonymised NHS electronic data captured as part of clinical care using the neonatal.net NHS platform (widely known as the "Badger system") from all babies admitted over 18 months (starting Nov 2011) throughout the country.

WHAT IS REQUIRED FROM NEONATAL UNITS IN THE UKNC-NEC STUDY GROUP?

Only to ensure completion of the following electronic fields on all neonatal admissions:

Static data
(Enter once on all babies)

- Birth weight
- Gestational age
- Sex
- Mother's race
- Antenatal steroids
- Gastrointestinal anomalies

Daily data
(Enter daily on all babies)

- Most recent weight
- Enteral feed (Type/ Fortifier/Volume)
- Medications (COX inhibitors/ Antibiotics)
- Lines in situ (Umbilical arterial line)
- Packed red cell transfusion

"Abdominal x-ray performed" ad-hoc form

- ONLY TO BE COMPLETED for babies who have an abdominal x-ray to investigate abdominal signs

If your unit would like to contribute to the study or would like further information, please contact :
Dr [redacted] Clinical Research Fellow, Neonatal Data Analysis Unit, Department of Medicine, Imperial College London, Chelsea & Westminster campus , Fulham Road, London SW10 9NH [redacted]

*** Medicines for Neonates Programme Senior Investigators:**
[redacted]

NDAU Neonatal Data Analysis Unit · Imperial College London · MfN Medicines for Neonates · Chelsea and Westminster Hospital NHS NHS Foundation Trust

UKNC-NEC STUDY POSTER VERSION 1

Version 1, July 2011

Chelsea and Westminster Hospital NHS
NHS Foundation Trust

Are you the parent or carer of a baby who has spent time on a neonatal unit?

If so we would love to hear from you!

In neonatal care, a lot of data is routinely collected from your baby. Some of this information could be very useful to health researchers but at the moment it is not automatically used for this purpose.

We would like you to answer our questionnaire, so we can hear your views on the subject: Who (if anyone) would you be happy to see your baby's data and for what sort of research?

We would love to hear from fathers as well as mothers who currently have or have recently had a baby in neonatal care. Users of languages other than English can be supported to take part in the survey.

What do I need to do to take part?

Please let one of the nurses know you are interested.

A research nurse will be able to tell you more about the survey, provide you with the questionnaire and help you fill it in if you would like.

There is also an information leaflet available to answer any questions you may have.

NDAU
Neonatal Data Analysis Unit

MANCHESTER 1824

More information
If you would like more information regarding the project please contact:
[redacted] Children's Research Nurse
Tel: [redacted]
Mobile [redacted]

Bliss
for babies born too soon, too small, too sick

What is the purpose of the study?

In neonatal care, a lot of data is routinely collected from your baby. A lot of this information could be very useful to health researchers but at the moment it is not automatically used for research purposes.

'Data' refers to all sorts of information that is collected, from birth weight to drugs administered, to the progress your baby is making and so on. This might be recorded in a database or on paper notes. We are asking only about information, not tissue samples etc.

'Research' refers to the process of collecting, ordering and evaluating information in order to provide further understanding, new knowledge and/or a basis for decision making and action or change.

Please note that we are not asking to use your baby's data. We are interested in hearing what you think. There are no 'right' or 'wrong' answers.

Where and when is the research taking place?

The survey is taking place in a number of neonatal units across England, in London, the North West and Yorkshire. It will run from Oct 2011 to Oct 2012.

What are the advantages of taking part?

The subject of data sharing for research purposes is high on the agenda in health care and so now is an opportunity to make your voice heard.

Are there any disadvantages of taking part?

The questionnaire will take 15-20 minutes of your time and some of the questions may be upsetting. As parents, we appreciate what a difficult time this is and we make it clear on the questionnaire which questions we think may cause upset. You are free to skip any questions that you do not wish to answer.

Is the questionnaire anonymous?

The information you give us will not be seen by anyone other than the researchers at the University of Manchester. You will have the option of supplying your name and address and letting us know if you want to be contacted in future regarding this or related research.

How many parents will be taking part in the study?

We are hoping to receive 1525 completed questionnaires altogether.

Am I eligible to take part in the study?

If you have, or have had, a baby in neonatal care, we would love to hear from you. We are keen for both mothers and fathers (and other carers if appropriate) to complete a questionnaire independently of one another. So both you and your partner (if you have one) can fill in a separate questionnaire.

What will happen if I don't want to carry on with the study?

If you decide to take part and then later change your mind, either before or during the study, you can withdraw. However, once the questionnaire has been filled in and sent back it will have been anonymised, therefore withdrawing will not be an option as we will not be able to trace the questionnaire back to you.

Who is organising and funding the study?

Funded by the National Institute for Health Research, it is part of a much bigger project that involves the Neonatal Data Analysis Unit (NDAU) based at Imperial College, London. This project is led by [redacted] and [redacted] at The University of Manchester, in partnership with Bliss.

This questionnaire has been designed by a group of parents (both mothers and fathers) who have all had babies in neonatal care in recent years. Whilst we have all spent time with our babies in neonatal care, we all have different views on who we would want to have access to our babies' data, depending for example on the purposes of the research and how/when we were asked.

Can I find out what the findings are from the research?

Yes. You will be given the option of providing us with your contact details so that we can post or e-mail you with the findings at the end of the study. Alternatively, they will be published on the Bliss website at www.bliss.org.uk

What if there is a problem?

If you have concerns about any aspect of this study, please ask to speak to the research nurses who will do their best to answer your questions.

If they are unable to resolve your concern or you wish to make a complaint regarding the study, please contact a University Research Practice and Governance Co-ordinator on [redacted] or [redacted]. Alternatively, email: [redacted]

Should any distress arise, the Bliss helpline [redacted] may be of use.

What do I have to do if I agree to take part?

If you are willing and able to take part, then a research nurse will ask you to sign a consent form. You will then be given a copy of the questionnaire.

You will be given the choice of filling in the questionnaire yourself or having a research nurse fill it in with you. The research nurse will ensure that anything you say is kept confidential.

Interpreter services are available should you or your partner require them.

Again, confidentiality is guaranteed. If you wish to complete the questionnaire on your own, you can do so on the unit. It should take around 15-20 minutes to fill in. A pre-paid envelope will be provided addressed to The University of Manchester so that the questionnaire can be sent back and it will not be seen by the staff at the hospital.

Why do the questions refer to one baby when I have had a multiple birth?

For consistency of style we use 'baby' in the singular. However we do appreciate that a number of parents completing the questionnaire will have more than one baby in neonatal care at this time. Please bear with us on this.

More Information

If you would like more information about the project please contact:

[redacted]esearch Nurse
Tel: [redacted]
Mobile: [redacted]
Email: [redacted]

Participant information sheet: Version 2, November 2011

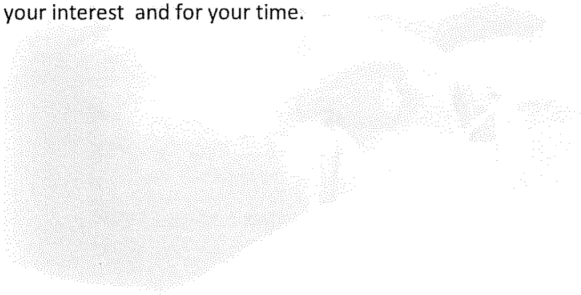
Chelsea and Westminster Hospital NHS
NHS Foundation Trust

Data Sharing in Neonatal Services

We would like to invite you to take part in a new research study from the University of Manchester and Bliss (the charity for babies born too soon, too small, too sick.) It is a questionnaire survey that has been designed by a group of parents, both mothers and fathers, who have all had babies in neonatal care in recent years.

Before you decide whether or not to participate, please read this booklet, as it explains why the research is being carried out and what it will involve. Feel free to ask us if there is anything that is not clear or you want to know more about.

Thank you for your interest and for your time.



MANCHESTER 1824

NDAU
Neonatal Data Analysis Unit

Bliss
for babies born too soon, too small, too sick

Appendix 6 Research ethics committee approvals

NHS
National Research Ethics Service
Oxfordshire REC A
2nd Floor, Astral House
Chaucer Business Park
Granville Way
Bicester
OX26 4JT

Telephone: [redacted]
Facsimile: [redacted]

12 May 2009

[redacted]
Neonatal Unit
Homerton University Hospital
Homerton Row
London
E9 6SR

[redacted]

Full title of study:	**The probiotic Bifidobacterium breve strain BBG-001 administered early to preterm infants to prevent infection, necrotising enterocolitis and death**
REC reference number:	**09/H0604/30**
Protocol number:	**1**
EudraCT number:	**2006-003445-17**

Thank you for your letter of 09 April 2009, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Vice-Chair, Ms Sara Owen.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

This Research Ethics Committee is an advisory committee to South Central Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. *Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.*

Sponsors are not required to notify the Committee of approvals from host organisations.

Clinical trial authorisation must be obtained from the Medicines and Healthcare products Regulatory Agency (MHRA).

The sponsor is asked to provide the Committee with a copy of the notice from the MHRA, either confirming no objection or giving grounds for objection, as soon as this is available.

Other conditions specified by the REC

1. Documents to be seen by participants should have all references to "PREFER" removed to avoid confusion. There were still references in some of the footers of the documents supplied.

2. It is a requirement of the Clinical Trials Regulations that any site not listed on the Application Form before the favourable ethical opinion is provided must be notified to the main REC as a Substantial Amendment. This will then by acknowledged by the REC.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Compensation Arrangements	Zurich Municipal	10 February 2009
Peer Review	Professor [redacted]	26 January 2009
Letter from Sponsor	Queen Mary, University of London	06 February 2009
Covering Letter		09 February 2009
Protocol	1	29 January 2009
Investigator CV	Professor [redacted]	06 February 2009
Application	Parts A-D	12 February 2009
Response to Request for Further Information		09 April 2009
Participant Consent Form	2	03 April 2009
Participant Information Sheet	2	03 April 2009
GP/Consultant Information Sheets	2	03 April 2009

Statement of compliance

This Committee is recognised by the United Kingdom Ethics Committee Authority under the Medicines for Human Use (Clinical Trials) Regulations 2004, and is authorised to carry out the ethical review of clinical trials of investigational medicinal products

The Committee is fully compliant with the Regulations as they relate to ethics committees and the conditions and principles of good clinical practice.

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review –guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email [redacted].

09/H0604/30	Please quote this number on all correspondence

With the Committee's best wishes for the success of this project

Yours sincerely

[redacted]

Chair

Email: [redacted]

Enclosures: "After ethical review – guidance for researchers"

Copy to: [redacted]

Clinical Trials Unit, MHRA

NHS
National Research Ethics Service
North West London REC 2
Royal Free Hospital NHS Trust
Royal Free Hospital
South House, Block A
Pond Street
London
NW3 2QG

Telephone: [redacted]
Facsimile: [redacted]

29 April 2010

[redacted]

[redacted]

Study Title: **Evaluating the reliability of standardised two-year neurodevelopmental data collected during NHS follow-up in children born preterm**
REC reference number: **10/H0720/35**
Protocol number: **1**

The Research Ethics Committee reviewed the above application at the meeting held on 21 April 2010. Thank you for attending to discuss the study.

Ethical opinion

The main ethical issue was the exclusion of non English speaking people from the study, but
Dr Huetas explained that apart from the cost of the translation, the neurodevelopment assessment tool being used had not been validated for other languages. The paediatric expert member of the committee was present for this item

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

The Committee has not yet been notified of the outcome of any site-specific assessment (SSA) for the non-NHS research site(s) taking part in this study. The favourable opinion does not therefore apply to any non-NHS site at present. I will write to you again as soon as

one Research Ethics Committee has notified the outcome of a SSA. In the meantime no study procedures should be initiated at non-NHS sites.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.

Sponsors are not required to notify the Committee of approvals from host organisations.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
REC application		22 March 2010
Protocol	1	18 March 2010
Investigator CV		
Participant Information Sheet	1	18 March 2010
Participant Consent Form	1	18 March 2010
Letter of invitation to participant	1	18 March 2010
GP/Consultant Information Sheets	1	18 March 2010
Student CV		

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email [redacted]

10/H0720/35	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

[redacted]

Chair

Email: [redacted]

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
"After ethical review – guidance for researchers" [SL-AR1 for CTIMPs, SL-AR2 for other studies]

Copy to: [redacted]
[R&D office for NHS care organisation at lead site]

North West London REC 2

Attendance at Committee meeting on 21 April 2010

Committee Members:

Name	*Profession*	*Present*	*Notes*
[redacted]	Statistician	Yes	
[redacted]	Co-ordinator	Yes	
[redacted]	Consultant Paediatrician	Yes	
[redacted]	Lecturer (Lay)	Yes	
[redacted]	Head of Pharmaceutical Services	No	
[redacted]	Lay Member	Yes	
[redacted]	Nuclear Medicine	Yes	
[redacted]	Clinical Trial Pharmacist	Yes	
[redacted]	Professor O & G, RFUCMS, Hampstead Campus	No	
[redacted]	Lay Member (Vice Chair)	Yes	
[redacted]	Senior Lecturer and Consultant Neurologist	No	
[redacted]	Chairman	Yes	
[redacted]	Committee Member (Lay)	Yes	
[redacted]	Nurse	No	
[redacted]	Clinical Trial Pharmacist	Yes	
[redacted]	Pharmacist	Yes	

South West London REC 3
Room 4W/12 4 Floor West
Charing Cross Hospital
Fulham Palace Road
London W6 8RF

Telephone [redacted]
Facsimile [redacted]

03 December 2010 (amended and reissued 7 February 2011)

[redacted]

[redacted]

Title of the Database: **National Neonatal Research Database**
REC reference: **10/H0803/151**

The Research Ethics Committee reviewed the above application at the meeting held on 24 November 2010. Thank you for attending to discuss the application.

Ethical opinion
You clarified that the application was requesting ethics approval for the creation and use of the database. The projects listed will use or are using the data and have already received REC approval. All studies will require ethics approval.

The Committee asked for clarification about whether ethics approval was also being sought for the studies detailed briefly in the application and in the protocol. You confirmed that the application was just for creation of the database. A new application would be submitted for each study considered as research. The database will also be used for service evaluation and ethics approval would not be requested for this.

You confirmed that applications for access of the database will be considered by a steering committee and that the database is a national resource and so will be available to everyone.

Members asked for clarification about the charge that will be levied to database users., and you stated that the project would not make a profit from this fee but it will be levied to cover the cost of the staff that manage the database because they do not have a grant for this.

No fee has been set at the moment but the fee charged to a recent study was based on the number of hours the database was used.

The Committee asked who would have ownership of the data. You confirmed that each Trust owns its own data and has access to its own data and it would not be necessary for a Trust to ask for access to its own data. However, problems may arise because of the form in which the data is released. Each Trust would have to manage its own data.

You stated that all parents were given an information sheet with details of all uses of the data held on the database. There is a provision which allows parents to opt out but none have done so as yet

The Committee stated that they believe the data will be used appropriately and only appropriate access will be allowed.

The members of the Committee present gave a favourable ethical opinion of the above research database on the basis described in the application form and supporting documentation.

Following further clarification being sought by you after receipt of this opinion letter, the Sub Committee provided the following further information:

This ethical approval extends to those projects listed in the application form, without the need to pursue further ethical approval. Any additional (research) projects (not listed) would need research ethics approval gained through a new research ethics approval application process.

Duration of ethical opinion
The favourable opinion is given for a period of five years from the date of this letter and provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully. The opinion may be renewed for a further period of up to five years on receipt of a fresh application. It is suggested that the fresh application is made 3-6 months before the 5 years expires, to ensure continuous approval for the research database.

Approved documents
The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
National Neonatal Research Database Items	1	11 October 2010
NIHR Award Letter		22 October 2008
REC application		07 October 2010
Covering Letter	from [redacted]	22 October 2010
Chief Investigator CV [redacted]		23 October 2010
NIHR Grant Application - Medicines for Neonates	RP-PG-0707	
Letter of Support - Dr [redacted]		07 October 2010
Protocol for Management of the Database	1	07 October 2010
Summary of Research Programme(s)	1	07 October 2010

Research governance
A copy of this letter is being sent to the R&D office responsible for Chelsea and Westminster NHS Foundation Trust.

Under the Research Governance Framework (RGF), there is no requirement for NHS research permission for the establishment of research databases in the NHS. Applications to NHS R&D offices through IRAS are not required as all NHS organisations are expected to have included management review in the process of establishing the database.

Research permission is also not required by collaborators at data collection centres (DCCs) who provide data under the terms of a supply agreement between the organisation and the database. DCCs are not research sites for the purposes of the RGF.

Database managers are advised to provide R&D offices at all DCCs with a copy of the REC application for information, together with a copy of the favourable opinion letter when available. All DCCs should be listed in Part C of the REC application.

NHS researchers undertaking specific research projects using data supplied by a database must apply for permission to R&D offices at all organisations where the research is conducted, whether or not the database has ethical approval.

Site-specific assessment (SSA) is not a requirement for ethical review of research databases. There is no need to inform Local Research Ethics Committees.

Membership of the Committee
The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance
The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review
Now that you have completed the application process please visit the National Research Ethics Service website > After Review

Here you will find links to the following:

a) Providing feedback. You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.
b) Annual Reports. Please refer to the attached conditions of approval.
c) Amendments. Please refer to the attached conditions of approval.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email [redacted]

10/H0803/151	Please quote this number on all correspondence

Yours sincerely

[redacted]

[redacted]
Chair
[redacted]

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*

Approval conditions

South West London REC 3

Attendance at Committee meeting on 24 November 2010

Committee Members:

Name	*Profession*	*Present*
[redacted]	Consultant Cardiologist	Yes
[redacted]	Senior Lecturer / Consultant Paediatrician	Yes
[redacted]	Consultant Obstetrician & Gynaecologist	Yes
[redacted]	Consultant Chest Physician	Yes
[redacted]	Lay Member	Yes
[redacted]	Clinical Nurse Specialist	Yes
[redacted]	Lay Member	Yes
[redacted]	Professor of Anesthesia	Yes
[redacted]	Consultant Radiologist	Yes
[redacted]	General Practitioner	Yes
[redacted]	Consultant Surgeon	Yes
[redacted]	Senior Lecturer in Medical Statistics	Yes
[redacted]	Pharmacist	Yes
[redacted]	Chaplain	Yes
[redacted]	Lay Member	Yes

Also in attendance:

Name	*Position (or reason for attending)*
[redacted]	Co-ordinator
[redacted]	Co-ordinator

NHS
Health Research Authority
NRES Committee London - Wandsworth
Room 4W/12 4 Floor West
Charing Cross Hospital
Fulham Palace Road
London W6 8RF
Telephone: [redacted]
Facsimile: [redacted]

13 January 2012

[redacted]

[redacted]

Title of the Database: National Neonatal Research Database
REC reference: 10/H0803/151
Amendment number: 1
Amendment date: 25 November 2011

The above amendment was reviewed by the Sub-Committee in correspondence.

Ethical opinion

The members of the Committee taking part in the review gave a favourable ethical opinion of the amendment on the basis described in the notice of amendment form and supporting documentation.

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Participant Information Sheet	4	16 November 2011
Notice of Substantial Amendment	1	25 November 2011
Covering Letter		25 November 2011

Membership of the Committee

The members of the Ethics Committee who took part in the review are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

10/H0803/151	Please quote this number on all correspondence

Yours sincerely

[redacted]

[redacted]
Chair

[redacted]

Enclosures: *List of names and professions of members who took part in the review*

Copy to: *R&D, Chelsea & Westminster NHS Foundation Trust*

NRES Committee London - Wandsworth

Attendance at Sub-Committee of the REC meeting on 03 January 2012

Committee Members:

Name	*Profession*	*Present*	*Notes*
[redacted]	Lay Member	Yes	
[redacted]	Lay Member	Yes	

NHS
National Research Ethics Service
NRES Committee London - Dulwich
Room 4W/12, 4th Floor
Charing Cross Hospital
Fulham Palace Road
London
W6 8RF

Telephone: [redacted]
Facsimile: [redacted]

28 September 2011

[redacted]

Dear [redacted]

Study title: **The National Neonatal Collaborative Necrotising Enterocolitis Study: Using operational clinical data captured electronically at the point of care for surveillance and research.**
REC reference: **11/LO/1430**
Protocol number: **CRO1531**

The Research Ethics Committee reviewed the above application at the meeting held on 21 September 2011. Thank you for attending to discuss the study.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non NHS sites

The Committee has not yet been notified of the outcome of any site-specific assessment (SSA) for the non-NHS research site(s) taking part in this study. The favourable opinion does not therefore apply to any non-NHS site at present. I will write to you again as soon as one Research Ethics Committee has notified the outcome of a SSA. In the meantime no study procedures should be initiated at non-NHS sites.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Covering Letter		17 August 2011
Investigator CV		
Letter from Sponsor		17 August 2011
Other: Student CV: [redacted]		17 August 2011
Protocol		18 August 2011
REC application		

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

11/LO/1430 **Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely,

[redacted]

[redacted]
Chair

Email: [redacted]

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*
"After ethical review – guidance for researchers"

Copy to: [redacted]

[redacted]

NRES Committee London - Dulwich

Attendance at Committee meeting on 21 September 2011

Committee Members:

Name	*Profession*	*Present*	*Notes*
[redacted]	Consultant Paediatrician	Yes	
[redacted]	Consultant Liver Intensivist	No	
[redacted]	Assistant Director of Pharmacy	Yes	
[redacted]	Barrister	No	
[redacted]	Head of Clinical Research Statistics	No	
[redacted]	Lay Member	Yes	
[redacted]	Professor Emeritus, Oral Pathology	No	
[redacted]	Consultant Rheumatologist	Yes	
[redacted]	Consultant Cardiologist	Yes	
[redacted]	Senior Nurse	Yes	
[redacted]	MHRN Service Users in Research Coordinator	Yes	
[redacted]	Stroke Research Coordinator	No	
[redacted]	Consultant Old Age Psychiatrist	No	
[redacted]	Coordinator	Yes	
[redacted]	Research Development Manager	Yes	
[redacted]	Research Nurse	Yes	
[redacted]	Consultant Midwife	Yes	

Written comments received from:

Name	*Position*
[redacted]	Consultant Liver Intensivist
[redacted]	Barrister

NHS
National Research Ethics Service
NRES Committee North West - Cheshire
Research Ethics Office
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: [redacted]
Facsimile:

21 November 2011

[redacted]

[redacted]

Study title: **Understanding parents' attitudes towards the use of nhs data for research purposes in the context of neonatal services**

REC reference: **11/NW/0765**

The Research Ethics Committee reviewed the above application at the meeting held on 09 November 2011.

Ethical opinion

1. The Committee queried whether the questionnaire being used in this study has been validated. Dr [redacted] clarified that is has been reviewed by the statistician and they have looked at the questions. Therefore it is not a common validated tool used widely for research of this type.

2. The Committee asked whether there is an element of bias in terms of somebody helping the participants to fill out the questionnaire. Dr [redacted] explained that the research nurses are trained in this area and are skilled at that practice; therefore they are best placed to do it and as it is a necessary thing to do they cannot think of another way around it.

3. The Committee questioned whether pressure will be put on participants to participate; how long will potential participants have been on the neonatal unit? Dr [redacted] commented that there will be a considerable variation and they will have to use their own judgement on whether an individual should be approached to participate.

4. The Committee commented that they will get a mixture of sick and well babies and felt that they will get a very different and varied response. Dr [redacted] commented that this is what they expect and hope occurs so it can be written up in the analysis.

5. The Committee felt that that there may be an element of pressure/coercion on the patient to take part. Dr [redacted] clarified that somebody from the research team will be involved in the asking rather than somebody from the direct care team.

6. The Committee asked whether the questionnaire will be anonymised. Dr [redacted] confirmed that they will be anonymised but parents have the option of giving their contact details if they wish to do so; this is if they wish to be considered for a separate research study or wish to be contacted for the results. The Committee asked why they are asking for consent because the return of the questionnaire would imply consent. Dr [redacted] explained that they have done this to be on the safe side.

7. The Committee queried the section 'What will happen if I don't want to carry on with the study?' and asked what would happen if the participant had submitted the questionnaire; would this mean that it could not be withdrawn as it would have been anonymised. Dr [redacted] agreed, once the questionnaire has been anonymised there is no way of tracing back therefore participants will not be able to withdraw after that point. Dr [redacted] agreed to reword the information sheet to reflect this.

8. Dr [redacted] clarified that if at any point a participant shows signs of distress and does not want to continue then the interview will be stopped. She explained that the research nurses are highly experienced and have worked with this group of patients before. She also added that the survey has been designed by parents so it should be user friendly. The Committee pointed out that the option of not wanting to answer any particular question should be stated at the top of the list rather than the bottom. Dr [redacted] agreed.

9. The Committee asked why participants would be asked about politics. Dr [redacted] explained that this is just to obtain a picture of what type of person will say 'yes' or 'no' and it is up to the participant if they wish to give this information.

10. The Committee asked whether this research is 'mother' biased. Dr [redacted] clarified that they are keen to involve fathers as much as possible in this research.

Dr [redacted] was thanked for attending and left the meeting.

The Committee considered Dr [redacted] responses.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, **subject to the conditions specified below.**

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

Other conditions specified by the REC

1. Please revise the information sheet under the heading 'What will happen if I don't want to carry on with the study?' as follows. 'If you decide to take part and then later change your mind, either before or during the study you can withdraw. However, once the questionnaire has been filled out and sent back it will have been anonymised therefore withdrawing will not be an option as we will not be able to trace the questionnaire back to you.'
2. Please revise the questionnaire as follows;
    a. Please move the option 'do not wish to answer' to the top of the list of options so that participants are aware that they do not have to answer any of the questions. This should be an option for all questions. It should be made clear at the beginning of the questionnaire in bold that if they do not wish to complete the form or answer any questions then they do not have to.
    b. Under question 10, there is a formatting error; the word 'least' needs to be replaced before the word 'happy' so it reads 'least happy'
    c. Under question 13, the options of 'least happy' and 'most happy' are missing and need to be included.
3. The Committee noted that on the project filter questions on IRAS it states that this is a project involving qualitative methods only. The Committee felt that this study involved a bit of both qualitative and quantitative methods and is also administering a questionnaire. Please clarify this; if it is the latter, then the IRAS form will need to be amended on the project filter questions to reflect the correct study type. Once amended and saved please submit the form with a new submission code (this is found on the bottom right hand corner of the page) to the co-ordinator for the file. The code should be slightly different to the one previously submitted. If you have any problems with this please contact the co-ordinator.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. Confirmation should also be provided to host organisations together with relevant documentation

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Advertisement	Flier - Version 1	01 July 2011
Advertisement	Poster - Version 1	01 July 2011
Covering Letter	email	20 October 2011
Evidence of insurance or indemnity	University of Manchester	17 October 2011
Investigator CV	[redacted]	
Letter from Sponsor	University of Manchester	17 October 2011
Participant Consent Form	1	01 July 2011
Participant Information Sheet: Information Booklet	1	01 July 2011
Protocol	1	01 September 2011
Questionnaire	1	01 July 2011
REC application	3.1	18 October 2011

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

11/NW/0765 **Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

[redacted]
Chair

Email: [redacted]

Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments
"After ethical review – guidance for researchers"

Copy to: [redacted] The University of Manchester

[redacted] Teaching Hospitals NHS Foundation Trust

NRES Committee North West - Cheshire

Attendance at Committee meeting on 09 November 2011

Committee Members:

Name	*Profession*	*Present*	*Notes*
[redacted]	Senior Lecturer	No	
[redacted]	GP	Yes	
[redacted]	Vicar	No	
[redacted]	Consultant ENT Surgeon	Yes	Chair
[redacted]	Consultant Clinical Psychologist	No	
[redacted]	Lay Member	Yes	
[redacted]	Consultant Member	No	
[redacted]		Yes	
[redacted]	Consultant Paediatrician	Yes	
[redacted]	University Lecturer in Health Research	Yes	
[redacted]	Pharmacist Member	Yes	
[redacted]	Consultant Member	Yes	
[redacted]	Lay member	Yes	
[redacted]	Lay Member	Yes	
[redacted]	Lay Member	No	

Also in attendance:

Name	*Position (or reason for attending)*
[redacted]	Co-ordinator
[redacted]	Assistant Co-ordinator

Appendix 7 Presentations arising from the Medicine for Neonates Programme

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. Bristol: Neonatal Nutrition Network meeting; September 2011.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. London: King's College Hospital London; September 2011.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. Taunton: Western Neonatal Network meeting; November 2011.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. Newcastle: Yorkshire Neonatal Network meeting; February 2012.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. London: NNAP/NDAU collaborators meeting; January 2012.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. Basildon: Basildon Hospital; February 2012.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. London: Neonatal Nutrition Network meeting; May 2012.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. Cambridge: East of England Neonatal Network meeting; May 2012.

Battersby C *Improving the Quality of Routinely Collected Electronic Data*. London: SIGNEC (Specialist Interest Group NEC) meeting; June 2012.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. Birmingham: West Midlands Neonatal Network research meeting; November 2012.

Battersby CW, Santhakumaran S, Modi N. *The UK Neonatal Collaborative Necrotising Enterocolitis Study: A Prospective Population-based Study Using the National Neonatal Research Database*. Scottish Informatics Programme Conference; April 2013.

Battersby C. *Improving the Quality of Routinely Collected Electronic Data*. London: 1st International SIGNEC conference (Specialist Interest Group NEC); July 2013.

Battersby C. *UK Neonatal Collaborative Necrotising Enterocolitis Study: A Prospective Population-based Study Using the National Neonatal Research Database*. St Andrews: SHIP conference; August 2013.

Battersby C. *UK Neonatal Collaborative Necrotising enterocolitis study: Feeding Practices in Babies Born Less than 33 Weeks in England*. Windsor: 21st European Neonatal Workshop; 17 September 2013.

Battersby C. *UK Neonatal Collaborative Necrotising Enterocolitis Study*. Cambridge: East of England Neonatal Network meeting; September 2013.

Battersby C. *Enteral Feed Exposures in Babies Born Less than 32 Weeks' Gestation*. London: Neonatal Society Autumn meeting; November 2013.

Battersby C. *The UKNC-NEC Study: Interim Results*. Medway Maritime Hospital; July 2014.

Foster V. *'We Just Want to Give Something Back . . .' Altruism and Data-sharing in Neonatal Services*. British Sociological Association Medical Sociology Annual Conference. Leicester: University of Leicester; September 2012.

Ibrahim B, Statnikov E, Gray D, Modi N, Saxena S. *Linking Electronic Records to Create a Birth Cohort of Infants Admitted to Neonatal Units in England*. London: Neonatal Society Autumn Meeting; November 2014.

Modi N. *Improving the Quality of Routinely Collected Electronic Data*. London: Neonatal Nutrition Network meeting; September 2011.

Murray J, Bottle A, Sharland M, Modi N, Aylin P, Majeed A, Saxena S. *Changes in the Severity and Age of RSV Bronchiolitis Hospital Admission Among Infants in England: A Population-based Birth Cohort Study*. London: Neonatal Society Autumn Meeting; November 2011.

Murray J, Bottle A, Sharland M, Modi N, Aylin P, Majeed A, Saxena S. *The Burden of RSV Bronchiolitis Among Infants in England: a Cohort Study*. London: European Society for Paediatric Infectious Diseases; June 2011.

Murray J, Saxena S, Majeed A, Modi N, Bottle A, Aylin P. *Creating a Birth Cohort to Examine RSV Bronchiolitis Hospital Admission Rates Among Term and Preterm Infants in England*. RCPCH Annual Meeting; May 2012.

Rodgers K. *Data-sharing in Neonatal Services Annual Stakeholder Meeting for MCRN*. Manchester: University of Manchester; June 2013.

Santhakumaran S. *Evaluating Mortality Rates for Neonatal Units Using Multiple Membership Models*. London: International Society of Clinical Biostatistics; September 2012.

Santhakumaran S. *The Neonatal Survival Prediction Tool: A New Resource for Clinicians, Managers, and Parents*. London: NNAP/NDAU collaborators meeting; February 2013.

Santhakumaran S. *Survival of Preterm Infants Admitted to Neonatal Care in England: A Population-based Study Using NHS Electronic Clinical Records*. York: RCPCH meeting; June 2013.

Santhakumaran S. *The NEC Care Bundle: Statistical Findings and Outcomes*. UK Neonatal Collaborative Necrotising Enterocolitis Study. Cambridge: East of England Neonatal Network meeting; September 2013.

Santhakumaran S. *A Regional Care Bundle Approach to Increase Maternal Breast Milk Use in Preterm Infants: Outcomes of the East of England Network Quality Improvement Project*. London: Neonatal Society Autumn Meeting; November 2013.

Statnikov Y, Santhakumaran S, Manktelow B, Modi N. *Intensive Care Provided by Non-level 3 Neonatal Units in England*. London: Neonatal Society Autumn Meeting; November 2009.

Statnikov Y, Santhakumaran S, Manktelow B, Modi N. *Surveillance of Necrotising Enterocolitis in England*. London: Neonatal Society Summer Meeting; June 2010.

Statnikov Y, Wong HS, Gray DR, Santhakumaran S, Modi N. *Screening for Retinopathy of Prematurity in English Neonatal Units*. London: BAPM Annual meeting; September 2012.

Statnikov Y. *National Neonatal Research Database British Intestinal Failure Working Group Meeting*. London: November 2012.

Statnikov Y. *A UK Neonatal Collaborative Online Portal NNAP/NDAU Collaborators Meeting*. London: February 2013.

Statnikov Y, Modi N. *Establishing a National Neonatal Research Database from Operational NHS Electronic Records*. Edinburgh Scottish Informatics Programme; August 2013.

Wong H, Huertas-Ceballos A, Cowan FM, Modi N. *Comparison of Two Parent-completed Questionnaires for the Identification of Children at Risk for Autism Spectrum Disorder in the Preterm Population*. Newcastle: Annual Meeting of the European Society for Paediatric Research; 2011.

Wong HS, Huertas-Ceballos A, Cowan FM, Modi N. *Evaluation of Early Childhood Social-communication Difficulties in Children Born Preterm Using the Quantitative Checklist for Autism In Toddlers (Awarded Prize for Best Presentation by a Trainee)*. London: Neonatal Society Spring Meeting; April 2012.

Wong HS, Huertas-Ceballos A, Cowan FM, Modi N. *Sociodemographic and Neonatal Factors Associated with Early Childhood Social-communication Difficulties in Children Born Preterm*. Canterbury: Neonatal Society Summer Meeting; June 2012.

Wong HS, Huertas-Ceballos A, Cowan FM, Modi N. *Sociodemographic and Neonatal Factors Associated with Early Childhood Social-communication Difficulties in Children Born Preterm*. Istanbul: Fourth Congress of the European Academy of Pediatric Societies; October 2012.

Wong HS, Santhakumaran S, Cowan FM, Modi N. *Predictive Validity of Early Developmental Assessments in Identifying School-age Cognitive Deficits in Children Born Preterm or Very Low Birthweight: Systematic Review and Meta-analysis*. Barcelona: Fifth Congress of the European Academy of Paediatric Societies; October 2014.

Appendix 8 Higher degrees awarded relating to the Medicines for Neonates Programme

Murray JC. *The Clinical Burden of Respiratory Syncytial Virus Bronchiolitis Among Infants in the United Kingdom*. PhD thesis. London: Imperial College London; 2013.

Watson S. *Economic and Healthcare Related Determinants of Infant Health at Birth*. PhD thesis. Coventry: University of Warwick; 2015.

Wong H. *Neurodevelopmental Outcomes of Children Born Preterm: Analyses into the Validity of Data Collection and Outcome Reports*. PhD thesis. London: Imperial College London; 2016.

Battersby C. *The UK Neonatal Collaborative Necrotising Enterocolitis (NEC) Study: Testing the Utility of Operational Clinical Data to Conduct Population Surveillance, Develop an Evidence-based Case-definition and Identify Risk Factors Associated with NEC*. PhD thesis. London: Imperial College London; 2017.

Santhakumaran S. *Statistical Implications of Centralised Care for Estimating Neonatal Unit Mortality Rates*. PhD thesis. London: Imperial College London; 2017.

Appendix 9 Studies and organisations using the National Neonatal Research Database

British Association of Perinatal Medicine

Revision of neonatal level of care definitions.

Royal College of Paediatrics and Child Health

National Neonatal Audit Programme Annual Reports 2008, 2009, 2010, 2011, 2012, 2013; 2014 in preparation (www.rcpch.ac.uk/improving-child-health/quality-improvement-and-clinical-audit/national-neonatal-audit-programme-nnap).

Department of Health and Social Care

Atlas of Variation in Healthcare for Children and Young People, 2012, 2013 (www.chimat.org.uk/variation#cmoreport).

Reducing Perinatal Brain Injury (from 2016).

HM Government

Data for Parliamentary questions (2013, 2014).

Health and Social Care Information Centre

Admissions with neonatal jaundice (from 2015).

NHS England

Impact of Greater Manchester Perinatal Services Re-configuration (2014).

Patient Safety Domain (Full Term Admissions to Newborn Care, from 2014; Neonatal Umbilical Venous Catheter insertions, 2014).

National Neonatal Clinical Reference Group (neonatal admissions annual mortality reports, from 2014).

Public Health England

Neonatal specialised activity metrics (from 2014).

Smoking in pregnancy (2015).

London Operational Delivery Network

Quarterly activity, clinical outcomes and benchmarking analyses for the London Neonatal Networks (from 2014).

World Health Organization

Data to inform the Every Newborn Action Plan (from 2014).

National and International Benchmarking and Quality Improvement Programmes

International Network for Evaluation of Outcomes of Neonates (iNeo): a quality improvement project based on collaborative comparisons of population-based international health care for neonates led by the University of Toronto (http://ineonetwork.org) (from 2013).

East of England Neonatal Networks regional care bundle to improve maternal breast milk use in preterm infants (www.unicef.org.uk/BabyFriendly/News-and-Research/Research/Neonatal/Impact-of-a-regional-care-bundle-on-maternal-breast-milk-use-in-preterm-infants) (2013–14).

Each Baby Counts: a national quality improvement programme led by the Royal College of Obstetricians and Gynaecologists to reduce the number of babies who die or are left severely disabled as a result of incidents occurring during term labour; cross-validation data will be provided from the NNRD (www.rcog.org.uk/eachbabycounts) (from 2014).

eNewborn: a pan-European preterm benchmarking platform led from Saint-Pierre University Hospital, Brussels (from 2015).

Medicines for Neonates Applied Health Research Programme

Lead institutions: Imperial College London, University of Manchester and NPEU (2009–15).

Downs in Neonates

Lead institutions: Queen Mary University of London and Hinchingbrooke Hospital NHS Trust (2012–15).

Neonatal Economics, Staffing and Clinical Outcomes Project

Lead institutions: University College London, University of Warwick, Imperial College London and University of Leicester (2012–15).

The right cot, at the right time, at the right place: providing a national demand/capacity model for neonatal care in England

Lead institutions: Peninsula Collaboration for Health Operational Research and Development, NIHR CLAHRC South West Peninsula and University of Exeter (2015–17).

Modelling care pathways in neonatal care: costs and consequences for the future

Lead institution: University of Leicester (2014–17).

PREVenting infection using antibiotic impregnated long lines (PreVail)

Lead institutions: Public Health England and Bradford Teaching Hospitals NHS Foundation Trust (2014–18).

Gentamicin, genetic variation and deafness in preterm children: the Mitogent Study

Lead institution: University College London (2015).

Medical Research Council

Preterm birth and neuropsychiatric genetic risk.

Lead institution: University of Cardiff (2016).

Clinician Science Award.

Lead institution: Imperial College London (2016).

National Institute for Health Research

Optimising service provision for preterm babies.

Lead institution: University of Leicester (2017).

Appendix 10 List of participating NHS trusts in England, and Neonatal Clinical Leads

1. Airedale General Hospital, Airedale NHS Trust

Dr Matthew Babirecki

2. Alexandra Hospital, Worcestershire Acute Hospitals NHS Trust

Dr Liza Harry

3. Arrowe Park Hospital, Wirral University Teaching Hospital NHS Foundation Trust

Dr Oliver Rackham

4. Barnet Hospital, Royal Free London NHS Foundation Trust

Dr Tim Wickham

5. Barnsley District General Hospital, Barnsley Hospital NHS Foundation Trust

Dr Sanaa Hamdan

6. Basildon Hospital, Basildon and Thurrock University Hospitals NHS Trust

Dr Aashish Gupta

7. Basingstoke and North Hampshire Hospital, Hampshire Hospitals NHS Foundation Trust

Dr Ruth Wigfield

8. Bassetlaw District General Hospital, Doncaster and Bassetlaw Hospitals NHS Foundation Trust

Dr L M Wong

9. Bedford Hospital, Bedford Hospital NHS Trust

Dr Anita Mittal

10. Birmingham City Hospital, Sandwell and West Birmingham Hospitals NHS Trust

Dr Julie Nycyk

11. Birmingham Heartlands Hospital, Heart of England NHS Foundation Trust

Dr Phil Simmons

12. Birmingham Women's Hospital, Birmingham Women's NHS Foundation Trust

Dr Alison Bedford-Russell

13. Bradford Royal Infirmary, Bradford Teaching Hospitals NHS Foundation Trust

Dr Sunita Seal

14. Broomfield Hospital, Chelmsford, Mid Essex Hospital Services NHS Trust

Dr Ahmed Hassan

15. Calderdale Royal Hospital, Calderdale and Huddersfield NHS Foundation Trust

Dr Karin Schwarz

16. Chelsea and Westminster Hospital, Chelsea and Westminster Hospital NHS Foundation Trust

Dr Mark Thomas

17. Chesterfield & North Derbyshire Hospital, Chesterfield Royal Hospital NHS Foundation Trust

Dr Aiwyne Foo

18. Colchester General Hospital, Colchester Hospital NHS Foundation Trust

Dr Karen Moss

19. Conquest Hospital, East Sussex Hospitals NHS Trust

Dr Jayaram Pai

20. Countess of Chester Hospital, Countess of Chester Hospital NHS Foundation Trust

Dr Stephen Brearey

21. Croydon University Hospital, Croydon Health Services

Dr John Chang

22. Cumberland Infirmary, North Cumbria University Hospitals NHS Trust

Dr Khairy Gad

23. Darent Valley Hospital, Dartford and Gravesham NHS Trust

Dr Abdul Hasib

24. Darlington Memorial Hospital, County Durham and Darlington NHS Foundation Trust

Dr Mehdi Garbash

25. Derriford Hospital, Plymouth Hospitals NHS Trust

Dr Nicci Maxwell

26. Dewsbury and District Hospital, Mid Yorkshire Hospitals NHS Trust

Dr David Gibson

27. Diana Princess of Wales Hospital, Northern Lincolnshire and Goole Hospitals NHS Foundation Trust

Dr Pauline Adiotomre

28. Doncaster Royal Infirmary, Doncaster and Bassetlaw Hospitals NHS Foundation Trust

Dr Jamal S Ahmed

29. Dorset County Hospital, Dorset County Hospital NHS Foundation Trust

Dr Abby Deketelaere

30. Ealing Hospital, London North West Health Care NHS Trust

Dr Ramnik Mathur

31. East Surrey Hospital, Surrey and Sussex Healthcare NHS Trust

Dr K Abdul Khader

32. Epsom General Hospital, Epsom and St Helier University Hospitals NHS Trust

Dr Ruth Shephard

33. Frimley Park Hospital, Frimley Park Hospital NHS Foundation Trust

Dr Abdus Mallik

34. Furness General Hospital, University Hospitals of Morecambe Bay NHS Trust

Dr Belal Abuzgia

35. George Eliot Hospital, George Eliot Hospital NHS Trust

Dr Mukta Jain

36. Gloucester Royal Hospital, Gloucestershire Hospitals NHS Foundation Trust

Dr Simon Pirie

37. Good Hope Hospital, Heart of England NHS Foundation Trust

Dr Phil Simmons

38. Great Western Hospital, Great Western Hospitals NHS Foundation Trust

Dr Stanley Zengeya

39. Guy's & St Thomas' Hospital, Guy's and St Thomas' NHS Foundation Trust

Dr Timothy Watts

40. Harrogate District Hospital, Harrogate and District NHS Foundation Trust

Dr C Jampala

41. Hereford County Hospital, Wye Valley NHS Trust

Dr Cath Seagrave

42. Hillingdon Hospital, The Hillingdon Hospital NHS Trust

Dr Michele Cruwys

43. Hinchingbrooke Hospital, Cambridgeshire Community Services NHS Trust

Dr Hilary Dixon

44. Homerton Hospital, Homerton University Hospital NHS Foundation Trust

Dr Narendra Aladangady

45. Hull Royal Infirmary, Hull and East Yorkshire Hospitals NHS Trust

Dr Peter Pairaudeau

46. Ipswich Hospital, Ipswich Hospital NHS Trust

Dr Matthew James

47. James Cook University Hospital, South Tees Hospitals NHS Trust

Dr M Lal

48. James Paget Hospital, James Paget University Hospitals NHS Foundation Trust

Dr Ambadkar

49. Kettering General Hospital, Kettering General Hospital NHS Foundation Trust

Dr Patti Rao

50. King George Hospital, Barking, Havering and Redbridge University Hospitals NHS Trust

Dr Khalid Mannan

51. King's College Hospital, King's College Hospital NHS Foundation Trust

Dr Ann Hickey

52. King's Mill Hospital, Sherwood Forest Hospitals NHS Foundation Trust

Dr Vibert Noble

53. Kingston Hospital, Kingston Hospital NHS Trust

Dr Nader Elgharably

54. Lancashire Women and Newborn Centre, East Lancashire Hospitals NHS Trust

Dr Meera Lama

55. Leeds Neonatal Service, Leeds Teaching Hospitals NHS Trust

Dr Lawrence Miall

56. Leicester General Hospital, University Hospitals of Leicester NHS Trust

Dr Jonathan Cusack

57. Leicester Royal Infirmary, University Hospitals of Leicester NHS Trust

Dr Venkatesh Kairamkonda

58. Leighton Hospital, Mid Cheshire Hospitals NHS Foundation Trust

Dr Jayachandran

59. Lincoln County Hospital, United Lincolnshire Hospitals NHS Trust

Dr Kollipara

60. Lister Hospital, East and North Hertfordshire NHS Trust

Dr J Kefas

61. Liverpool Women's Hospital, Liverpool Women's NHS Foundation Trust

Dr Bill Yoxall

62. Luton and Dunstable Hospital, Luton and Dunstable Hospital NHS Foundation Trust

Dr Sarah Skinner

63. Macclesfield District General Hospital, East Cheshire NHS Trust

Dr Gail Whitehead

64. Manor Hospital, Walsall Hospitals NHS Trust

Dr Bashir Jan Muhammad

65. Medway Maritime Hospital, Medway NHS Foundation Trust

Dr Aung Soe

66. Milton Keynes General Hospital, Milton Keynes Hospital NHS Foundation Trust

Dr I Misra

67. New Cross Hospital, The Royal Wolverhampton Hospitals NHS Trust

Dr Tilly Pillay

68. Newham General Hospital, Barts Health

Dr Imdad Ali

69. Norfolk & Norwich University Hospital, Norfolk and Norwich University Hospitals NHS Foundation Trust

Dr Mark Dyke

70. North Devon District Hospital, North Devon Healthcare NHS Trust

Dr Michael Selter

71. North Manchester General Hospital, The Pennine Acute Hospitals NHS Trust

Dr Nagesh Panasa

72. North Middlesex University Hospital, North Middlesex University Hospital NHS Trust

Dr Lesley Alsford

73. Northampton General Hospital, Northampton General Hospital NHS Trust

Dr Subodh Gupta

74. Northwick Park Hospital, London North West Health Care NHS Trust

Dr Richard Nicholl

75. Nottingham City Hospital, Nottingham University Hospitals NHS Trust

Dr Steven Wardle

76. Ormskirk District General Hospital, Southport and Ormskirk Hospital NHS Trust

Dr Tim McBride

77. Oxford University Hospitals, Horton Hospital, Oxford University Hospitals NHS Trust

Dr Naveen Shettihalli

78. Oxford University Hospitals, John Radcliffe Hospital, Oxford University Hospitals NHS Trust

Dr Eleri Adams

79. Peterborough City Hospital, Peterborough and Stamford NHS Foundation Trust

Dr Seif Babiker

80. Pilgrim Hospital, United Lincolnshire Hospitals NHS Trust

Dr Margaret Crawford

81. Pinderfields General Hospital, Mid Yorkshire Hospitals NHS Trust

Dr David Gibson

82. Poole General Hospital, Poole Hospital NHS Foundation Trust

Dr Minesh Khashu

83. Princess Alexandra Hospital, The Princess Alexandra Hospital NHS Trust

Dr Caitlin Toh

84. Princess Anne Hospital, Southampton University Hospitals NHS Trust

Dr M Hall

85. Princess Royal Hospital, Brighton and Sussex University Hospitals NHS Trust

Dr P Amess

86. Princess Royal University Hospital, King's College Hospital NHS Foundation Trust

Dr Elizabeth Sleight

87. Queen Alexandra Hospital, Portsmouth Hospitals NHS Trust

Dr Charlotte Groves

88. Queen Charlotte's Hospital, Imperial College Healthcare NHS Trust

Dr Sunit Godambe

89. Queen Elizabeth Hospital, Gateshead, Gateshead Health NHS Foundation Trust

Dr Dennis Bosman

90. Queen Elizabeth Hospital, The Queen Elizabeth Hospital King's Lynn NHS Trust

Dr Susan Rubin

91. Queen Elizabeth Hospital, Woolwich, Lewisham and Greenwich NHS Trust

Dr Banjoko

92. Queen Elizabeth the Queen Mother Hospital, East Kent Hospitals University NHS Trust

Dr Rfidah

93. Queen's Hospital, Burton on Trent, Burton Hospitals NHS Foundation Trust

Dr A Manzoor

94. Queen's Hospital, Romford, Barking, Havering and Redbridge University Hospitals NHS Trust

Dr Khalid Mannan

95. Rosie Maternity Hospital, Addenbrookes Cambridge University Hospitals NHS Foundation Trust

Dr Angela D'Amore

96. Rotherham District General Hospital, Rotherham NHS Foundation Trust

Dr MacFarlane

97. Royal Albert Edward Infirmary, Wrightington, Wigan and Leigh NHS Foundation Trust

Dr Vibha Sharma

98. Royal Berkshire Hospital, Royal Berkshire NHS Foundation Trust

Dr Peter De Halpert

99. Royal Bolton Hospital, Royal Bolton Hospital NHS Foundation Trust

Dr Paul Settle

100. Royal Cornwall Hospital, Royal Cornwall Hospitals NHS Trust

Dr Paul Munyard

101. Royal Derby Hospital, Derby Teaching Hospitals NHS Foundation Trust

Dr Gitika Joshi

102. Royal Devon & Exeter Hospital, Royal Devon and Exeter NHS Foundation Trust

Dr Vaughan Lewis

103. Royal Hampshire County Hospital, Hampshire Hospitals NHS Foundation Trust

Dr D Schapira

104. Royal Lancaster Infirmary, University Hospitals of Morecambe Bay NHS Trust

Dr Joanne Fedee

105. Royal Oldham Hospital, The Pennine Acute Hospitals NHS Trust

Dr Natasha Maddock

106. Royal Preston Hospital, Lancashire Teaching Hospitals NHS Foundation Trust

Dr Richa Gupta

107. Royal Shrewsbury Hospital, Shrewsbury and Telford Hospital NHS Trust

Dr Deshpande

108. Royal Surrey County Hospital, The Royal Surrey County Hospital NHS Trust

Dr Charles Godden

109. Royal Sussex County Hospital, Brighton and Sussex University Hospitals NHS Trust

Dr P Amess

110. Royal United Hospital, Royal United Hospital Bath NHS Trust

Dr Stephen Jones

111. Royal Victoria Infirmary, Newcastle Upon Tyne Hospitals Foundation Trust

Dr Alan Fenton

112. Russells Hall Hospital, Dudley Group of Hospitals NHS Foundation Trust

Dr Mahadevan

113. Salisbury District Hospital, Salisbury NHS Foundation Trust

Dr Nick Brown

114. Scarborough General Hospital, York Teaching Hospitals NHS Foundation Trust

Dr Kirsten Mack

115. Scunthorpe General Hospital, Northern Lincolnshire and Goole Hospitals NHS Foundation Trust

Dr Pauline Adiotomre

116. South Tyneside District Hospital, South Tyneside NHS Foundation Trust

Dr Rob Bolton

117. Southend Hospital, Southend University Hospital NHS Foundation Trust

Dr A Khan

118. Southmead Hospital, North Bristol NHS Trust

Dr Paul Mannix

119. St George's Hospital, St George's Healthcare NHS Trust

Dr Charlotte Huddy

120. St Helier Hospital, Epsom and St Helier University Hospitals NHS Trust

Dr Salim Yasin

121. St Mary's Hospital, Isle of Wight Healthcare NHS Trust

Dr Sian Butterworth

122. St Mary's Hospital, London, Imperial College Healthcare NHS Trust

Dr Sunit Godambe

123. St Mary's Hospital, Manchester, Central Manchester University Hospitals NHS Foundation Trust

Dr Ngozi Edi-Osagie

124. St Michael's Hospital, University Hospitals Bristol NHS Foundation Trust

Dr David Harding

125. St Peter's Hospital, Ashford and St Peter's Hospitals NHS Trust

Dr Peter Reynolds

126. St Richard's Hospital, Western Sussex Hospitals NHS Trust

Dr Nick Brennan

127. Stepping Hill Hospital, Stockport NHS Foundation Trust

Dr Carrie Heal

128. Stoke Mandeville Hospital, Buckinghamshire Hospitals NHS Trust

Dr Sanjay Salgia

129. Sunderland Royal Hospital, City Hospitals Sunderland NHS Foundation Trust

Dr Majd Abu-Harb

130. Tameside General Hospital, Tameside Hospital NHS Foundation Trust

Dr Jacqeline Birch

131. Taunton and Somerset Hospital, Taunton and Somerset NHS Foundation Trust

Dr Chris Knight

132. The Jessop Wing, Sheffield, Sheffield Teaching Hospitals NHS Foundation Trust

Dr Simon Clark

133. The Royal Free Hospital, Royal Free London NHS Foundation Trust

Dr Vivienne Van Sommen

134. The Royal London Hospital, Constance Green, Barts Health

Dr Nandiran Ratnavel

135. Torbay Hospital, South Devon Healthcare NHS Foundation Trust

Dr Mala Raman

136. Tunbridge Wells Hospital, Maidstone and Tunbridge Wells NHS Trust

Dr Hamudi Kisat

137. University College Hospital, University College London Hospitals NHS Foundation Trust

Dr Sara Watkin

138. University Hospital Coventry, University Hospitals Coventry and Warwickshire NHS Trust

Dr Kate Blake

139. University Hospital Lewisham, Lewisham and Greenwich NHS Trust

Dr Jauro Kuna

140. University Hospital of North Durham, County Durham and Darlington NHS Foundation Trust

Dr Mehdi Garbash

141. University Hospital of North Staffordshire, University Hospitals of North Midlands NHS Trust

Dr Kate Palmer

142. University Hospital of North Tees, North Tees and Hartlepool NHS Foundation Trust

Dr B Reichert

143. University Hospital of South Manchester, University Hospital of South Manchester NHS Foundation Trust

Dr Gopi Vemuri

144. Victoria Hospital, Blackpool, Fylde and Wyre Hospitals NHS Foundation Trust

Dr Chris Rawlingson

145. Wansbeck General Hospital, Northumbria Healthcare NHS Trust

Dr Alan Fenton

146. Warrington Hospital, Warrington and Halton Hospitals NHS Foundation Trust

Dr Delyth Webb

147. Warwick Hospital, South Warwickshire General Hospitals NHS Trust

Dr Semeer Kallaroth

148. Watford General Hospital, West Hertfordshire Hospitals NHS Trust

Dr Sankara Narayanan

149. West Cumberland Hospital, North Cumbria University Hospitals NHS Trust

Dr Mithun Urs

150. West Middlesex University Hospital, West Middlesex University Hospital NHS Trust

Dr Elizabeth Eyre

151. West Suffolk Hospital, West Suffolk Hospital NHS Trust

Dr Ian Evans

152. Wexham Park Hospital, Heatherwood and Wexham Park Hospitals NHS Foundation Trust

Dr Rekha Sanghavi

153. Whipps Cross University Hospital, Barts Health

Dr Caroline Sullivan

154. Whiston Hospital, St Helens and Knowsley Teaching Hospitals NHS Trust

Dr Laweh Amegavie

155. Whittington Hospital, The Whittington Hospital NHS Trust

Dr Wynne Leith

156. William Harvey Hospital, East Kent Hospitals University NHS Trust

Dr Vimal Vasu

157. Worcestershire Royal Hospital, Worcestershire Acute Hospitals NHS Trust

Dr Andrew Gallagher

158. Worthing Hospital, Western Sussex Hospitals NHS Trust

Dr Katia Vamvakiti

159. Yeovil District Hospital, Yeovil District Hospital NHS Foundation Trust

Dr Megan Eaton

160. York District Hospital, York Teaching Hospitals NHS Foundation Trust

Dr Guy Millman

Appendix 11 Medicines for Neonates Steering Committee

Professor Michael Goldacre (chairperson): independent member.

Professor Andrew Wilkinson (deputy chairperson): independent member.

Mrs Jane Abbott (later Ms Zoe Chivers): investigator.

Professor Deborah Ashby: investigator.

Professor Peter Brocklehurst: investigator.

Professor Kate Costeloe: investigator.

Professor Elizabeth Draper: investigator.

Mrs Jacquie Kemp: investigator.

Professor Azeem Majeed: investigator.

Professor Neena Modi: lead investigator.

Professor Stavros Petrou: investigator.

Professor Alys Young: investigator.

Appendix 12 Other funding sources contributing to this research

Unrestricted donations from

Abbott Laboratories (Maidenhead, UK): £35,000.

Nutricia Research Foundation (Schiphol, the Netherlands): £15,000.

GE Healthcare (Amersham, UK): £1000.

Grant to support the use of routinely collected, standardised, electronic clinical data for audit, management and multidisciplinary feedback in neonatal medicine.

Department of Health and Social Care: £135,494.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library